

**Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.**

**Are Spectral Estimators Useful for Implementing Long-Run
Restrictions in SVARs?**

Elmar Mertens

2010-09

NOTE: Staff working papers in the Finance and Economics Discussion Series (FEDS) are preliminary materials circulated to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors. References in publications to the Finance and Economics Discussion Series (other than acknowledgement) should be cleared with the author(s) to protect the tentative character of these papers.

Are Spectral Estimators Useful for Implementing Long-Run Restrictions in SVARs?*

Elmar Mertens[†]
Federal Reserve Board

July 2009

*This paper is based on the second chapter of my dissertation written at the University of Lausanne. An earlier draft of this paper has been written while visiting the Federal Reserve Bank of Minneapolis. I would like to thank the Bank and its staff for their kind hospitality. In particular, I would like to thank V.V. Chari, Patrick Kehoe and Ellen McGrattan for discussions and sharing their computer codes. In addition I am grateful for discussions with Lawrence Christiano, Bill Dupor, Mark Watson as well as seminar participants at the University of Basel, the University of Zurich, the Young Swiss Economist Meetings, and the European Economics Association meeting 2007 in Budapest and the members of my dissertation committee, Jean-Pierre Danthine, Philippe Bacchetta, Robert G. King, and Peter Kugler. All remaining errors are of course mine. This research has been carried out within the National Center of Competence in Research “Financial Valuation and Risk Management” (NCCR FINRISK). NCCR FINRISK is a research program supported by the Swiss National Science Foundation.

[†]For correspondence: Elmar Mertens, Board of Governors of the Federal Reserve System, Washington D.C. 20551. email elmar.mertens@frb.gov. Tel.: +(202) 452 2916. The views in this paper do not necessarily represent the views of the Federal Reserve Board, or any other person in the Federal Reserve System or the Federal Open Market Committee. Any errors or omissions should be regarded as those solely of the author.

Are Spectral Estimators Useful for Implementing Long-Run Restrictions in SVARs?

Abstract

No, not really, since spectral estimators suffer from small sample and misspecification biases just as VARs do. Spectral estimators are no panacea for implementing long-run restrictions.

In addition, when combining VAR coefficients with non-parametric estimates of the spectral density, care needs to be taken to consistently account for information embedded in the non-parametric estimates about serial correlation in VAR residuals. This paper uses a spectral factorization to ensure a correct representation of the data's variance. But this cannot overcome the fundamental problems of estimating the long-run dynamics of macroeconomic data in samples of typical length.

JEL Classification: C32, C51, E17, E32

Keywords: Structural VAR, Long-Run Identification, Non-parametric Estimation, Spectral Factorization

1 Introduction

VARs have been criticized for failures in estimating the responses to long-run shocks. A crucial element for long run identification is the spectral density at zero-frequency, also known as “long-run variance”. But OLS estimates of VAR coefficients are concerned with minimizing forecast error variance, not estimating the long run variance. This has motivated Christiano, Eichenbaum, and Vigfusson (2006a, 2006b), henceforth “CEV”, to propose a new way of estimating structural VARs using a combination of OLS and a non-parametric estimator. They argue that their estimator virtually eliminates the bias associated with the standard OLS estimator.

This paper shows that non-parametric estimates of the spectral density, henceforth called “spectral estimators”, are no panacea for the implementation of long-run restrictions in finite sample. Macroeconomic time series display a fair amount of persistence, posing two serious challenges for long-run identification. First, an accurate representation of the true model typically requires a VAR with a high lag order, much higher than what is affordable in a sample of typical length and resulting in a sizable truncation bias (Chari, Kehoe, and McGrattan 2008, henceforth “CKM”). Second, there is the small sample bias in estimated coefficients known from Hurwicz (1950), which becomes ever more severe the smaller the sample, and the more persistent the data. As will be shown, both issues affect not only VARs in the time domain, but also spectral estimators in the frequency domain.

The conventional VAR technique as well as different combinations with spectral estimators are evaluated in the context of a simple two-shock RBC model, which has also been used by CEV and CKM. When using the various procedures to estimate the response of hours to technology, or to decompose the variance of fluctuations in output or hours, none of the procedures clearly dominates the others.

Furthermore, CEV do not consider some conceptual pitfalls in combining VAR coefficients with spectral estimates. Non-parametric estimates of the spectral density allow for non-*iid* residuals in the finite-order VAR, which is good since the underlying model is likely of infinite order. In what may be called “mixing and matching”, the CEV approach plugs these estimates into the

standard VAR formula alongside with coefficients from the finite-order VAR. This approach uses the extra information about omitted lags in the VAR to compute the long-run responses of variables to shocks—but not when mapping these back into impact responses. To retain a consistent representation of the data, that would however be necessary. Otherwise, the total variance of the data is misrepresented. In the simulations reported here, this misrepresentation is quantitatively relevant. As a related issue, when the relationship between forecast errors and structural shocks is inverted with the CEV coefficients, one obtains a time series which is identical to the shock estimates from OLS up to a scale factor. All in all, this is of concern for any researcher wanting to adopt the CEV strategy.

The CEV framework is amended here by recognizing that the non-parametric estimate contains information about omitted lags in the VAR. This misspecification has been stressed by CKM, Erceg, Guerrieri, and Gust (2005), Ravenna (2007) and Cooley and Dywer (1998). The adjusted procedure retains the OLS estimates and fills up the omitted lags with a spectral factorization of the spectral density's non-parametric estimate. By construction, this adjusted SVAR—in fact an SVARMA—matches the sample variance of the data just as OLS does. Overall, this corrected procedure suffers from the same basic problems as the other long-run identification methods: truncation and small sample bias.

The remainder of this paper is structured as follows: Section 2 describes the model economy against which the various estimation routines will be evaluated. Section 3 describes the various SVAR methods, including a new spectral factorization procedure. Section 4 presents the Monte Carlo results and Section 5 concludes the paper.

2 A Model Economy

This section describes a simple model economy, which will be used to illustrate and quantify the issues associated with various long-run identification schemes. None of the conceptual concerns related to spectral estimates raised in Section 3 will be specific to this model. The model is identical

to the two-shock economy used by CKM and CEV.

The model is a common one-sector RBC economy driven by two shocks: First, a unit root shock to technology, z_t . This is the permanent shock to be estimated by the VAR. Second, a transitory non-technology shock, τ_{lt} , which drives a wedge between private household's labor-consumption decision.

The representative household maximizes his lifetime utility over (per-capita) consumption, c_t , and labor services, l_t

$$E_0 \sum_{t=0}^{\infty} (\beta(1 + \gamma))^t u(c_t, l_t)$$

and faces the budget constraint $c_t + (1 + \gamma)k_{t+1} - (1 - \delta)k_t = (1 - \tau_{lt})w_t l_t + r_t k_t + T_t$ where k_t is the per-capita stock of capital, w_t the wage rate, r_t the rental rate of capital, T_t are lump sum taxes, γ is the growth rate of population, δ the depreciation rate of capital ($\gamma > 0$, $0 \leq \delta \leq 1$ and $\beta < 1$).

The non-technology shock τ_{lt} is an exogenous labor tax. As discussed by CKM, it need not be literally interpreted as a tax levy, but stands in for the effects of a variety of non-technology shocks introduced into second-generation RBC models. Mechanically, it distorts the first-order condition for consumption and labor. It works similar to a stochastic preference shock to the Frisch elasticity of labor supply. Chari, Kehoe, and McGrattan (2007) show how this labor “wedge” can be understood more generally as the reduced form process for more elaborate distortions, such as sticky wages.

The production function $F(k_t, Z_t l_t)$ is constant returns to scale, where Z_t is labor-augmenting technological progress. Firms are static and maximize profits $F(k_t, Z_t l_t) - w_t l_t - r_t k_t$. Per-capita output equals production, $y_t = F(k_t, Z_t l_t)$, and the economy's resource constraint is $y_t = c_t + (1 + \gamma)k_{t+1} - (1 - \delta)k_t$. The exogenous drivers follow linear stochastic processes:

$$\begin{aligned} \log Z_t &= \mu_z + \log Z_{t-1} + \sigma_z \varepsilon_t^Z \\ \log \tau_{l,t+1} &= (1 - \rho_l) \bar{\tau}_l + \rho_l \log \tau_{l,t-1} + \sigma_l \varepsilon_t^l \end{aligned}$$

where ε_t^Z and ε_t^l are *iid* standard-normal random variables. They are the technology shock, respec-

tively labor shock. ρ_l measures the persistence of the transitory labor tax. The scale factors σ_z and σ_l determine their relative importance in the model. (μ_z is the drift in log-technology and $\bar{\tau}_t$ is the average tax rate.)

The calibration is identical to the baseline model of CEV, which uses parameter values familiar from the business cycle literature. Utility is specified as $u(c, l) = \log c + \psi \log(1 - l)$ (consistent with balanced growth) and the production function is Cobb-Douglas $F(k, l) = k^\theta l^{1-\theta}$ with a capital share of $\theta = 0.33$. The labor preference parameter is set to $\psi = 2.5$. On an annualized basis, the calibration sets the depreciation rate to 6%, the rate of time preferences to 2% and population growth to 1%.¹ Following CEV, the transitory shock is calibrated as an AR(1) with persistence $\rho_l = 0.986$. This calibration is identical to the values used by CKM except for their values of $\phi = 1.6$ and $\rho_l = 0.95$.

The model economy is calibrated over different ratios in the variance of transitory to permanent shocks, σ_l^2/σ_z^2 , which translate into different assumptions about the share of output fluctuations explained by technology shocks.² As a benchmark, maximum-likelihood estimates of CEV obtained from fitting the model to U.S. post-war data imply that around two-thirds of the bandpass-filtered variance in output can be attributed to technology shocks.³ The bandpass filter employed throughout this paper considers only fluctuations with durations between two-and-a-half and eight years, which is consistent with the NBER definitions of Burns and Mitchell (1946).

Data is simulated for samples of length $T = 180$, corresponding to 45 years of quarterly data; identical to the simulations of CKM and CEV. Following CEV and CKM, bivariate VARs are estimated using simulated data of the (log) growth rate of labor productivity and hours worked; $X_t = \begin{bmatrix} \Delta \log(y_t/l_t) & \log l_t \end{bmatrix}^T$.⁴ For each simulated sample, the lag length of the VAR(p) is cho-

¹The drift in technology is set to 0.4% and the average “labor tax” is set to 24.2% per quarter.

²CKM extensively document how different ratios in the variance of transitory to permanent shocks, σ_l^2/σ_z^2 , affect the performance of standard VARs both in population and in small sample. McGrattan (2005) shows that in the limit, $\sigma_l/\sigma_z \rightarrow 0$, a finite order VAR (even a $p = 1$) in productivity growth and hours recovers the true responses—though the true system does not have a finite-order VAR representation. In this special case the model reduces to a standard, one-shock RBC model.

³To be precise, CEV estimate $\sigma_l = 0.0056^2$, $\sigma_z = 0.00953$ corresponding to a technology share of 67.5%.

⁴In addition to this “LSVAR” specification, CKM run also VARs with quasi-differenced hours. This replaces the second VAR element l_t with $(1 - \alpha L) l_t$ ($\alpha \in \{0; 0.999\}$). Depending on α , this captures popular (but also contested)

sen by minimizing the Schwartz Information Criterion (SIC), typically picking small values close to one.⁵ When computing population moments, a VAR(1) is used. For each calibration, 1,000 samples are simulated.

When looking at data simulated from this model, two statistics are of particular interest for this paper: How do hours worked respond to a technology shock? What is the share of fluctuations due to technology shocks? These questions are typically asked by empirical researchers trying to evaluate predictions from business cycle models with SVARs, such as Gali (1999) or Christiano, Eichenbaum, and Vigfusson (2004).

3 Long-Run Identification in VARs

The linearized solution to the model described in the previous section is only one example from a wider class of linear dynamic models to which the SVAR methods discussed here can be applied. None of the issues discussed in this section will be specific to the model described above. An economic model from this class is supposed to specify a VAR representation for a stationary vector of observable variables⁶ X_t :

$$X_t = B(L)X_{t-1} + e_t \quad (1)$$

where $B(L)$ is a polynomial in the lag-operator L , $B(L) = \sum_{k=1}^{\infty} B_k L^{k-1}$ whose roots lie all outside the unit circle and the innovations are *iid*, $e_t \sim iid(0, \Omega)$.

In principle, the model prescribes an infinite order VAR. When $B_k = 0$ for $k > p$ this nests the case of a finite order VAR. But as noted by Cooley and Dywer (1998), many interesting models

specifications: On the one hand the “LSVAR” with hours in levels and $\alpha = 0$ and on the other hand the “QDSVAR” with $\alpha = 0.999$, which approximates a VAR with differenced hours without introducing a unit MA root. The quasi-differencing is discussed in more detail by CKM as well as Marcet (2005), Gali and Rabanal (2004) and Christiano, Eichenbaum, and Vigfusson (2003).

⁵Results are insensitive to using other information criteria, such as the Akaike criterion (AIC). In general, AIC is known for picking higher values of p compared to SIC. For this lab economy, AIC has been found to pick lag lengths of up to $p = 6$ with an average of $p = 2$.

