A Reliable and Testable Alternative to Long-run Restrictions in Structural VAR models[*]

Alain Guay[†]

Florian Pelgrin[‡]

This version: October 31, 2015

Do not quote without permission

Abstract

This paper proposes a new identification method for structural VAR models based on frequency interval restrictions. In so doing, we use the methodology of Carrasco and Florens (2000), the generalization of GMM for a continuum of moment conditions, in the case of the asymptotic least squares method, and we thus propose a new estimator, namely the continuum asymptotic least squares estimator (C-ALS). Our new methodology allows to obtain consistent estimates of impulse responses and reliable confidence intervals in contrast to usual long-run restrictions. Moreover the imposed restrictions can be tested formally and it offers a data-driven procedure that can assess formally the relevance of the imposed identifying restrictions. Finally, we provide some new results using extensive Monte Carlo simulations and an application regarding the hours-productivity debate.

Keywords: SVARs, Frequency domain, Continuum of moments.

[*]The first author gratefully acknowledges financial support of Fonds pour la Formation de Chercheurs et l'aide à la Recherche (FCAR). We would also like to thank Jean-Pierre Florens, René Garcia, John Galbraith, Stéphane Gregoir, Patrick Fve, Benoît Perron, Simon Van Norden, Roberto Pancrazi, Eric Renault and Eva Ortega for comments on an earlier draft of the paper.

[†]Université du Québec à Montréal, CIRPÉE and CIREQ, *e-mail*: `guay.alain@uqam.ca`.

[‡]EDHEC Business School, *e-mail*: `florian.pelgrin@edhec.edu`.

1 Introduction

Following the seminal works of Sims (1980a, 1980b), economists have been widely concerned with studying sources of economic fluctuations in the context of vector autoregressive models (VARs). Indeed, VAR analysis makes possible to study the average response of the variables of interest to a given one-time structural shock, to provide historical decompositions that capture the cumulative effect of each structural shock to each variable dynamics, to construct forecast error variance decompositions to assess the average contribution of each structural shock to the variability of data, and to generate forecast scenarios conditional on some sequences of future structural shocks.[1] While the ability of VAR models as descriptive and/or forecasting tools is well established, structural interpretation of VAR models is still subject to effervescent debates, and for good reasons. Moving from atheoretical VAR models to structural VAR models requires identifying assumptions that rest on economic theory (among others)—VAR results cannot be interpreted independently of a more structural macroeconomic model (Cooley and Leroy, 1985; Bernanke, 1986).

Among the different methodologies that impose *sufficient* identifying restrictions, Blanchard and Quah (1989), and Gali (1999) exploit long-run (permanent) effects of shocks with the appeal that they might be more consistent with business cycles models.[2] However, the ability of long-run structural VARs to uncover the dynamic response of macroeconomic variables to structural shocks has been questioned, and for good reasons.[3] Notably, as shown by the long-lasting debate regarding the sources of macroeconomic fluctuations, one key issue of a long-run identification scheme in the structural vector autoregressive (hereafter, VAR) literature is the estimation of the long-run multipliers of the reduced-form specification. A first line of criticisms rests on several points addressed by Sims (1971, 1972). Notably the sum of the regression coefficients might be distorted when there is a specification error, and long-run restrictions do involve a reliable estimate of the sum of the coefficients in distributed lag regressions (e.g., vector autoregressions). All in all, this sum is critical to estimate even if the individual coefficients are reasonably precisely estimated. Indeed the true VAR of the data has infinite lags whereas the econometrician can only use a finite number of lags in the estimation procedure. In other words, since the reduced form VAR model is characterized *a priori* by an infinite parametrization, it can be shown that convergence of the estimates of the model's parameters is not

[1] See Kilian (2013).

[2] Other identifying schemes include short-run, signs restrictions, medium-term identification (Uhlig, 2004) or restrictions in the frequency domain (Wen, 2001).

[3] As pointed out by Kilian (2013), four main concerns have been addressed about the reliability of long-run restrictions. First, long-run structural VAR models require some accurate estimates of the impulse responses at the infinite horizon. Second, responses in long-run VAR models are identified only up to their sign. Third, identification is often achieved with the use of I(0) but (quite) persistent variables that might distort some statistical properties. Fourth, results might be sensitive to the level or first-difference specification of some variables.

sufficient to guarantee the convergence of some functions of those parameters (Sims, 1971, 1972, Pötscher, 2002)—pointwise convergence does not imply (locally) uniform convergence. Consequently, a given norm, and especially the L_2-norm of least squares methods, between the true parameters and the estimated parameters of the VAR can go to zero while the sums of the coefficients converge to different limits. This results from the fact that functions of a lag distribution (e.g., the sum of the coefficients) are in general discontinuous with respect to the metric implied by least-squares estimation. This so-called unreliability problem is well-known since the seminal papers of Sims (1971, 1972) and the contributions of Faust (1996), Faust and Lepper (1997), Pagan and Robertson (1998), and Pötscher (2002).

In this perspective, Faust and Lepper (1997) make important related criticisms of the long-run identification strategy. More specifically, unreliable long-run effects of shocks in finite samples are transferred on estimates of other model parameters through the long run identification scheme and thus any test of the null hypothesis that the kth coefficient of an autoregressive polynomial in the SVAR equals zero is not consistent, i.e., the test has significance level greater than or equal to maximum power. In particular, as is shown in Faust (1996), this result stems from the null measure of the spectrum at zero frequency implied by long-run restrictions. Say differently, the convergence to normality of the point estimate is not uniform. Since the unreliability of the long run effect estimator is transferred to the estimator of the dynamic multipliers of the structural shocks, one cannot compute asymptotically correct confidence intervals for impulse responses functions.

In light of the difficulty to implement long-run restrictions paired with the widespread support for such restrictions, we offer an alternative identification and estimation approach that rests on sufficient identifying restrictions not only at the zero-frequency but also on its neighborhood, or more generally on a given frequency interval. This allows us to provide an overidentification test of the imposed identifying restrictions and to assess on which frequency intervals these restrictions might be satisfied. In so doing, we make use of two ingredients. The first one, which has been suggested by Faust and Lepper (1997), Faust (1998), and Pötscher (2002), is to impose restrictions on the effect of these shocks at non-zero frequencies and not only at the zero-frequency. Thus the common long-run restrictions at frequency zero, $\omega = 0$ are replaced by restrictions on $\omega = 0$ and its neighborhood, say $\omega \in (\underline{\omega}, \overline{\omega})$. For example, shocks might have no impact to cycles lasting more than 80 quarters. By considering restrictions on the long run and close to the long run, we preserve the appeal of the standard Blanchard-Quah method. The resulting set of restrictions now has a measure strictly greater than zero in the frequency domain. Consequently, an accurate approximation from the point of view of least-squares fit involves an accurate approximation over the area given by the frequency band of interest for identification purposes. Therefore the distance between the true vector autoregressive process and its approximation cannot go to zero while the distance over the interval does not go to zero. Inference for this kind of restrictions is then reliable. Following the terminology of Pötscher (2002), the problem is no longer ill-posed.

From an economic point of view, implementing such restrictions is quite intuitive. On the one hand, it means that identification is achieved by assuming that the effects of some shocks die out after a finite time period. On the other hand, VAR analysis generally provides evidence that the contribution of shocks is the most important at intermediate horizons of about two to ten years, meaning that there is some information content at intermediate to nearly long-run frequencies. Moreover, SVAR evidence also outlines that almost all shocks does not explain a given variable (using the variance decomposition) and that the effects of these shocks (using impulse responses functions) are not statistically different from zero in the medium-to-long term. Therefore there is a room to identify some (permanent) structural shocks by minimizing the contribution of other shocks while imposing restrictions on some nearly long-run frequencies. This is precisely what our approach is doing. Say differently, there is no reason to think that some given (permanent) shocks might be exactly identified by short-run, medium-term, or long-run identification schemes: it might come close to but not exactly to the (true) unobserved shocks. Using information close to the long-run component might help consistently identifying such shocks. Finally, this approach is similar in spirit to the medium term approach developed by Faust (1998), Uhlig (2004), Francis et al. (2005, 2014), Francis and DiCecio (2010). For instance, Francis et al. (2014) identify a technology shock that maximizes the share of the forecast-error variance in productivity growth and hours at various frequencies (i.e., using different frequency windows) in a bivariate VAR model. At the same time, it is worth noting that our approach might be used to imposed both medium-term and nearly long-run restrictions. Indeed, one might consider restrictions on a frequency interval that does not contain the zero-frequency. Such identifying restrictions are consistent with some recent papers that outline the contribution of the medium-term component of fluctuations (e.g., Comin and Gertler, 2006; Pancrazi, 2015).

Taking such frequencies restrictions, the second key ingredient is to rewrite the identification problem as an asymptotic least squares problem in the frequency domain.[4] In particular, this leads to a continuum of moment conditions. Indeed, identifying restrictions over the frequency window constitute the set of moment conditions, namely a system of nonlinear equations that depends on the structural coefficients and the (first-step) estimator of the reduced form VAR model. Solving this problem thus requires using the

[4]For the asymptotic least squares approach, see Gourieroux and Monfort (1995), and Gourieroux et al. (1985).

methodology proposed by Carrasco and Florens (2000), and Carrasco et al. (2007). In this respect, we propose a new efficient estimator (C-ALS estimator) that exploits the seminal work of Carrasco and Florens (2000) in the case of the asymptotic least squares method (Gourieroux et al., 1985; Gourieroux and Monfort, 1995) with a continuum of restrictions.[5] One major advantage of this approach is that it allows for selecting the frequency window (i.e., the reliability of the imposed restrictions), thereby permitting us to conduct a data-driven procedure in order to assess frequency intervals on which the imposed restrictions might be satisfied. This sharply contrasts with existing literature and this provides, to the best of our knowledge, the first method in which identifying restrictions are testable, especially in the case of an original just-identified VAR specification.[6] Note that our approach does not only restrict to structural VAR models but can be applied in more general frameworks that involve an asymptotic least squares problem with a continuum of constraints, as for instance any generalized spectral estimation problem of Euler-based equations in a structural model (e.g., Berkowitz, 2001).

A significant literature aims at circumventing the unreliability problem and proposing new identification schemes or estimation approaches in the frequency domain. Christiano, Eichenbaum and Vigfusson (2006a, 2006b) propose to combine a non-parametric estimation of the long-run variance-covariance matrix and a parametric estimation of the reduced-form VAR. The consistency of their approach has been discussed in Chari et al. (2005, 2008) and Mertens (2012). Uhlig (2004) proposes using the k-step ahead forecast revision at some fixed finite horizon.[7] As pointed out by Francis et al. (2014), this choice may significantly distort the results and they thus propose deriving a Max Share approach that utilizes a maximization routine for horizons up to and including k. However, no testing framework is provided in both papers. Finally, instead of first estimating an autoregressive reduced-form VAR model and then obtaining the corresponding moving average representation, Dupor and Kiefer (2008) propose estimating directly the multivariate moving average representation using the local projections method of Jorda (2005).[8]

The rest of the paper is organized as follows. Section 2 reviews long-run identifying restrictions. Then we discuss the unreliability problem and propose an alternative identification scheme. Section 3 presents

[5]The implementation is available as a package for Matlab. Therefore it is hoped that our methodology for structural VAR identification as well as the availability of the associated programs will allow researchers to apply our results in a large number of situations.

[6]Overidentification tests have been proposed when the number of restrictions is greater than required in the identification procedure (see Bernanke and Mihov, 1998).

[7]Uhlig (2004) also presents a method based on the cumulative effect of shocks.

[8]Kilian and Kim (2011) have raised some issues in the Jorda's projection method, especially regarding the existence of biased impulse-response functions estimates.

the (optimal) C-ALS estimator using the methodology proposed by Carrasco and Florens (2000). Section 4 provides two applications of our methodology, namely a bivariate structural VAR and the identification of a single structural shock in the presence of a m-variate VAR specification. Section 5 presents some Monte-Carlo simulations whereas Section 6 proposes an application regarding the technology-hours debate in a bivariate structural VAR specification. The last section contains concluding comments and future extensions. Proofs are gathered in Appendix.

2 Identification and unreliability of long-run structural VAR models

In this section, we first introduce preliminary notation and provide an overview of the long-run identification scheme of structural VAR models. Then we discuss the fact that long-run effects of shocks are generally unreliable in finite samples and thus the nature of the (statistical) problem, namely that an accurate approximation of the true data generating process from the point of view of least-squares fit does not imply an accurate approximation of long-run effects. This then allows us to propose a new identification strategy in the frequency domain.

2.1 Long-run identification scheme

It is assumed that a K-dimensional multiple time series $X_1, X_2, \cdots, X_T$ with $X_t = (X_{1t}, \cdots, X_{Kt})'$ is available and that these variables are stationary in first difference.[9] Accordingly, $\Delta X_t = (\Delta X_{1t}, \cdots, \Delta X_{Kt})'$ is a $(K \times 1)$ random vector and $\Delta X = (\Delta X_1, \cdots, \Delta X_T)$ is a $(K \times T)$ random matrix. We have a sample of size T for each of the K variables for the same sample period. To simplify the notation, presample values for each variable are assumed to be available. Furthermore all deterministic regressors have been suppressed for notational convenience. ΔX_t is generated by a stationary, stable, reduced-form $VAR(p)$ process:

$$\Delta X_t = \Phi_1 \Delta X_{t-1} + \cdots \Phi_p \Delta X_{t-p} + u_t \tag{2.1}$$

where the Φ_i are fixed $(K \times K)$ coefficient matrices, $u_t = (u_{1t}, \cdots, u_{Kt})'$ is a K-dimensional white noise or innovation process, that is, $E(u_t) = 0_{K\times 1}$, $E(u_t u_t') = \Sigma_u$ is nonsingular and $E(u_t u_s') = 0_{K\times K}$ for $s \neq t$.

The corresponding reduced-form vector moving average representation is defined to be:

$$\Delta X_t = \sum_{s=0}^{\infty} C_s u_{t-s} = C(L) u_t \tag{2.2}$$

[9] We rule out the case of VAR models with integrated and cointegrated variables.

where $C(L) = \sum_{s=0}^{\infty} C_s L^s$, L is the lag operator, and $C_0 = I_K$ (with I_K the identity matrix of order K), $C_s = \sum_{j=1}^{s} C_{s-j}\Phi_j$, and $C(1) = \sum_{s=1}^{\infty} C_s = [I_K - \Phi_1 - \cdots \Phi_p]^{-1}$.

On the other hand, the structural VAR model can be written as:

$$\Delta X_t = \sum_{s=0}^{\infty} A_s \epsilon_{t-s} = A(L)\epsilon_t \tag{2.3}$$

where $A(L) = \sum_{s=0}^{\infty} A_s L^s$, $A_0 = A(0)$, $A(1) = \sum_{s=0}^{\infty} A_s$, and ϵ_t is a random $K \times 1$ vector of structural shocks with $E(\epsilon_t) = 0$ and $E(\epsilon_t \epsilon_t') = \Gamma_\epsilon$. A common identification assumption is $\Gamma_\epsilon = I_K$.

Taking equations (2.2) and (2.3), the error terms of the reduced-form model are related to the structural shocks as follows:[10]

$$u_t = A(0)\epsilon_t, \tag{2.4}$$

and thus

$$C(0)\Sigma_u C(0)' \equiv \Sigma_u = A(0)A(0)'. \tag{2.5}$$

Then the central question is how to recover the elements of $A(0)$ from consistent estimates of the reduced-form parameters. Following Blanchard and Quah (1989), one idea is to impose restrictions on long-run responses of variables to shocks (in the presence of unit roots in some variables of the VAR specification but not in others). For instance, most economists agree that demand shocks are neutral in the long run, whereas productivity shocks are not. In so doing, using the long-run variance-covariance matrix of (2.5) and the long-run covariance matrix of the structural form, one obtains

$$C(1)\Sigma_u C(1)' = A(1)A(1)'. \tag{2.6}$$

The key observation is that Σ_u and $C(1) = [I_K - \Phi_1 - \cdots \Phi_p]^{-1}$ can be estimated from data. Therefore if one puts enough restrictions on $A(1)$, the remaining elements of $A(1)$ can be pinned down using numerical procedures. More specifically, since the left-hand side expression represents a variance-covariance matrix, only $K(K-1)/2$ restrictions on $A(1)$ are required to satisfy the order condition in the case of exact identification.[11] The matrix $A(0)$ is then recovered by the relationship: $C(1)A(0) = A(1)$.

[10]For a more general presentation, see Amisano and Giannini (1997), and Kilian (2013).

[11]If exclusion restrictions on A(1) are recursive, then a standard Choleski decomposition might be applied to $C(1)\Sigma_u C(1)'$.

2.2 The unreliability problem

Combining (2.5) and (2.6) sheds some light on the unreliability problem of long-run SVAR models. Indeed, since

$$C(1)A(0) = A(1), \tag{2.7}$$

and

$$A(L) \quad = \quad C(L)A(0) = C(L)C(1)^{-1}A(1), \tag{2.8}$$

the long-run identification scheme conducts to reliable inference if and only if the $A(1)$ is consistently estimated in finite samples and especially the lag order p is not misspecified. Otherwise, any inconsistent estimate of $A(1)$ leads to unreliable long-run effects of shocks (in finite samples). This in turn is transferred to the estimates of the dynamic multipliers of the structural shocks by virtue of Eq. (2.8).[12] In particular, one cannot form asymptotically correct confidence intervals for impulse responses of each structural shock and there is no consistent test that an individual impulse response coefficient is zero (Faust and Leeper, 1997). The fundamental issue is that the true data generating process may have an infinite-ordered VAR representation with $\Phi_0(L) = \sum_{s=1}^{\infty} \Phi_{0,k} L^k$ and thus the infinite sequence $\Phi_0 = \{\Phi_{0,1}, \Phi_{0,2}, \cdots\}$ must be approximated by a finite sequence $\widetilde{\Phi}_p = \{\Phi_1, \cdots, \Phi_p\}$ (i.e., a misspecified VAR model). Such finite-parameter approximations to infinite lag distributions have been studied extensively by Sims (1971, 1972) and Pötscher (2002), especially for least-squares criterion.[13] An accurate approximation from the point of view of least-squares fit does not imply an accurate approximation of the long run effect.[14] This means that convergence of the sequence $\tilde{\Phi}_p$ is not sufficient to guarantee the convergence of some functions of those parameters (Sims, 1971,1972; Pötscher, 2002) as pointwise convergence does not imply (locally) uniform convergence. More specifically, functions of a lag distribution (e.g., the sum of coefficients) are in general discontinuous with respect to the metric implied be least-squares estimation.[15] Say differently, the best least-squares approximation of Φ_0, Φ_p, might be arbitrarily close (w.r.t. L_2-norm) whereas $\Phi(1)$ and $\tilde{\Phi}_p(1)$ are arbitrarily far apart and thus converge to different limits. This stems also from the fact that the least-squares criterion at a single frequency admits a zero Lebesgue measure. From a practical point of view, it turns out that standard errors of estimates or the coefficient of determination might approach their optimum values in arbitrarily large samples while the estimated sum of coefficients remains arbitrarily far from their true values. Inference

[12] Using Monte Carlo simulations, Erceg et al. (2005) and Chari et al. (2008) study the extent of these small-sample estimation problems.

[13] A similar argument can be found in Christiano et al. (2006a).

[14] See Faust (1996,1999) for an application of this result to unit root tests and confidence intervals for points on spectrum.

[15] The functional $S_{\Delta X} \to S_{\Delta X}(0)$, with $S_{\Delta X}$ the spectrum of the stochastic process (ΔX_t), is highly discontinuous w.r.t. L_2-distance. This makes the problem fall into the category of ill-posed problem (Sims, 1972; Pötscher, 2002).

based on the sum of coefficients is then highly unreliable unless Φ is in fact contained in Φ_p, and not only close to it (Pötscher, 2002).[16]

2.3 Imposing restrictions on a continuum of frequencies

As suggested by Faust and Leeper (1997), and Pötscher (2002), an alternative, which solves the unreliability problem, is to make use of restrictions on a frequency interval.[17] Taking that any frequency-based restrictions can be written for $\omega \in [-\pi, \pi]$:

$$C(e^{-i\omega})\Sigma_u \overline{C(e^{-i\omega})}' = A(e^{-i\omega})\overline{A(e^{-i\omega})}', \tag{2.9}$$

our methodology consists of imposing restrictions on a continuum of frequencies and thus on the matrix $A(e^{-i\omega})$. In contrast to long-run restrictions, our proposed identification strategy rather considers restrictions (in the frequency domain) on the whole interval $\omega \in (\underline{\omega}, \overline{\omega})$. Say differently, common long run restrictions can be replaced by restrictions on $\omega = 0$ and its neighborhood.[18] This corresponds to the restriction that a given structural shock has no impact at the long-run as well as for cycles lasting close to the long-run, say more than 80 quarters. These restrictions imply that the spectral density of the variable of interest resulting from the structural shock is zero (or close as possible to zero) for the whole interval frequency and not on a single point. This means that we consider local averages of the spectral density which are continuous with respect to the L_2-norm (Pötscher, 2002). Hence this set of restrictions has a Lebesgue measure strictly greater than zero and an accurate approximation from the point of view of least-squares fit will involve an accurate approximation over the area given by the band $(\underline{\omega}, \overline{\omega})$ of interest for identification purpose. In particular, the distance between the true vector autoregressive process and its approximation cannot go to zero while the distance over the interval does not go to zero.

At the same time, it is worth noticing that the zero frequency might not belong to the frequency interval of interest. This corresponds to medium term restrictions or the medium-frequency component of fluctuations (Uhlig, 2003; Comin and Gertler, 2006; Pancrazi, 2015). Indeed, as pointed out by Uhlig (2003), VAR

[16] Note that it might occur regardless of how large the sample size is.

[17] Different strategies can be proposed to circumvent this issue. A first strategy is to restrict the behavior of Φ. On the one hand, one may restrict Φ to vanish for t larger than some number T. However, there may not exist convincing reasons to restrict Φ in such a way. On the other hand, one may also limit the oscillatory behavior of Φ in a neighborhood around frequency zero. In the univariate case, Pötscher (2002) shows that such limitations are less than innocuous. For instance, this excludes autoregressive process or moving average process of unknown order. In fact, assuming a fixed known dimension of the autoregressive process limits the oscillatory behavior sufficiently to retrieve standard asymptotic results. However, such restrictions might be difficult to justify from a theoretical point of view.

[18] Pötscher (2002, p. 1054) proposes a similar solution in another context to the aforementioned ill-posed problem.

analysis generally leads to the results that the contributions of shocks is the most important at intermediate horizons (of about two to ten years), and that each variable dynamics is almost not explained by only one (different) structural shock. Therefore a given structural shock might not be exactly identified by neither short-run, medium-run, nor long-run identification scheme; it might come close to but not exactly. In this respect, imposing restrictions on a frequency interval that does not contain $\omega = 0$ might also provide some further information for identification.[19]

Using the vech operator, expression (2.9) writes for all $\omega \in (\underline{\omega}, \overline{\omega})$:

$$vech\left(C(e^{-i\omega})\Sigma_u \overline{C(e^{-i\omega})}'\right) - vech\left(C(e^{-i\omega})A(0)A(0)'\overline{C(e^{-i\omega})}'\right) = \mathbf{0_d} \tag{2.10}$$

where $C(e^{-i\omega})A(0) = A(e^{-i\omega})$ and $\mathbf{0_d}$ is a $d \times 1$ vector of zeroes.

Then, as explained in Section 3, our approach proceeds as follows. In a first step, estimation of the reduced-form VAR(p) model leads to some consistent estimates of both the variance-covariance matrix and the matrices of the infinite moving average representation. In a second step, we make use of the previous restrictions in order to get some estimates of $a_0 = vec(A(0))$. In so doing, our approach is similar to the asymptotic least squares method developed by Gourieroux et al. (1985), and Gourieroux and Monfort (1995).[20] The key point is that the second step rests on a continuum of restrictions so that their results cannot be used in a straightforward way. To circumvent this issue, we rather rely on Carrasco and Florens (2000) and Carrasco et al. (2007), namely the generalization of the GMM estimator for a continuum of moment conditions.

As a final remark, it is worth noticing that the system of nonlinear equations is overidentified even if the original structural VAR model for a given frequency, say $\omega = 0$, is just-identified. This allows us for proposing a testing procedure in order to assess the reliability of the imposed restrictions.

3 Asymptotic Least Squares in the frequency domain

In this section, we propose a general asymptotic least squares estimator in the presence of a continuum of restrictions (hereafter, C-ALS estimator), and thus proceed in two steps. First, we define the class of C-ALS

[19]Note that medium-term restrictions might be imposed by assuming that the effect of a given shock in a particular time-horizon (chosen by the economist) should be zero outside a frequency interval. In this case, the integrated spectrum can be used to define the relevant medium-term restrictions. Taking the corresponding moment conditions, the methodology proposed in this paper might be used to proceed with such an identification scheme. Moreover such restrictions might be combined with the nearly long-run restrictions considered in this paper. We leave both points for future research.

[20]For an overview of the asymptotic least squares method, see the technical appendix.

estimators for every sequence of random bounded linear operators. Second, the optimal C-ALS estimator is presented. Finally, a test of overidentification and a data-driven procedure for the choice of the frequency interval are discussed.

3.1 Notation

Let $\widehat{\beta}_T$ denote a first-step M-estimator defined by:

$$\widehat{\beta}_T = \arg\min_{\beta \in \mathcal{B}} Q_T(Z_T, \beta) \tag{3.11}$$

where $\beta_0 = P_0 \lim_{T\to\infty} \widehat{\beta}_T$ is the true unknown value of the instrumental parameters and P_0 is the true unknown probability distribution of the data generating process, and Q_T is the (sample) objective function. Under standard regularity conditions $\sqrt{T}\left(\widehat{\beta}_T - \beta_0\right) \xrightarrow{d} \mathcal{N}(0, \Omega)$ and $\Omega = \lim_{T\to\infty} \text{Var}\left(\sqrt{T}\widehat{\beta}_T\right)$ under P_0.

Taking this first-step estimator, we consider a system of J constraints defined on a continuum of frequencies (e.g., Eq. 2.10). These constraint functions are complex-valued and indexed by a parameter ω taking its values in R^d for any arbitrary $d \geq 1$

$$g(a_0, \widehat{\beta}_T, \omega) = 0,$$

where $g(\cdot, \cdot, \omega)$ takes its values in $H = (L^2(\varphi, [-\pi, \pi]))^J$, a Hilbert space with the inner product $< ., . >$ and the norm $\| \cdot \|$. $L^2(\varphi, [-\pi, \pi]) \equiv L^2(\varphi)$ is the space of complex valued functions that are square integrable with respect to the probability density function φ of a distribution for ω. [21] Let S denote a bounded linear operator defined on $(L^2(\varphi, [-\pi, \pi]))^J$ or a subspace of $(L^2(\varphi, [-\pi, \pi]))^J$ and $\overline{g(\cdot, \cdot, \omega)}$ denote the complex conjugate of $g(\cdot, \cdot, \omega)$.

3.2 The class of C-ALS estimators

In order to define the C-ALS estimator, we need to replace the common sequence of symmetric positive definite matrices in the GMM literature by a converging sequence of random bounded linear operators (Carrasco and Florens, 2000; Carrasco et al., 2007). Hence, for a given sequence, say S_T, converging to S, the ALS estimator for a continuum of constraints (hereafter, the C-ALS estimator) is defined from:

$$\widehat{a}_T(S_T) = \arg\min_{a \in \mathcal{A}} \left\| S_T^{1/2} g(a, \widehat{\beta}_T, \omega) \right\|^2.$$

Therefore the C-ALS estimator renders the constraints, $g(a, \widehat{\beta}_T, \omega) = 0$ for $\omega \in I_\omega = [\underline{\omega}, \overline{\omega}]$, as close as possible to zero by using the metric associated with the inner product defined by S_T. For instance, if ω

[21] All assumptions are given in Appendix 1.

belongs to the interval $[\underline{\omega}, \overline{\omega}]$, the C-ALS estimator is given by

$$\hat{a}_T(S_T) = \arg\min_{a \in \mathcal{A}} \int_{\underline{\omega}}^{\overline{\omega}} \int_{\underline{\omega}}^{\overline{\omega}} g(a, \widehat{\beta}_T, \omega_1)' S_T(\omega_1, \omega_2) \overline{g(a, \widehat{\beta}_T, \omega_2)} d\omega_1 d\omega_2.$$

The following proposition provides its asymptotic properties.

Proposition 3.1. *Suppose that Assumptions A.1 to A.10 hold true and that S_T denote a sequence of random bounded linear operators converging to S, the C-ALS estimator associated with S_T is a solution $\hat{a}_T(S_T)$ to the problem:*

$$\widehat{a}_T = \arg\min_{a \in \mathcal{A}} \| S_T^{1/2} g(a, \widehat{\beta}_T, \omega) \|^2. \tag{3.12}$$

The C-ALS estimator exists and weakly converges to a_0:

$$\widehat{a}_T \xrightarrow{p} a_0.$$

Moreover it is asymptotically normally distributed:

$$\sqrt{T}\left(\widehat{a}_T(S_T) - a_0\right) \xrightarrow{l} N(0, Q(S))$$

with

$$\begin{aligned} Q(S) \quad = \quad & \left\| S^{1/2} \frac{\partial g}{\partial a'}(a_0, \beta_0, \omega) \right\|^{-2} \left\langle S^{1/2} \frac{\partial g}{\partial a'}(a_0, \beta_0, \omega), \left(S^{1/2} K \overline{S^{1/2}} \right) S^{1/2} \frac{\partial g'}{\partial a}(a_0, \beta_0, \omega) \right\rangle \\ & \times \left\| S^{1/2} \frac{\partial g}{\partial a'}(a_0, \beta_0, \omega) \right\|^{-2} \end{aligned}$$

where $K : (L^2(\varphi))^J \to (L^2(\varphi))^J$ is an Hilbert-Schmidt operator such that for all f_l in $L^2(\varphi)$:

$$(Kf)(\omega_1) = \left(\sum_{l=1}^{J} \int E^{P_0} \left[k_{jl}(\omega_1, \omega_2) \right] f_l(\omega_2) \varphi(\omega_2) d\omega_2 \right)_{j=1,\dots,J}$$

with

$$k_{jl}(\omega_1, \omega_2) = \frac{\partial g_j}{\partial \beta'}(a_0, \beta_0, \omega_1) \Omega \overline{\frac{\partial g_l}{\partial \beta}(a_0, \beta_0, \omega_2)'}.$$

Proof: See Appendix 2.

Several points are worth commenting. First, the previous proposition is an implication of the following functional convergence (as $T \to \infty$):

$$\frac{\partial g}{\partial \beta'}(a_0, \widehat{\beta}_T, \omega) \sqrt{T}(\widehat{\beta}_T - \beta_0) \Rightarrow N(0, K)$$

where $N(0, K)$ is the Gaussian random vector of $(L^2(\varphi))^J$.[22] Second, the Hilbert-Schmidt operator K is not invertible on the full reference space but has a finite dimensional closed range (at most) equal to p—the dimension of the (asymptotic) variance-covariance matrix of the parameter vector β. This differs from the framework of Carrasco and Florens (2000) and Carrasco et al. (2007) in which the inverse of K is not bounded and the range of K is not closed.[23] Accordingly, the Moore-Penrose inverse operator K^{-1} is not bounded and the solution $K^{-1}f$ to a Fredholm equation of the first kind $K\Phi = f$ is not continuous in f; K does not admit a generalized inverse over the entire Hilbert space of reference. Consequently, to guarantee the stability of the solution, Carrasco and Florens (2000) replace the operator K by some nearby operator (e.g., using a Tikhonov regularization).[24] In contrast, since the operator K depends on the first step estimation through $\widehat{\beta}_T$ in our approach, the range of K, denoted $R(K)$, is then known and $R(K) = dim(\beta)$. Note however that the derivation of the (Moore-Penrose) generalized inverse of K can be cumbersome in finite samples due to the presence of nearly zero eigenvalues and thus a regularization matrix cannot be precluded in our framework (see further).

Third, since the range of K equals (at most) p, the number of its eigenvalues different from zero is finite. Using Carrasco and Florens (2000), one has the following decomposition

$$\langle Kf \rangle (\omega_1) = \sum_{i=1}^{p} \mu_i \gamma_{i,T}(\omega_1) \langle f' \gamma_{i,T} \rangle = \sum_{i=1}^{p} \mu_i \gamma_{i,T}(\omega_1) \int f(\omega_2)' \gamma_{i,T}(\omega_2) \varphi(\omega_2) d\omega_2.$$

where μ_i, $i = 1, \ldots, p$ denotes the p eigenvalues of K different from zero and γ_i the corresponding vector of orthonormalized eigenfunctions.

This implies that the Moore-Penrose generalized inverse of K, K^{-1}, satisfies:

$$\left\langle K^{-1} f \right\rangle (\omega_1) = \sum_{i=1}^{p} \frac{1}{\mu_i} \gamma_{i,T}(\omega_1) \langle f' \gamma_{i,T} \rangle .$$

and that a consistent estimator of K^{-1} can be obtained as follows.

Proposition 3.2. *Let $\widehat{a}_T^1$ denote a first-step consistent estimator of a_0. A consistent estimator of the Moore-Penrose generalized inverse is defined to be:*

$$\left\langle K_T^{-1} f \right\rangle (\omega_1) = \sum_{i=1}^{p} \frac{1}{\mu_{i,T}} \gamma_{i,T}(\omega_1) \langle f' \gamma_{i,T} \rangle$$

where $\gamma_{i,T}$ is given by:

$$\frac{\partial g}{\partial \beta'} (\widehat{a}_T^1, \widehat{\beta}_T, \omega_1) \Omega_T^{1/2} D_h,$$

[22] See Carrasco and Florens (2000), remark 2 p. 803.

[23] For a general discussion of linear inverse problems in econometrics, see Carrasco, Florens and Renault (2007).

[24] See also Carrasco (2012).

the eigenvalues $\mu_{i,T}$ are those of the $p \times p$ matrix:

$$\int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\widehat{a}_T^1, \widehat{\beta}_T, \omega_2)} \frac{\partial g}{\partial \beta'}(\widehat{a}_T^1, \widehat{\beta}_T, \omega_2) \Omega_T^{1/2} \varphi(\omega_2) d\omega_2,$$

and the matrices $D = [D_1 \cdots D_p]$ and $\Gamma = [\gamma_1 \cdots \gamma_p]$ satisfy

$$\int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\widehat{a}_T^1, \widehat{\beta}_T, \omega_2)} \frac{\partial g}{\partial \beta'}(\widehat{a}_T^1, \widehat{\beta}_T, \omega_2) \Omega_T^{1/2} \varphi(\omega_2) d\omega_2 D = D\Gamma.$$

where Ω_T is a consistent estimate of Ω.

Proof: See Appendix 2.

3.3 The optimal C-ALS estimator

Using Proposition 3.1, it can be shown that choosing $S_T = K_T^{-1/2}$ leads to the estimator of minimum variance and thus the optimal C-ALS estimator.[25]

Proposition 3.3. *Let K_T denote a consistent estimator of K and K_T^{-1} a Moore-Penrose generalized inverse estimator of K^{-1}. The optimal C-ALS estimator of a_0 is obtained by*

$$\widehat{a}_T = \arg\min_{a \in \mathcal{A}} \|K_T^{-1/2} g(a, \widehat{\beta}_T, \omega)\|^2 \tag{3.13}$$

and $\widehat{a}_T$ is consistent and asymptotically normally distributed:

$$\sqrt{T}\left(\widehat{a}_T(K_T^{-1/2}) - a_0\right) \xrightarrow{l} N\left(0, \left\| \frac{\partial g}{\partial a'}(a_0, \beta_0) \right\|_K^{-2}\right)$$

Proof: See Appendix 2.

Taking the eigenfunctions and eigenvalues decomposition of K_T, the optimal C-ALS objective function to minimize can be rewritten as:

$$\|K_T^{-1/2} g(a, \widehat{\beta}_T, \omega)\|^2 = \sum_{i=1}^{p} \frac{1}{u_{i,T}^{(p)}} \left\langle g(a, \widehat{\beta}_T, \omega), \phi_{i,T}^{(p)}(\omega) \right\rangle^2.$$

However, this computation can be burdensome, particularly for large p. As in Carrasco et al. (2007), we propose a simple expression of the objective function in terms of vectors and matrices.

Proposition 3.4. *A simplified expression for the objective function of the C-ALS problem is given by :*

$$\widehat{a}_T = \arg\min_{a \in \mathcal{A}} \underline{s}(a, \widehat{\beta}_T)' \widetilde{W}_T^2 \overline{\underline{s}(a, \widehat{\beta}_T)}$$

[25] Intuitively, S is chosen such that $SK\bar{S}$ is equal to the identity operator. See Carrasco and Florens (2000), remark 4 p. 804.

where $\widetilde{W}_T$ *is a generalized inverse of* W_T *such that* $\widetilde{W}_T W_T = I_p$ *and*

$$W_T = \int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\widehat{a}_T^1, \widehat{\beta}_T, \omega)} \frac{\partial g}{\partial \beta'}(\widehat{a}_T^1, \widehat{\beta}_T, \omega) \Omega_T^{1/2} \varphi(\omega_2) d\omega_2$$

is the $p \times p$*-matrix and*

$$\underline{s}(a, \hat{\beta}_T) = \int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\widehat{a}_T^1, \widehat{\beta}_T, \omega)} g(a, \widehat{\beta}_T, \omega) \varphi(\omega) d\omega$$

is a p*-vector. When the matrix* W *is of full rank, then* $\widetilde{W} = W^{-1}$*.* □

Proof: See Appendix 2.

As explained before, one key issue is that the matrix W might not be of full rank p in finite samples. As proposed by Carrasco and Florens (2000) in a C-GMM context, a generalized inverse of W_T might be obtained through a regularization à la Tikhonov.

Proposition 3.5. *A simplified expression for the regularized objective function of the C-ALS problem is given by :*

$$\widehat{a}_T = \arg\min_{a \in \mathcal{A}} \underline{s}(a, \widehat{\beta}_T)' [\eta_T I_T + W_T^2]^{-1} \overline{\underline{s}(a, \widehat{\beta}_T)}$$

where the regularization parameter η_T *goes to zero at a suitable rate (see Carrasco et al., 2007; Carrasco, 2012).*

3.4 Test of overidentification

Using Carrasco and Florens (2000), a test of overidentification can also be performed using the following J-statistic.

Proposition 3.6. *The overidentification test is given by:*

$$J_T = \| \sqrt{T} K_T^{-1/2} g(\widehat{a}, \widehat{\beta}_T, \omega) \|^2$$

The statistic J_T *is distributed as a chi-square with p degrees of freedom.*

In the presence of a rank-order deficiency, Proposition 3.5 can be used to define a regularized version of the test statistic. More specifically, for a given value of η_T, the J_T statistic converges to a weighted Chi-squared distribution given by:

$$T \underline{s}(\widehat{a}_T, \widehat{\beta}_T)' [\eta_T I_T + \widehat{W}^2]^{-1} \overline{\underline{s}(\widehat{a}_T, \widehat{\beta}_T)} \xrightarrow{d} \sum_{j=1}^{p} \frac{\lambda_j^2}{\lambda_j^2 + \eta_T} \chi_j^2(1) \tag{3.14}$$

where the random variables $\chi_j^2(1)$ are (centered) Chi-squared distributed with one degree of freedom and the λ_j terms are the eigenvalues of the matrix W that are replaced by their sample counterparts (see Arellano, Hansen and Sentana, 2012; Vuong, 1989).

3.5 Data-driven procedure for the frequency interval

The next question to address is the determination of the interval $I_\omega = (\underline{\omega}, \overline{\omega})$ over which one might impose and assess the reliability of the identifying restrictions. Without loss of generality, we consider the class of increasing symmetric intervals of ω around zero, i.e. $I_\omega = (-\omega, \omega)$ and we use the information criteria-based methodology of Hall et al. (2007) in order to propose a statistical criterion which selects the largest interval I_ω that might guarantee consistent estimation of $\widehat{a}_T$. In so doing, we select the interval by minimizing the Valid Interval Selection Criterion (VISC) defined by:

$$\hat{\omega}_T = \underset{\omega \in \mathcal{C}(\omega)}{\operatorname{argmin}} \mathtt{VISC}_T(\omega)$$

where $\mathcal{C}(\omega)$ is the class of symmetric intervals around zero for the identification scheme of interest and

$$\mathtt{VISC}_T(\omega) = J_T(\omega) - h(|w|)\kappa_T \tag{3.15}$$

where $h(|w|)\kappa_T$ is a deterministic penalty that is an increasing function of the length of the interval. Proposition 3.7 shows that $\hat{\omega}_T$ converges in probability to the unique ω_0 that chooses the maximal bound for a valid consistent estimation of $\hat{a}_T$.

Proposition 3.7. *Suppose that (1) There exists a lower bound ω_{lb} such that the restrictions are respected for the interval $(\omega_{lb}, \omega_{lb})$, (2) $\omega_{max} = (-\omega_0, \omega_0)$, and (3) $h(\cdot)$ is strictly increasing and $\kappa_T \to \infty$ as $T \to \infty$ with $\kappa_T = o(T)$. Then the estimator $\hat{\omega}_T$ defined as the solution of the criterion (3.15) converges in probability to ω_0.*

Note that the first assumption imposes that the restrictions are valid for at least an interval with minimal length characterized by the lower bound ω_{lb}. The second assumption ensures that the interval $(-\omega, \omega)$ is uniquely identified. The latter imposes assumption on the penalty terms that guarantee the validity of the criterion. The SIC-type penalty term $((h|\omega|) = 2\omega$ and $\kappa_T = \ln(T))$ and the Hannan-Quinn-type penalty term $((h|\omega|) = 2\omega$ and $\kappa_T = ln(ln(T)))$ satisfies this assumption while the AIC-type penalty term $((h|\omega|) = 2\omega$ and $\kappa_T = 2)$, does not.

4 Asymptotic least squares in the frequency domain for structural VAR models

In this section, we make use of the general results of Section 3 and show how they can be applied to identify structural VAR models with frequency-based restrictions. Then we discuss two applications involving some simplifications of the general procedure: (i) the identification of structural shocks for bivariate VAR model and (ii) the identification of a given structural shocks in the case of a m-variable structural VAR specification.

4.1 General framework

Let θ_0 denote the parameter vector of interest, with $\theta_0 = (\beta_0', a_0')'$ where $\beta_0 = (vec(\Phi_p)', vech(\Sigma)')'$ and $a_0 = vec\,(A(0))$. Let $Z_t = \left(\Delta X_t', \ldots, \Delta X_{t-p}'\right)'$. The estimator $\widehat{\beta}_T$ is obtained by minimizing the sum of squared residuals of the reduced-form VAR.

The estimator $\widehat{a}_T$ is obtained as a C-ALS estimator defined in (3.13) by using the frequency constraints between the estimated coefficients of the moving average representation resulting from the VAR estimation and the variance-covariance estimator of the residuals, for all $\omega \in (\underline{\omega}, \overline{\omega})$.:

$$vech\left(\widehat{C}(e^{-i\omega})\widehat{\Sigma}\overline{\widehat{C}(e^{-i\omega})}' - \widehat{C}(e^{-i\omega})A(0)A(0)'\overline{\widehat{C}(e^{-i\omega})}'\right) = 0. \tag{4.16}$$

Accordingly, the C-ALS estimator of $a_0 = vec(A(0))$ is defined to be the solution of the following minimization problem:

$$\widehat{a}_T = \arg\min_{a \in \mathcal{A}} \|K_T^{-1/2} g(a, \widehat{\beta}_T, \omega)\|.$$

Then inference on parameters of the matrix $A(0)$ can be performed using the asymptotic distribution of the C-ALS estimator. An overidentifying test can also be conducted for the validity of the restrictions on the interval frequency with the statistic presented above. In so doing, the first order partial derivatives of the constraints with respect to the parameters of interest must be determined (e.g., see Proposition 3.4) and thus are stated in the following proposition.

Proposition 4.1. *Let a_0 and $\hat{\beta}_T$ denote $a_0 \equiv vec(A_0) \in \mathbb{R}^{K^2}$ and $\hat{\beta}_T \equiv (\hat{\alpha}_T', \hat{\sigma}_T')' \in \mathbb{R}^{K^2p+K(K+1)/2}$ where $\hat{\alpha}_T = vec(\hat{\Phi}_p)$ and $\hat{\sigma}_T = vech(\hat{\Sigma}_T)$. The first-order partial derivatives of the moment conditions,*

$$g(a, \hat{\beta}_T, \omega) = vech\left(\widehat{C}(e^{-i\omega})\widehat{\Sigma}_u\overline{\widehat{C}(e^{-i\omega})}' - \widehat{C}(e^{-i\omega})A_0A_0'\overline{\widehat{C}(e^{-i\omega})}'\right)$$

with respect to a and $\hat{\beta}_T$ are respectively given by:

$$\frac{\partial g}{\partial a'}(a, \hat{\beta}_T, \omega) = -L\left[\left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})A_0\right)P_{K,K} + \left(\widehat{C}(e^{-i\omega})A_0 \otimes \widehat{C}(e^{-i\omega})\right)\right],$$

and

$$\frac{\partial g}{\partial \beta'}(a, \hat{\beta}_T, \omega) = \begin{pmatrix} \frac{\partial g}{\partial \alpha'}(a, \hat{\beta}_T, \omega) & \frac{\partial g}{\partial \sigma'}(a, \hat{\beta}_T, \omega) \end{pmatrix} \in \mathcal{M}_{K(K+1)/2, K^2p+K(K+1)/2}$$

with

$$\begin{aligned}\frac{\partial g}{\partial \alpha'}(a, \hat{\beta}_T, \omega) &= L\left\{\left[I_K \otimes C(e^{-i\omega})\left(\Sigma_u - A_0A_0'\right)\right]P_{K,K} + C(e^{-i\omega})\left(\Sigma_u - A_0A_0'\right) \otimes I_K\right\} \\ &\quad \times \left(\overline{C(e^{-i\omega})}' \otimes C(e^{-i\omega})\right)\left(e_p' \otimes I_{K^2}\right) \\ \frac{\partial g}{\partial \sigma'}(a, \hat{\beta}_T, \omega) &= L\left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)D_K\end{aligned}$$

where L is an $\left(\frac{K(K+1)}{2} \times K^2\right)$ elimination matrix, D_K is an $\left(K^2 \times \frac{K(K+1)}{2}\right)$ duplication matrix, $P_{m,n}$ is an $(mn \times mn)$ commutation matrix, I_K is the identity matrix of order K, and $e_p = (1, \cdots, 1)'$ is a $(p \times 1)$ vector.

Proof: See Appendix 2.

4.2 Bivariate VAR

As a first application, we consider a bivariate VAR model. The structural model (respectively, the reduced form) of the joint time series $\Delta X_t = (\Delta x_{1t}, \Delta x_{2t})'$ is defined by (2.1) (respectively, (2.2)). Let $\widehat{\Sigma}_T$ and $A(0)$ denote

$$\widehat{\Sigma}_T = \begin{pmatrix} \widehat{\sigma}_{11} & \widehat{\sigma}_{12} \\ \widehat{\sigma}_{12} & \widehat{\sigma}_{22} \end{pmatrix} \quad \text{and} \quad A(0) = \begin{pmatrix} A^0_{11} & A^0_{12} \\ A^0_{21} & A^0_{22} \end{pmatrix}.$$

Suppose that the second shock has no impact on the first variable over a symmetric interval around zero $[\underline{\omega}, \overline{\omega}]$ with $\underline{\omega} = -\overline{\omega}$. At the limit, the interval shrinks to $\omega = 0$, and thus to the usual Blanchard-Quah long-run restriction. The next proposition shows how a consistent estimator of $a_0 = vec(A(0))$ can be obtained through a C-ALS estimator.

Proposition 4.2. *Consider the following identification constraint $\forall \omega \in [\underline{\omega}, \overline{\omega}]$,*

$$C_{11}(e^{-i\omega})A^0_{12} + C_{12}(e^{-i\omega})A^0_{22} = 0.$$

The constraints are defined by

$$\begin{aligned} g(\tilde{a}_{12}, \widehat{\beta}_T, \omega) &= \widehat{C}_{11}(e^{-i\omega})\tilde{a}_{12} + \widehat{C}_{12}(e^{-i\omega}) = 0 \\ &= \sum_{k=0}^{\infty} \left[\widehat{C}_{11,k} e^{-i\omega k} \tilde{a}_{12} + \widehat{C}_{12,k} e^{-i\omega k}\right] = 0 \end{aligned}$$

$\forall \omega \in [\underline{\omega}, \overline{\omega}]$ where $\tilde{a}_{12} = A^0_{12}/A^0_{22}$ and $\Sigma = A(0)A(0)'$. Then the matrix $A(0)$ is locally identified (up to a sign restriction).

Proof : See Appendix 2.