⁶For notational convenience, but without loss of generality, X_t represents the demeaned variables, which is equivalent to including a constant in a VAR using the original data.

have only infinite order VAR representations. In the remainder of this paper the true VAR representation is always assumed to be of infinite order. The linearized solution to the model described in Section 2 has such an infinite order VAR representation; details are shown in Appendix B.

For the identification of structural shocks, there has to be an invertible one-to-one mapping from innovations e_t to the structural shocks ε_t driving the underlying model—such as technology, monetary policy errors, exogenous government spending etc.:

$$e_t = A_0 \varepsilon_t \quad (2)$$

where A_0 is square and $|A_0| \neq 0$. Fernández-Villaverde et al. (2007) derive conditions when a linear dynamic model has an invertible VAR representation.⁷ (These are summarized in Appendix B.) This paper considers only cases where these conditions are satisfied, though possibly only in an infinite order VAR representation. The same applies to the situations studied by CKM, CEV as well as Erceg, Guerrieri, and Gust (2005). Excluding the complications arising from non-invertibilities allows to focus on problems owing to small sample bias and the finite order approximations of the VAR.⁸

It will be handy to introduce the notation

$$C(L) \equiv (I - B(L)L)^{-1} = \sum_{k=0}^{\infty} C_k L^k \quad \text{where } C_0 = I \quad (3)$$

for the non-structural moving average (VMA) coefficients of $X_t = C(L)e_t$. The structural moving average representation for X_t is then $X_t = A(L)\varepsilon_t$ with $A(L) = C(L)A_0$.

In the spirit of CEV and CKM, only one of the structural shocks will be identified. For concreteness, let it be the first one, denoted ε_t^z , and call it “technology shock”. Think of the first element of X_t as being a growth rate (a difference in logs), like the change in labor-productivity

⁷Fernandez-Villaverde, Rubio-Ramirez, and Sargent (2005) give examples of interesting models where the conditions are satisfied and where not. For all calibrations considered, the model of Section 2 satisfies the condition of Fernández-Villaverde et al. (2007).

⁸See for example Giannone and Reichlin (2006) on the non-invertibility problem.

(Gali 1999) or output growth (Blanchard and Quah 1989). The identifying assumption is then that only the technology shock has a permanent effect on the level of the first element of X_t . This restricts the matrix of long-run coefficients, $A(1) = \sum_{i=0}^{\infty} A_i$:

$$A(1) = C(1)A_0 = \begin{bmatrix} \bar{a}_{11} & 0 & \dots & 0 \\ \# & \# & \dots & \# \end{bmatrix} \quad \text{and} \quad \bar{a}_{11} > 0 \quad (4)$$

This restriction holds exactly in the linearized solution to the model described in Section 2.

A key object for implementing this constraint is the spectral density of X_t . The spectral density at frequency ω is defined as $S_X(\omega) \equiv \sum_{k=-\infty}^{\infty} E(X_t X_{t-k}^T) e^{-i\omega k} = C(e^{-i\omega}) \Omega C(e^{-i\omega})^T$ where i is the imaginary unit.⁹ $A(1)$ factors the spectral density of X_t at frequency zero:

$$A(1)A(1)^T = C(1)\Omega C(1)^T = S_X(0) \quad (5)$$

One way to compute the first column of A_0 is by recovering $A(1)$ from the Cholesky decomposition of $S_X(0)$. (This is the unique lower triangular factorization of a positive definite matrix.¹⁰):

$$A(1) = \text{chol} \{S_X(0)\}$$

CEV show that the restriction in (4) uniquely pins down the first column of A_0 and the Cholesky factorization is one possible implementation.¹¹

The long-run coefficients can then be mapped into the matrix of impact responses using the

⁹Throughout this paper, transposes are complex conjugate.

¹⁰The spectral density $S_X(0) = C(1)\Omega C(1)^T$ is strictly positive definite when the variance covariance matrix of the forecast errors, Ω , is nonsingular. $S_X(0)$ inherits positive definiteness from Ω since $C(1)$ is nonsingular. $I - B(1) = C(1)^{-1}$ exists because of the assumed stationarity of the VAR process.

¹¹Under the restrictions stated in (4), $A(1)$ can generally be described as

$$A(1) = \text{chol} \{S_X(0)\} \begin{bmatrix} 1 & 0 \\ 0 & W \end{bmatrix} \quad \text{for some } W \text{ such that } WW^T = I$$

In the lab economy described later, the VAR will be bivariate and the forecast errors e_t are a linear combination of only two shocks. Knowing the technology shock will then also identify the second shock up to its sign, $|W| = 1$.

VAR dynamics encoded in the polynomial of lag coefficients $B(L)$:

$$A_0 = (I - B(1)) A(1) \quad (6)$$

3.1 OLS Implementation with Finite-Order VAR

Since the VAR innovations in (1) are assumed to be white noise, they satisfy the OLS normal equations $EX_{t-k}e_t^T = 0$ ($\forall k$). And in principle, the coefficients B_k could be estimated from least squares projections of X_t on its infinite past. In practice, an empirical implementation can only work with a finite lag length. Henceforth $B(L)^{\text{OLS}}$ denotes a lag polynomial of finite order $p < \infty$:

$$\begin{aligned} B(L)^{\text{OLS}} &\equiv \sum_{k=1}^p B_k^{\text{OLS}} L^{k-1} \\ v_t^{\text{OLS}} &\equiv X_t - B(L)^{\text{OLS}} X_{t-1} \\ \Omega_v^{\text{OLS}} &\equiv E [v_t^{\text{OLS}} (v_t^{\text{OLS}})^T] \end{aligned} \quad (7)$$

where the normal equations are imposed for all lags $k \leq p$

$$EX_{t-k}(v_t^{\text{OLS}})^T = 0 \quad (8)$$

The associated VMA is $C(L)^{\text{OLS}} \equiv (I - B(L)^{\text{OLS}} L)^{-1}$. Only stable VARs are considered, formally this requires all roots of $C(L)^{\text{OLS}}$ to be outside the unit-circle.

The standard procedure is to assume uncorrelated residuals, v_t^{OLS} . Following Blanchard and Quah (1989), the long run restriction (4) is implemented based on an estimate of the spectral density at frequency zero constructed from the OLS estimates. Impact coefficients are computed

by plugging these estimates into (6):

$$S_X(0)^{\text{OLS}} = C(1)^{\text{OLS}} \Omega_v^{\text{OLS}} (C(1)^{\text{OLS}})^T \quad (9)$$

$$A_0^{\text{OLS}} = (I - B(1)^{\text{OLS}}) \text{chol} \{S_X(0)^{\text{OLS}}\} \quad (10)$$

Using a finite-order VAR when the data has been generated from an infinite order process induces a truncation bias into the estimates. In this case, the OLS assumption of uncorrelated forecast errors v_t^{OLS} is violated, which is an example of what Cooley and Dywer criticized as an “auxiliary” (but not innocuous) assumption. This truncation bias arises even when the true population moments of the data generating process were known. Applied to data generated from a business cycle model the truncation bias in SVARs can be substantial, as shown by Cooley and Dywer (1998), Erceg, Guerrieri, and Gust (2005), Ravenna (2007) or CKM. The truncation bias is also sizable for data from the model described Section 2 as will be seen in Figure 3 below.

3.2 CEV: Combining OLS with Spectral Estimate

CEV propose an alternative estimator for the matrix of impact coefficients. This new estimator uses a mixture of the OLS estimates of $B(1)$ and a non-parametric estimator for $S_X(0)$. The procedure is motivated by observing that OLS projections construct $B(L)^{\text{OLS}}$ not necessarily with regard to $B(1)$ but in order to minimize the forecast error variance Ω_v^{OLS} . Following Sims (1972), the least-squares objective seeks OLS coefficients which minimize the average distance between themselves and the true $B(e^{-i\omega})$, weighted by the spectral density of X_t , which may or may not be large at the zero frequency:¹²

$$\min_{B_1^{\text{OLS}}, \dots, B_p^{\text{OLS}}} \Omega_v^{\text{OLS}} = \Omega_v + \int_{-\pi}^{\pi} (B(e^{-i\omega}) - B(e^{-i\omega})^{\text{OLS}}) S_X(\omega) (B(e^{-i\omega}) - B(e^{-i\omega})^{\text{OLS}})^T d\omega \quad (11)$$

¹²This can be derived by applying the definition of spectral density to $v_t^{\text{OLS}} = v_t + [B(L) - B(L)^{\text{OLS}}]X_{t-1}$ and computing the variance Ω_v^{OLS} as the integral over the spectral density.

Accordingly, OLS will try to set $B(1)^{\text{OLS}}$ close to $B(1)$ only if the data's spectrum is high at the zero frequency and $S_X(0)^{\text{OLS}}$ need not be the best possible estimate for the spectral density at frequency zero.

Instead of using $S_x(0)^{\text{OLS}}$, CEV employ a spectral estimator of $S_X(0)$ to construct $A(1)$. In Christiano, Eichenbaum, and Vigfusson (2006a), they consider two estimators, one based on Newey and West (1987) and the other on Andrews and Monahan (1992). Both are based on truncated sums of autocovariance matrices. To ensure positive definiteness, these are weighted by a Bartlett kernel. Where Newey-West sums over the (sample) autocovariances of X_t ,

$$S_X(0)^{\text{NW}} = \sum_{k=-b}^b \left(1 - \frac{|k|}{b+1}\right) E [X_t X_{t-k}^T] \quad (12)$$

Andrews-Monahan uses first the VAR to prewhiten the data and then sums over the residual autocovariances:

$$S_X(0)^{\text{AM}} = C(1)^{\text{OLS}} S_v^{\text{NW}}(0) (C(1)^{\text{OLS}})^T \quad (13)$$

$$\text{where } S_v(0)^{\text{NW}} = \sum_{k=-b}^b \left(1 - \frac{|k|}{b+1}\right) E [v_t^{\text{OLS}} (v_{t-k}^{\text{OLS}})^T] \quad (14)$$

b is a truncation parameter, also known as “bandwidth”, which will be discussed in more detail below.¹³ The Andrews-Monahan estimator nests the OLS case when $b = 0$.

The new CEV estimator computes the long-run coefficients from the non-parametric density estimate. Combined with the OLS lag coefficients, CEV obtain their impact coefficients as

$$A_0^{\text{CEV-AM}} = (I - B(1)^{\text{OLS}}) A(1)^{\text{AM}} \quad (15)$$

$$\text{where } A(1)^{\text{AM}} = \text{chol} \{S_X(0)^{\text{AM}}\} \quad (16)$$

Impulse responses are $A(L)^{\text{CEV-AM}} = C(L)^{\text{OLS}} A_0^{\text{CEV-AM}}$. Using the Newey-West estimator, impact

¹³As elsewhere in this section, estimators have been written in terms of population moments, $E [v_t^{\text{OLS}} (v_{t-k}^{\text{OLS}})^T]$. In empirical applications, the population moments are replaced by sample moments. For some variable Z_t , the sample moment is $E_T Z_t \equiv 1/T \sum_{t=1}^T Z_t$.

coefficients are computed as $A_0^{\text{CEV-NW}} = (I - B(1)^{\text{OLS}}) \text{chol} \{S_X(0)^{\text{NW}}\}$. For brevity, the remainder of this section will mostly refer to the Andrews-Monahan estimator, with similar arguments holding for the Newey-West estimator. Section 4 presents simulations using both estimators.

The bandwidth choice b is critical in estimating spectra, akin to choosing the lag order of a VAR. Bandwidth choice has been shown to be more important using other weighting schemes than the Bartlett kernel (Newey and West 1994).¹⁴ For a consistent estimator, b can grow with the sample size but at a smaller rate. CEV use a fixed and fairly large value of $b = 150$ in a sample of 180 observations.¹⁵

Theoretically, the prewhitening of Andrews-Monahan is appealing since it removes spikes from the spectral density of X_t which make spectral estimation difficult (Priestley 1981, Chapter 7). Andrews and Monahan (1992) and Newey and West (1994) find the prewhitening to fare better in Monte Carlo studies than the original Newey-West estimator. Christiano, Eichenbaum, and Vigfusson (2006a) find no clearly superior choice between the two and Christiano, Eichenbaum, and Vigfusson (2006b) proceed to use only the Newey-West estimator.

To minimize the mean-squared error (MSE) of spectral estimates, the bandwidth selection schemes of Andrews (1991) and Newey and West (1994) can be used. However, constructing an MSE optimal estimator of the spectrum does not necessarily translate into an MSE optimal estimate of coefficients like $A_0^{\text{CEV-AM}}$ or $A_0^{\text{CEV-NW}}$. Their MSE depends not solely on the MSE of $S_X(0)^{\text{AM}}$ but—amongst others—on bias and standard error of the spectrum in ways which are specific to the data generating process.¹⁶ Hence, the bandwidth selection scheme of Newey and West (1994) may serve as a useful starting point for bandwidth choice, but it is not necessarily optimal for the purpose of estimating impulse response or variance shares.