In this case, the minimization of the objective function (4.17) can be replaced by a simple two-step procedure that exploits the linearity of the constraints with respect to a_{12}. First, a two-step C-ALS estimator of $\tilde{a}_{12,T}$ can be derived by using the identity operator as a kernel operator K_T in the minimization problem:[26]

$$\widehat{\tilde{a}}_{12,T} = \arg\min_{a \in \mathcal{A}} \|K_T^{-1/2} g(\tilde{a}_{12}, \widehat{\beta}_T, \omega)\|. \tag{4.17}$$

[26]To some extent, this corresponds to a standard two-step GMM procedure.

Interestingly, this first-step estimator is the one proposed by Wen (2001). Indeed,

$$\begin{aligned}\widehat{\tilde{a}}^{1}_{12,T} &= \arg\min \int_{\underline{\omega}}^{\overline{\omega}}\int_{\underline{\omega}}^{\overline{\omega}} \overline{g(\tilde{a}_{12},\widehat{\beta}_T,\omega_i)}' I_{\{\omega_i=\omega_j\}} g(\tilde{a}_{12},\widehat{\beta}_T,\omega_j) d\omega_i d\omega_j \\ &= \arg\min \int_{\underline{\omega}}^{\overline{\omega}} |g(\tilde{a}_{12},\widehat{\beta}_T,\omega)|^2 d\omega.\end{aligned}$$

with $I_{\{.\}}=1$ for $\omega_i=\omega_j$ and zero otherwise. It can be shown that the estimator is given by:

$$\widehat{\tilde{a}}^{1}_{12,T} = -\frac{\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\widehat{C}_{11,k}\widehat{C}_{12,j}\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega}{\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\widehat{C}_{11,k}\widehat{C}_{11,j}\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega}. \tag{4.18}$$

In a second, taking the estimator of $\tilde{a}_{12}$ and the identifying restrictions, the matrix $A(0)$ can be rewritten as:

$$\tilde{A}(0) = \begin{pmatrix} A^0_{11} & A^0_{22}\widehat{\tilde{a}}_{12,T} \\ A^0_{21} & A^0_{22}\end{pmatrix},$$

the estimator $\widehat{a}_T$ is obtained as the solution of the nonlinear system of equations:

$$\widehat{a}_T = \arg\min_{a\in\mathcal{A}} vech\left(\widehat{\Sigma}-\tilde{A}(0)\tilde{A}(0)'\right).$$

Using Proposition 3.4, one obtains the corresponding objective function and the explicit expression of $\widehat{\tilde{a}}_{12,T}$.

Proposition 4.3. *The solution of the minimization for the C-ALS objective function problem corresponding to the bivariate VAR model is*

$$\widehat{\tilde{a}}_{12,T} = -\frac{\widehat{A}'_T(\widehat{W}^2_T)^{-1}\widehat{B}_T}{\widehat{A}'_T(\widehat{W}^2_T)^{-1}\widehat{A}_T}$$

where $\widehat{A}_T$, $\widehat{B}_T$ and $\widehat{W}_T$ are given by:

$$\widehat{A}_T = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\left(\frac{\partial\widehat{C}_{11,k}}{\partial\beta'}\widehat{\tilde{a}}^{1}_{12,T}+\frac{\partial\widehat{C}_{12,k}}{\partial\beta'}\right)\widehat{\Omega}^{1/2}\right)'\widehat{C}_{11,j}\right]\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega$$

$$\widehat{B}_T = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\left(\frac{\partial\widehat{C}_{11,k}}{\partial\beta'}\widehat{\tilde{a}}^{1}_{12,T}+\frac{\partial\widehat{C}_{12,k}}{\partial\beta'}\right)\widehat{\Omega}^{1/2}\right)'\widehat{C}_{12,j}\right]\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega,$$

and

$$\widehat{W}_T = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\left(\frac{\partial\widehat{C}_{11,k}}{\partial\beta'}\widehat{\tilde{a}}^{1}_{12,T}+\frac{\partial\widehat{C}_{12,k}}{\partial\beta'}\right)\widehat{\Omega}^{1/2}\right)'\left(\left(\frac{\partial\widehat{C}_{11,j}}{\partial\beta'}\widehat{\tilde{a}}^{1}_{12,T}+\frac{\partial\widehat{C}_{12,j}}{\partial\beta'}\right)\widehat{\Omega}^{1/2}\right)\right]\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega.$$

where $\widehat{\tilde{a}}^{1}_{12,T}$ is a first-step estimator and $\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega = \frac{1}{j-k}\left[\sin((j-k)\omega)\right]_{\underline{\omega}}^{\overline{\omega}}$ for $(j-k)\neq 0$ and $\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega = \overline{\omega}-\underline{\omega}$ for $j=k$. For a symmetric interval $(\overline{\omega},-\overline{\omega})$, $\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega = \frac{2}{j-k}\sin((j-k)\overline{\omega})$ for $j\neq k$. $\square$

Proof: See Appendix 2.

In this respect, our method can be interpreted intuitively as finding the linear combination of the reduced-form shocks such that the contribution of the second structural shock to the first variable is minimized or is as close as possible to zero.

4.3 Identification of the technology shock with a m-variable VAR

As a second application, we consider the identification of a single structural shock in a VAR containing more than two variables. Without loss of generality, we assume that the structural shock of interest is the first one. For instance, this corresponds to the identification of a technology shock in a multivariate VAR without requiring the identification of other shocks (see Christiano et al., 2006b; Francis and Ramey, 2009). For sake of completeness, we first provide the common sign and (long-run) exclusion restrictions and then turns to the frequency identifying restrictions in order to apply the C-ALS estimator.

Following Christiano et al. (2006a), the dynamic effects of the first structural shock can be computed by identifying only $A(0)_1$, the first column of $A(0)$, since combining a sign restriction and zero restrictions on the long-run impact does uniquely identify the vector $A(0)_1$. In this respect, one needs to impose the $m-1$ zero-restrictions:

$$C(1)A(0) = F = \begin{bmatrix} f_{11} & 0_{1\times(m-1)} \\ F_{21} & F_{22} \end{bmatrix}$$

paired with the sign restriction that the element $(1,1)$ of the F matrix is positive. In addition, imposing that only the first structural shock has a positive long-run impact on the first variable yields the following specification of the long-run variance-covariance matrix:[27]

$$C(1)A(0)A(0)'C(1)' = FF' = \begin{bmatrix} f_{11}^2 & d_{11}F_{21}' \\ F_{21}f_{11} & F_{21}F_{21}' + F_{22}F_{22}' \end{bmatrix} = C(1)\Sigma C(1)'.$$

This implies that f_{11}^2 is equal to the element (1,1) of the matrix $C(1)\Sigma C(1)'$ and that F_{21} is equal to the corresponding elements of the matrix $C(1)\Sigma C(1)'$ divided by f_{11}. Since the first column of the matrix $C(1)A(0)$, denoted F_1, is known, the column vector $A(0)_1$ is uniquely identified by the relation $A(0)_1 =$

[27] Note that many matrices $A(0)$ are comfortable with these restrictions but the first column of each of these matrices $A(0)_1$ is the same (see Christiano et al., 2006b).

$C(1)^{-1}F_1$. Therefore the moment conditions can be written as:

$$\begin{aligned} f_{11}^2 &= [C(1)\Sigma C(1)']_{11} \\ F_{21}f_{11} &= [C(1)\Sigma C(1)']_{k1,\ k=2,\ldots,m} \end{aligned}$$

where $[C(1)\Sigma C(1)']_{k1}$ is the element $(k,1)$ of the matrix $C(1)\Sigma C(1)'$.

Taking the previous specification (including the sign restriction), we are now in a position to set the continuum of moment conditions for $\omega \in [\underline{\omega}, \overline{\omega}]$ (Proposition 4.4).

Proposition 4.4. *Consider the following identification constraints* $\sum_{j=1}^{m} C_{1j}(e^{-i\omega})A_{jk}^0 = 0$ *for* $k = 2,\ldots,m$. *The constraints* $g(a,\widehat{\beta}_T,\omega) = \left(g_1(a,\widehat{\beta}_T,\omega)\ \ g_2(a,\widehat{\beta}_T,\omega)\ \ \ldots\ \ g_m(a,\widehat{\beta}_T,\omega)\right)'$ *corresponding to (3.14) are defined by*

$$\begin{aligned} g_1(a,\widehat{\beta}_T,\omega) &= \left|[C(e^{-i\omega})A(0)]_{11}\right|^2 - \left[C(e^{-i\omega})\Sigma\overline{C(e^{-i\omega})}'\right]_{11} \\ g_k(a,\widehat{\beta}_T,\omega) &= \left[C(e^{-i\omega})A(0)\right]_{k1}[A(0)'\overline{C(e^{-i\omega})}']_{11} - \left[C(e^{-i\omega})\Sigma\overline{C(e^{-i\omega})}'\right]_{k1} \end{aligned}$$

for $k = 2,\ldots,m$ *and* $\forall\omega \in [\underline{\omega},\overline{\omega}]$ *uniquely identified the first column of the matrix* $A(0)$ *up to a sign restriction.*

Appendix 3 provides more primitive expressions for the constraint functions in Proposition 4.4. Moreover, using Proposition 4.1, we can write explicitly the objective function of the C-ALS problem as in Proposition 3.4. Note finally that the interpretation is the same as in the case of a bivariate VAR model: one seeks to find the best linear combination of the reduced-form shocks such that the contribution of the (m-1) other structural shocks to the first variable is minimized.

5 Monte Carlo simulations

In this section, we illustrate the usefulness of our approach by diagnosing several frequency identification schemes featured recently in the literature, and thus provide some Monte Carlo simulations to assess the finite sample performances of the C-ALS estimator and the overidentification test. We assume that the data generating process is a bivariate VAR(1) model (with different parameter configurations) in which the first variable, $X_{1,t}$, is nonstationary and thus written in first-difference and the second variable, $X_{2,t}$, is a weakly stationary process:

$$\Delta X_{1,t} = \rho_{11}\Delta X_{1,t-1} + (\rho_{12}+\delta)X_{2,t-1} - \rho_{12}X_{2,t-2} + \epsilon_{1,t} \tag{5.19}$$

$$X_{2,t} = \rho_{22}X_{2,t-1} + \rho_{21}\Delta X_{1,t-1} + b_{21}\epsilon_{1,t} + \epsilon_{2,t} \tag{5.20}$$

where the vector $\epsilon_t = (\epsilon_{1,t}, \epsilon_{2,t})'$ represents some structural shocks, with $\epsilon_t \sim N(0, I_2)$. The parameter δ controls the magnitude of the long-run effect of the second shock $\epsilon_{2,t}$ on the first variable $X_{1,t}$. When $\delta = 0$, only the first shock has a long-run impact on the first variable. To some extent, the corresponding specification can be viewed as the one often encountered in the macro literature in order to identify a permanent shock, e.g., the identification of a technology shock with some measures of (labor or total) productivity and hours worked.

Using Eq. (5.19) and Eq. (5.19), we generate 1,000 samples of size $T = 200$ observations—a sample size often encountered in applied macro works—and the effect of initial conditions is controlled by including 200 pre-sampled observations that are subsequently discarded in the estimation. For each repetition, the lag order is set to its true value so that results are interpreted free of any lag order misspecification issue.[28] Our method denoted `Freq`, which is based on the two-step C-ALS procedure (Section 4.2.), is compared with three approaches. The first one, denoted `LR`, is a standard long-run identification scheme à la Blanchard-Quah, i.e. we only impose the identification constraint at $\omega = 0$.[29] The second alternative is the one proposed by Wen (2001) and it corresponds to the first-step C-ALS estimator defined in Eq. 4.18 when the kernel operator is defined to be the identity operator. The last alternative, denoted `DO`, is the procedure of DiCecio and Owyang (2010), which identifies the structural shock of interest that maximizes (or minimizes) the share of the forecast-error variance of a given variable in the frequency domain (see also Francis et al., 2014).

With the exception of the `LR` method, we consider the following fixed symmetric frequency intervals $\omega_n = (-\frac{2\pi}{n}, \frac{2\pi}{n})$ for $n = 30, 60, 90, 120, 180, 210,$ and 240 quarters and thus do not implement any data-driven procedure to determine the optimal frequency interval.[30] Results are then assessed along three dimensions. On the one hand, we compute the initial impact of each structural shock to each variable and determine the corresponding mean absolute bias and root mean squared errors (RMSE). On the other hand, we provide the cumulative mean absolute bias and RMSE for $h \in [0, H]$, with $H = 4$, 8, and 12, by using the impulse response functions.[31] More specifically, the cumulative mean absolute bias is defined as $cmd(H) = \sum_{h=0}^{H} |irf_h(model) - irf_h(svar)|$ where H denotes the selected horizon, $irf_h(model)$ the impulse response at horizon h from the model defined by (5.19) and (5.20), and $irf_h(svar) = (1/N) \sum_{j=1}^{N} irf_h(svar)^j$ the average impulse response function over the N simulation experiments obtained from a SVAR model. Say

[28] Several robustness exercises, which are available upon request, have been experimented to control for the lag order misspecification. All in all, our results remain unchanged and our estimator performs better than the competing estimators.

[29] We also implement the methodology of Christiano et al. (2006a), i.e. a nonparametric approach to estimate the zero-frequency spectral density (with a Bartlett or Andrews-Monahan kernel). However, our Monte Carlo results, which are available upon request, provide almost no support for such a correction.

[30] Simulation results of the data-driven procedure are available upon request.

[31] Results for larger horizons are consistent with those presented in the sequel.

differently, the cumulative mean absolute bias is a measure of the area between the impulse response function up to a certain horizon H and the horizontal axis. Finally, we compare the true impulse response function of the second variable relative to the first structural shock with the estimated impulse response functions provided by the competing methods.

In all simulation experiments, it turns out that the matrix W is generally not of full rank so that we make use of a generalized inverse through a regularization method and the objective function is defined as in Eq. 3.5. Indeed one main difficulty of solving the Moore-Penrose pseudo-inverse of W without regularization stems from the fact that the matrix W has tiny positive singular values and this leads to severe numerical instability due to round-off errors and unstable behavior of the solution; these problems being more and more accurate as the frequency interval shrinks toward zero and the unreliability problem thus occurs. To control the magnitude of the smaller eigenvalues of W, the penalization parameter of the regularization matrix in Eq. 3.5 is based on a truncated singular value decomposition. More specifically, we determine the truncation number of smaller eigenvalues, say k, in the eigenvalue decomposition such that the cumulative contribution of the first k eigenvalues (factors) explains more than 99.99% of the total variability of the matrix W and thus set $\eta_T = \lambda_k$ where λ_k is the truncated minimal eigenvalue.[32] In so doing, we preserve the most important part of the information contained in the constraint functions in order to compute the optimal weighting matrix.

In the sequel, we consider two sets of Monte-Carlo simulations.[33] First, we assume that only the first structural shock has a permanent effect on the first variable. Second, we proceed with a misspecified exclusion restriction in the sense that both shocks have a permanent effect whereas we do only impose that the first structural shock matters permanently for the first variable. In both cases, we experiment different parameter configurations.[34]

In the first set of experiments, the vector of parameters $(\rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, b, \delta)$ is given by (0, 0, 0.2, 0.9, 0.2, 0) and (0, 0.3, 0.2, 0.9, 0.2, 0), respectively. In the first case, when $\rho_{1,1} = \rho_{12} = \delta = 0$, the first variable, $X_{1,t}$, is a random walk and the second variable is a persistent stationary process. Therefore the variance contribution of the second structural shock to the first variable is equal to zero irrespective of the frequency interval under consideration and, the long-run restriction is satisfied *a fortiori* irrespective of ω_n. Figure 1

[32]The optimal rule for the choice of the regularization parameter is beyond the scope of this paper.

[33]Technical appendix provides all of the technical computational details. Matlab codes are also available on request.

[34]The technical appendix provides all values of the (cumulative) mean absolute value and (cumulative) RMSE for all parameter vectors and ω_n.

reports the mean absolute bias (left panel) and RMSE (right panel) of the contemporaneous effect of each structural shock to each variable dynamics at different frequencies ω_n. Except for the LR method, note that the lowest frequency is computed for $n = 20,000$. Two points are worth commenting. First, an eye inspection of Figure 1 shows that the mean absolute bias and RMSE curves of the frequency-based approaches are generally below the solid line that represents the results of the LR approach.[35] More specifically, when focusing on the contemporaneous impact, the LR approach is clearly dominated by the first-step method and our Freq methodology in terms of both mean absolute bias and RMSE irrespective of the dynamic response of the first variable to the permanent structural shock. This also holds true when comparing the LR and DO approach, with the exception of the effect of the first structural shock on the first variable. Second, the Freq approach outperforms other methods for both statistical criteria. Notably it turns out that the mean absolute bias differences between the Freq methods and other alternatives are substantial irrespective of the frequency. In this respect, the Freq estimator leads to a significant bias reduction while being more efficient.

Regarding the cumulative absolute bias between the average response in SVARs and the true response and the cumulative RMSE up to twelve periods for $n = 30, 60, 90, 120, 180, 210, 240$, and 20,000, Figure 2 provides also supportive evidence for the Freq approach. Indeed the cumulative bias and RMSE performances of our two-step procedure are better than those of the competing approaches when studying the dynamic effect of the second variable to each structural shock. Moreover, the Freq methodology displays less (respectively, more) cumulative bias and RMSE for tiny (respectively, wider) intervals around $\omega = 0$ in the case of the cumulative response of the first variable to the second structural shock, especially relative to the DO and first-step estimator. At the same time, the magnitude of the cumulative mean absolute bias remains rather small.

[Insert Figures 1 and 2 around here]

To further contrast the different approaches, Figure 3 (respectively, Figure 4) reports the true and estimated impulse response function of the second variable relative to the first structural shock as well as the confidence intervals in the case of the frequency interval $\omega_{30} = (-\frac{2\pi}{30}, \frac{2\pi}{30})$ (respectively, $\omega_{240} = (-\frac{2\pi}{240}, \frac{2\pi}{240})$).[36]. Interestingly, the impulse responses for the LR restriction mimics the empirical results for the impact of a technology shock on hours worked when the hours series is included in level in the VAR (see Christiano, Eichenbaum and Vigfusson, 2006a). Notably the response is positive at the impact and declines toward zero, and the confidence interval contains zero at all horizons. This is also the case for the first-step and DO

[35] Tables are also provided in the Technical Appendix available upon request.

[36] Obviously, the impulse responses are the same for the LR restriction in Figure 2 and 3.

estimators that display narrower confidence intervals than the one of the `LR` method. On the other hand, the `Freq`-based impulse response function is more precise and one can reject the hypothesis that the effect of the first shock on the second variable is equal to zero up to an horizon $H = 20$ for the frequency interval ω_{30} and up to an horizon $H = 10$ for ω_{240}. It is worth noticing that a wider frequency interval sizeably increases the precision of the proposed two-step C-ALS estimator.

In order to assess the reliability of our estimator, we implement the overidentification test (Proposition 3.6) and especially its regularized version in Eq. 3.14. As explained before, one issue is the regularization of the W matrix and the fact that the test cannot be implemented if the original weighting matrix displays a large rank-deficiency (i.e., the proportion of smaller eigenvalues is too large). For instance, in the case of ω_{30} (respectively, ω_{240}), the overidentification test can be performed for 690 (respectively, 133) out of the 1,000 simulations. As reported in Panel A of Table 1, the test is conservative under the null hypothesis irrespective of the frequency interval.

[Insert Figures 3 and 4, and Table 1 around here]

To evaluate the robustness of the previous results, we now proceed with the second parameter configuration. When $\rho_{12} = .3$ and $\delta = 0$, and other parameters are left unchanged, the long-run restriction is still valid—the second structural shock does not have a permanent effect on the first variable—but the exclusion restriction does not strictly hold on the whole interval ω_n. At the same time, the proportion of the variance of the first variable explained by the second shock remains close to zero for the frequency intervals under consideration (see Panel B of Table 1) and thus this case can be interpreted as a local alternative to a standard long-run restriction à la Blanchard-Quah. For instance, in the case of ω_{30}, the second shock explains 4 % of the first variable variance and this declines to 2 % for ω_{240}. In this respect, Figure 5 and 6 provide support that our methodology clearly outperforms other methods in terms of (cumulative) mean absolute bias and (cumulative) RMSE. Again here, the reduction of both the bias and the RMSE is quite substantial. With the exception of the effect of the first shock to the first variable, the Freq estimator is dominated by the `LR` restriction and the first-step estimators in term of bias while the Freq estimator is doing slightly better in terms of RMSE. Moreover, as already observed in the benchmark parameter vector, Figures 7 and 8 show that the impulse response function of the first structural shock to the second variable are much closer to the true ones irrespective of the frequency interval under consideration (e.g., ω_{30} or ω_{240}), and that the `Freq`-based method leads to narrow confidence bands at the impact. Finally, Panel B of Table 1 shows that the overdentification test has some power to reject an alternative close to the null when the number of principal components is superior to one.

[Insert Figures 5 to 8 around here]

We now turn to the second set of Monte Carlo simulations in which the long-run identifying restriction is misspecified in the sense that the second structural shock on top of the first one has some impacts on the first variable—the contribution of the second shock being driven by δ. As a first parameter combination, it is assumed that $\rho_{12} = .3$ and $\delta = 0.1$, and other parameters are left unchanged. As a result, the variance contribution of the second structural shock on the first variable ranges between 32% over 30 quarters and 41% over 240 quarters (Table 1). Looking at Figures 9 and 10, three points are worth commenting. First, the `Freq` estimator strongly dominates the first-step, `DO` and `LR` approaches and displays a small (cumulative) mean absolute bias for $n \leq 150$ meanwhile the first-step and `DO` estimators are only slightly improving relative to the standard `LR` approach. Notably, as to be expected, the mean absolute bias of the two-step C-ALS estimator is close to zero for the widest interval ω_{30} while it increases with decreasing intervals. This comes from the fact that the variance contribution of the second shock to the first variable augments when the frequency interval becomes smaller and smaller and that the (regularized) minimization problem of the `Freq` estimator seeks to find the optimal linear combination of the reduced-form shocks such that the contribution of the second structural shock to the first variable is minimized. Second, the bias reduction of the `Freq` estimator is achieved with a lower (cumulative) RMSE relative to other estimators. Finally, all of these results are robust irrespective of the horizon H.

[Insert Figures 9 and 12 around here]

In this respect, the `Freq` estimator matches better the true impulse response function of the first structural shock on the second variable. Indeed, as illustrated in Figure 11 for ω_{30}, the discrepancy between the true impulse response function and the one obtained from the `Freq` estimator is rather small whereas those of other methods display a significant bias at very short horizons and are below the lower bound of the confidence interval after a few quarters. In contrast, for ω_{240}, the Freq estimator is also outside the confidence interval for the first four periods but it remains still more accurate than the other estimators (Figure 12). This relative good performance of our proposed estimator when moment conditions are violated might come from the fact that the minimization solution yields a pseudo-true value that is the closest to the true value and that the contribution of the second structural is minimized. Finally, Panels C and D of Table 1 show that the J-test has good power properties for this case. As a robustness check, we assume that $\delta = 0.05$ (Figure 13) and then $\delta = 0.2$ (Figure 14). When $\delta = 0.05$ (respectively, $\delta = 0.2$), the contribution of the second structural shock on the first variable is reduced (respectively, reinforced), we do observe the same qualitative results and similar power performances for ω_{30}.

To summarize, our Monte Carlo simulations provide evidence that the two-step C-ALS estimator outperforms the other methods in terms of both (cumulative) mean absolute bias and RMSE. Contrasting the true impulse responses with those of the competing methods shows also that the two-step C-ALS estimator is more reliable and precise. At the same time, the proposed J-test behaves nicely in the presence of local alternatives and misspecified identifying restrictions.

6 Application

In this section, we provide an application regarding the technology-hours debate in light of the contribution of Francis and Ramey (2009). Indeed structural vector autoregressive models yield conflicting results on the effect of technology shocks on hours, generally due to the assumed data generating process for the hours (per capita) measures. As a consequence, the predominant role of technology shock as the main source behind movements in macro data has been sharply challenged since the appraisal of Gali (1999). On the one hand, using a first difference specification of the hours measure, structural VAR models predict a decline of hours in response to a positive technological shock (e.g., Galí (1999) or Francis and Ramey, 2005), opposite of that implied by Real Business Cycles models.[37] On the other hand, entered in level, hours rise in response to a positive technological shock and the standard result at the core of the long-standing RBC model emerges (Christiano et al, 2006a). To go one step further, Francis and Ramey (2009) argue that one potential explanation of these conflicting results is that the standard measure of hours per capita and productivity have significant low-frequency movements and these movements may conduct to misleading results in the level-based specification of a structural VAR model.