¹⁴Alternative weighting schemes are for example discussed by Priestley (1981) and Phillips, Sun, and Jin (2006).

¹⁵The discussion of CEV suggests that this choice is supposed to be compatible with consistency—essentially promising that this bandwidth choice would barely grow as longer data samples become available. Watson (2006) regards it as a practically untruncated and inconsistent estimator.

¹⁶To be specific, $A_0^{\text{CEV-AM}}$ and $A_0^{\text{CEV-NW}}$ are functions of a spectral estimate and OLS estimates of the VAR. Analogously to arguments employed by Sun, Phillips, and Jin (2008) in the context of constructing confidence intervals, the MSEs of $A_0^{\text{CEV-AM}}$ and $A_0^{\text{CEV-NW}}$ could be approximated to a second order—holding the OLS estimates fixed—by a linear combination of bias and variance of the spectral estimate whose weights depend on the true values of the VAR. Furthermore, the MSEs of $A_0^{\text{CEV-AM}}$ and $A_0^{\text{CEV-NW}}$ will also depend on the covariance between $S_X(0)^{\text{AM}}$, respectively $S_X(0)^{\text{NW}}$, and the OLS estimates.

The simulations reported below use both the optimal bandwidth selection scheme of Newey and West (1994) and the large bandwidth choice of CEV. The former tend to pick fairly small bandwidths. For the various calibrations of the model economy considered here, the optimal bandwidth of the Newey-West estimator is typically close to ten, and the average for the Andrews-Monahan estimator is about four. In simulations not reported here, spectral estimates with intermediate bandwidth choices displayed performance characteristics which were intermediate between what is shown here for these two choices here.

3.3 Conceptual Problems with the CEV Procedure

The CEV procedure is motivated by dissatisfaction with $B(1)^{OLS}$. In conventional SVAR implementations, this estimate is needed for two purposes: First, to construct the long run responses $A(1)$ as in (9), and second in order to map $A(1)$ back into impact responses A_0 as in (10). CEV replace $B(1)^{OLS}$ with a spectral estimate in the first step, but not in the second. This creates a non-negligible inconsistency in representing the overall dynamics of the VAR.

By plugging a spectral estimate into their SVAR computations, CEV have weakened the OLS assumption of uncorrelated residuals without fully accounting for its consequences. As a result, the impact coefficients of CEV will in general not reproduce the forecast error variance of the VAR, which is at the heart of variance computations. These and other consequences are illustrated here. The next sub-section shows how a spectral factorization could be used to incorporate spectral estimates into a VAR model while retaining an internally consistent model of the data.

The spectral estimates embody information about correlation in the VAR residuals v_t^{OLS} . As can be seen from (14), the Andrews-Monahan estimator is constructed from autocovariances of the VAR residuals v_t^{OLS} . Obviously, $b > 0$ expresses a concern about serially correlated residuals. The Newey-West estimator $S_X^{NW}(\omega)$ also embodies concerns about serially correlated VAR residuals since it implies a spectrum for the VAR residuals, which is generally not constant across

frequencies.¹⁷

Under the premise that the true model has only an infinite-order representation, it is indeed very plausible that the residuals from a VAR(p) will be correlated. In the spirit of Andrews and Monahan (1992), the VAR could then be viewed as having merely prewhitened the data. But typically, researchers fit the lag length of their VARs until the point where estimated residuals do not display any significant correlation. Employing a spectral estimate like (14) beyond this point implies a belief that there is still useful information to be gleaned from the estimated residuals—or in other words, it implies a distrust against the lag-selection criteria being chosen for the VAR. By allowing for residual dynamics poses, a researcher risks of overfitting the data, which may still reduce bias in the estimated spectra, but at the expense of a higher standard error.¹⁸ Against this backdrop, the assertion of Christiano, Eichenbaum, and Vigfusson (2006a) that the impulse responses computed from their procedure have “smaller bias, smaller means square error” appear even more striking—and as will be seen, these properties do neither extend to the wider set of calibrations studied below nor to other SVAR statistics like variance decompositions.

A researcher adopting the CEV strategy wants $S_X(0)^{AM} \neq S_X(0)^{OLS}$ and thus $S_v(0)^{NW} \neq \Omega_v^{OLS}$.¹⁹ As a direct implication, the impact coefficients of CEV do not reproduce the forecast error variance of the VAR, $A_0^{CEV-AM}(A_0^{CEV-AM})^T \neq \Omega_v^{OLS}$. When computing the total variance of the data by summing over the conditional variations implied by the SVAR, the CEV procedure would not match the unconditional variance of the data either.²⁰ This mismatch in the unconditional variances

¹⁷ $S_X(\omega)^{NW}$ generalizes (12) to the case of non-zero frequencies with $\Gamma_k \equiv E[X_t X_{t-k}^T]$:

$$S_X(\omega)^{NW} = \Gamma_0 + \sum_{k=1}^b \left(1 - \frac{|k|}{b+1}\right) (\Gamma_k e^{-i\omega k} + (\Gamma_k)^T e^{i\omega k})$$

and the implied spectrum of VAR residuals is $(I - B(e^{-i\omega})^{OLS} e^{-i\omega}) S_X(\omega)^{NW} (I - B(e^{-i\omega})^{OLS} e^{-i\omega})^T$

¹⁸In the simulations reported below, lag length is chosen separately for each sample based on the generally conservative Schwartz Information Criterion, typically choosing a lag length of one in the application presented here.

¹⁹This follows from comparing (15) and (16) with (9) and (10).

²⁰The impulse-responses $A(L)^{CEV-AM} = C(L)^{OLS} A_0^{CEV-AM}$ imply the following variance measure

$$\text{Var } X_t^{CEV-AM} = \sum_{k=0}^{\infty} C_k^{OLS} A_0^{CEV-AM} (A_0^{CEV-AM})^T (C_k^{OLS})^T \neq \sum_{k=0}^{\infty} C_k^{OLS} \Omega_v^{OLS} (C_k^{OLS})^T = \text{Var } X_t$$

The last step holds because of the normal equations (8) regardless of whether v_t^{OLS} is *iid* or not.

of VAR what is implied by the impulse responses of CEV occurs both in population as well as in small sample. A similar argument applies to the Newey-West variant of the CEV procedure, where a researcher seeks $S_X(0)^{NW} \neq S_X(0)^{OLS}$ and thus $(I - B(1)^{OLS})S_X(0)^{NW}(I - B(1)^{OLS})^T \neq \Omega_v^{OLS}$.

[Figure 1 about here.]

For the model economy described in Section 2, Figure 1 illustrates the mismatch in the variance of output growth. In small sample, the variances implied by the CEV procedure are only about half as big as the OLS sample moments. As can be seen in the right panel of Figure 1, this occurs both when using the Newey-West or the Andrews-Monahan variant of the procedure and regardless of the share of fluctuations explained by technology shocks. As depicted in the left panel of the figure, the mismatch is qualitatively different, but also sizable when applying the procedure to the population moments of the model while using a lag length of $p = 1$ and spectral bandwidth of $b = 150$.

The CEV procedure is motivated by concerns about the ability of OLS estimate to correctly capture the low-frequency dynamics of the data. But implicitly, differences between spectra estimated from OLS and the non-parametric methods are not attributed to misspecified dynamics, but rather to the VAR's forecast error variance. However, the accuracy of estimating Ω_v^{OLS} has so far not been doubted. In fact, getting a good estimate for forecast error variance is precisely the objective of OLS projections—see (11) above. Still, the CEV procedure deviates from previous contributions to the SVAR literature where identification is defined as a search over the space of matrices \hat{A}_0 satisfying $\hat{A}_0 \hat{A}_0^T = \Omega_v^{OLS}$.²¹

Finally, a researcher might want to re-construct structural shocks based on (2) as $\varepsilon_t^{CEV-AM} = (A_0^{CEV-AM})^{-1} v_t^{OLS}$ and compare them against $\varepsilon_t^{OLS} = (A_0^{OLS})^{-1} v_t^{OLS}$. She will be troubled noticing that the estimated technology shocks are perfectly correlated:²²

$$(\varepsilon_t^z)^{CEV-AM} = \frac{\bar{a}_{11}^{OLS}}{\bar{a}_{11}^{AM}} \cdot (\varepsilon_t^z)^{OLS} \quad \text{and} \quad (\varepsilon_t^z)^{CEV-NW} = \frac{\bar{a}_{11}^{OLS}}{\bar{a}_{11}^{NW}} \cdot (\varepsilon_t^z)^{OLS}$$

²¹See for example Faust (1998), Canova and de Nicolò (2003) or Uhlig (2005).

²²Recall from (4) that \bar{a}_{11} is the top element of $A(1)$.

This holds for any pair of matrices A_0^1 and A_0^2 constructed from (4) and (6) using $B(1)^{\text{OLS}}$ and a $A(1)$ satisfying the zero restrictions (4). Under those conditions the top rows of $(A(1)^{\text{OLS}})^{-1}$, $(A(1)^{\text{AM}})^{-1}$ and $(A(1)^{\text{NW}})^{-1}$ are identical up to a scaling.²³

Since CEV were only concerned with impulses-responses and A_0 , the problem does not show up in their analysis. The construction of estimated shocks is however often used by researchers, for instance in order to plot historical decompositions or when identifying several shocks (see for example Altig et al. (2004)).

3.4 Correct Identification with Spectral Factorization

To overcome the problems with the CEV procedure discussed above, it is necessary to parse out dynamics of v_t^{OLS} implied by the spectral estimates. Also when the true model has an infinite order VAR representation, OLS projections of X_t on a finite number of lags are well defined in the sense of satisfying the projection equations (8) for $k \leq p$, but the residuals v_t^{OLS} are not *iid*. In general, the residuals follow a moving average representation:

$$v_t^{\text{OLS}} = e_t + D_1 e_{t-1} + D_2 e_{t-2} + D_3 e_{t-3} + \dots = D(L)e_t$$

$$\text{with spectral density } S_v(\omega) = D(e^{-\omega})\Omega D(e^{-\omega})^T \quad (17)$$

$$\text{where } D(L) = (I - B(L)^{\text{OLS}}L)C(L) \quad (18)$$

[Figure 2 about here.]

CKM and CEV discuss a truncation bias which is hard to detect based on VAR lag-length selection procedures. In terms of the moving average $D(L)$, their results can be read as finding $D_i \approx 0$ but $D(1) \neq I$. This can also be illustrated in the model economy described in Section 2.

Figure 2 plots the population values of the cumulated sums $\sum_{k=0}^K D_k$ when $p = 1$ for different

²³Both CEV and OLS use $B(1)^{\text{OLS}}$ in computing $A_0^{-1} = A(1)^{-1}(I - B(1))^{-1}$ and except for the top left element, the first row of $A(1)^{-1}$ is full of zeros. Applying a standard result for inverting partitioned matrices (Magnus and Neudecker 1988, p. 11), the long run restriction (4) places the same zero restrictions on $A(1)^{-1}$ as it does on $A(1)$. This applies equally to $A(1)^{\text{CEV-AM}}$, $A(1)^{\text{CEV-NW}}$ and $A(1)^{\text{OLS}}$. Finally, the top left element of $A(1)^{-1}$ equals $1/\bar{a}_{11}$.

calibrations of the share of fluctuations in output explained by technology shocks. (Results are similar for other values of p .) At each lag, the increments are small and close to zero, but summing over many lags leads to $D(1) \neq I$.

Many moving average representations can be consistent with a given spectrum. But only one of them is invertible. As will be shown next, $D(L)$ is invertible and can be uniquely identified with a spectral factorization of $S_v(\omega)$.

Proposition 1 (Invertibility of $D(L)$). *When the underlying model has a fundamental VAR representation as in (1), and the OLS-VAR is stable, the moving average polynomial $D(L)$ defined in (18) has all its roots outside the unit-circle.*

Proof. The proof is straightforward since $(I - B(L))^{-1} = C(L) = (I - B(L)^{\text{OLS}})^{-1}D(L)$ has all roots outside the unit circle and the same has been assumed for the VMA of the $\text{VAR}(p)$, $C(L)^{\text{OLS}} = (I - B(L)^{\text{OLS}}L)^{-1}$. \square

It is straightforward to recover $D(L)$ from $S_v(\omega)$ with a spectral factorization. The ‘‘canonical spectral factorization’’ is a classic theorem in linear quadratic control, assuring existence and uniqueness of an invertible $D(L)$ and a positive definite Ω consistent with (17).

Theorem 1 (Spectral Factorization, (Hannan 1970)). *Suppose a variable v_t has autocovariances $\Gamma_k \equiv E[v_t v_{t-k}^T] = (\Gamma_{-k})^T$ and a spectral density*

$$S_v(\omega) \equiv \sum_{k=-q}^q \Gamma_k e^{-ik\omega} \quad \forall \quad \omega \in [-\pi, \pi]$$

which is non-singular at each frequency ($|S_v(\omega)| \neq 0 \forall \omega$), as well as being non-zero at the zero frequency, $S_v(0) \neq 0$. There is a unique factorization of $S_v(\omega)$ into

$$S_v(\omega) = D(e^{-ik\omega}) \Omega D(e^{-ik\omega})^T$$

where Ω is positive definite and $D(z)$ is a q 'th order polynomial $D(z) = I + \sum_{k=1}^q D_k z^k$ which has all its roots outside the unit circle.

The theorem factors a spectrum constructed from a finite number of autocovariances into a finite-order MA. For an empirical application, a finite q has of course to be chosen. But when applying the spectral factorization to the population objects of the true model (1), it remains to consider that $D(L)$ is in general an MA(∞). However, since the processes for X_t and v_t^{OLS} are stationary, their autocovariances and MA-coefficients vanish for large lags (Hamilton 1994, Chapter 3.A). Analogous to Sims (1972), a spectral factorization with an arbitrarily large but finite q can arbitrarily well approximate the true spectrum and true $D(L)$. Alternatively the true $D(L)$ can be thought of as being the limit of applying Theorem 1 to an ever increasing sequence of q 's.

For a correct identification of the structural shocks, the true impact coefficients (6) can be written in terms $B(L)^{\text{OLS}}$ and $D(L)$ as

$$A(1) = \text{chol} \{ (I - B(1)^{\text{OLS}})^{-1} S_v(0) (I - B(1)^{\text{OLS}})^{-T} \} \quad (19)$$

$$A_0 = D(1)^{-1} (I - B(1)^{\text{OLS}}) A(1) \quad (20)$$

CEV construct $A(1)^{\text{AM}}$ according to (19) while using the spectral estimate $S_v(0)^{\text{NW}}$. But they ignore the residual dynamics captured by $D(1)$ in (20) when mapping $A(1)^{\text{AM}}$ back into the impact coefficients. As illustrated in Figure 2, $D(1)$ is typically not a diagonal matrix in the model economy, far from equal to the identity matrix. Ignoring the residual captured by $D(L)$ is the source of the variance misrepresentation discussed in the previous subsection.

To combine VAR coefficients and spectral estimates in an internally consistent fashion, a spectral factorization must be used. The spectral factorization of $S_v(\omega)^{\text{NW}}$ yields a unique and invertible MA(b), denoted $D(L)^{\text{SF-AM}}$, and an innovations variance matrix $\Omega^{\text{SF-AM}}$. The superscript ‘‘SF-AM’’ indicates that these are calculated from the residual spectrum employed by the Andrews-Monahan estimator $S_X(\omega)^{\text{AM}}$. Impact coefficients are then

$$\begin{aligned} A_0^{\text{SF-AM}} &= (D(1)^{\text{SF-AM}})^{-1} (I - B(1)^{\text{OLS}}) A(1)^{\text{AM}} \\ &= (D(1)^{\text{SF-AM}})^{-1} A_0^{\text{CEV-AM}} \end{aligned}$$

Sayed and Kailath (2001) survey a number of different algorithms for performing spectral factorizations. The computations reported here use a reliable and efficient algorithm from Li (2005), based on a state space representation of the moving average process of v_t^{OLS} . (Details are given in Appendix A.)

In contrast to the CEV procedure, the spectral factorization is consistent with the variance of the data, in sample as well as in population.

Proposition 2. *By construction, estimates of $A_0^{\text{SF-AM}}$ and $D(L)^{\text{SF-AM}}$ factor the spectral density $S_v(\omega)^{\text{NW}}$ and thus reproduce the variance of the VAR residuals $\Omega_v^{\text{OLS}} = \int_{-\pi}^{\pi} S_v(\omega)^{\text{NW}} d\omega$. As a corollary, this preserves also the unconditional variance of X_t .*

Proof. The spectral estimate is

$$S_v(\omega)^{\text{NW}} = \Gamma_0 + \sum_{k=1}^b \left(1 - \frac{|k|}{b+1}\right) (\Gamma_k e^{-i\omega k} + (\Gamma_k)^T e^{i\omega k})$$

and the result follows from $\int_{-\pi}^{\pi} e^{-i\omega k} d\omega = 0$, regardless of whether population moments $\Gamma_k = E v_t^{\text{OLS}} (v_{t-k}^{\text{OLS}})^T$ or sample moments $\hat{\Gamma}_k = \frac{1}{T} \sum_{t=k}^T v_t^{\text{OLS}} (v_{t-k}^{\text{OLS}})^T$ are used. The corollary follows from the normal equations of the VAR, which enforce $\text{Var } X_t = \sum_{k=0}^{\infty} C_k^{\text{OLS}} \Omega_v^{\text{OLS}} (C_k^{\text{OLS}})^T$. \square

A spectral factorization can also be applied directly to the Newey-West estimate of the data's spectrum, $S_X \omega^{\text{NW}}$, yielding coefficients for the VMA of X_t , $C(L)^{\text{SF-NW}}$ and innovation variance $\Omega^{\text{SF-NW}}$. Following (6), impact coefficients and impulse responses can then be computed as

$$\begin{aligned} A_0^{\text{SF-NW}} &= (C(1)^{\text{SF-NW}})^{-1} A(1)^{\text{NW}} \\ A(L)^{\text{SF-NW}} &= C(L)^{\text{SF-NW}} A_0^{\text{SF-NW}} \end{aligned}$$

These impulse responses do not involve any VAR coefficients. Analogously to Proposition 2, their construction preserves the variance of the data.

4 SVARs applied to Data from Lab Economy

The previous section described several schemes for imposing the long-run restriction (4) on the data. The conventional method, going back to Blanchard and Quah (1989), uses OLS estimates of a VAR. The recently proposed procedure of CEV combines this with a non-parametric estimate of the spectral density at frequency zero. This procedure has been criticized above for its lack of internal consistency. Finally, this paper proposed a new method, combining OLS estimates and spectral estimators in an internally consistent way. This method relies on a spectral factorization (“SF”) to uncover the dynamics implied by the non-parametric spectral estimators.

These procedures are applied here to data simulated from the model economy described in Section 2. The same data generating process has also been used by CEV and CKM.²⁴ For the CEV and SF methods, there are two variants depending on whether the spectral estimators of Newey and West (1987) or Andrews and Monahan (1992) are used. This section reports results for both.

Mimicking conditions faced by empirical researchers, “small” samples with 180 observations are simulated. In small sample, two distinct issues arise. First, there is truncation bias in VARs and spectral estimators arising from the need to specify a finite lag length p , respectively a finite bandwidth b . As discussed in Section 2, lag length is determined individually for each draw with an information criterion and spectral bandwidth is fixed at 150. In addition, alternative results are reported using the bandwidth selection procedure of Newey and West (1994) for Newey-West spectra. (See Section 3.2 for further discussion of bandwidth selection.)

Second, there is the small sample bias in estimated parameters known from Hurwicz (1950).²⁵ To isolate the pure truncation effects from the Hurwicz bias, the identification procedures are not only applied to simulated data, but also to VARs and spectral estimates constructed from the model’s true population moments.²⁶

²⁴As discussed in Section 2, the calibrations employed here are identical to the setting of CEV—except for considering a wider range of the technology share in output fluctuations. As discussed above as well, the CKM experiments differ slightly in their choice of ψ and ρ_l .

²⁵This bias is particularly acute the smaller the sample and the higher the persistence of the data. It is pertinent in this example, since calibrating the model to match salient features of U.S. data requires a high degree of persistence in the non-technology shock, ρ_l .

²⁶In the case of the spectral estimators, this means evaluating (12) and (13) at true population moments, instead of

The procedures are evaluated in terms of their capability to uncover two statistics typically of interest to applied researchers. Following CKM and CEV, estimates of the response of hours to a technology shock are computed. For brevity, the discussion is limited on impact coefficients A_0 , since all methods compute impulse responses from $C(L)^{OLS}$ and their estimates of A_0 (except when factorizing $S_X(\omega)^{NW}$). In addition, the share of fluctuations in output and hours due to technology shocks is estimated. As it is typical in the business cycle literature, these shares are computed after filtering out any fluctuations which do not correspond to cycles with a duration between two-and-a-half and eight years.²⁷ Two criteria are reported to assess the goodness of estimates: Bias and Root Mean Square Error (RMSE), both expressed as percentages relative to the true value known from the model.²⁸

The results show that all procedures are subject to substantial truncation and small biases and none works like a panacea. Different methods display different strengths and weaknesses. The claims by Christiano, Eichenbaum, and Vigfusson (2006a) of “smaller bias, smaller means square error” associated with their procedure do neither generalize to a wider range of model calibrations nor do they extend from the estimation of impact responses to variance shares.

[Figure 3 about here.]

Effects from the truncation and the small sample bias can offset each other. This is the case when estimating the impact of technology on hours. The left column in Figure 3 shows how impact responses are overestimated in population whereas the simulated bias shown in the middle column of the figure is lower (more negative). This simulated bias displays the total effect from truncation and Hurwicz bias. The OLS method has the largest population bias and it is only partially offset by the Hurwicz bias. The two spectral methods suffer from substantially smaller truncation bias,

sample autocovariances, while keeping the Bartlett weights and the truncation at the chosen bandwidth (here: $b = 150$, respectively $b = 15$ when comparing against simulations using the bandwidth selection procedure of Newey and West (1994)). The computation of VARs from population moments is equally straightforward, and details are described in Appendix B.

²⁷The variance computations are explained in Appendix C.

²⁸Denoting the estimated parameter as θ and its estimate as $\hat{\theta}$, relative bias is computed as $E(\hat{\theta} - \theta)/\theta \cdot 100\%$. The RMSE is defined as $RMSE = \sqrt{E(\hat{\theta} - \theta)^2} = \sqrt{(E\hat{\theta} - \theta)^2 + \text{Var} \hat{\theta}}$ and it is converted into a percentage error using $RMSE/\theta \cdot 100\%$. In both cases, expectations are computed from the arithmetic average over 1,000 simulations.

and depending on the simulated importance of technology shocks, the total bias can be either negative or positive. Coincidentally, the upwards bias in SF-AM and SF-NW is exactly offset around technology shares of about two thirds, corresponding to the range of MLE estimates of CEV and CKM for U.S. data. (Similarly for CEV-AM, but not CEV-NW.) However, results are different for other calibrations of the technology share, which cautions strongly against extrapolating from a particular result to different data sets and different applications.

Unless the true share of technology shocks is very large, the RMSE of estimated impact coefficients are very large, often surpassing more than 100% of the true value. Interestingly, the RMSE do not differ much across the different methods, as can be seen in the right-most column of Figure 3. If anything, SF-NW is outperforming CEV-NW on bias, at the expense of a worse RMSE. This is likely due to an overfitting of the residual dynamics by SF-NW.

[Figure 4 about here.]

Turning to the estimated variance shares of output and hours shown in Figure 4, the relative performance of the various methods looks quite different. The panels in the top row of the figure show bias and RMSE for variance decompositions of output, the bottom row for variance decompositions of hours. For this figure, spectral densities have been estimated with the Andrews-Monahan estimator. Results are broadly similar when using the Newey-West estimator (see Figure A.1 in the separate appendix with additional results.)

Strikingly, for technology shares in output, bias and RMSE are very similar when using either OLS or CEV. The mismatch in total variance discussed in Section 3.3, does not seem to distort the computations of relative variance measures in this case. But, the two methods differ when decomposing the variance of hours. Bias and RMSE in the variance decomposition of hours are an order of magnitude larger than for output, cautioning very strongly against neglecting small sample issues when comparing SVAR estimates against model predictions. Moreover, the variance decompositions of hours provide a useful counterexample against disregarding OLS methods altogether, since OLS dominates the spectral methods both in terms of simulated bias and RMSE for all calibrations considered here. All in all, these results underline how truncation and Hurwicz bias

interact with the different methods in ways which are hard to anticipate for an empirical researcher who does not know the true dynamics of the data.

The results presented in Figures 3 and 4 are based on a large and fixed spectral bandwidth of $b = 150$. A separate appendix with additional results shows that the results are similar for Newey-West spectra when their bandwidth has been chosen by the automatic bandwidth selection procedure of Newey and West (1994). Compared to the case of a large and fixed bandwidth, only two differences stand out. Estimating technology shares from a direct factorization of the Newey West spectrum perform worse compared to the large bandwidth case, both in terms of bias and RMSE, unless technology accounts for less than two thirds of business cycle fluctuations in output (see Figure A.1 in the appendix). Furthermore, the RMSE of impact coefficients estimated with CEV-NW is almost flat at around two thirds of the true value, independently of the true technology share (Figure A.2). Applying the automatic bandwidth selection for the residual spectra of the Andrews-Monahan estimator yields bandwidths close to zero, such that the results are mostly indistinguishable from the OLS estimates (Figure A.3).

5 Conclusions

In finite sample, truncation bias and Hurwicz bias pose fundamental problems when identifying structural shocks from restrictions on the long-run behavior of the data. These issues are present in the time domain when working with a VAR, as well as in the frequency domain when working with spectral estimators. Basically, the same estimates of the data's autocovariances are employed for constructing non-parametric estimates of the spectrum as well as for computing OLS coefficients. In both cases, truncation bias arises since there are only as many sample autocovariances as there are data points. And due to the Hurwicz bias, variance estimates tend to be biased downwards the smaller the sample and the larger the persistence of the data—again affecting both OLS estimates of VAR coefficients as well as non-parametric estimates of the spectral density.