More specifically, Francis and Ramey (2009) show that demographic trends and sectoral allocation are important sources of low-frequency movements in hours worked and labor productivity. Consequently, labor productivity might be driven by two permanent shocks, the technology shock and the demographic shock, and thus the usual long-run restriction of hours-productivity VAR models might be violated. To circumvent this problem, Francis and Ramey (2009) propose using new measures of hours worked per capita and labor productivity that are more comfortable with the imposed long-run restriction(s). Taking the adjusted series, it turns out that hours worked now respond negatively in the short run, and then become slightly positive after a year for a structural VAR model in which the adjusted hours worked per capita series is included

[37] While standard unit root tests can not reject the presence of an unit root for hours worked series, most dynamic macroeconomic models with standard preference specifications imply that the hours worked per capita should be stationary in the absence of permanent structural changes in government spending, labor income taxes and preferences (see Francis and Ramey, 2009).

in level.[38] In this respect, a more complete test of their results asks: (1) Is there any evidence that only technology shocks have a long-run effect on labor productivity using unadjusted hours and productivity measures? If not, How effective is the technology shock identified with the adjusted series?

To this end, we conduct structural bivariate VAR analysis in which the first variable is labor productivity and the second variable (in level) is subsequently the standard hours per capita measure (private business hours per capita) and the adjusted hours series constructed by Francis and Ramey (2009). Starting from the two-step procedure defined in Section 4.2. (Proposition 4.3.), we implement the overidentification test to assess the reliability of the identifying restrictions. In so doing, we proceed as follows. We estimate reduced-form VAR models in which the hour series enters in level. As in Francis and Ramey (2009), the sample period is 1948Q1-2007Q4 and the lag order is set to 4. To compare the results of our approach with those of Francis and Ramey (2009), the identifying constraints are imposed over the frequency interval $\omega_{240} = (-\frac{2\pi}{240}, \frac{2\pi}{240})$. Finally, standard error bands are 95 percent confidence bands based on bootstrap standard errors with 1,000 replications.

As reported in Figure 15, using the standard `LR` restriction, Panel A shows that (unadjusted) private business hours per capita respond significantly, with the exception of the initial period, and positively in the short run to a positive technological shock and then decreases at intermediate to long horizons. In contrast, using our approach, (unadjusted) private business hours per capita initially decrease, and then respond positively in the short run (after one year) before gradually decreasing toward zero in the medium-to-long term. Moreover none of the effect of the technological shock is statistically different from zero. As pointed out by Francis and Ramey (2009), one explanation of this apparent discrepancy is that the identifying assumption, namely the technological shock does explain alone the long-run effect on labor productivity for the unadjusted hours series, is misspecified. To shed some light on this issue, we perform our identification test and find that the J_T statistic, which is equal to 19.23, is greater than the simulated critical values 6.21 and 7.75, respectively at 5% and 10%. Consequently, our proposed overidentification test clearly rejects the hypothesis that there exists only a shock that have a permanent effect on the labor productivity when using the unadjusted series of hours.

[Insert Figure 15 around here]

On the other hand, as reported in Panel B of Figure 15, both methods lead to the same shape of the impulse response function, with the exception of the initial effect, when the VAR specification contains the adjusted

[38] Results are also consistent with a first-difference specification.

series of hours. More specifically, there is a statistically significant negative effect of the technological shock on (adjusted) hours worked over the first periods in the case of our methodology whereas those effects are not statistically different from zero using the standard LR method. Note that the LR results of Panels A and B are consistent with those of Francis and Ramey (2009). Interestingly, the J_T statistic, which is now equal to 1.27, is lower than the simulated critical values at 5% (5.87) and 10% (5.33). Say differently, this provides some support of the argument of Francis and Ramey (2009): the adjusted hours worked series for demographic and sectoral changes is now compatible with the hypothesis that only the technology shock has long-run effect on labor productivity.

Therefore to answer our two questions, the evidence that only the technological shock has a long-run effect on labor productivity is weak and correcting the hours series for demographic and sectoral changes is more consistent with the usual long-run restriction and leads to a negative effect of a technological shock.

7 Conclusion

In this paper, we propose a new identification scheme and the corresponding estimation method using frequency interval restrictions for structural VAR models. In so doing, the usual long-run identifying restrictions (i.e., when $\omega = 0$) have been replaced by those on $\omega = 0$ and its neighborhood, i.e. $\omega \in (\underline{\omega}, \overline{\omega})$ so that the resulting set of restrictions now has a measure strictly greater than zero in the frequency domain (Faust and Lepper (1997), Hauser and al. (1999) and Pötscher (2002)). Using the the methodology of Carrasco and Florens (2000) and Carrasco and al. (2004), we derive a continuum asymptotic least squares estimator, the C-ALS estimator, that allows to obtain reliable estimates of the dynamic responses of macroeconomic variables to structural shocks and also to assess formally the relevance of the imposed restrictions over either a given set of frequencies or a data-driven selected interval. Monte Carlo simulations argue in favor of our approach with respect to competing methods. Finally, our application regarding the hours-productivity debate provides some new insights and, especially, the relevance of the argument of Francis and Ramey (2009).

From an empirical point of view, interesting avenues of research might concern the identification and reliability of news' shocks (Beaudry and Portier, 2006; Barsky and Sims, 2011; Beaudry et al., 2013), the assessment of the long-run neutrality (super-neutrality) of money or the long-run Fisher relation, or the effect of technological shocks in m-variate SVAR specifications (Christiano et al., 2006a). On the other hand, the derivation of the optimal rule for the choice of the regularization parameter and the extension to SVAR models with integrated and cointegrated variables (Lütkepohl, 2007; Lütkepohl and Velinov, 2014) deserve, among others, some future attention from a theoretical point of view.

Appendix 1: Assumptions

Assumption A.1. The stochastic process Z_t is a $N \times 1$-vector of random variables. Z_t is stationary and α-mixing with coefficients α_j that satisfy $\sum_{j=1}^{\infty} j^2 \alpha_j < \infty$. The true unknown probability distribution of the $\{Z_t\}_{t=1}^{T}$ is denoted P_0 which belongs to a family of probability distribution.

Assumption A.2. φ is the p.d.f. of a distribution that is absolutely continuous with respect to Lebesgue measure on R^d and admits all its moments. $\varphi(\tau) > 0$ for all $\tau \in R^d$. $L^2(\varphi, (-\pi, \pi)) \equiv L^2(\varphi)$ is the Hilbert space of complex-valued functions that are square integrable with respect to φ:

$$L^2(\pi) = \{g : R^d \to C | \int |g(\omega)|^2 \varphi(\omega) d\omega < \infty\}$$

Given this definition of the Hilbert Space, one can define $\langle ., . \rangle$ and $\|.\|$ the inner product and the norm on $L^2(\varphi)$. The inner product is

$$\langle f, h \rangle = \int f(\tau) \overline{h(\omega)} \varphi(\omega) d\omega$$

where $\overline{h(\omega)}$ denotes the complex conjugate of $h(\omega)$. If $f = (f_1, \cdots, f_N)'$ and $h = (h_1, \cdots, h_N)'$ are vectors of functions of $L^2(\pi)$, the inner product $\langle f, h' \rangle$ is the $N \times N$ matrix with (i, j) element defined by $\int f_i(\omega) \overline{h_j(\omega)} \varphi(\omega) d\omega$.

Assumption A.3. It exists a sequence of estimators $\widehat{\beta}_T$ such that

$$\widehat{\beta}_T = \arg \min_{\beta \in \mathcal{B}} Q_T(Z_T, \beta)$$

and $\beta_0 = P_0 \lim_{T \to \infty} \widehat{\beta}_T$ is the true unknown value of the instrumental parameters. Under standard regularity conditions $\sqrt{T} \left(\widehat{\beta}_T - \beta_0 \right) \xrightarrow{d} \mathcal{N}(0, \Omega)$ and $\Omega = P_0 \lim_{T \to \infty} Var\left(\sqrt{T} \widehat{\beta}_T \right)$.

Assumption A.4. The function

$$g(a_0, \beta_0, \omega) = 0$$

$\forall \omega \in R^d$, φ— almost everywhere, has a unique solution a_0 which is an interior point of $\mathcal{A}$ a compact set and a_0 and b_0 denotes the unknown value under P_0.

Assumption A.5. (i) g is a measurable function from $\mathcal{R}^d \times \mathcal{A} \times \mathcal{B}$ into C.
(ii) $g_t(a, \beta, \omega)$ is continuously differentiable with respect to a and β and $g(a, \beta, \omega) \in L_\infty(\varphi \otimes P_0)$ where $L_\infty(\varphi \otimes P_0)$ is the set of measurable bounded functions of (ω, P_0).
(iii) $\sup_{a \in \mathcal{A}} \|g(a, \beta, \omega) - g(a_0, \beta, \omega)\| = O_p\left(\frac{1}{\sqrt{T}}\right)$ for all $\beta \in \mathcal{B}$ and $\omega \in \mathcal{R}^d$.
(iv) $\sup_{a \in \mathcal{A}_0} \|\partial g(a, \beta, \omega)/\partial a' - \partial g(a_0, \beta, \omega)/\partial a'\| = O_p\left(\frac{1}{\sqrt{T}}\right)$ for all $\beta \in \mathcal{B}$ and $\omega \in \mathcal{R}^d$ where $\mathcal{A}_0$ is some neighborhood about a_0.
(v) $\sup_{\beta \in \mathcal{B}_0} \|\partial g(a, \beta, \omega)/\partial \beta' - \partial g(a, \beta_0, \omega)/\partial \beta'\| = O_p\left(\frac{1}{\sqrt{T}}\right)$ for all $a \in \mathcal{A}$ and $\omega \in \mathcal{R}^d$ where $\mathcal{B}_0$ is some neighborhood about β_0.

Assumption A.6. Let S be a nonrandom bounded linear operator defined on $\mathcal{D}(S) \subset L^2(\varphi)$. The operator S does not depend on a but may depend on a_0 and $g(a, \beta, \omega) \in \mathcal{D}(S), \forall a$ and $\forall \beta$ under P_0.

Assumption A.7. Let $N(S)$ denote the null space of S, $N(S) = f \in L^2(\varphi) | Sf = 0$. Suppose that $g(a_0, b_0, \omega) \in N(S)$ implies $g(a_0, b_0, \omega) = 0$.

Assumption A.8. Let S_T be a sequence of bounded linear operators converging to S defined on $\mathcal{D}(S_T) \subset L^2(\varphi)$. Suppose that $g(a, \widehat{\beta}_T, \omega) \in \mathcal{D}(S_T)$, $\forall a \in \mathcal{A}$ and that $\|S_T g(a, \widehat{\beta}_T, \omega)\|$ is a continuous function of a. $\partial g(a, \beta, \omega)/\partial a \in \mathcal{D}(S)$ for all $a \in \mathcal{A}$, and $g(a, \beta, \omega)/\partial b \in \mathcal{D}(S)$ for all $\beta \in \mathcal{B}_0$ under P_0.

Assumption A.9. The matrix $\|Sg(a_0, b_0, \omega)\|^2$ is positive definite and symmetric which implies that $dim(a) \leq dim(g)$.

Assumption A.10.

$$\frac{\partial g}{\partial \beta'}(a_0, \widehat{\beta}_T, \omega)\sqrt{T}(\widehat{\beta}_T - \beta_0) \Rightarrow \xi \sim \mathcal{N}(0, K)$$

as $T \to \infty$ in $L^2(\pi)$ where $N(0, K)$ is the Gaussian random element of $L^2(\pi)$ with the covariance operator $K : L^2(\pi) \to L^2(\pi)$ satisfying

$$(Kf)(\omega_1) = \int E^{P_0} k(\omega_1, \omega_2) f(\omega_2) \varphi(\omega_2) d\omega$$

for all f in $L^2(\varphi)$ where under P_0

$$k(\omega_1, \omega_2) = \frac{\partial g}{\partial \beta'}(a_0, \beta_0, \omega_1)\Omega\overline{\frac{\partial g'}{\partial \beta}(a_0, \beta_0, \omega_2)}.$$

Here $\xi \in D(S)$.

Appendix 2

Proof of Proposition 3.1

The estimator is given by

$$\widehat{a}_T = \arg\min_{a\in\mathcal{A}} \left\| S_T^{1/2} g(a,\widehat{b}_T) \right\|$$

where S_T is a sequence of random bounded linear operators. To simplify the notation, we denote $g(a,b) = g(a,b,w)$.

A mean value expansion yields:

$$g(\widehat{a}_T,\beta_0) = g(a_0,\beta_0) + \frac{\partial g}{\partial a'}(\bar{a},\beta_0)(\widehat{a}_T - a_0)$$

where $\bar{a}$ is on the line segment joining $\widehat{a}_T$ and a_0.

Differentiating the objective function with respect to a leads to:

$$\left\langle S_T^{1/2}\tfrac{\partial g}{\partial a'}\left((\widehat{a}_T,\widehat{\beta}_T\right), S_T^{1/2} g(\widehat{a}_T,\widehat{\beta}_T)\right\rangle = 0$$

$$\iff \quad \left\langle S_T^{1/2}\tfrac{\partial g}{\partial a'}\left((\widehat{a}_T,\widehat{\beta}_T\right), S_T^{1/2}\left\{ g(a_0,\beta_0) + \tfrac{\partial g}{\partial a'}(\bar{a},\widehat{\beta}_T)(\widehat{a}_T - a_0) + \tfrac{\partial g}{\partial \beta'}(\widehat{a}_T,\bar{\beta})(\widehat{\beta}_T - \beta_0)\right\}\right\rangle = 0$$

where $\bar{\beta}$ is on the line segment joining $\widehat{\beta}_T$ and β_0.

Using the linearity of the operator and $g(a_0,\beta_0) = 0$, we obtain:

$$\widehat{a}_T - a_0 = -\left\langle S_T^{1/2}\frac{\partial g}{\partial a'}(\widehat{a}_T,\widehat{\beta}_T), S_T^{1/2}\frac{\partial g}{\partial a}(\bar{a},\widehat{\beta}_T)\right\rangle^{-1}\left\langle S_T^{1/2}\frac{\partial g}{\partial a'}(\widehat{a}_T,\widehat{\beta}_T), S_T^{1/2}\frac{\partial g}{\partial \beta'}(\widehat{a}_T,\beta)(\widehat{\beta}_T - \beta_0)\right\rangle.$$

Since $\widehat{a}_T \xrightarrow{p} a_0$ (to show), $\widehat{\beta}_T \xrightarrow{p} \beta_0$ and under the assumption that $\|S_T - S\| \to 0$ in probability

$$\sqrt{T}(\widehat{a}_T - a_0) = -\left\langle S^{1/2}\frac{\partial g}{\partial a}(a_0,\beta_0), S^{1/2}\frac{\partial g}{\partial a'}(a_0,\beta_0)\right\rangle^{-1}\left\langle S^{1/2}\frac{\partial g}{\partial a'}(a_0,\beta_0), S^{1/2}\frac{\partial g}{\partial \beta'}(a_0,\beta_0)\sqrt{T}(\widehat{\beta}_T - \beta_0)\right\rangle + o_p(1)$$

Using Assumption A.10, one has

$$S_T^{1/2}\frac{\partial g}{\partial \beta'}(a_0,\widehat{\beta}_T)\sqrt{T}(\widehat{\beta}_T - \beta_0) \to Z$$

and $Z = \mathcal{N}(0, S^{1/2}K\overline{S^{1/2}}')$. Then, we get

$$\sqrt{T}(\widehat{a}_T - a_0) = -\left\langle S^{1/2}\frac{\partial g}{\partial a'}(a_0,\beta_0), S^{1/2}\frac{\partial g}{\partial a}(a_0,\beta_0)\right\rangle^{-1}\left\langle S^{1/2}\frac{\partial g}{\partial a'}(a_0,\beta_0), Z\right\rangle + o_p(1).$$

Using the previous result,

$$\left\langle S^{1/2}\frac{\partial g}{\partial a'}(a_0,\beta_0), Z\right\rangle \sim \mathcal{N}\left(0, \left\langle S^{1/2}\frac{\partial g}{\partial a'}(a_0,\beta_0), (S^{1/2}K\overline{S^{1/2}}')S^{1/2}\frac{\partial g}{\partial a'}(a_0,\beta_0)\right\rangle\right).$$

The result for the asymptotic distribution for a given sequence of of random linear operators S_T follows.

Proof of Proposition 3.2

Since K is bounded in our framework, we can consider the following estimator of K_T^{-1}:

$$\left\langle K_T^{-1}f\right\rangle(\omega_1) = \sum_{i=1}^{p}\frac{1}{\mu_{i,T}}\gamma_{i,T}(\omega_1)\left\langle f'\gamma_{i,T}\right\rangle$$

where eigenvalues and eigenfunctions can be computed by solving a linear system. Indeed let $(\widehat{K}_T f)(\omega_1)$ denote

$$(\widehat{K}_T f)(\omega_1) = \left(\sum_{l=1}^{J}\int \widehat{k}_{jl,T}(\omega_1,\omega_2) f_l(\omega_2)\varphi(\omega_2)d\omega_2\right)_{j=1,\ldots,J}$$

with

$$\widehat{k}_{jl,T}(\omega_1,\omega_2)=\frac{\partial g_j}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_1)\Omega_T\overline{\frac{\partial g_l'}{\partial\beta}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)}$$

where $\widehat{a}_T^1$ is a consistent first step estimator of a_0. Then $(\widehat{K}_T f)(\omega_1)$ can be written as:

$$(\widehat{K}_T f)(\omega_1) \quad = \quad \frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_1)\Omega_T^{1/2}\int\Omega_T^{1/2}\overline{\frac{\partial g'}{\partial\beta}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)}f(\omega_2)\varphi(\omega_2)d\omega_2.$$

In this case, $R(K_T)$ is the space spanned by $\frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)\Omega_T^{1/2}$ with rank at most equals to p. The eigenfunction γ_p is necessarily of the form $\frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)\Omega_T^{1/2}D_i$ where the matrix D_i is of dimension $p\times 1$ and $D=[D_1 \;\; D_2 \;\; \ldots \;\; D_p]$ where D is of dimension $p\times p$. For the matrix of eigenfunctions $\Gamma=[\gamma_1 \;\; \ldots\gamma_p]$, we have

$$\begin{aligned}< K_T\gamma > (\omega_1) \quad &= \quad \frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_1)\Omega_T^{1/2}\int\Omega_T^{1/2}\overline{\frac{\partial g'}{\partial\beta}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)}\frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)\Omega_T^{1/2}D\varphi(\omega_2)d\omega_2\\ &= \quad \frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_1)\Omega_T^{1/2}D\Gamma.\end{aligned}$$

The matrix D and Γ are then the eigenvectors and the eigenvalues associated to:

$$\int\Omega_T^{1/2}\overline{\frac{\partial g'}{\partial\beta}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)}\frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)\Omega_T^{1/2}\varphi(\omega_2)d\omega_2 D=D\Gamma.$$

More specifically, the eigenvectors $D_h, h=1,\ldots,p$ and the corresponding eigenvalues γ_h solve the following system of p equations:

$$\int\Omega_T^{1/2}\overline{\frac{\partial g'}{\partial\beta}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)}\frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)\Omega_T^{1/2}\pi(\omega_2)d\omega_2 D_h=\gamma_h D_h.$$

In particular, the solutions $D=[D_1 \;\; D_2 \;\; \ldots \;\; D_p]$ and Γ are the eigenvectors and eigenvalues of the $p\times p$ matrix defined by:

$$\int\Omega_T^{1/2}\overline{\frac{\partial g'}{\partial\beta}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)}\frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_2)\Omega_T^{1/2}\varphi(\omega_2)d\omega_2,$$

and the eigenvectors γ^h are then given by $\frac{\partial g}{\partial\beta'}(\widehat{a}_T^1,\widehat{\beta}_T,\omega_1)\Omega_T^{1/2}D_h$.

Proof of Proposition 3.3

After imposing that $S^{1/2}=K^{-1/2}$, one obtains

$$\left\langle K^{-1/2}\frac{\partial g}{\partial a'}(a_0,\beta_0),Z\right\rangle\sim\mathcal{N}\left(0,\left\langle\frac{\partial g}{\partial a'}(a_0,\beta_0),K^{-1}\frac{\partial g'}{\partial a}(a_0,\beta_0)\right\rangle\right).$$

Consequently, using $S^{1/2}=K^{-1/2}$,

$$\begin{aligned}\lim_{T\to\infty}Var\left(\sqrt{T}(\widehat{a}_T-a_0)\right) \quad &= \quad \left\langle\frac{\partial g}{\partial a'}(a_0,\beta_0),K^{-1}\frac{\partial g'}{\partial a}(a_0,\beta_0)\right\rangle^{-1}\left\langle\frac{\partial g}{\partial a'}(a_0,\beta_0),K^{-1}\frac{\partial g'}{\partial a}(a_0,\beta_0)\right\rangle\\ &\quad\times\left\langle\frac{\partial g}{\partial a'}(a_0,\beta_0),K^{-1}\frac{\partial g'}{\partial a}(a_0,\beta_0)\right\rangle^{-1}\\ &= \quad \left\|\frac{\partial g}{\partial a'}(a_0,\beta_0)\right\|_K^{-2}.\end{aligned}$$

Proof of Proposition 3.4

The C-ALS estimator is defined as the solution of the following problem:

$$\begin{aligned}\widehat{a}_T=\arg\min_{a\in\mathcal{A}}\|K_T^{-1/2}g(a,\widehat{\beta}_T,\omega)\|^2\\ \Longleftrightarrow\;\widehat{a}_T=\arg\min_{a\in\mathcal{A}}\left\langle K_T^{-1}g(a,\widehat{\beta}_T,\omega),g(a,\widehat{\beta}_T,\omega)\right\rangle.\end{aligned}$$

We can rewrite this objective function as:

$$\widehat{a}_T=\arg\min_{a\in\mathcal{A}}\left\langle K_T^{-1}g(a,\widehat{\beta}_T,\omega),K_TK_T^{-1}g(a,\widehat{\beta}_T,\omega)\right\rangle.$$

For sake of notation, $g(a, \widehat{\beta}_T, \omega) \equiv g(\omega)$. Let h_T denote $h_T(\omega) = K_T^{-1} g(\omega)$, the objective function is thus given by:

$$\left\langle h(\omega), K_T K_T^{-1} h(\omega) \right\rangle$$

where

$$\langle K_T h \rangle (\omega) = \frac{\partial g}{\partial \beta'}(\omega) \Omega_T^{1/2} \int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\omega_1)} h(\omega_1) d\omega_1.$$

This yields

$$\langle h(\omega), K_T h(\omega) \rangle = \int h(\omega_1) \frac{\partial g}{\partial \beta'}(\omega_1) \Omega_T^{1/2} d\omega_1 \int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\omega_2)} h(\omega_2) d\omega_2.$$

Using the notation,

$$b = \int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\omega)} h(\omega) d\omega,$$

the objective function is then defined by $b'b$.

After multiplying $\langle K_T h \rangle (\omega)$ by $\Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\omega_1)}$ and integrating, one obtains:

$$\int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\omega)} \frac{\partial g}{\partial \beta'}(\omega) \Omega_T^{1/2} d\omega \int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\omega_1)} h(\omega_1) d\omega_1$$

$$= \int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\omega_1)} g(\omega_1) d\omega = s.$$

using $K_T h(\omega) = g(\omega)$. Denoting $W = \int \Omega_T^{1/2} \overline{\frac{\partial g'}{\partial \beta}(\omega)} \frac{\partial g}{\partial \beta'}(\omega) \Omega_T^{1/2} d\omega$, we obtain: $Wb = s$. Now suppose that there exists a generalized inverse of the matrix W denoted $\widetilde{W}$ such that $\widetilde{W}W = I_p$. Then $b = \widetilde{W}s$ and the objective function can be rewritten as $s'\widetilde{W}^2 s$. This provides the result. When W is of rank equal to p, then $\widetilde{W} = W^{-1}$.

Proof of Proposition 3.7

Suppose there exists a lower bound ω_{lb} such that for this lower bound $J_T(\omega_{lb}) = O_p(1)$. The restrictions are then asymptotically valid for the interval $(-\omega_{lb}, \omega_{lb})$. Now, there exists two possible cases for which $|\omega| \neq |\omega_0|$. First, consider the case where $|\omega| > |\omega_0|$. For this case, $J_T(\omega) \to \infty$ while $J_T(\omega_0) = O_p(1)$. Thus $VISC_T(\omega_0) - VISC_T(\omega) \xrightarrow{p} -\infty$. The criterion selects the interval $(-\omega_0, \omega_0)$ with a probability going to one when T is going to ∞. For the second case, $|\omega| < |\omega_0|$ which implies that both $J_T(\omega)$ and $J_T(\omega_0)$ are $O_p(1)$. Since $|\omega| < |\omega_0|$ and by Assumption 4.3, $-h(|\omega_0|\kappa_T + h(|\omega_0|\kappa_T \to -\infty$ which implies $VISC_T(\omega_0) - VISC_T(\omega) \xrightarrow{p} -\infty$. By combining the two results, the criterion selects ω_0 with a probability going to one when T diverges toward ∞ for all $\omega \neq \omega_0$.