Thus, spectral estimates offer no panacea against the truncation and small sample problems

known from OLS. At best, by allowing for additional dynamics, they might improve upon OLS in terms of bias, but by overfitting the data, this comes at the expense of increasing RMSE.

The performance of different estimators appears to be very specific to the underlying model and its calibration, making it hard to predict, which procedure would do well in future applications using new data. Even for a given calibration, when a method performs better in terms of one model statistic, say impact coefficients, this does not necessarily translate into better performance for another statistic, like a variance share. Going forward, it would be more suitable to compare SVAR estimates (from any procedure), against the small sample predictions, not the true moments, of a specific model as in Cogley and Nason (1995), Kehoe (2006), Dupaigne, Feve, and Matheron (2007) and Dupaigne and Feve (2009).

Appendix

A Spectral Factorization Method

Spectral factorization has a long tradition in the fields of linear quadratic control, robust estimation and control as surveyed for example by Whittle (1996).²⁹ Theorem 1 has been adapted from Hannan (1970, p. 66). The original theorem allows for unit roots in $D(L)$. The version stated above has been slightly strengthened by excluding the case of zero power in the spectral density at zero-frequency, to ensure the invertibility of the MA(b).³⁰

In the context of this paper, $S_v(\omega)$ will be the spectral density of $v_t^{\text{OLS}} = D(L)e_t$ where $Ee_t e_t^T = \Omega = A_0 A_0^T$. We will be using non-parametric estimates of $S_v(\omega)$ based on weighted sums of the sample autocovariance function as described in Section 3.2.³¹

Theorem 1 requires $S_v(\omega)$ to be non-singular. This can be understood as requiring that the autocovariances need to decay sufficiently fast in relation to the number of MA lags. For example, in the scalar case and with $q = 1$, the first-order autocorrelation to be matched with a MA(1) cannot be larger than 0.5 in absolute value.³²

Algorithms for implementing the factorization go back to Whittle (1963) and have recently been surveyed by Sayed and Kailath (2001). The simulations reported here use the algorithm of Li (2005), which is based on a state space representation of v_t and performed very reliably.³³ The remainder of this appendix describes the algorithm in more detail.

Suppose v_t follows an MA(q) as above. To represent it in a state space system, define the state

²⁹For a reference in the context of economics see Hansen and Sargent (2007, 2005).

³⁰Suppose that $S(0) \neq 0$. Since Ω is positive definite, it follows that $D(1) \neq 0$. All roots of $D(z)$ are thus outside the unit circle and $D(L)$ is an invertible MA(b).

³¹The Γ_k from Theorem 1 are a smoothed version of the sample autocovariance since they are the coefficients of an inverse Fourier transform of the Newey-West estimate of the spectral density.

³²Given a covariance γ_0 and first-order autocovariance γ_1 , the spectrum equals $s(\omega) = \gamma_0 \cdot (1 + 2\gamma_1 \cos(\omega))$. And $|s(\omega)| \neq 0$ requires $|\gamma_1/\gamma_0| < 0.5$.

³³The paper of Li also shows how to reduce the number of iterations by stacking the MA(q) into first order form, however this comes at the cost of inverting larger matrices in the Riccati iterations which proved to be numerically less stable in the simulations computed for this paper.

vector $s_t = E \left\{ \left[\begin{array}{cccc} v_t & v_{t+1} & \dots & v_{t+q-1} \end{array} \right]^T \middle| v^{t-1} \right\}$ where v^{t-1} is the entire history of realizations of v_t up to time $t - 1$. Li then constructs the following state space system

$$\begin{aligned} s_{t+1} &= A s_t + D e_t \\ v_t &= C s_t + e_t \end{aligned}$$

$$A = \begin{bmatrix} 0_m & I_m & 0_m & \dots & 0_m \\ 0_m & 0_m & I_m & 0_m & \dots & 0_m \\ \vdots & & & \ddots & & \vdots \\ 0_m & \dots & & 0_m & I_m \\ 0_m & \dots & & 0_m & 0_m \end{bmatrix} \quad D = \begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_q \end{bmatrix} \quad C = \begin{bmatrix} I_m & 0_m & \dots & 0_m \end{bmatrix}$$

where I_m and 0_m are the $m \times m$ identity matrix, respectively the $n \times n$ zero matrix.

What is needed is a mapping from the autocovariances of v_t , Γ_k , to the state space objects. The objects of interest are the matrix D containing the stacked MA coefficients D_i as well as the variance $\Omega = E e_t e_t^T$ of the innovations process. To obtain this mapping, it is useful to stack the autocovariances into a matrix $M = \begin{bmatrix} \Gamma_1^T & \Gamma_2^T & \vdots & \Gamma_q^T \end{bmatrix}^T$

Li (2005, Theorem 2) shows that the variance-covariance matrix of the states $\Psi \equiv E s_t s_t^T$ solves the Riccati equation $\Psi = A \Psi A^T + (M - A \Psi C^T)(\Gamma_0 - C \Psi C^T)^{-1}(M - A \Psi C^T)^T$ and that the MA(q) coefficients can be recovered as $D = (M - A \Psi C^T)(\Gamma_0 - C \Psi C^T)^{-1}$ and $\Omega = \Gamma_0 - C \Psi C^T$. As shown by Li (2005), the above Riccati equation can be solved recursively, starting from $\Psi^{(0)} = 0$ and iterating over $\Psi^{(n+1)} = A \Psi^{(n)} A^T + (M - A \Psi^{(n)} C^T)(\Gamma_0 - C \Psi^{(n)} C^T)^{-1}(M - A \Psi^{(n)} C^T)^T$ since $\Psi = \lim_{n \rightarrow \infty} \Psi^{(n)}$ and $\Psi^{(n+1)} \geq \Psi^{(n)}$.

At the end of each factorization computed for this paper, it has been verified that the factorization produces an invertible MA(q) polynomial, which matches the original spectral density. In all simulations, this held up to machine accuracy.

B VARs Implied by Lab Economy

This section outlines how to derive the following: First, values from the lab economy for true VAR objects like A_0 , $A(1)$, $B(1)$, and the autocovariances of X_t . Second, population coefficients of finite-order VARs implied by the lab economy.³⁴

The linearized solution to the lab economy described in Section 2 yields a state space model for labor productivity growth and hours

$$X_t = \begin{bmatrix} \Delta \log(y_t/l_t) \\ \log l_t \end{bmatrix} = \mathcal{C}Z_t \quad \text{with} \quad Z_t = \mathcal{A}Z_{t-1} + \mathcal{B}\varepsilon_t \quad (21)$$

State vector and shock vector are:

$$Z_t = \begin{bmatrix} \hat{k}_t & \varepsilon_t^z & \tau_{l,t}^z & \hat{k}_{t-1} & \varepsilon_{t-1}^z & \tau_{l,t-1}^z \end{bmatrix}^T \quad \varepsilon_t = \begin{bmatrix} \varepsilon_t^z & \varepsilon_t^l \end{bmatrix}^T \quad (22)$$

where \hat{k}_t is the log-deviation of detrended capital from its steady state, $\tau_{l,t}$ and ε_t^z are the labor wedge and the growth rate in technology. (Z_t includes also lagged variables due to the presence of labor productivity growth in X_t .)

The computation of the matrices \mathcal{A} , \mathcal{B} and \mathcal{C} is straightforward, please see CKM for a detailed presentation.

True VAR objects

The decomposition in section 4 uses the following objects of the true process: A_0 , $A(1)$, $B(1)$ as well as the autocovariances of X_t . Their computation from the state space is straightforward since true impulse responses and spectrum are given by $A(L) = \mathcal{C}(I - \mathcal{A}L)^{-1}\mathcal{B}$ and $S_X(\omega) = A(e^{-i\omega})A(e^{-i\omega})^T$. The impact coefficients $A_0 = \mathcal{C}\mathcal{B}$ are apparent from (21). Recalling equation

³⁴For this specific two-shock economy, details can also be found in McGrattan (2005). For general state space models details can be found in Fernandez-Villaverde, Rubio-Ramirez, and Sargent (2005). To simplify the VAR notation, X_t has been demeaned prior to the analysis.

(2), this also pins down the covariance matrix of the forecast errors $\Omega = \mathcal{C}\mathcal{B}\mathcal{B}\mathcal{C}^T$.

In order to map forecast errors into structural shocks, A_0 must obviously be square and invertible. Furthermore, Fernandez-Villaverde, Rubio-Ramirez, and Sargent (2005) show that invertibility requires the eigenvalues of $\mathcal{A} - \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\mathcal{A}$ to be strictly less than one in modulus, which is satisfied for all calibrations considered here.

The non-structural moving average representation of X_t is $X_t = A(L)A_0^{-1}e_t = C(L)e_t$. From (3), the coefficients of the non-structural VAR(∞) representation of the model can be obtained by inverting this moving average, yielding $B(L)L = I - C(L)^{-1}$.

The autocovariances $EX_tX_{t-k}^T$ can be directly computed from the state space model. The covariance matrix of the states $EZ_tZ_t^T \equiv \Omega$ is obtained as the solution to a discrete Lyapunov equation: $\Omega = \mathcal{A}\Omega\mathcal{A}^T + \mathcal{B}\mathcal{B}^T$ and the autocovariances of X_t are $EX_tX_{t-k}^T = \mathcal{C}\mathcal{A}^k\Omega\mathcal{C}^T$.

VAR(p) coefficients in population

Finite-order VAR(p) can be computed as projections of X_t on a finite number of its past values, $X_{t-1} \dots X_{t-p}$. In line with the notation of the main text, population coefficients of a VAR(p) are denoted with a superscript “OLS”.

$$X_t = B(L)^{\text{OLS}}X_{t-1} + v_t^{\text{OLS}}$$

The coefficients of the lag polynomial $B(L)^{\text{OLS}} = \sum_{i=0}^{p-1} B_i^{\text{OLS}}L^i$ solve the OLS normal equations

$$E \left(X_t - \sum_{i=0}^{p-1} B_i^{\text{OLS}}X_{t-1-i} \right) X_{t-j}^T = 0 \quad \forall j = 1 \dots p$$

which are evaluated using the autocovariance matrices of X_t whose computations are described in the preceding paragraph. For instance if $p = 1$, $B_1^{\text{OLS}} = (EX_tX_{t-1}^T) (EX_tX_t^T)^{-1}$. Detailed formulas for higher VARs can be found in Fernandez-Villaverde, Rubio-Ramirez, and Sargent (2005).

Chari, Kehoe, and McGrattan (2005, Proposition 1) show that the VAR representation of X_t in the model is of infinite order and residuals from a VAR(p) will not be martingales. By construction, the projection residuals v_t^{OLS} are orthogonal to X_{t-1}, \dots, X_{t-p} , but they are not orthogonal to the complete history of X_t . The moving average representation of the forecast errors $v_t^{\text{OLS}} = D(L)e_t$ is easily constructed from $D(L) = (I - B(L)^{\text{OLS}}L) (I - B(L)L)^{-1}$.

Variance equation

Even though the VAR(p) residuals v_t^{OLS} are not *iid*, the usual variance equation is still applicable. For notational convenience, take the case of a VAR(1), $X_t = B_1^{\text{OLS}}X_{t-1} + v_t^{\text{OLS}}$. The normal equations imply

$$\begin{aligned} \text{Var } X_t &= B_1^{\text{OLS}} (\text{Var } X_t) (B_1^{\text{OLS}})^T + \Omega_v^{\text{OLS}} & (23) \\ &= \sum_{k=0}^{\infty} (B_1^{\text{OLS}})^k \Omega_v^{\text{OLS}} ((B_1^{\text{OLS}})^k)^T \\ &= \sum_{k=0}^{\infty} C_k^{\text{OLS}} \Omega_v^{\text{OLS}} (C_k^{\text{OLS}})^T \end{aligned}$$

The second line is obtained by recursive substitution of $\text{Var } X_t$ and the third line follows from the construction of moving-average coefficients of a VAR(1), $C_k^{\text{OLS}} = (B_1^{\text{OLS}})^k$. The argument is easily extended to VARs with higher lag orders by using their companion form.

C Bandpass-Filtered Variance Share from SVARs

This appendix describes how to compute the share of bandpass-filtered fluctuations attributed to technology shocks from a set of SVAR parameters, $\hat{B}(L)$ and \hat{A}_0 . The bandpass filter employed here considers only cycles with a duration between two-and-a-half and eight years. Denoting the bandpass-filtered level of (log) output \tilde{y}_t , its variance can be easily computed from the transfer function

$$T_y(\omega) = \begin{bmatrix} (1 - e^{-i\omega})^{-1} & 1 \end{bmatrix} \left(I - \hat{B}(e^{-i\omega})e^{-i\omega} \right)^{-1} \hat{A}_0$$

Depending on the identification scheme, \hat{A}_0 corresponds to A_0 (true value) or A_0^{OLS} , $A_0^{\text{CEV-AM}}$, $A_0^{\text{CEV-NW}}$, $A_0^{\text{SF-AM}}$ or $A_0^{\text{SF-NW}}$ and $\hat{B}(L)$ corresponds to the true (infinite order) polynomial $B(L)$ or its finite-order counterpart $B(L)^{\text{OLS}}$. When using the spectral factorization of $S_X(\omega)^{\text{NW}}$, the inverse of $I - \hat{B}(e^{-i\omega})e^{-i\omega}$ is replaced by $C^{\text{SF-NW}}(e^{-i\omega})$.