Proof of Proposition 4.1

It is worth noting that:

$$g(a, \hat{\beta}_T, \omega) = L\text{vec}\left(\widehat{C}(e^{-i\omega}) \widehat{\Sigma}_u \overline{\widehat{C}(e^{-i\omega})}' - \widehat{C}(e^{-i\omega}) A_0 A_0' \overline{\widehat{C}(e^{-i\omega})}' \right)$$

where L denotes the following $\left(\frac{K(K+1)}{2} \times K^2 \right)$ elimination matrix:

$$\begin{bmatrix} I_K & 0_{K,K} & \cdots & 0_{K \times K} \\ 0_{K-1 \times K} & [0_{K-1 \times 1}, I_{K-1}] & \cdots & 0_{K-1 \times K} \\ \vdots & \ddots & \ddots & \vdots \\ 0_{1 \times K} & 0_{1 \times K} & \cdots & [0_{1 \times K-1}, 1] \end{bmatrix}.$$

Therefore,

$$\begin{aligned} \frac{\partial g}{\partial \alpha'}(a, \hat{\beta}_T, \omega) &= \frac{\partial}{\partial \alpha'} \left(L\text{vec}\left(\widehat{C}(e^{-i\omega}) \widehat{\Sigma}_u \overline{\widehat{C}(e^{-i\omega})}' - \widehat{C}(e^{-i\omega}) A_0 A_0' \overline{\widehat{C}(e^{-i\omega})}' \right) \right) \\ &= L \frac{\partial}{\partial \alpha'} \text{vec}\left(\widehat{C}(e^{-i\omega}) \widehat{\Sigma}_u \overline{\widehat{C}(e^{-i\omega})}' - \widehat{C}(e^{-i\omega}) A_0 A_0' \overline{\widehat{C}(e^{-i\omega})}' \right) \end{aligned}$$

i.e.,

$$\frac{\partial g}{\partial \alpha'}(a,\hat{\beta}_T,\omega) \quad = \quad L\left[\frac{\partial}{\partial \alpha'}\text{vec}\left(\widehat{C}(e^{-i\omega})\widehat{\Sigma}_u\overline{\widehat{C}(e^{-i\omega})}'\right) - \frac{\partial}{\partial \alpha'}\text{vec}\left(\widehat{C}(e^{-i\omega})A_0A_0'\overline{\widehat{C}(e^{-i\omega})}'\right)\right].$$

Using product rules for differentiation, the first right-hand side term in brackets is given by:

$$\frac{\partial}{\partial \alpha'}\text{vec}\left(\widehat{C}(e^{-i\omega})\widehat{\Sigma}_u\overline{\widehat{C}(e^{-i\omega})}'\right) = \left(I_K \otimes \widehat{C}(e^{-i\omega})\Sigma_u\right)\frac{\partial \text{vec}\left(\overline{C(e^{-i\omega})}'\right)}{\partial \alpha'} + \left(\widehat{C}(e^{-i\omega})\Sigma_u \otimes I_K\right)\frac{\partial \text{vec}\left(\widehat{C}(e^{-i\omega})\right)}{\partial \alpha'}$$

with

$$\frac{\partial \text{vec}\left(\overline{C(e^{-i\omega})}'\right)}{\partial \alpha'} = P_{K,K}\frac{\partial \text{vec}\left(\widehat{C}(e^{-i\omega})\right)}{\partial \alpha'}$$

where $P_{K,K}$ is a commutation matrix $(K^2 \times K^2)$ given by $P_{K,K} = \sum_{i=1}^{K}\left(\iota_i \otimes I_K \otimes \iota_i'\right)$, with ι_i denotes the i-th column of I_K. Since

$$\begin{aligned}
\frac{\partial \text{vec}(\widehat{C}(e^{-i\omega}))}{\partial \alpha'} &= \frac{\partial \text{vec}(\widehat{C}(e^{-i\omega}))}{\partial \text{vec}(C(e^{-i\omega})^{-1})'}\frac{\partial \text{vec}(C(e^{-i\omega})^{-1})}{\partial \alpha'} \\
&= -\left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)\frac{\partial}{\partial \alpha'}\text{vec}\left(I_K - \Phi_1 - \cdots \Phi_p\right) \\
&= \sum_{i=1}^{p}\left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)\frac{\partial \text{vec}(\Phi_i)}{\partial \alpha'}
\end{aligned}$$

and

$$\frac{\partial \text{vec}(\Phi_i)}{\partial \beta'} \quad = \quad \begin{pmatrix} 0_{K^2\times K^2(i-1)} & I_{K^2} & 0_{K^2\times K^2(p-i)} \end{pmatrix},$$

one gets:

$$\frac{\partial \text{vec}(\widehat{C}(e^{-i\omega}))}{\partial \alpha'} = \left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)\left(e_p' \otimes I_{K^2}\right)$$

with $e_p = (1,\cdots,1)'$ is and $(p\times 1)$ vector, and $e_p' \otimes I_{K^2} = \sum_{i=1}^{p}\frac{vec(\Phi_i)}{\partial \alpha'}$. In this respect,

$$\frac{\partial}{\partial \alpha'}\text{vec}\left(\widehat{C}(e^{-i\omega})\widehat{\Sigma}_u\overline{\widehat{C}(e^{-i\omega})}'\right) = L\left[\left(I_K \otimes C(e^{-i\omega})\Sigma_u\right)P_{K,K} + \left(C(e^{-i\omega})\Sigma_u \otimes I_K\right)\right]\left(\overline{C(e^{-i\omega})}' \otimes C(e^{-i\omega})\right)\left(e_p' \otimes I_{K^2}\right)$$

Using the previous expression, the second right-hand side term in brackets is given by:

$$\begin{aligned}
\frac{\partial}{\partial \alpha'}\text{vec}\left(\widehat{C}(e^{-i\omega})A_0(a)A_0'(a)\overline{\widehat{C}(e^{-i\omega})}'\right) &= L\left[\left(I_K \otimes C(e^{-i\omega})A_0(a)A_0'(a)\right)P_{K,K} + \left(C(e^{-i\omega})A_0(a)A_0'(a) \otimes I_K\right)\right] \\
&\quad \times \left(\overline{C(e^{-i\omega})}' \otimes C(e^{-i\omega})\right)\left(e_p' \otimes I_{K^2}\right).
\end{aligned}$$

Finally,

$$\frac{\partial g}{\partial \alpha'}(a,\hat{\beta}_T,\omega) \quad = \quad L\left\{\left[I_K \otimes C(e^{-i\omega})\left(\Sigma_u - A_0A_0'\right)\right]P_{K,K} + C(e^{-i\omega})\left(\Sigma_u - A_0A_0'\right)\otimes I_K\right\}\left(\overline{C(e^{-i\omega})}' \otimes C(e^{-i\omega})\right)\left(e_p' \otimes I_{K^2}\right).$$

On the other hand,

$$\frac{\partial g}{\partial \sigma'}(a,\hat{\beta}_T,\omega) \quad = \quad L\frac{\partial}{\partial \sigma'}\text{vec}\left(\widehat{C}(e^{-i\omega})\Sigma_u\overline{\widehat{C}(e^{-i\omega})}'\right)$$

with:

$$\begin{aligned}
\frac{\partial}{\partial \sigma'}\text{vec}\left(\widehat{C}(e^{-i\omega})\Sigma_u\overline{\widehat{C}(e^{-i\omega})}'\right) &= \left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)\frac{\partial}{\partial \sigma'}\text{vec}(\Sigma_u) \\
&= \left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)D_K\frac{\partial}{\partial \sigma'}\text{vech}(\Sigma_u)
\end{aligned}$$

where D_K is an $(K^2 \times \frac{K(K+1)}{2})$ duplication matrix. Therefore,

$$\begin{aligned}
\frac{\partial}{\partial \sigma'}\text{vec}\left(\widehat{C}(e^{-i\omega})\Sigma_u\overline{\widehat{C}(e^{-i\omega})}'\right) &= \left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)D_K I_{K(K+1)/2} \\
&= \left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)D_K
\end{aligned}$$

and

$$\frac{\partial g}{\partial \sigma'}(a, \hat{\beta}_T, \omega) \quad = \quad L\left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right) D_K.$$

The partial derivative of g with respect to a' is given by:

$$\begin{aligned}
\frac{\partial g}{\partial a'}(a, \hat{\beta}_T, \omega) &= -L\frac{\partial}{\partial a'}vec\left(\widehat{C}(e^{-i\omega})A_0A_0'\overline{\widehat{C}(e^{-i\omega})}'\right) \\
&= -L\left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)\frac{\partial}{\partial a'}vec\left(A_0A_0'\right) \\
&= -L\left(\overline{\widehat{C}(e^{-i\omega})}' \otimes \widehat{C}(e^{-i\omega})\right)\left[(I_K \otimes A_0)\frac{\partial vec(A_0')}{\partial a'} + (A_0 \otimes I_K)\frac{\partial vec(A_0)}{\partial a'}\right]
\end{aligned}$$

with:

$$\begin{aligned}
\frac{\partial}{\partial a'}vec\left(A_0\right) &= I_{K^2} \\
\frac{\partial}{\partial a'}vec\left(A_0'\right) &= P_{K,K}\frac{\partial}{\partial \beta'}vec(A_0)
\end{aligned}$$

and $P_{K,K}$ is defined above. □

Proof of Proposition 4.3

The objective function of the C-ALS problem of a bivariate VAR model is based on the W and $\underline{s}(\cdot,\cdot)$ matrices, with

$$\underline{s}(\tilde{a}_{12}, \hat{\beta}_T,) = \int_{\underline{\omega}}^{\overline{\omega}} \sum_{k=0}^{\infty}\left[\left(\frac{\partial \widehat{C}_{11,k}}{\partial \beta'}e^{i\omega k}\widehat{a}^1_{12,T} + \frac{\partial \widehat{C}_{12,k}}{\partial \beta'}e^{i\omega k}\right)\Omega^{1/2}\right]' \sum_{j=0}^{\infty}\left[\widehat{C}_{11,j}e^{-i\omega j}\tilde{a}_{12} + \widehat{C}_{12,j}e^{-i\omega j}\right]d\omega$$

$$= \int_{\underline{\omega}}^{\overline{\omega}} \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\widehat{a}^1_{12,T}\frac{\partial \widehat{C}_{11,k}}{\partial \beta'}\Omega^{1/2}\right)'\widehat{C}_{11,j}\tilde{a}_{12} + \left(\widehat{a}^1_{12,T}\frac{\partial \widehat{C}_{11,k}}{\partial \beta'}\Omega^{1/2}\right)'\widehat{C}_{12,j} + \left(\frac{\partial \widehat{C}_{12,k}}{\partial \beta'}\Omega^{1/2}\right)'\widehat{C}_{11,j}\tilde{a}_{12} + \left(\frac{\partial \widehat{C}_{12,k}}{\partial \beta'}\Omega^{1/2}\right)'\widehat{C}_{12,j}\right] \times \exp((k-j)\omega)d\omega$$

$$= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\widehat{a}^1_{12,T}\frac{\partial \widehat{C}_{11,k}}{\partial \beta'}\Omega^{1/2}\right)'\widehat{C}_{11,j}\tilde{a}_{12} + \left(\widehat{a}^1_{12,T}\frac{\partial \widehat{C}_{11,k}}{\partial \beta'}\Omega^{1/2}\right)'\widehat{C}_{12,j} + \left(\frac{\partial \widehat{C}_{12,k}}{\partial \beta'}\Omega^{1/2}\right)'\widehat{C}_{11,j}\tilde{a}_{12} + \left(\frac{\partial \widehat{C}_{12,k}}{\partial \beta'}\Omega^{1/2}\right)'\widehat{C}_{12,j}\right] \times \int_{\underline{\omega}}^{\overline{\omega}} \cos((j-k)\omega)d\omega$$

and

$$\begin{aligned}
\widehat{W}_T &= \int_{\underline{\omega}}^{\overline{\omega}} \sum_{k=0}^{\infty}\left[\left(\frac{\partial \widehat{C}_{11,k}}{\partial \beta'}e^{i\omega k}\widehat{a}^1_{12,T} + \frac{\partial \widehat{C}_{12,k}}{\partial \beta'}e^{i\omega k}\right)\Omega^{1/2}\right]' \sum_{j=0}^{\infty}\left[\left(\frac{\partial \widehat{C}_{11,j}}{\partial \beta'}e^{i\omega k}\widehat{a}^1_{12,T} + \frac{\partial \widehat{C}_{12,j}}{\partial \beta'}e^{i\omega k}\right)\Omega^{1/2}\right]d\omega \\
&= \int_{\underline{\omega}}^{\overline{\omega}} \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\left(\frac{\partial \widehat{C}_{11,k}}{\partial \beta'}\widehat{a}^1_{12,T} + \frac{\partial \widehat{C}_{12,k}}{\partial \beta'}\right)\Omega^{1/2}\right)'\left(\left(\frac{\partial \widehat{C}_{11,j}}{\partial \beta'}\widehat{a}^1_{12,T} + \frac{\partial \widehat{C}_{12,j}}{\partial \beta'}\right)\Omega^{1/2}\right)\right]\exp((k-j)\omega)d\omega \\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\left(\frac{\partial \widehat{C}_{11,k}}{\partial \beta'}\widehat{a}^1_{12,T} + \frac{\partial \widehat{C}_{12,k}}{\partial \beta'}\right)\Omega^{1/2}\right)'\left(\left(\frac{\partial \widehat{C}_{11,j}}{\partial \beta'}\widehat{a}^1_{12,T} + \frac{\partial \widehat{C}_{12,j}}{\partial \beta'}\right)\Omega^{1/2}\right)\right]\int_{\underline{\omega}}^{\overline{\omega}} \cos((j-k)\omega)d\omega
\end{aligned}$$

.

and the last equality holds by the symmetry of the interval around zero. The objective function can be then rewritten as:

$$\underline{s}(\tilde{a}_{12}, \widehat{\beta}_T)'\left(\widehat{W}^2\right)^{-1}\overline{\underline{s}(\tilde{a}_{12}, \widehat{\beta}_T)} = (\widehat{A}_T\tilde{a}_{12} + \widehat{B}_T)'(\widehat{W}_T^2)^{-1}(\widehat{A}_T\tilde{a}_{12} + \widehat{B}_T) \tag{7.21}$$

where $\widehat{A}_T$, $\widehat{B}_T$ and $\widehat{W}_T$ are given by:

$$\widehat{A}_T = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\frac{\widehat{a}^1_{12,T}\partial\widehat{C}_{11,k}}{\partial \beta'}\widehat{\Omega}^{1/2}\right)'\widehat{C}_{11,j} + \left(\frac{\partial \widehat{C}_{12,k}}{\partial \beta'}\widehat{\Omega}^{1/2}\right)'\widehat{C}_{11,j}\right]\int_{\underline{\omega}}^{\overline{\omega}} \cos((j-k)\omega)d\omega$$

$$\widehat{B}_T = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\frac{\widehat{\tilde{a}}^1_{12,T}\partial\widehat{C}_{11,k}}{\partial\beta'}\widehat{\Omega}^{1/2}\right)'\widehat{C}_{12,j} + \left(\frac{\partial\widehat{C}_{12,k}}{\partial\beta'}\widehat{\Omega}^{1/2}\right)'\widehat{C}_{12,j}\right]\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega,$$

and

$$\widehat{W}_T = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\left[\left(\left(\frac{\partial\widehat{C}_{11,k}}{\partial\beta'}\widehat{\tilde{a}}^1_{12,T} + \frac{\partial\widehat{C}_{12,k}}{\partial\beta'}\right)\widehat{\Omega}^{1/2}\right)'\left(\widehat{\Omega}^{1/2}\left(\frac{\partial\widehat{C}_{11,j}}{\partial\beta'}\widehat{\tilde{a}}^1_{12,T} + \frac{\partial\widehat{C}_{12,j}}{\partial\beta'}\right)\right)\right]\int_{\underline{\omega}}^{\overline{\omega}}\cos((j-k)\omega)d\omega.$$

The minimizer of the objective function 7.21 is given by $\widehat{\tilde{a}}_{12,T} = -\frac{\widehat{A}'_T(\widehat{W}^2_T)^{-1}\widehat{B}_T}{\widehat{A}'_T(\widehat{W}^2_T)^{-1}\hat{A}_T}$. □

Appendix 3

In the Appendix, we derive the explicit expression for the matrices $E[C(e^{-i\omega})\varepsilon_t\varepsilon_t'\overline{C(e^{-i\omega})}' = C(e^{-i\omega})\Sigma\overline{C(e^{-i\omega})}'$. We first consider the case with three variables: Let us denote

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix}$$

and

$$A(0) = \begin{pmatrix} A^0_{11} & A^0_{12} & A^0_{13} \\ A^0_{21} & A^0_{22} & A^0_{23} \\ A^0_{31} & A^0_{32} & A^0_{33} \end{pmatrix}.$$

It follows that:

$$\widetilde{C} \equiv C(e^{-i\omega})A(0)A(0)'\overline{C(e^{-i\omega})}' = \begin{bmatrix} \widetilde{C}_{11} & \widetilde{C}_{12} & \widetilde{C}_{13} \\ \widetilde{C}_{21} & \widetilde{C}_{22} & \widetilde{C}_{23} \\ \widetilde{C}_{31} & \widetilde{C}_{32} & \widetilde{C}_{33} \end{bmatrix}$$

The nine elements of the $\widetilde{C}$ matrix are given by

$$\begin{aligned}
\widetilde{C}_{11} &= \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{1j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{jk}\big), \widetilde{C}_{12} = \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{1j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{2j}(e^{i\omega})A^0_{jk}\big) \\
\widetilde{C}_{13} &= \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{1j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{3j}(e^{i\omega})A^0_{jk}\big), \widetilde{C}_{21} = \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{2j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{jk}\big) \\
\widetilde{C}_{22} &= \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{2j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{2j}(e^{i\omega})A^0_{jk}\big), \widetilde{C}_{23} = \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{2j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{3j}(e^{i\omega})A^0_{jk}\big) \\
\widetilde{C}_{31} &= \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{3j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{jk}\big), \widetilde{C}_{32} = \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{3j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{2j}(e^{i\omega})A^0_{jk}\big) \\
\widetilde{C}_{33} &= \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{3j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{3j}(e^{i\omega})A^0_{jk}\big).
\end{aligned}$$

More generally, one can write (for m variables):

$$\widetilde{C} = [\widetilde{C}_{pq}]_{1\le p,q\le m}$$

where

$$\widetilde{C}_{pq} = \sum_{k=1}^{m}\big(\sum_{j=1}^{m} C_{pj}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{m} C_{qj}(e^{i\omega})A^0_{jk}\big).$$

In the same respect, the nine elements of the $\widehat{C}_t = C(e^{-i\omega})\Sigma\overline{C(e^{-i\omega})}'$ are given by

$$\begin{aligned}
\hat{C}_{11} &= \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{1j}(e^{-i\omega})\sigma_{lj}C_{1l}(e^{i\omega})\right), \hat{C}_{12} = \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{1j}(e^{-i\omega})\sigma_{lj}C_{2l}(e^{i\omega})\right) \\
\hat{C}_{13} &= \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{1j}(e^{-i\omega})\sigma_{lj}C_{3l}(e^{i\omega})\right), \hat{C}_{21} = \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{2j}(e^{-i\omega})\sigma_{lj}C_{1l}(e^{i\omega})\right) \\
\hat{C}_{22} &= \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{2j}(e^{-i\omega})\sigma_{lj}C_{2l}(e^{i\omega})\right), \hat{C}_{23} = \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{2j}(e^{-i\omega})\sigma_{lj}C_{3l}(e^{i\omega})\right) \\
\hat{C}_{31} &= \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{3j}(e^{-i\omega})\sigma_{lj}C_{1l}(e^{i\omega})\right), \hat{C}_{32} = \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{3j}(e^{-i\omega})\sigma_{lj}C_{2l}(e^{i\omega})\right) \\
\hat{C}_{33} &= \sum_{l=1}^{3}\left(\sum_{l=1}^{3} C_{3j}(e^{-i\omega})\sigma_{lj}C_{3l}(e^{i\omega})\right).
\end{aligned}$$

More generally, one can write (for m variables)

$$\hat{C} = [\hat{C}_{pq}]_{1\leq p,q\leq m}$$

where

$$\hat{C}_{pq} = \sum_{l=1}^{m}\left(\sum_{j=1}^{m} C_{pj}(e^{-i\omega})\sigma_{lj}C_{ql}(e^{i\omega})\right)$$

In the case of Proposition 4.4 with three imposed restrictions $\forall\omega \in [\underline{\omega}, \overline{\omega}]$ are:

$$\begin{aligned} C_{11}(e^{-i\omega})A^0_{12} + C_{12}(e^{-i\omega})A^0_{22} + C_{13}(e^{-i\omega})A^0_{32} &= 0 \\ C_{11}(e^{-i\omega})A^0_{13} + C_{12}(e^{-i\omega})A^0_{23} + C_{13}(e^{-i\omega})A^0_{33} &= 0. \end{aligned}$$

These constraints implies for the first colum of the matrix $\widehat{C}$:

$$\begin{aligned} \widetilde{C}_{11} &= \big(\sum_{j=1}^{3} C_{1j}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{j1}\big), \widetilde{C}_{21} = \big(\sum_{j=1}^{3} C_{2j}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{j1}\big) \\ \widetilde{C}_{31} &= \big(\sum_{j=1}^{3} C_{3j}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{j1}\big). \end{aligned}$$

More generally, in the case of m variables, one obtains

$$\widetilde{C}_{p1} = \big(\sum_{j=1}^{m} C_{kj}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{m} C_{1j}(e^{i\omega})A^0_{j1}\big)$$

for $p = 1, \ldots, m$ or equivalently

$$\widetilde{C}_{p1} = \sum_{l=1}^{m}\left(\sum_{j=1}^{m} C_{pj}(e^{-i\omega})A_{j1}A_{l1}C_{1l}(e^{i\omega})\right)$$

Therefore, we have the vector of moment conditions for $m = 3$

$$g(a, \beta, \omega) = \begin{pmatrix} g_1(a, \beta, \omega) = \widetilde{C}_{11} - \widehat{C}_{11} \\ g_2(a, \beta, \omega) = \widetilde{C}_{21} - \widehat{C}_{21} \\ g_3(a, \beta, \omega) = \widetilde{C}_{31} - \widehat{C}_{31} \end{pmatrix}$$

where

$$\begin{aligned} g_1(a, \beta, \omega) &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{1q,j}\big(A^0_{p1}A^0_{q1} - \sigma_{pq}\big)cos((k-j)\omega) \\ g_2(a, \beta, \omega) &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{2q,j}\big(A^0_{p1}A^0_{q1} - \sigma_{pq}\big)cos((k-j)\omega) \\ g_3(a, \beta, \omega) &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{3q,j}\big(A^0_{p1}A^0_{q1} - \sigma_{pq}\big)cos((k-j)\omega). \end{aligned}$$

This can be easily generalized in the case of m variables.
In the case in which we impose a lower triangular matrix for the long-run impact, $\forall\omega \in [\underline{\omega}, \overline{\omega}]$:

$$\begin{aligned} C_{11}(e^{-i\omega})A^0_{12} + C_{12}(e^{-i\omega})A^0_{22} + C_{13}(e^{-i\omega})A^0_{32} &= 0 \\ C_{11}(e^{-i\omega})A^0_{13} + C_{12}(e^{-i\omega})A^0_{23} + C_{13}(e^{-i\omega})A^0_{33} &= 0 \\ C_{21}(e^{-i\omega})A^0_{13} + C_{22}(e^{-i\omega})A^0_{23} + C_{23}(e^{-i\omega})A^0_{33} &= 0. \end{aligned}$$