Using $\underline{\omega} = \frac{2\pi}{8 \cdot 12}$ and $\bar{\omega} = \frac{2\pi}{2.5 \cdot 12}$ the bandpass-filtered variance is

$$\text{Var } \tilde{y}_t = \int_{\underline{\omega}}^{\bar{\omega}} T_y(\omega) T_y(\omega)^T d\omega$$

and the share of fluctuations attributed to technology shocks is the ratio $(\text{Var } \tilde{y}_t | \varepsilon_t^z) / (\text{Var } \tilde{y}_t)$, where $\text{Var } \tilde{y}_t | \varepsilon_t^z$ conditions only on fluctuations attributed to technology shocks.

$$\text{Var } \tilde{y}_t | \varepsilon_t^z = \int_{\underline{\omega}}^{\bar{\omega}} T_y(\omega) \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} T_y(\omega)^T d\omega$$

Similar computations yield the variance shares for hours, when using the transfer function $T_l(\omega) = \begin{bmatrix} 0 & 1 \end{bmatrix} \left(I - \hat{B}(e^{-i\omega})e^{-i\omega} \right)^{-1} \hat{A}_0$.

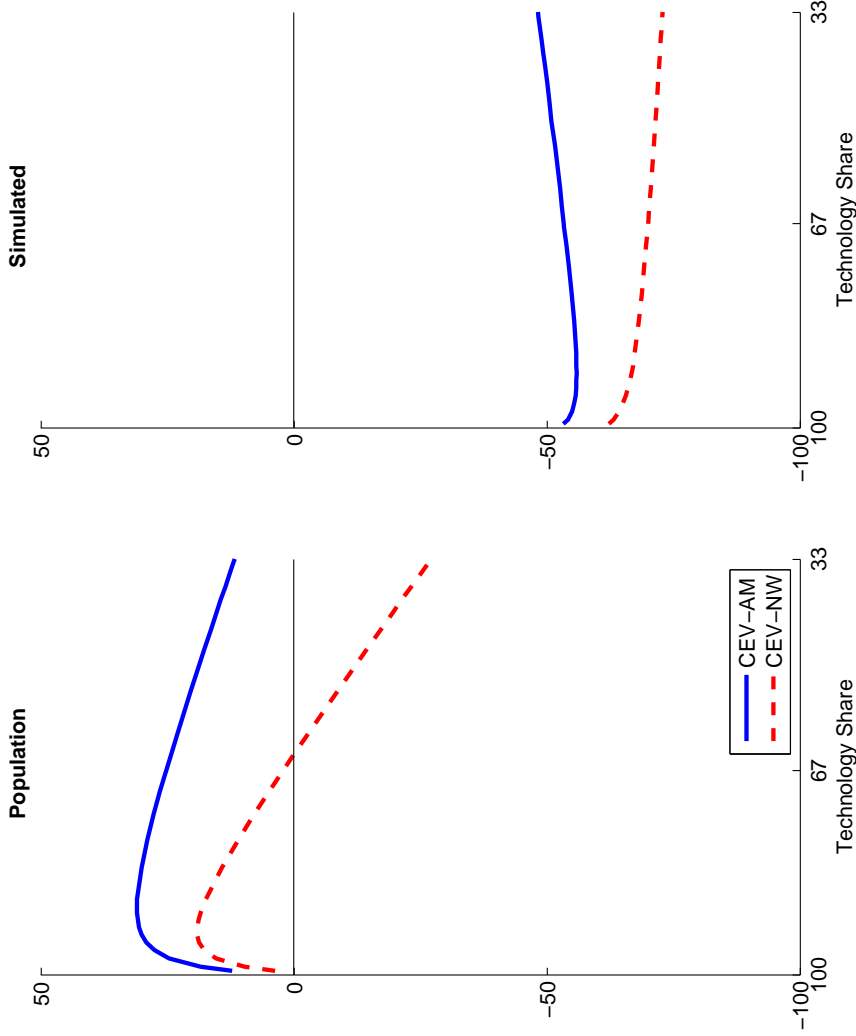
References

- Acemoglu, Daron, Kenneth Rogoff, and Michael Woodford, eds. 2006. *NBER Macroeconomics Annual 2006*. Cambridge, MA: MIT Press.
- Altig, David, Lawrence J. Christiano, Martin Eichenbaum, and Jesper Linde. 2004, December. "Firm-Specific Capital, Nominal Rigidities and the Business Cycle." Northwestern University, mimeo.
- Andrews, Donald W K. 1991. "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica* 59 (3): 817–58 (May).
- Andrews, Donald W K, and J Christopher Monahan. 1992. "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator." *Econometrica* 60 (4): 953–66 (July).
- Blanchard, Olivier Jean, and Danny Quah. 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *The American Economic Review* 79 (4): 655–673 (September).
- Burns, Arthur F., and Wesley C. Mitchell. 1946. *Measuring Business Cycles*. NBER.
- Canova, Fabio, and Gianni de Nicolò. 2003. "On the sources of business cycles in the G-7." *Journal of International Economics* 59 (1): 77–100 (January).
- Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2005, July. "A critique of structural VARs using business cycle theory." Staff Report 364, Federal Reserve Bank of Minneapolis.
- . 2007. "Business Cycle Accounting." *Econometrica* 75 (3): 781–836 (May).
- Chari, V.V., Patrick J. Kehoe, and Ellen R. McGrattan. 2008. "Are structural VARs with long-run restrictions useful in developing business cycle theory?" *Journal of Monetary Economics* 55 (8): 1337–1352 (November).
- Christiano, Lawrence J., Martin Eichenbaum, and Robert Vigfusson. 2003, July. "What Happens After a Technology Shock?" Working Paper 9819, NBER.
- . 2004. "The Response of Hours to a Technology Shock: Evidence based on a Direct Measure of Technology." *Journal of the European Economic Association* 2 (2–3): 381–395 (April–May).
- . 2006a. "Alternative Procedures for Estimating Vector Autoregressions identified with Long-Run Restrictions." *Journal of the European Economic Association* 2 (2–3): 475–483 (April–May).
- . 2006b. "Assessing Structural VARs." In Acemoglu, Rogoff, and Woodford 2006, Chapter 1.
- Cogley, Timothy, and James M Nason. 1995. "Output Dynamics in Real-Business-Cycle Models." *The American Economic Review* 85 (3): 492–511.
- Cooley, Thomas F., and Mark Dywer. 1998. "Business cycle analysis without much theory A look at structural VARs." *Journal of Econometrics* 83 (1-2): 57–88 (March-April).
- Dupaigne, Martial, and Patrick Feve. 2009. Technology shocks around the world. *Review of Economic Dynamics*, forthcoming.

- Dupaigne, Martial, Patrick Feve, and Julien Matheron. 2007. "Avoiding pitfalls in using structural VARs to estimate economic models." *Review of Economic Dynamics* 10 (2): 238–255 (April).
- Erceg, Christopher J., Luca Guerrieri, and Christopher Gust. 2005. "Can Long-Run Restrictions Identify Technology Shocks?" *Journal of the European Economic Association* 3 (6): 1237–1278 (December).
- Faust, Jon. 1998. "The robustness of identified VAR conclusions about money." *Carnegie-Rochester Conference Series on Public Policy* 49 (December): 207–244.
- Fernandez-Villaverde, Jesus, Juan Rubio-Ramirez, and Thomas J. Sargent. 2005, June. "A, B, C's (and D)'s for Understanding VARs." Technical Working Paper 0308, NBER.
- Fernández-Villaverde, Jesús, Juan F. Rubio-Ramirez, Thomas J. Sargent, and Mark W. Watson. 2007. "ABCs (and Ds) of Understanding VARs." *American Economic Review* 97 (3): 1021–1026 (June).
- Gali, Jordi. 1999. "Technology, Employment, and the Business Cycle: Do Technology Shocks explain Aggregate Fluctuations?" *The American Economic Review* 89 (1): 249–271.
- Gali, Jordi, and Pau Rabanal. 2004, July. "Technology Shocks and Aggregate Fluctuations: How well does the RBC Model fit postwar U.S. data?" NBER Working Paper 10636.
- Giannone, Domenico, and Lucrezia Reichlin. 2006. "Does information help recovering structural shocks from past observations?" *Journal of the European Economic Association* 4 (2-3): 455–465 (April-May).
- Hamilton, James D. 1994. *Time-Series Analysis*. Princeton, NJ: Princeton University Press.
- Hannan, E. J. 1970, June. *Multiple Time Series*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons.
- Hansen, Lars P., and Thomas J. Sargent. 2005, September. "Recursive Models of Dynamic Linear Economies." unpublished Manuscript.
- . 2007. *Robustness*. Princeton University Press.
- Hurwicz, Leonid. 1950. "Least-Squares Bias in Time Series." Chapter XV of *Statistical Inference in Dynamic Economic Models*, edited by Tjalling C. Koopmans, Volume 10 of *Cowles Commission for Research in Economics*, 365–383. New York: John Wiley & Sons, Inc.
- Kehoe, Patrick J. 2006. "How to advance theory with structural VARs: use the Sims-Cogley-Nason approach." Staff Report 379, Federal Reserve Bank of Minneapolis.
- Li, Lei M. 2005. "Factorization of moving-average spectral densities by state-space representations and stacking." *Journal of Multivariate Analysis* 96 (2): 425–438 (October).
- Magnus, Jan R., and Heinz Neudecker. 1988. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley Series in Probability and Statistics. John Wiley & Sons.
- Marcet, Albert. 2005. Overdifferencing VAR's is OK. *mimeo*.
- McGrattan, Ellen. 2005, May. "Appendix: A Critique of Structural VARs Using Real Business Cycle Theory." Federal Reserve Bank of Minneapolis, *mimeo*.
- Newey, Whitney K, and Kenneth D West. 1994. "Automatic Lag Selection in Covariance Matrix Estimation." *Review of Economic Studies* 61 (4): 631–53.