These constraints can be plugged in the $\widetilde{C}$ matrix, so that we end up with the following expression (for $m=3$):

$$\begin{aligned}
\widetilde{C}_{11} &= \big(\sum_{j=1}^{3} C_{1j}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{j1}\big), \widetilde{C}_{12} = \big(\sum_{j=1}^{3} C_{1j}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{3} C_{2j}(e^{i\omega})A^0_{j1}\big)\\
\widetilde{C}_{13} &= \big(\sum_{j=1}^{3} C_{1j}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{3} C_{3j}(e^{i\omega})A^0_{j1}\big), \widetilde{C}_{21} = \big(\sum_{j=1}^{3} C_{2j}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{j1}\big)\\
\widetilde{C}_{22} &= \sum_{k=1}^{2}\big(\sum_{j=1}^{3} C_{2j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{2j}(e^{i\omega})A^0_{jk}\big), \widetilde{C}_{23} = \sum_{k=1}^{2}\big(\sum_{j=1}^{3} C_{2j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{3j}(e^{i\omega})A^0_{jk}\big)\\
\widetilde{C}_{31} &= \big(\sum_{j=1}^{3} C_{3j}(e^{-i\omega})A^0_{j1}\big)\big(\sum_{j=1}^{3} C_{1j}(e^{i\omega})A^0_{j1}\big), \widetilde{C}_{32} = \sum_{k=1}^{2}\big(\sum_{j=1}^{3} C_{3j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{2j}(e^{i\omega})A^0_{jk}\big)\\
\widetilde{C}_{33} &= \sum_{k=1}^{3}\big(\sum_{j=1}^{3} C_{3j}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{3} C_{3j}(e^{i\omega})A^0_{jk}\big).
\end{aligned}$$

More generally, one can write (for m variables):

$$\widetilde{C} = [\widetilde{C}_{pq}]_{1\leq p,q\leq m}$$

where (if $p \leq q$)

$$\widetilde{C}_{pq} = \sum_{k=1}^{p}\big(\sum_{j=1}^{m} C_{pj}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{m} C_{qj}(e^{i\omega})A^0_{jk}\big)$$

and (if $p > q$)

$$\widetilde{C}_{pq} = \sum_{k=1}^{q}\big(\sum_{j=1}^{m} C_{pj}(e^{-i\omega})A^0_{jk}\big)\big(\sum_{j=1}^{m} C_{qj}(e^{i\omega})A^0_{jk}\big)$$

Let denote $C_{ij}(e^{-i\omega}) = \sum_{k=0}^{\infty} C_{ij,k}e^{-i\omega k}$ and $C_{ij}(e^{i\omega}) = \sum_{k=0}^{\infty} C_{ij,k}e^{i\omega k}$, where $e^{-i\omega k} = cos(\omega k) - isin(\omega k)$. We can rewrite the respective elements of the $\widetilde{C}$ and $\widehat{C}$ matrices (using the previous contraint). For example, let consider the (1,1) element of the $\widetilde{C}$ matrix, one has:

$$\begin{aligned}
\widetilde{C}_{11} &= C_{11}(e^{-i\omega})C_{11}(e^{i\omega})(A^0_{11})^2 + C_{12}(e^{-i\omega})C_{12}(e^{i\omega})(A^0_{21})^2 + C_{13}(e^{-i\omega})C_{13}(e^{i\omega})(A^0_{31})^2\\
&+\big(C_{11}(e^{-i\omega})C_{12}(e^{i\omega}) + C_{12}(e^{-i\omega})C_{11}(e^{i\omega})\big)A^0_{21}A^0_{11} + \big(C_{11}(e^{-i\omega})C_{13}(e^{i\omega}) + C_{13}(e^{-i\omega})C_{11}(e^{i\omega})\big)A^0_{31}A^0_{11}\\
&+\big(C_{12}(e^{-i\omega})C_{13}(e^{i\omega}) + C_{13}(e^{-i\omega})C_{12}(e^{i\omega})\big)A^0_{31}A^0_{21}
\end{aligned}$$

with (among others) $C_{11}(e^{-i\omega})C_{11}(e^{i\omega}) = C_{11}(e^{i\omega})C_{11}(e^{-i\omega}) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{11,k}C_{11,j}cos((j-k)\omega)$ and $C_{11}(e^{-i\omega})C_{12}(e^{i\omega})+ C_{12}(e^{-i\omega})C_{11}(e^{i\omega}) = 2\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{11,k}C_{11,j}cos((j-k)\omega)$, so that (using the real part)

$$\begin{aligned}
\widetilde{C}_{11} &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{1q,j}A^0_{p1}A^0_{q1}cos((k-j)\omega), \widetilde{C}_{12} = \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{2q,j}A^0_{p1}A^0_{q1}cos((k-j)\omega)\\
\widetilde{C}_{13} &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{3q,j}A^0_{p1}A^0_{q1}cos((k-j)\omega), \widetilde{C}_{22} = \sum_{v=1}^{2}\sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{2p,k}C_{2q,j}A^0_{pv}A^0_{qv}cos((k-j)\omega)\\
\widetilde{C}_{23} &= \sum_{v=1}^{2}\sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{2p,k}C_{3q,j}A^0_{pv}A^0_{qv}cos((k-j)\omega), \widetilde{C}_{33} = \sum_{v=1}^{3}\sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{3p,k}C_{3q,j}A^0_{pv}A^0_{qv}cos((k-j)\omega).
\end{aligned}$$

This can be generalized as follows:

$$\widetilde{C} = [\widetilde{C}_{p'q'}]_{1\leq p',q'\leq m}$$

where (since $p' \leq q'$)

$$\widetilde{C}_{p'q'} = \sum_{v=1}^{p'}\sum_{p=1}^{m}\sum_{q=1}^{m}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{p'p,k}C_{q'q,j}A^0_{pv}A^0_{qv}cos((k-j)\omega).$$

We obtain similar expression for the $\widehat{C}$ matrix:

$$\begin{aligned}
\widehat{C}_{11} &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{1q,j}\sigma_{pq}cos((k-j)\omega), \widehat{C}_{12} = \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{2q,j}\sigma_{pq}cos((k-j)\omega) \\
\widehat{C}_{13} &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{3q,j}\sigma_{pq}cos((k-j)\omega), \widehat{C}_{22} = \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{2p,k}C_{2q,j}\sigma_{pq}cos((k-j)\omega) \\
\widehat{C}_{23} &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{2p,k}C_{3q,j}\sigma_{pq}cos((k-j)\omega), \widehat{C}_{33} = \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{3p,k}C_{3q,j}\sigma_{pq}cos((k-j)\omega).
\end{aligned}$$

This can also be generalized as follows

$$\widehat{C} = [\widehat{C}_{p'q'}]_{1\le p',q'\le m}$$

where (since $p' \le q'$)

$$\widehat{C}_{p'q'} = \sum_{p=1}^{m}\sum_{q=1}^{m}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{p'p,k}C_{q'q,j}\sigma_{pq}cos((k-j)\omega).$$

Therefore, we have the vector of moment conditions for $m = 3$

$$g(a,\beta,\omega) = (g_i(a,\beta,\omega))_{i=1,\cdots,9}$$

where

$$\begin{aligned}
g_1(a,\beta,\omega) &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{1q,j}\big(A^0_{p1}A^0_{q1} - \sigma_{pq}\big)cos((k-j)\omega) \\
g_2(a,\beta,\omega) &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{2q,j}\big(A^0_{p1}A^0_{q1} - \sigma_{pq}\big)cos((k-j)\omega) \\
g_3(a,\beta,\omega) &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{1p,k}C_{3q,j}\big(A^0_{p1}A^0_{q1} - \sigma_{pq}\big)cos((k-j)\omega) \\
g_4(a,\beta,\omega) &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{2p,k}C_{2q,j}\big(\sum_{v=1}^{2} A^0_{pv}A^0_{qv} - \sigma_{pq}\big)cos((k-j)\omega) \\
g_5(a,\beta,\omega) &= \sum_{p=1}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{2p,k}C_{3q,j}\big(\sum_{v=1}^{2} A^0_{pv}A^0_{qv} - \sigma_{pq}\big)cos((k-j)\omega) \\
g_6(a,\beta,\omega) &= \sum_{p=1}^{3}\sum_{q=1}^{3}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{3p,k}C_{3q,j}\big(\sum_{v=1}^{3} A^0_{pv}A^0_{qv} - \sigma_{pq}\big)cos((k-j)\omega) \\
g_7(a,\beta,\omega) &= \sum_{k=0}^{\infty}\sum_{p=1}^{3}\big(C_{1p,k}A^0_{p2}\big)cos(k\omega), g_8(a,\beta,\omega) = \sum_{k=0}^{\infty}\sum_{p=1}^{3}\big(C_{1p,k}A^0_{p3}\big)cos(k\omega), g_9(a,\beta,\omega) = \sum_{k=0}^{\infty}\sum_{p=1}^{3}\big(C_{2p,k}A^0_{p3}\big)cos(k\omega).
\end{aligned}$$

This can be generalized as follows. The first $\frac{m(m+1)}{2}$ moment conditions are given by (using the lexicographic order: $i \equiv (p',q') = (1,1),(1,2),...,(1,m),(2,2),...,(2,m),...,(m,m)$):

$$\widetilde{g}(a,\beta,\omega) = \big(g_i\big)_{i=(p',q')}$$

where (since $p' \le q'$)

$$g_i(a,\beta,\omega) = \sum_{p=1}^{m}\sum_{q=1}^{m}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} C_{p'p,k}C_{q'q,j}\big(\sigma_{pq} - \sum_{v=1}^{p'} A^0_{pv}A^0_{qv}\big)cos((k-j)\omega).$$

The last $\frac{m(m+1)}{2} - m = \frac{m(m-1)}{2}$ moment conditions correspond to the identifying constraints, which are given by (using the lexicographic order $i \equiv (p',q') = (1,2),(1,3),...,(1,m),(2,3),...,(2,m),....,(m-1,m)$):

$$g_{i=(p',q')}(a, \beta\omega) = \sum_{k=0}^{\infty} \sum_{p=1}^{m} (C_{p'p,k} A^0_{pq'}) cos(k\omega). \square$$

References

[1] Amisano, G., and C. Giannini (1997), Topics in Structural VAR Econometrics, Springer Verlag, second edition.

[2] Arellano, M, Hansen, L.P., and E. Sentana (2012), "Underidentification?", *Journal of Econometrics*, vol. 170(2), 256-280.

[3] Barsky, R. B., and E. R. Sims (2011), "News Shocks and Business Cycles", *Journal of Monetary Economics*, vol. 58(3), 273289.

[4] Beaudry, P., and F. Portier (2006), "Stock Prices, News, and Economic Fluctuations", *American Economic Review*, vol. 96(4), 12931307.

[5] Beaudry, P., Portier, F., and A. Seymen (2013), "Comparing Two Methods for the Identification of New Shocks", Discussion paper No. 13-110, Center for European Economic Research.

[6] Berkowitz, J. (2001), "Generalized Spectral Estimation of the Consumption-Based Asset Pricing Model", *Journal of Econometrics*, vol. 104, 269-288.

[7] Bernanke, B. (1986), "Alternative Explanations of the Money-Income Correlation", *Carnegie Rochester Conference Series on Public Policy*, vol. 25(0), 49-99.

[8] Bernanke, B., and I. Mihov (1998), "Measuring Monetary Policy", *Quarterly Journal of Economics*, vol. 113(3), 869-902.

[9] Blanchard, O.J. and D. Quah (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review*, vol. 79, 655-673.

[10] Carrasco, M. (2012), A Regularization Approach to the Many Instruments Problem, *Journal of Econometrics*, vol. 170(2), 383398

[11] Carrasco, M., Chernov, M., Florens, J.P., and E. Ghysels (2007), "Efficient Estimation of General Dynamic Models with a Continuum of Moment Conditions", Journal of Econometrics, vol. 140(2), 529-573.

[12] Carrasco, M., and J.P. Florens (2000), "Generalization of GMM to a continuum of moment conditions", *Econometric Theory*, vol. 16, 797-834.

[13] Carrasco, M., Florens, J.P. and E. Renault (2007), Linear Inverse Problems in Structural Econometrics: Estimations based on Spectral Decomposition and Regularization, edited in Handbook of Econometrics, vol. 6 (chapter 77), part B, 56335751.

[14] Chari, V., Kehoe, P., and E. McGrattan (2008), "Are Structural VARs with long-run Restrictions Useful in Developing Business Cycle Theory?", *Journal of Monetary Economics*, vol. 55(8), 1337-1352.

[15] Chari, V., Kehoe, P., and E. McGrattan (2005), "Critique of Structural VARs Using Real Business Cycle Theory", Federal Reserve Bank of Minneapolis, Working Paper, No. 631.

[16] Christiano, L.J., Eichenbaum, M. and R. Vigfusson (2006a), "Assessing Structural VARs", NBER Macroeconomics Annual, vol. 21., 1-106.

[17] Christiano, L.J., Eichenbaum, M. and R. Vigfusson (2006b), "Alternative Procedures for Estimating Vector Autoregressions Identified with Long-run Restrictions", *Journal of the European Economic Association*, vol. 4(2-3), 475-483.

[18] Comin, D., and M. Gertler (2006), "Medium-Term Business Cycles", *American Economic Review*, vol. 96(3), 523-551.

[19] Cooley, T., and S. LeRoy (1985), "Atheoretical Macroeconomics: A Critique", *Journal of Monetary Economics*, vol. 16(3).

[20] DiCecio, R., and M.T. Owyang (2010), "Identifying technology shocks in the frequency domain", Working Papers 2010-025, Federal Reserve Bank of St. Louis.

[21] Dupor, B., and L. Kiefer (2008), "Executing Long-run Restrictions", Manuscript.

[22] Erceg, Guerrieri and Gust (2005), "Can Long-run Restrictions Identify Technology Shocks?", *Journal of the European Economic Association*, vol. 3(6), 1237-78.

[23] Faust, J. (1996), "Near Observational Equivalence and Theoretical Size Problems with Unit Root Tests", *Econometric Theory*, vol. 12, 724-731.

[24] Faust, J. (1998), "The Robustness of Identified VAR Conclusions about Money", *Carnegie-Rochester Conference Series ob Public Policy*, vol. 49(0), 207-44.

[25] Faust, J. (1999), "Conventional Confidence Intervals for Points on Spectrum Have Confidence Level Zero", *Econometrica*, vol. 67, 629-637.

[26] Faust, J., and E.M. Leeper (1997), "When Do Long-Run Identifying Restrictions Give Reliable Results?", *Journal of Business & Economic Statistics*, vol. 15, 345-353.

[27] Francis, N., Owyang, M.T., Roush, J.E., and R. DiCecio (2014), "A Flexible Finite-Horizon Alternative to Long-Run Restrictions with an Application to Technology Shocks," *The Review of Economics and Statistics*, vol. 96(3), 638-647.

[28] Francis, N., and V. Ramey (2009), "Measures of Per Capita Hours and Their Implications for the Technology-Hours Debate", *Journal of Money, Credit and Banking*, vol. 41(6), 1071-1097.

[29] Francis, N., and V. Ramey (2005), "Is the Technology-driven Real Business Cycle Hypothesis Dead? Shocks and Aggregate Fluctuations Revisited", *Journal of Monetary Economics*, vol. 52, 1379-1399.

[30] Galí, J. (1999), "Technology, Employment and the Business Cycle: So Technology Shocks Explain Aggregate Productivity?", *American Economic Review,* vol. 41, 1201-1249.

[31] Gourieroux, C., and A. Monfort (1995), Statistics and Econometric Modems: Volume 1, Cambridge University Press.

[32] Gourieroux, C., Monfort, A., and A. Trognon (1985), "Moindres Carrés Asymptotiques", *Annales de l'INSEE*, vol. 58, 91-121.

[33] Hall, A.R., Inoue, A., Nason, J.M., and B. Rossi (2012), "Information criteria for impulse response function matching estimation of DSGE models", *Journal of Econometrics*, vol. 170(2), 499-518.

[34] Hauser, M.P., Pötscher, B., and E. Reschenhofer (1999), "Measuring Persistence in Aggregate Output: ARMA models, Fractionally Integrated ARMA Models and Nonparametric Procedures", *Empirical Economics*, vol. 24, 243-269.

[35] Jordà, Ò. (2005), "Estimation and Inference of Impulse Responses by Local Projections", *American Economic Review*, vol. 95(1), 161-182.

[36] Kilian, L. (2013), Structural Vector Autoregressions, in: N. Hashimzade and M.A. Thornton (eds.), Handbook of Research Methods and Applications in Empirical Macroeconomics, Cheltenham, UK: Edward Elgar, 2013, pp. 515-554.

[37] Kilian, L., and Y.J. Kim (2011), "How Reliable Are Local Projection Estimators of Impulse Responses?", *The Review of Economics and Statistics*, vol. 93(4), 1460-1466.

[38] Lütkepohl, H. (2007), New Introduction to Multiple Time Series Analysis, Springer-Verlag, Berlin.

[39] Lütkepohl, H. and A. Velino (2014), "Structural Vector Autoregressions: Checking Identifying Long-run Restrictions via Heteroskedasticity, Working Paper No. 1356, DIW Berlin.

[40] Mertens, E. (2012), Are Spectral Estimators Useful for Long-run Restrictions in SVARs?", *Journal of Economic Dynamics and Control*, vol. 36(12), 1983-1844.

[41] Pagan, A.R., and J.C. Robertson (1998), "Structural Models Of The Liquidity Effect", *The Review of Economics and Statistics*, vol. 80(2), 202-217.

[42] Pancrazi, R. (2015), "The Heterogeneous Great Moderation", *European Economic Review*, vol. 74, 207-228.

[43] Pötscher, B.M. (2002), "Lower Risk Bounds and Properties of Confidence Sets for Ill-Posed Estimation Problems with Applications to Spectral Density and Presistence Estimation, Unit Roots, and Estimation of Long Memory Parameters," *Econometrica,* **70,** 1035-1065.

[44] Sims, C. (1971), "Distributed Lag Estimation When the paarmeter Space is Explicitly Infinite-Dimensional", *Annals of Mathematical Statistics*, vol. 42, 1622-1636.

[45] Sims, C. (1972), "The Role of Approximation Prior Restrictions in Distributed Lag Estimation", *Journal of the American Statistical Association*, vol. 67, 169-175.

[46] Sims, C. (1980a), "Macroeconomics and Reality", *Econometrica*, vol. 48(1), 1-48.

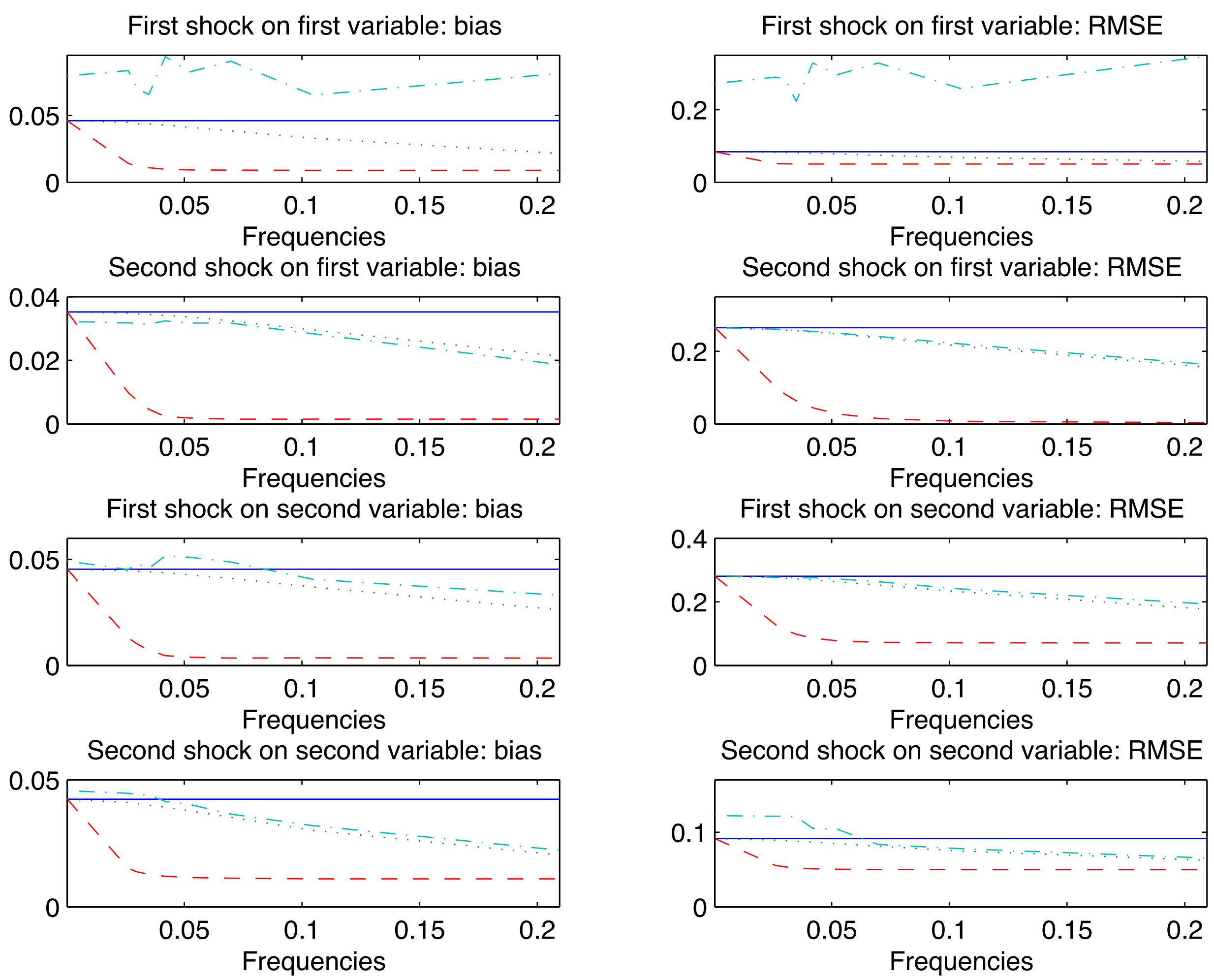
[47] Sims, C. (1980b), "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered", *American Economic Review*, vol. 70(2), 250-57.

[48] Sims, C. (1986), "Are Forecasting Models Usable for Policy Analysis?", *Federal Reserve Bank of Minneapolis Quartely Review*, vol. 10(1), 2-16.

[49] Uhlig, H. (2004), "Do Technology Shocks Lead to a Fall in Total Hours Worked?", *Journal of the European Economic Association*, vol. 2(2-3), 361-371.

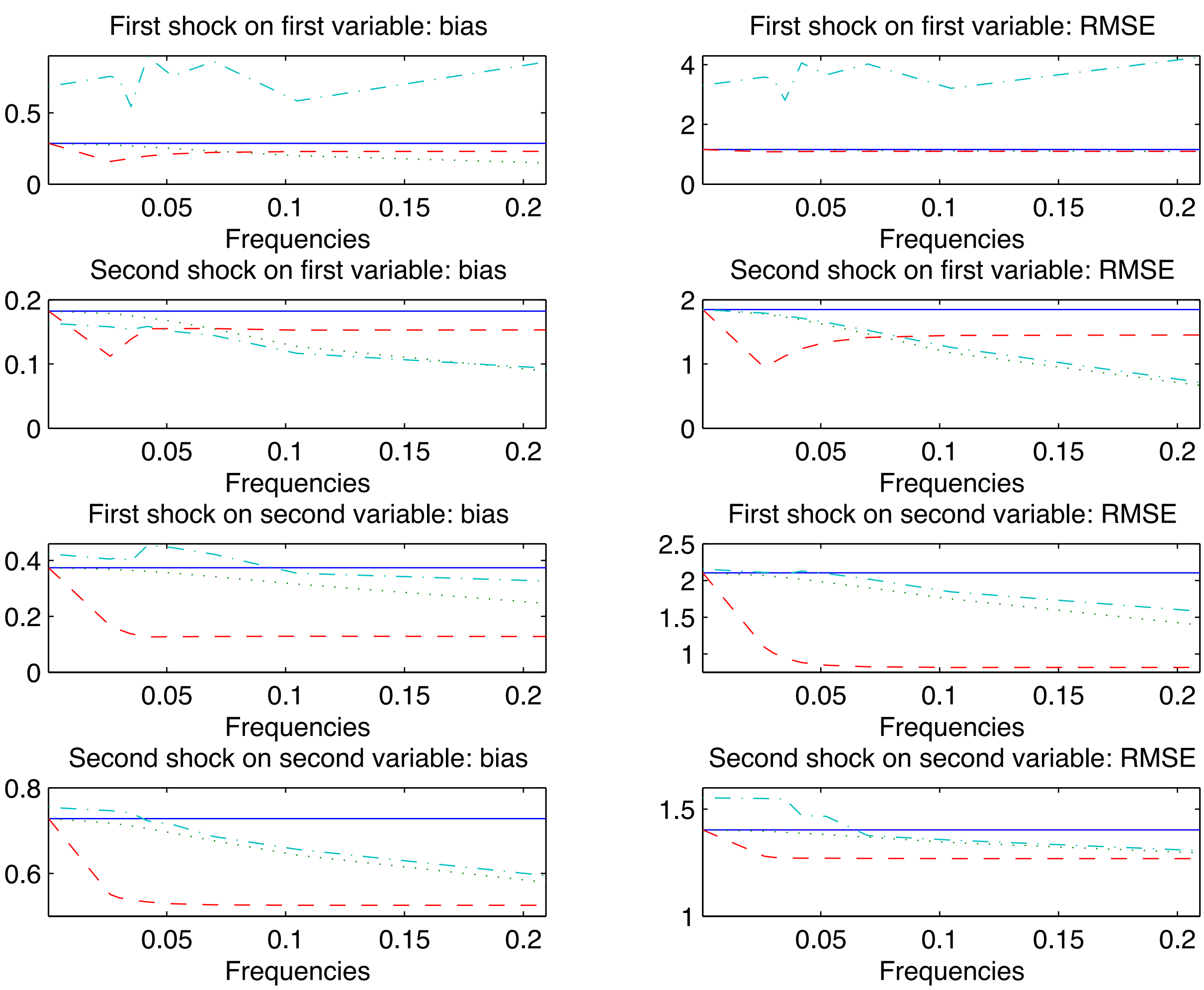
[50] Wen, Y. (2001), "A Generalized Method of Impulse Identification",*Economic Letters*, vol. 73, 367-374.