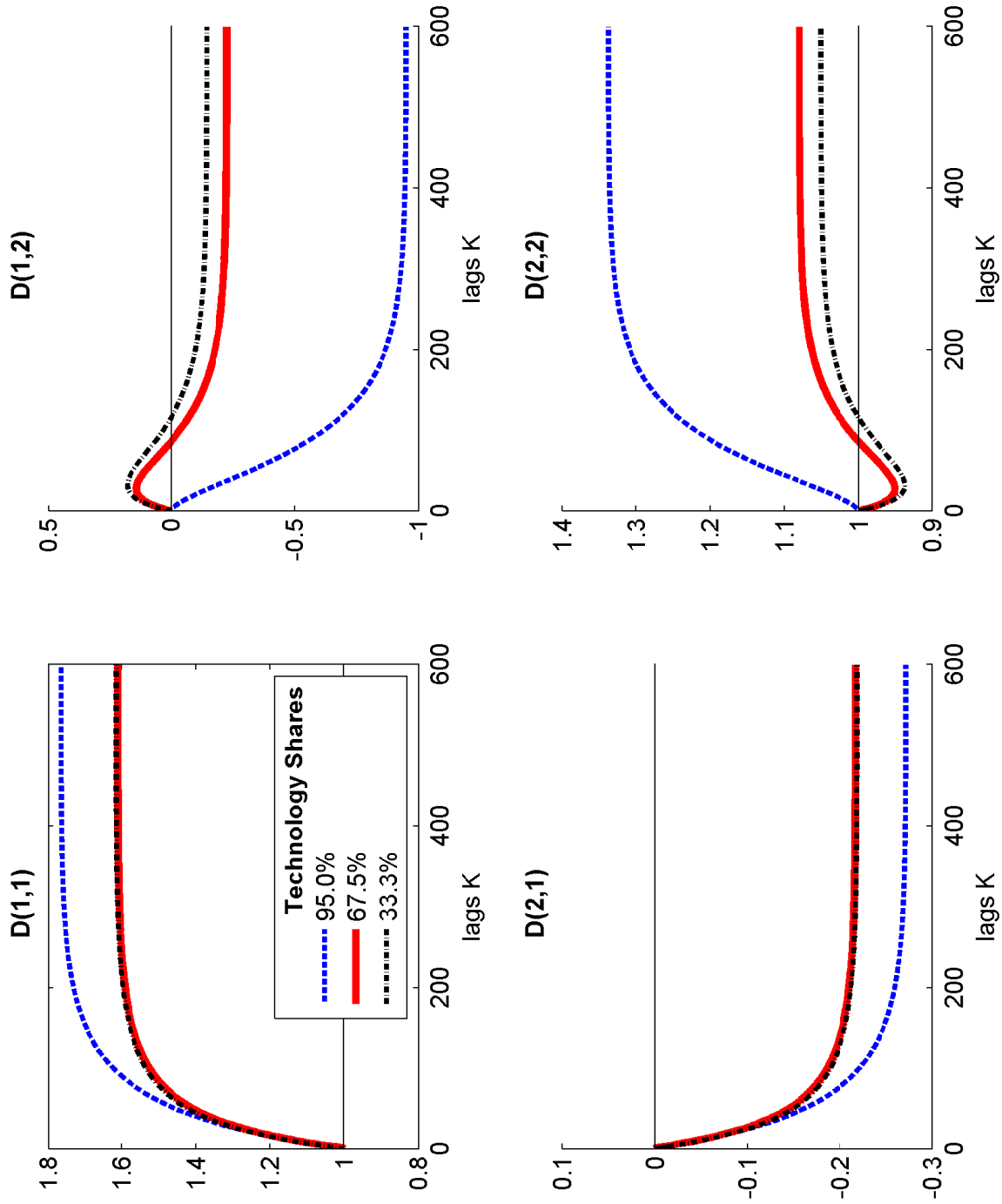
- Newey, W.K., and K.D. West. 1987. "A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica* 55:703–708.
- Phillips, Peter C. B., Yixiao Sun, and Sainan Jin. 2006. "Spectral Density Estimation And Robust Hypothesis Testing Using Steep Origin Kernels Without Truncation." *International Economic Review* 47 (3): 837–894 (August).
- Priestley, M.B. 1981. *Spectral Analysis and Time Series*. Probability and Mathematical Statistics. Amsterdam: Elsevier Academic Press.
- Ravenna, Federico. 2007. "Vector autoregressions and reduced form representations of DSGE models." *Journal of Monetary Economics* 54 (7): 2048–2064 (October).
- Sayed, A.H., and T. Kailath. 2001. "A Survey of Spectral Factorization Methods." *Numerical Linear Algebra with Applications* 8:467–496.
- Sims, Christopher A. 1972. "The Role of Approximate Prior Restrictions in Distributed Lag Estimation." *Journal of the American Statistical Association* 67 (337): 169–175 (March).
- Sun, Yixiao, Peter C. B. Phillips, and Sainan Jin. 2008. "Optimal Bandwidth Selection in Heteroskedasticity-Autocorrelation Robust Testing." *Econometrica* 76 (1): 175–194 (January).
- Uhlig, Harald. 2005. "What are the effects of monetary policy on output? Results from an agnostic identification procedure." *Journal of Monetary Economics* 52 (2): 381–419 (March).
- Watson, Mark W. 2006. "Comment on Christiano, Eichenbaum, and Vigfusson's "Assessing Structural VARs"." In Acemoglu, Rogoff, and Woodford 2006, Chapter 1.
- Whittle, P. 1963. "On the Fitting of Multivariate Autoregressions, and the Approximate Canonical Factorization of a Spectral Density Matrix." *Biometrika* 50 (1/2): 129–134 (June).
- Whittle, Peter. 1996. *Optimal Control, Basics and Beyond*. Wiley-Interscience Series in Systems and Optimization. Chichester: John Wiley & Sons.

Figure 1: Percentage Errors in Variance of Output Growth Computed from CEV Procedure



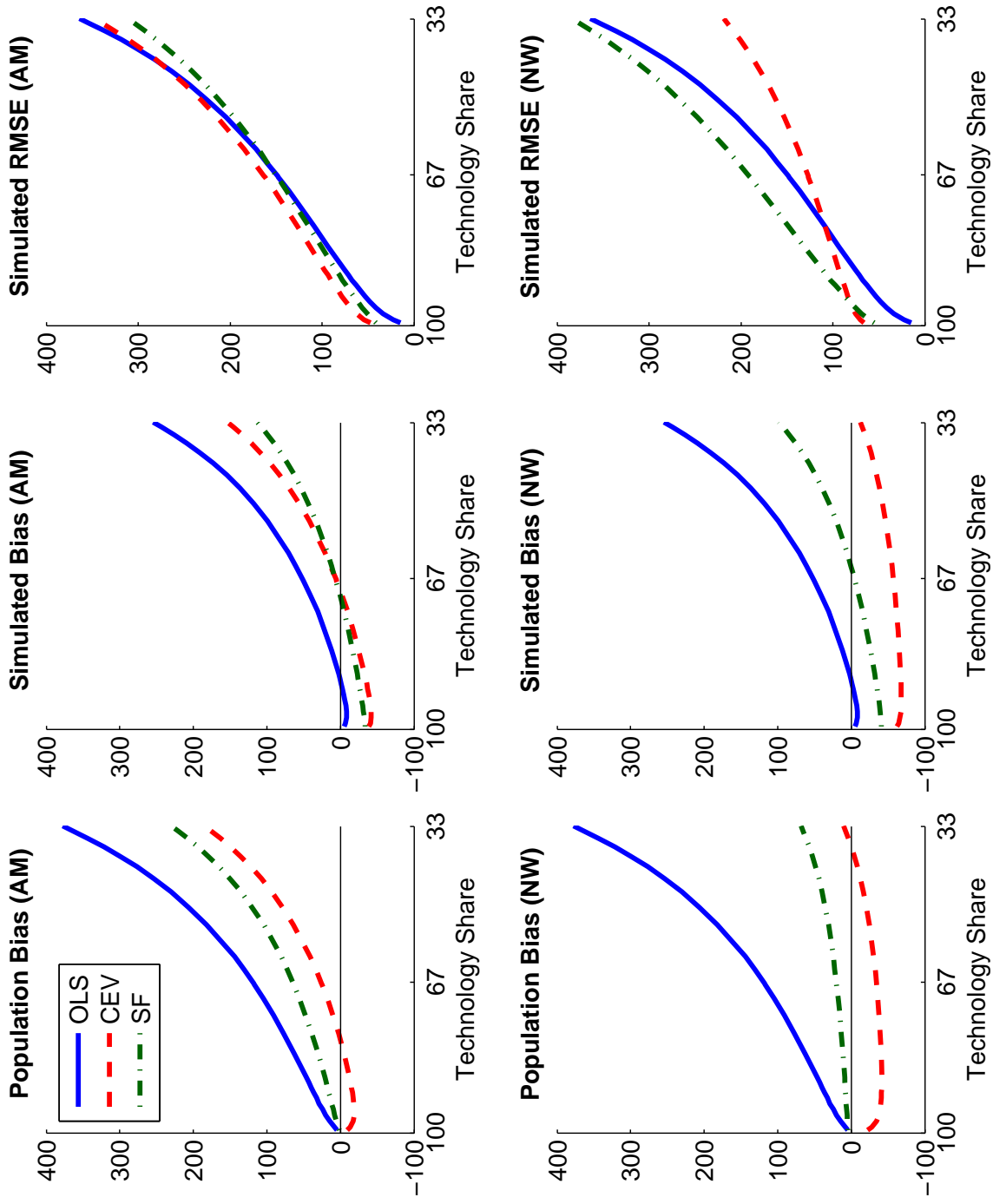
Note: Percentage errors of variance in output growth, relative to true moments (left panel, “Population”) or OLS sample moments (right panel, “Simulated”). Errors are computed for different calibrations of the importance of technology shocks in the data generating process. “Technology share” on the x-axis is the percentage of output variability due to technology shocks at business cycle frequencies (cycles with durations between two-and-a-half and eight years) in the data generating process. Variances denoted “CEV-NW” and “CEV-AM” are computed from the impulse responses of CEV’s SVAR procedure using the Newey-West or Andrews-Monahan estimators respectively. The left panel, labeled “Population”, uses variances computed from applying the SVAR procedure to the true population moments of the model. This panel documents mismatches arising from finite VAR order and spectral bandwidth. The right panel, labeled “Simulations”, reports small means computed from 1,000 simulations. (The variances are not bandpass-filtered.)

Figure 2: $D(1) \neq I$



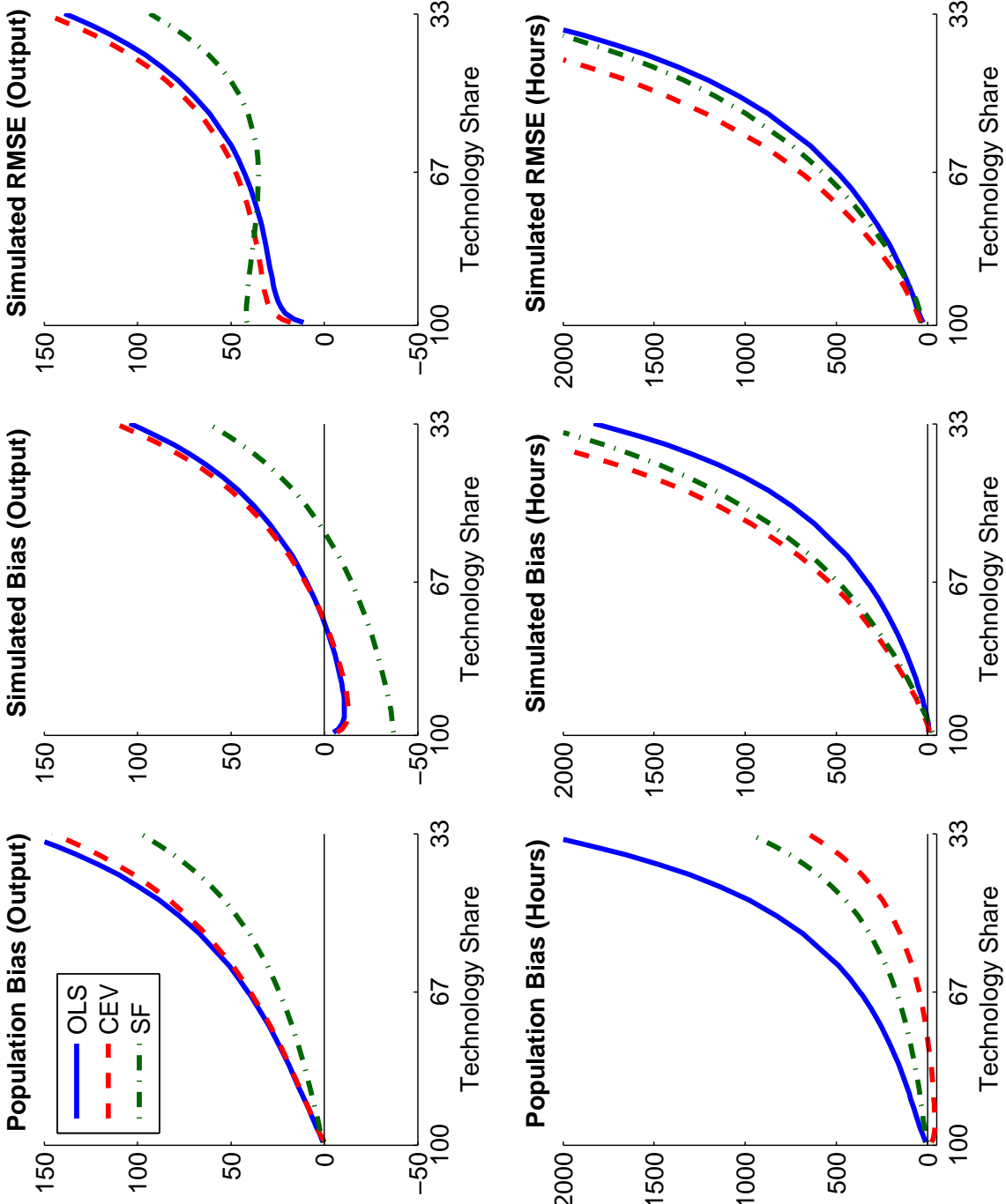
Note: Each panel shows an element of $\sum_{k=0}^K D_k$ for different lags K . $D(L)$ is computed from population moments of a VAR(1) applied to the model economy of Section 2 for three different percentages of the bandpass-filtered variability in output explained by technology shocks. A technology share of 67.5% (solid line) corresponds to the maximum likelihood estimates of CEV.

Figure 3: Impact Response of Hours to Technology Shock



Note: Percentage points relative to the model's true impact response of hours to a technology shock. Top row based on Andrews-Monahan estimator for CEV and SF, bottom row using Newey-West estimators; both with a fixed bandwidth of $b = 150$. "Technology share" on the x-axis is the percentage of output variability due to technology shocks at business cycle frequencies (cycles with durations between two-and-a-half and eight years) in the data generating process.

Figure 4: Technology Shares in Fluctuations of Output and Hours



Note: Percentage points relative the model's true technology share. Top row reports bias and RMSE for variance decomposition of output, bottom row for hours. Andrews-Monahan estimators of the spectral density used for CEV and SF with a fixed bandwidth of $b = 150$. (Similar results reported for Newey-West in Figure A.1). "Technology share" on the x-axis is the percentage of output variability due to technology shocks at business cycle frequencies (cycles with durations between two-and-a-half and eight years) in the data generating process.