Figure 1: Bias and RMSE at the impact: $\rho_{12} = 0$ and $\delta = 0$



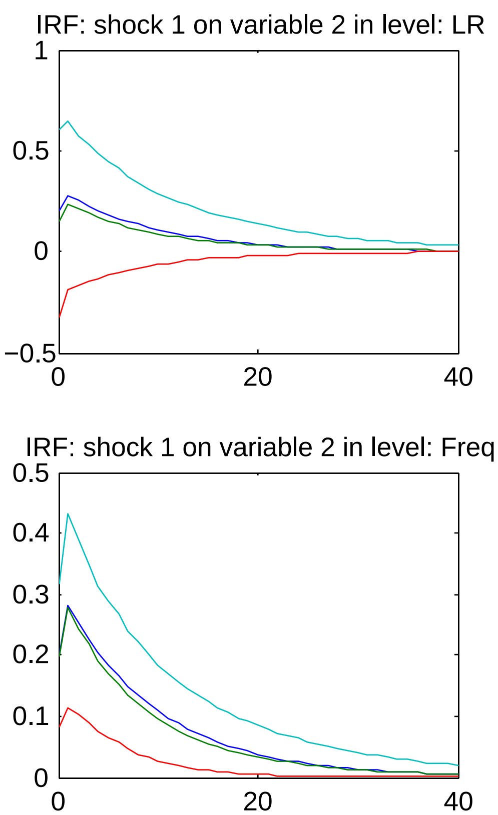
Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent, respectively, LR, Freq, DO and first-step estimators.

Figure 2: Cumulative Bias and RMSE up to 12 quarters: $\rho_{12} = 0$ and $\delta = 0$



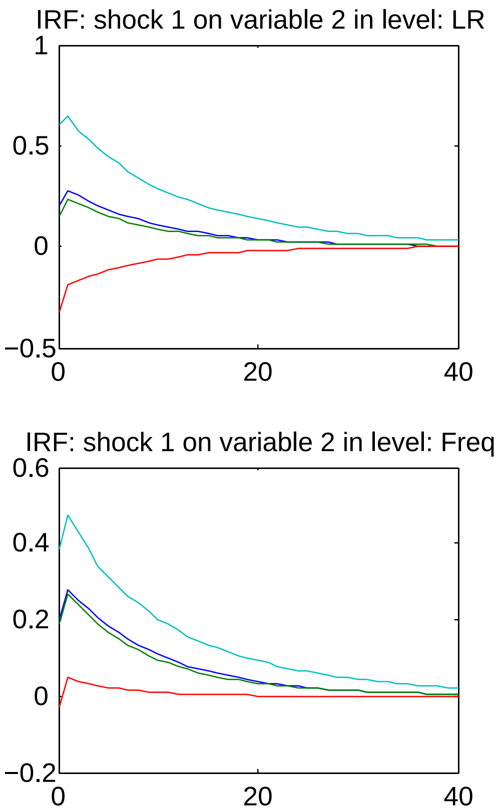
Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent, respectively, LR, Freq, DO and first-step estimators.

Figure 3: Impulse Responses for the first shock on second variable: ω_{30}, $\rho_{12} = 0$ and $\delta = 0$



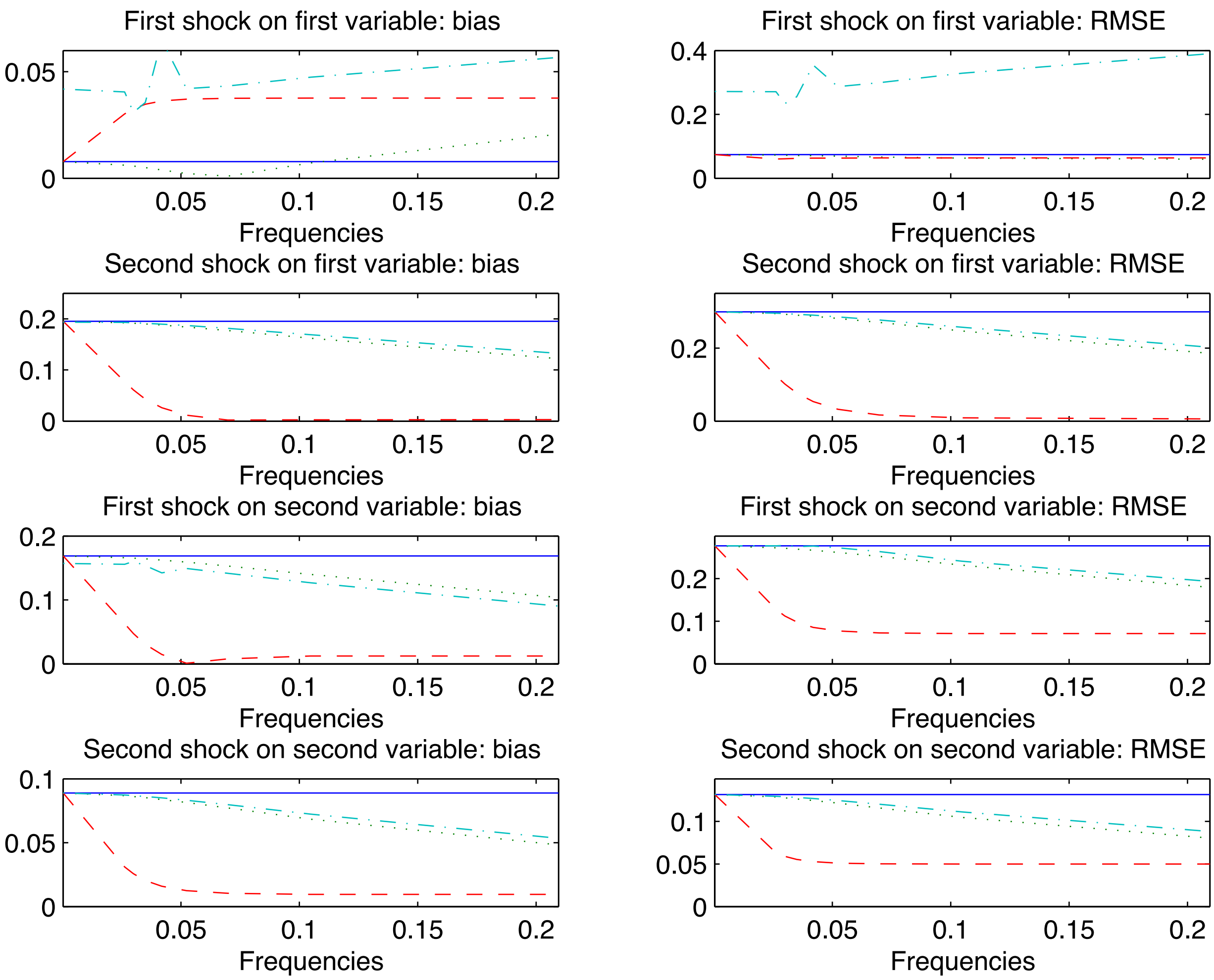
Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Figure 4: Impulse Responses for the first shock on second variable: ω_{240}, $\rho_{12} = 0$ and $\delta = 0$



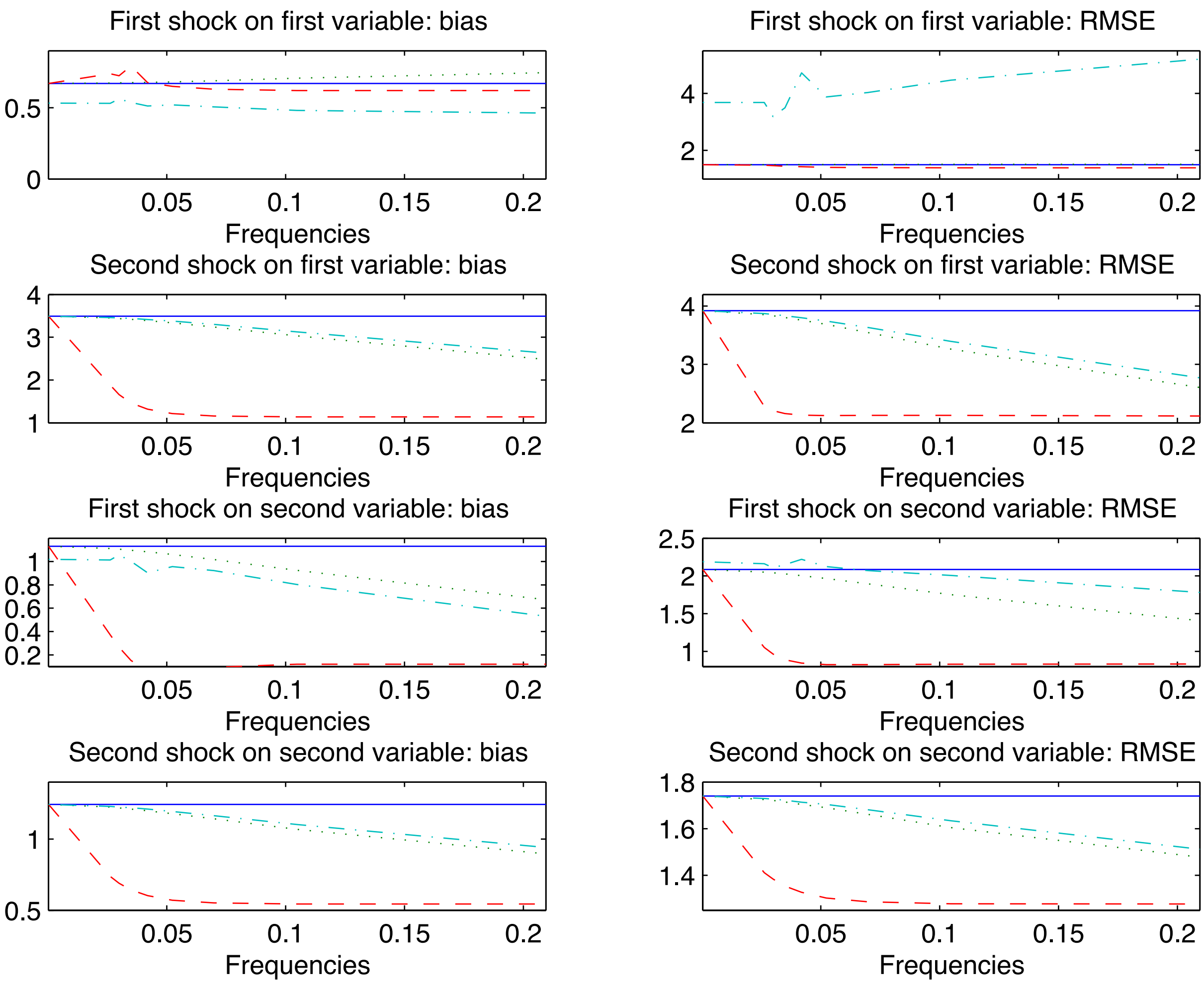
Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Figure 5: Bias and RMSE at the impact: $\rho_{12} = .3$ and $\delta = 0$



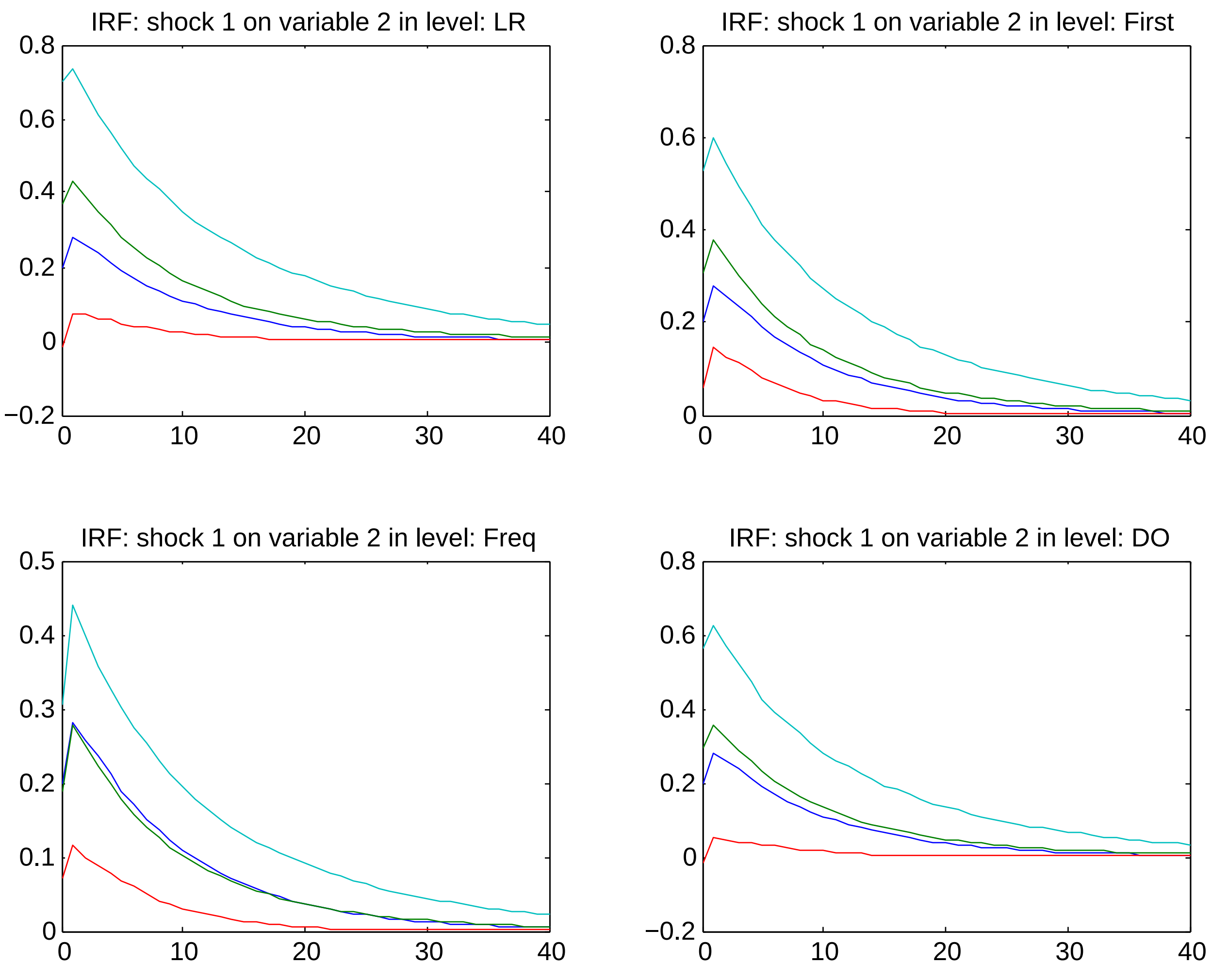
Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent, respectively, LR, Freq, DO and first-step estimators.

Figure 6: Cumulative Bias and RMSE up to 12 quarters: $\rho_{12} = .3$ and $\delta = 0$



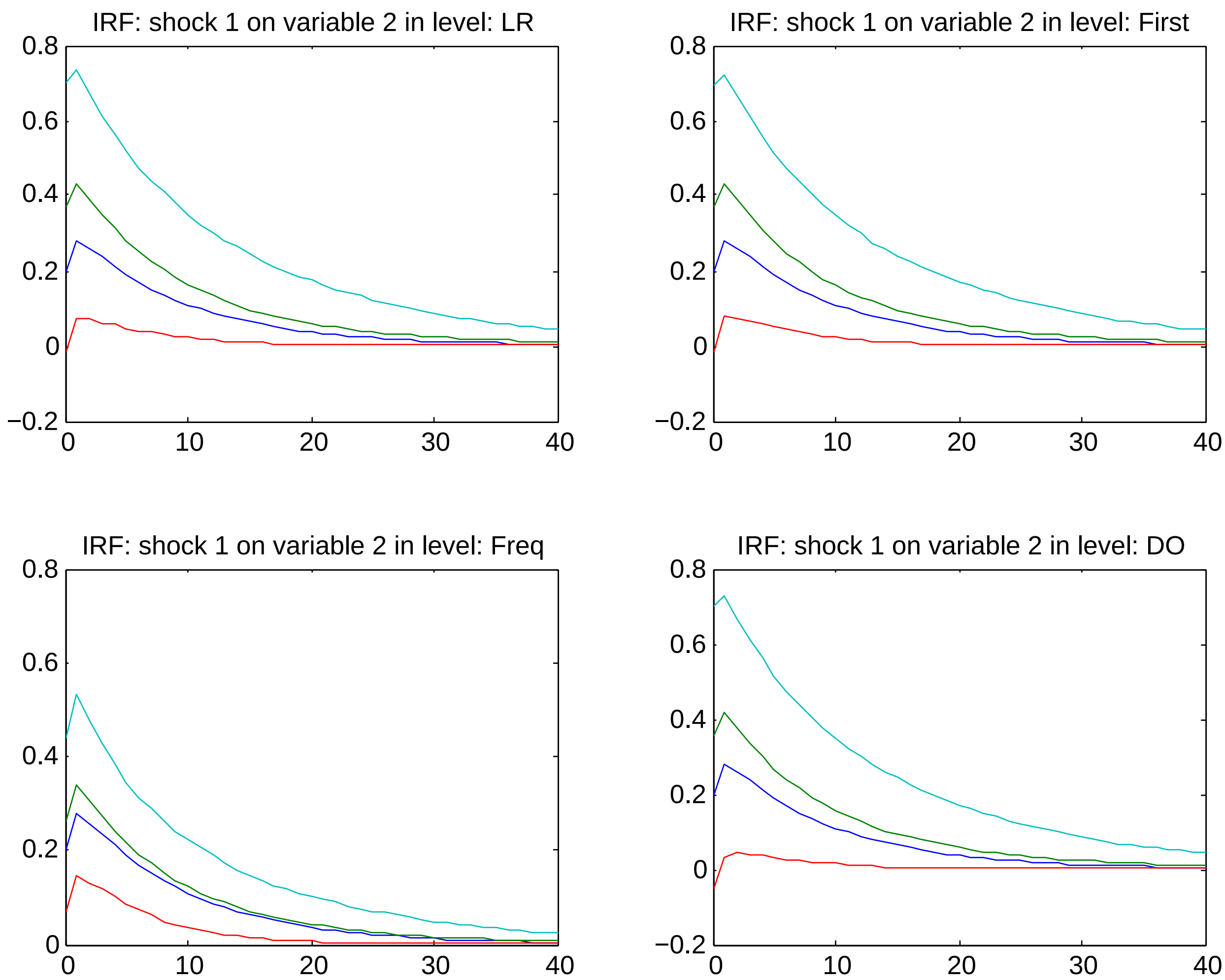
Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent, respectively, LR, Freq, DO and first-step estimators.

Figure 7: Impulse Responses for the first shock on second variable: ω_{30}, $\rho_{12} = .3$ and $\delta = 0$



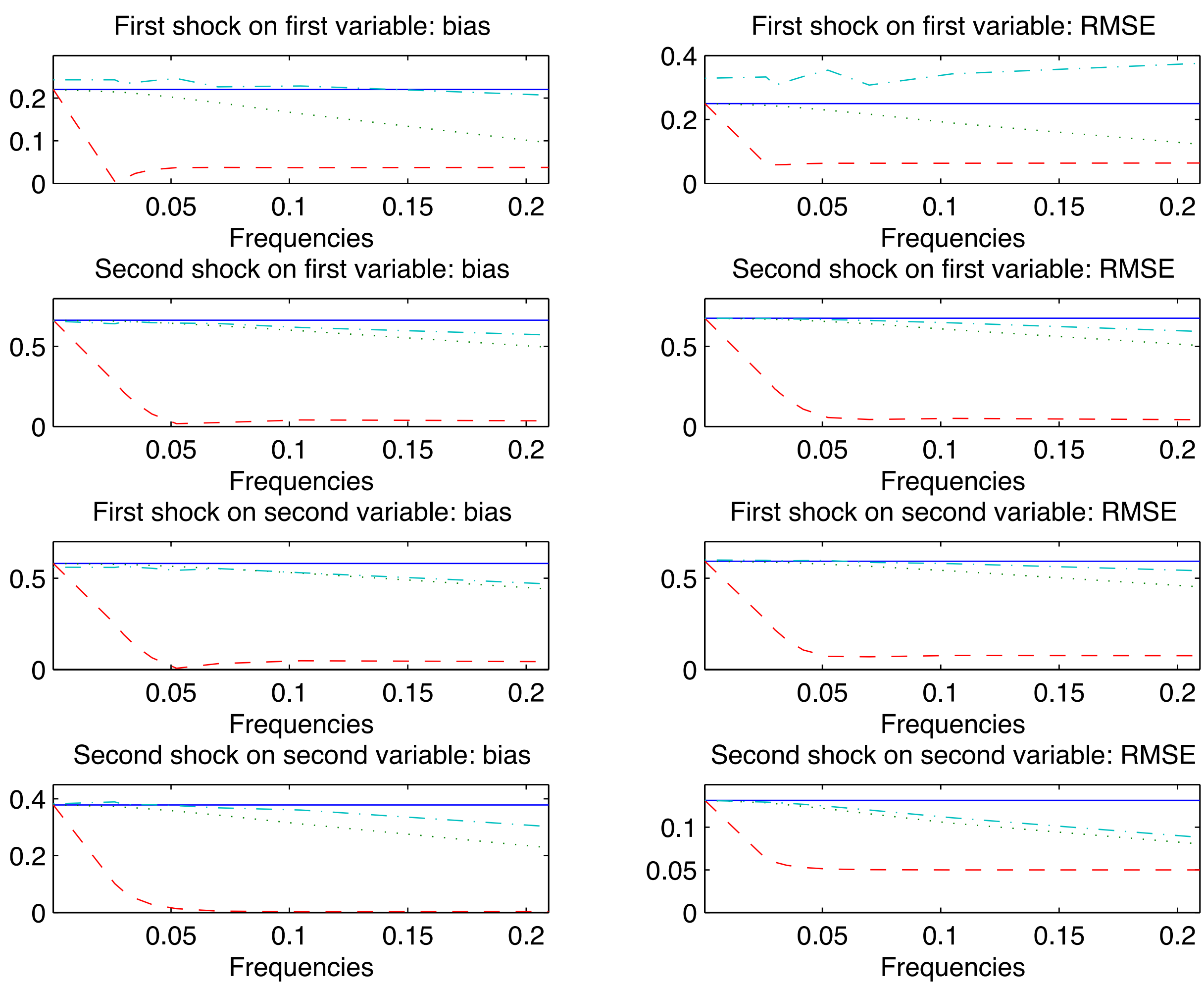
Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Figure 8: Impulse Responses for the first shock on second variable: ω_{240}, $\rho_{12} = .3$ and $\delta = 0$