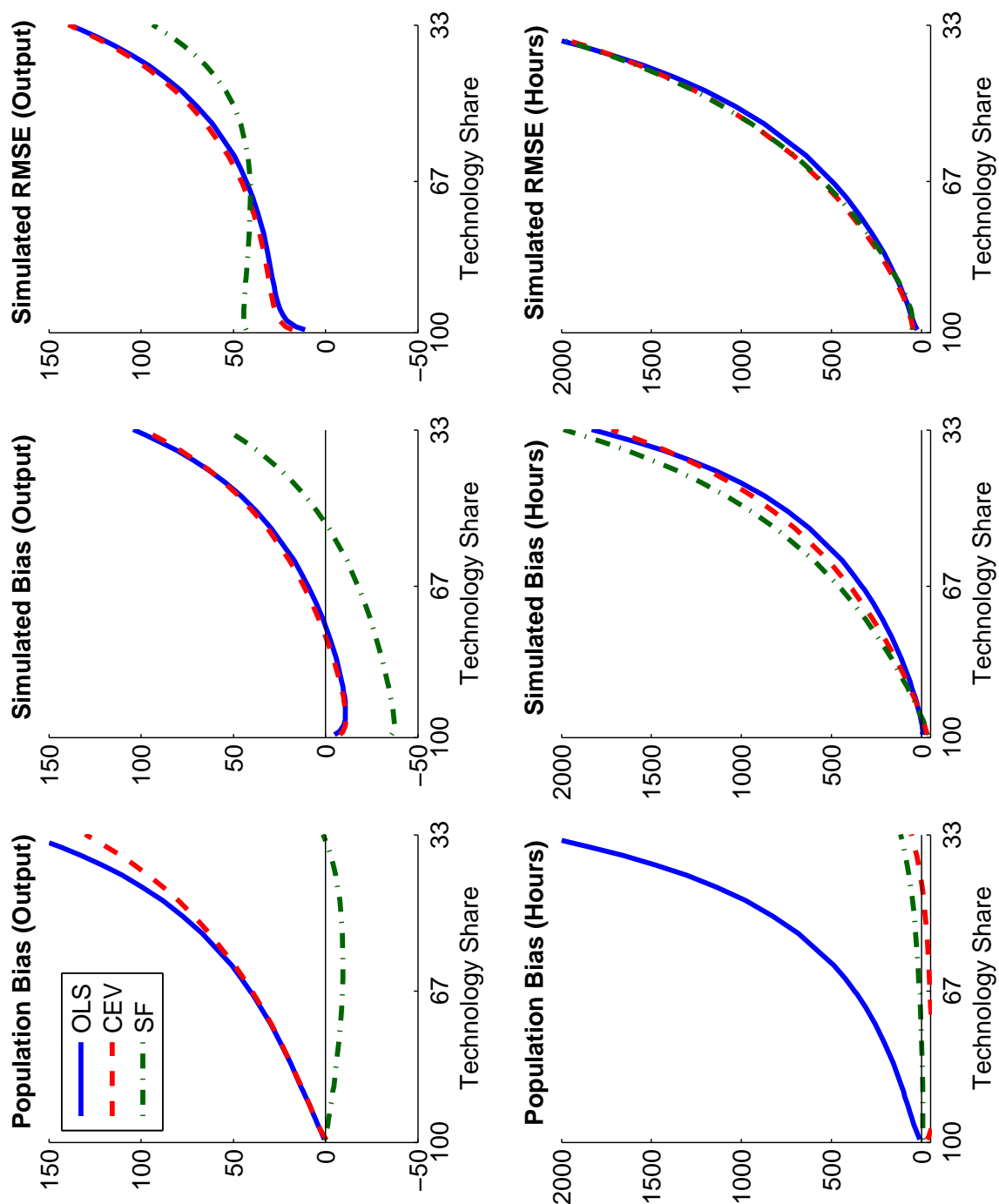
ADDITIONAL RESULTS

ARE SPECTRAL ESTIMATORS USEFUL FOR IMPLEMENTING LONG-RUN RESTRICTIONS IN SVARS?

List of Figures

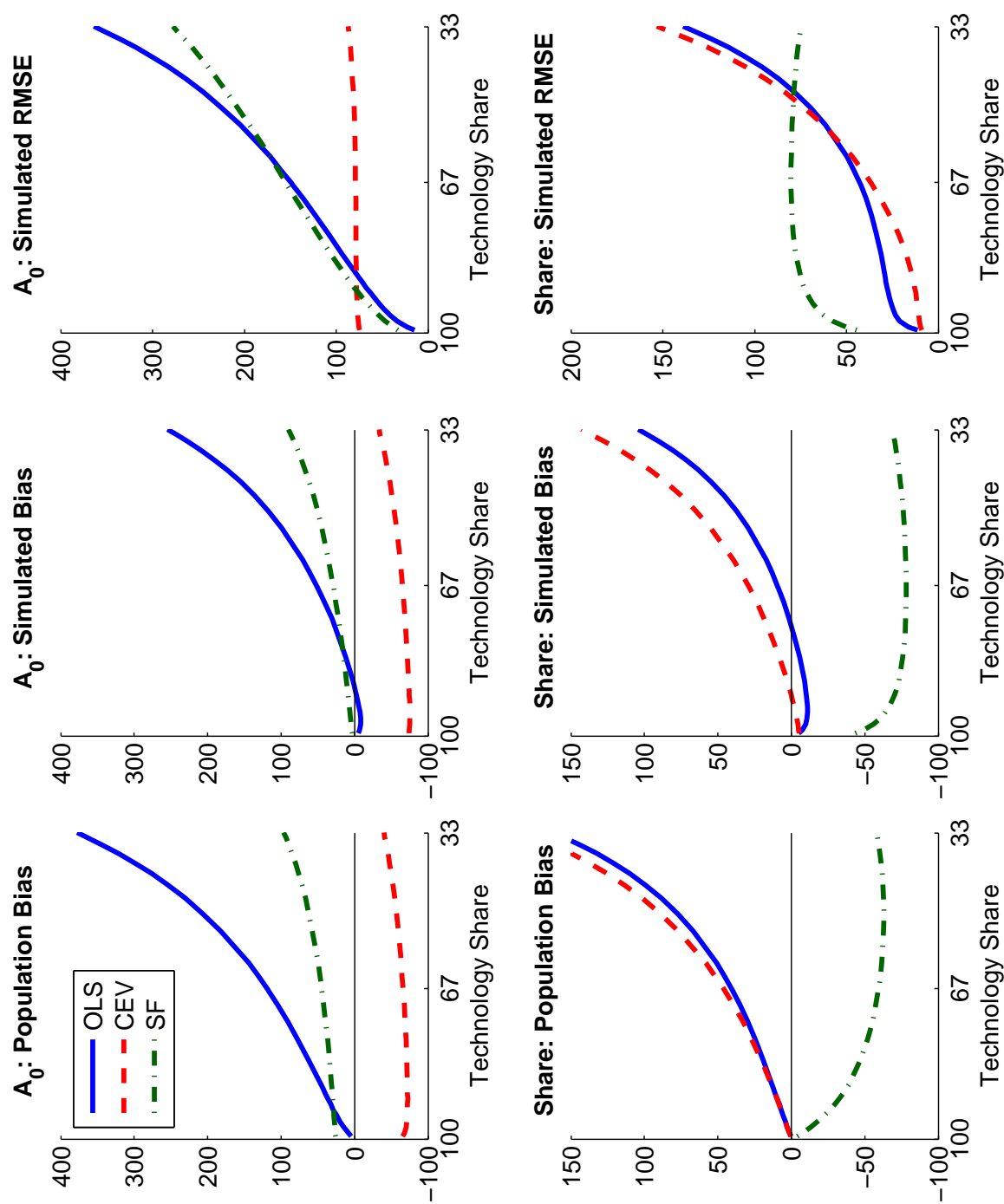
A.1	Technology Shares in Fluctuations of Output and Hours (Newey-West)	ii
A.2	Simulations with Automatic Bandwidth Selection (Newey-West)	iii
A.3	Simulations with Automatic Bandwidth Selection (Andrews-Monahan)	iv

Figure A.1: Technology Shares in Fluctuations of Output and Hours (Newey-West)



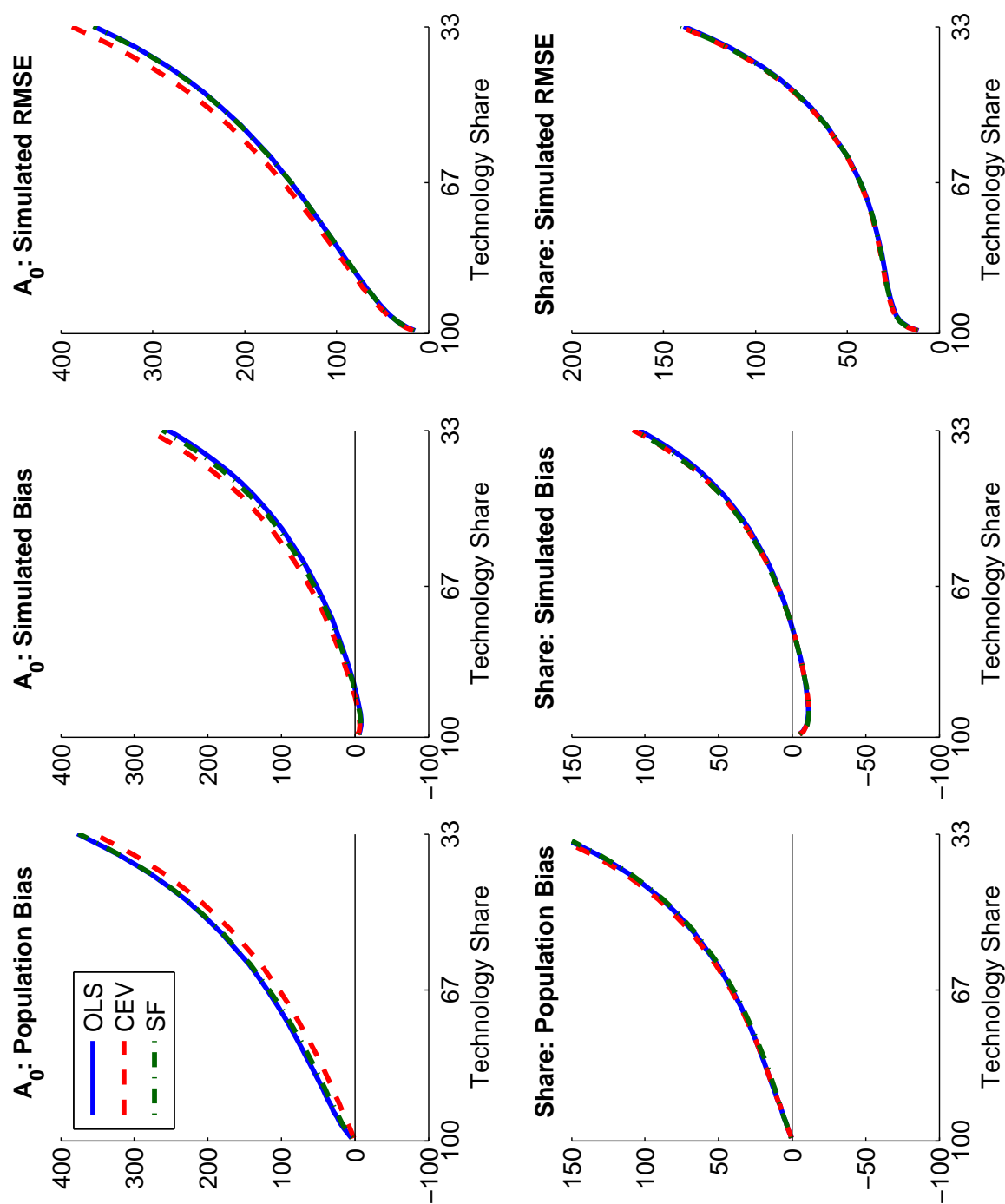
Note: Percentage points relative the model's true technology share. Top row reports bias and RMSE for variance decomposition of output, bottom row for hours. Newey-West estimators of the spectral density used for CEV and SF with a fixed bandwidth of $b = 150$. "Technology share" on the x-axis is the percentage of output variability due to technology shocks at business cycle frequencies (cycles with durations between two-and-a-half and eight years) in the data generating process.

Figure A.2: Simulations with Automatic Bandwidth Selection (Newey-West)



Note: Estimated impact responses of hours to technology (top row) and technology share (bottom row) when using the Newey-West estimator with automatic bandwidth selection (Newey and West, 1994). “Technology share” on the x-axis is the percentage of output variability due to technology shocks at business cycle frequencies (cycles with durations between two-and-a-half and eight years) in the data generating process.

Figure A.3: Simulations with Automatic Bandwidth Selection (Andrews-Monahan)



Note: Estimated impact responses of hours to technology (top row) and technology share (bottom row) when using the Andrews-Monahan estimator with automatic bandwidth selection (Newey and West, 1994). “Technology share” on the x-axis is the percentage of output variability due to technology shocks at business cycle frequencies (cycles with durations between two-and-a-half and eight years) in the data generating process.