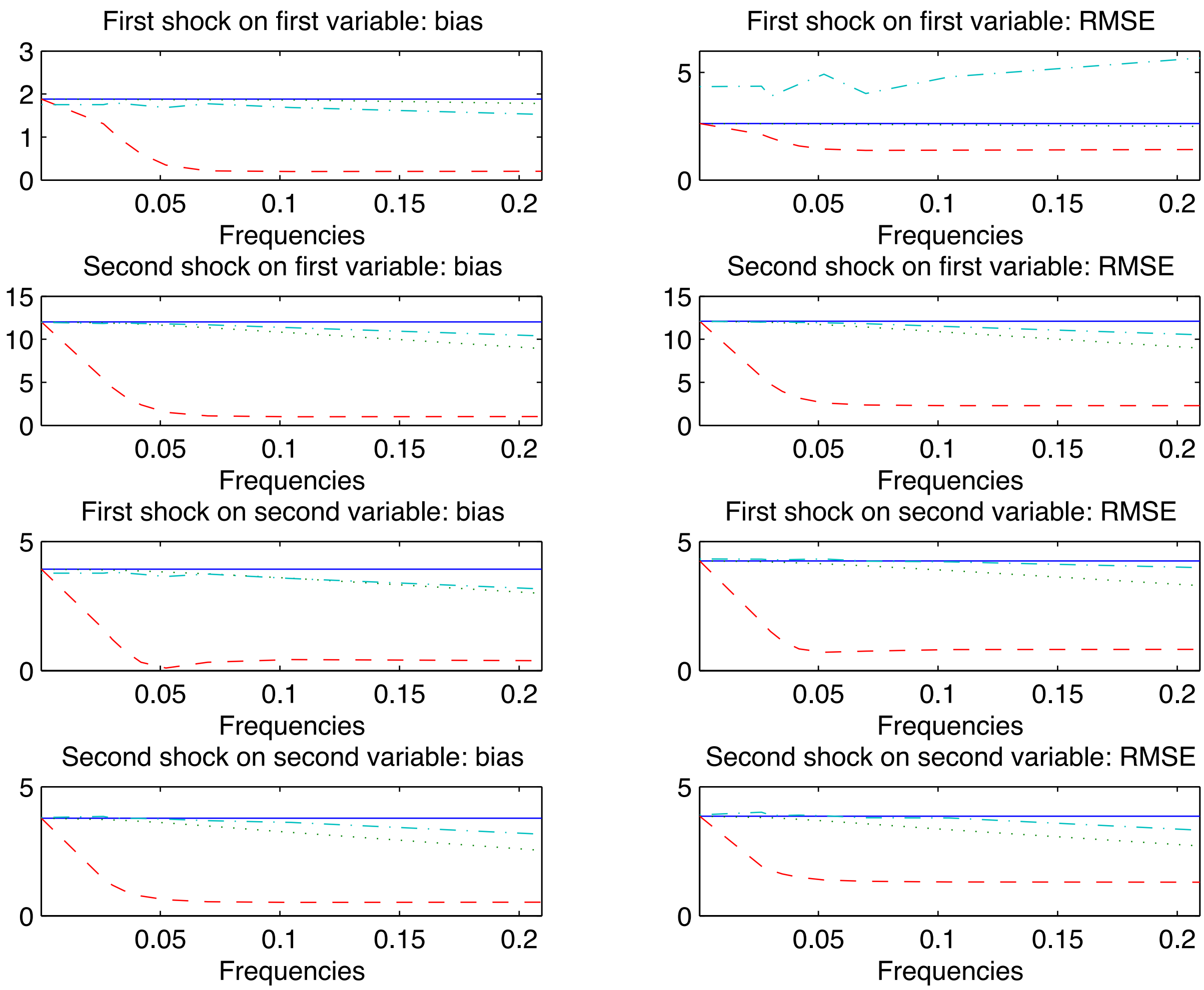
Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Figure 9: Bias and RMSE at the impact: $\rho_{12} = .3$ and $\delta = .1$



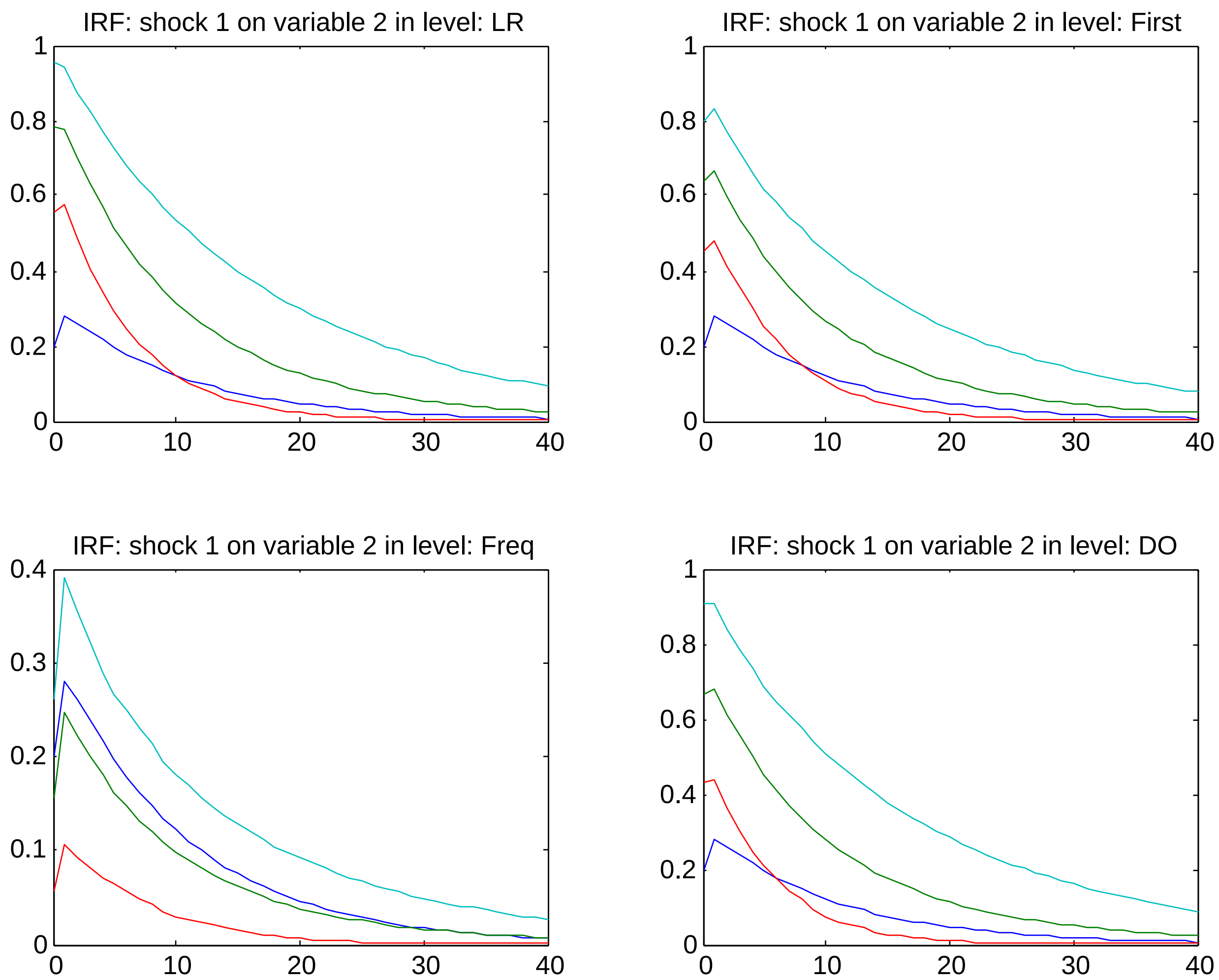
Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent, respectively, LR, Freq, DO and first-step estimators.

Figure 10: Cumulative Bias and RMSE up to 12 quarters: $\rho_{12} = .3$ and $\delta = .1$



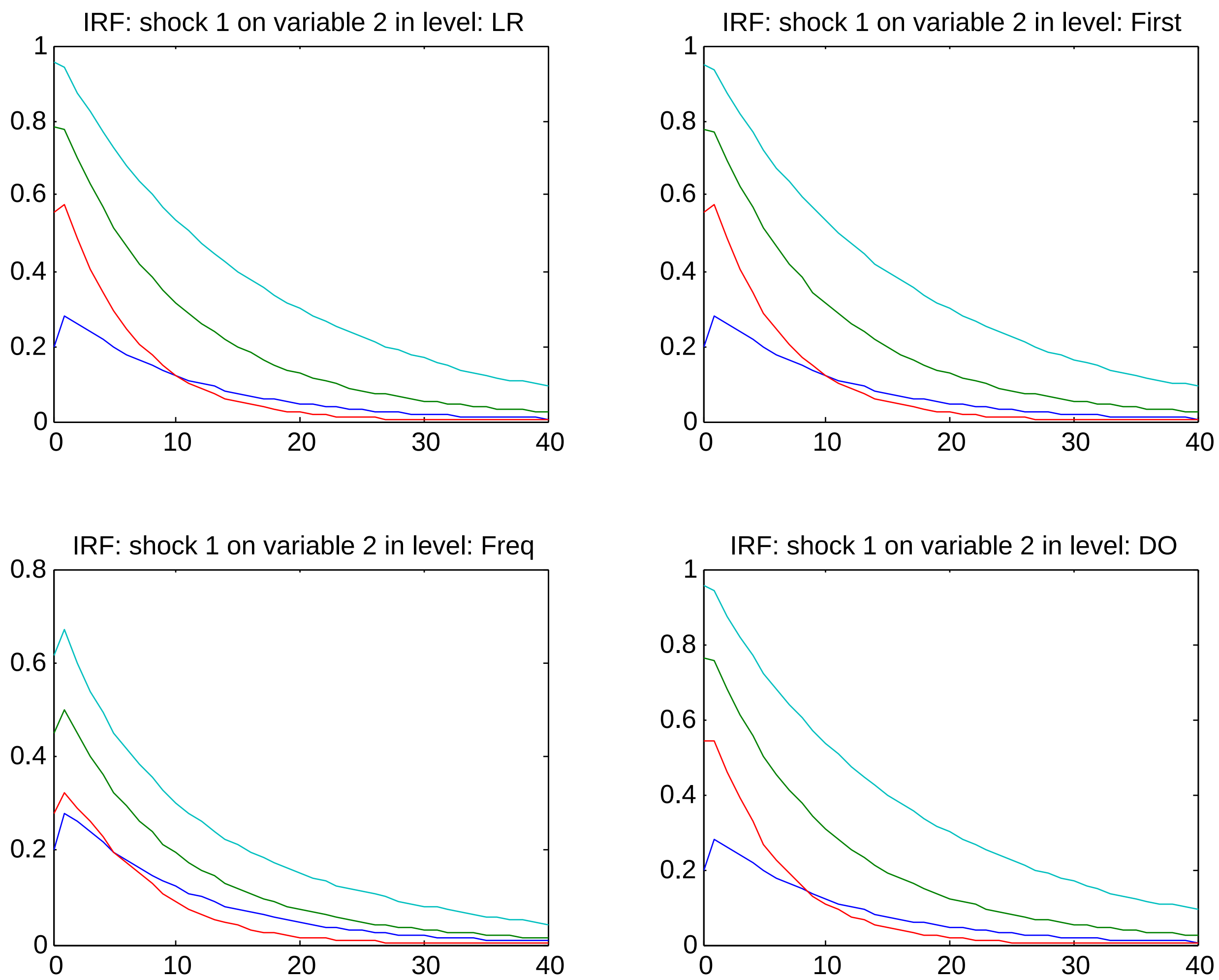
Note: The solid line, long dashed line, dash-dotted line, and dotted lines represent, respectively, LR, Freq, DO and first-step estimators.

Figure 11: Impulse Responses for the first shock on second variable: ω_{30}, $\rho_{12} = .3$ and $\delta = .1$



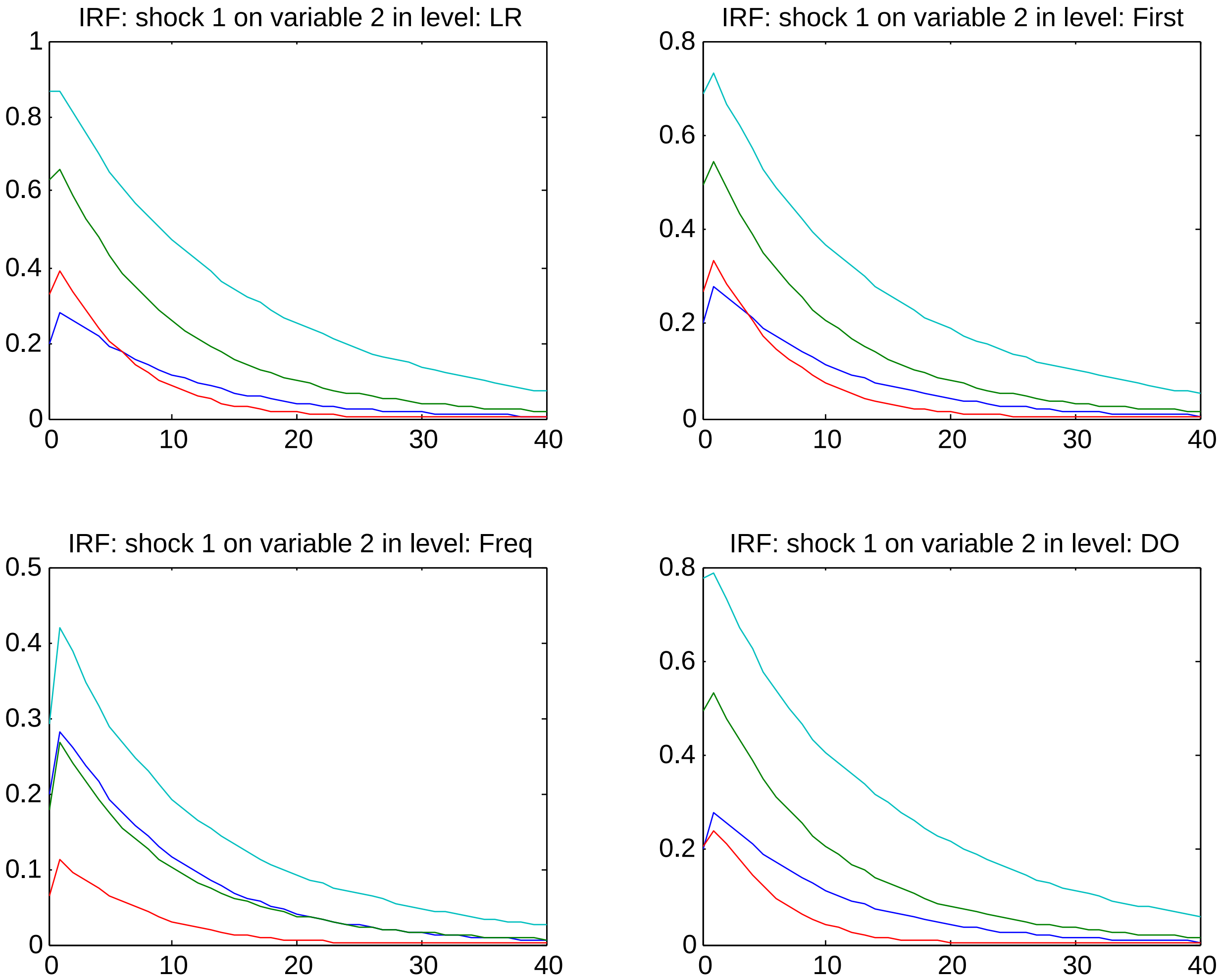
Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Figure 12: Impulse Responses for the first shock on second variable: ω_{240}, $\rho_{12} = .3$ and $\delta = .1$



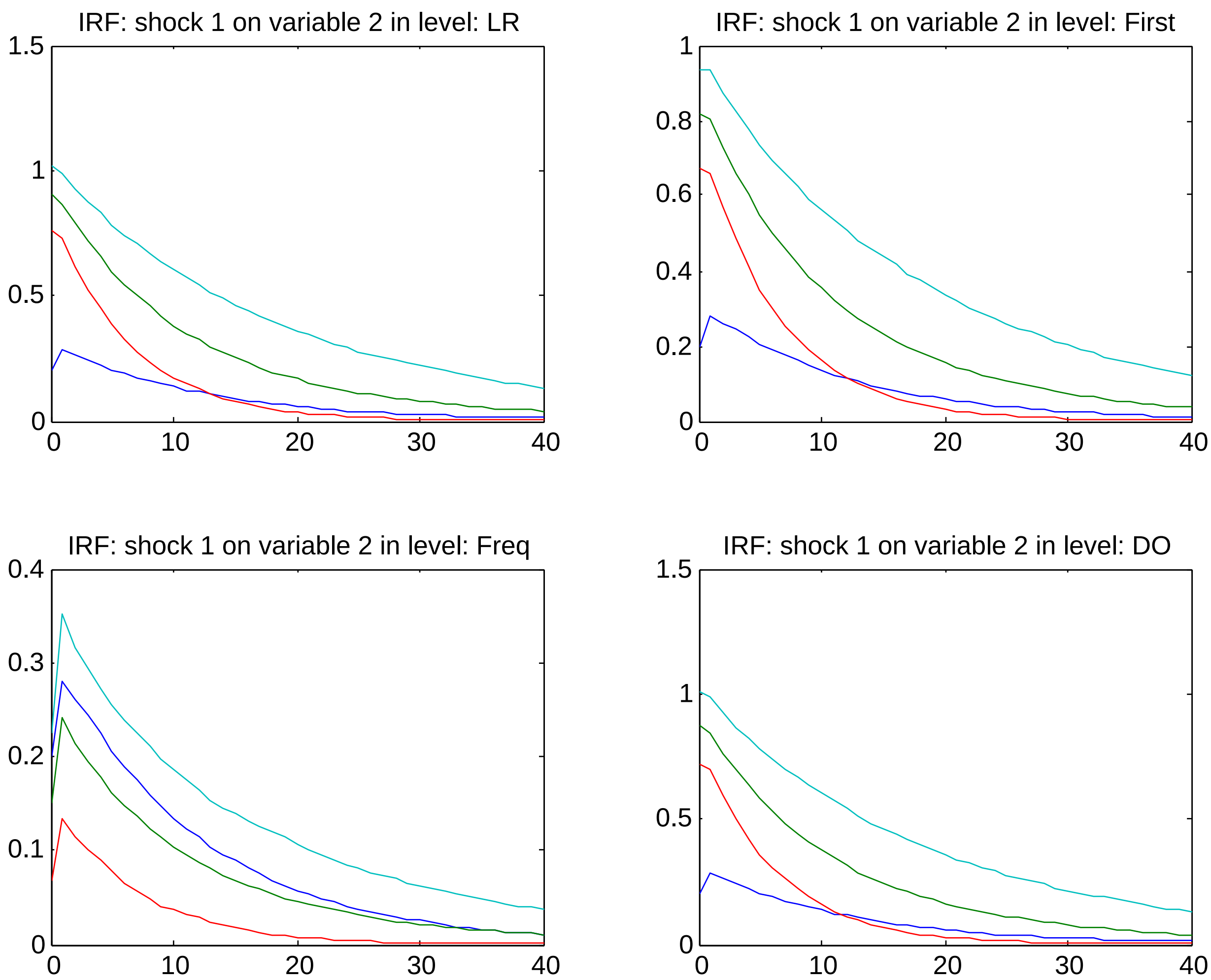
Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Figure 13: Impulse Responses for the first shock on second variable: ω_{30}, $\rho_{12} = .3$ and $\delta = .05$



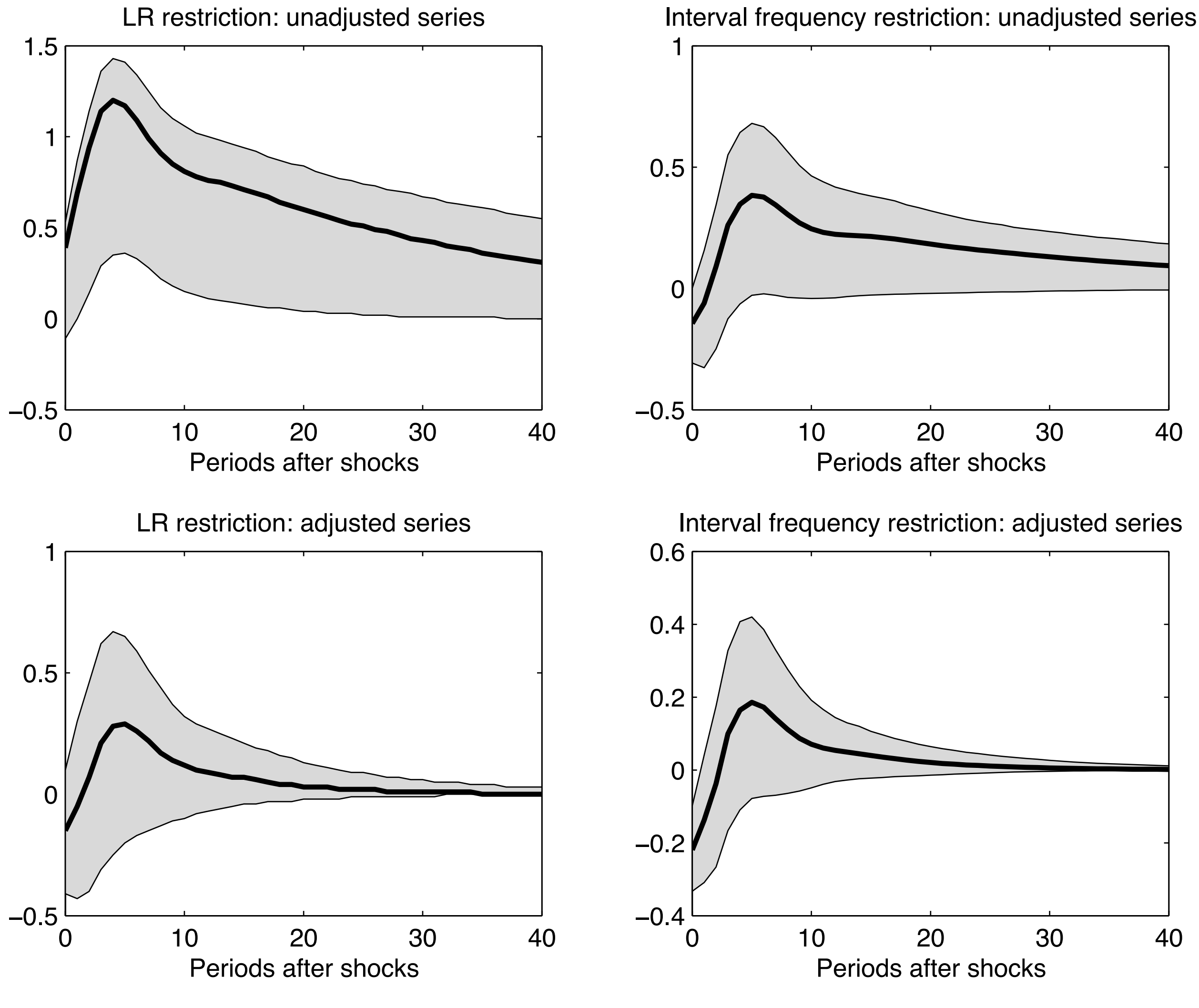
Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Figure 14: Impulse Responses for the first shock on second variable: ω_{30}, $\rho_{12} = .3$ and $\delta = .2$



Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Figure 15: Impulse responses for the technology shock on hours worked



Note: Confidence intervals are based the 95–percentile from 1000 Monte–Carlo experiments.

Table 1: J-test

	Quarters							
	30	60	90	120	150	180	210	240
	Panel A: $\rho_{12} = 0$ and $\delta = 0$							
(a)	.690	.600	.501	.410	.321	.236	.180	.133
(b)	0	0	0	0	0	0	0	0
5 %	1.16	1.50	2.00	1.71	2.18	1.27	1.11	1.50
10 %	2.61	3.67	4.99	6.59	7.17	6.36	5.56	3.01
	Panel B: $\rho_{12} = .3$ and $\delta = 0$							
(a)	.743	.627	.569	.463	.357	.274	.191	.132
(b)	.0405	.0204	.0114	.0071	.0048	.0034	.0026	.0020
5 %	5.65	6.55	8.44	9.72	11.76	12.77	11.52	9.09
10 %	9.02	10.57	13.71	16.20	20.73	21.53	23.04	21.97
	Panel C: $\rho_{12} = .3$ and $\delta = .1$							
(a)	.999	.997	.994	.982	.940	.870	.785	.664
(b)	.3246	.3736	.3908	.3983	.4022	.4044	.4058	.4067
5 %	90.29	90.97	91.75	92.46	92.02	90.34	87.01	84.48
10 %	93.69	94.08	94.77	95.52	96.81	96.55	96.82	94.73
	Panel D: $\rho_{12} = .3$ and $\delta = .05$							
(a)	.975	.950	.914	.857	.772	.664	.565	.446
(b)	.1399	.1567	.1633	.1664	.1680	.1689	.1695	.1699
5 %	46.67	49.68	52.95	55.19	56.09	54.52	51.33	46.41
10 %	55.69	58.63	62.25	66.51	70.08	71.23	69.03	67.49

Note: The frequency intervals under investigation are: $\omega_n = (-\frac{2\pi}{n}, \frac{2\pi}{n})$ for $n = 30, 60, 90, 120, 180, 210, 240$ quarters. In rows indexed by (a), numbers are the percentage of times that the number of principal components is greater than one. The numbers in the arrow indexed by (b) are the proportion of the variance explained by the second shock for the first variable in the frequency interval of interest.