

# Unit Roots, Cointegration and Pre-Testing in VAR Models\*

Nikolay Gospodinov<sup>†</sup>  
Concordia University,  
CIRANO and CIREQ

Ana María Herrera<sup>‡</sup>  
University of Kentucky

Elena Pesavento<sup>§</sup>  
Emory University

April 2013

## Abstract

This chapter investigates the robustness of impulse response estimators to near unit roots and near cointegration in VAR models. We compare estimators based on VAR specifications determined by pre-tests for unit roots and cointegration as well as unrestricted VAR specifications in levels. Our main finding is that the impulse response estimators obtained from the levels specification tend to be most robust when the magnitude of the roots is not known. The pre-test specification works well only when the restrictions imposed by the model are satisfied. Its performance deteriorates even for small deviations from the exact unit root for one or more model variables. We illustrate the practical relevance of our results through simulation examples and an empirical application.

---

\*We would like to thank the participants at the 12th Annual Advances in Econometrics Conference for useful feedback. We are particularly indebted to Lutz Kilian for numerous detailed comments and insightful suggestions that substantially improved the presentation of the chapter. Nikolay Gospodinov gratefully acknowledges financial support from FQRSC and SSHRC.

<sup>†</sup>Department of Economics, Concordia University, Montreal, QC H3G 1M8, Canada; e-mail: [nikolay.gospodinov@concordia.ca](mailto:nikolay.gospodinov@concordia.ca)

<sup>‡</sup>Department of Economics, University of Kentucky, Lexington, KY 40506-0034, USA; e-mail: [anherrera@uky.edu](mailto:anherrera@uky.edu)

<sup>§</sup>Department of Economics, Emory University, Atlanta, GA 30322-2240, USA; e-mail: [epesave@emory.edu](mailto:epesave@emory.edu)

# 1 Introduction

Most economic variables that are used for policy analysis and forecasting are characterized by high persistence and possibly nonstationary behavior. It is common practice in applied work to subject these series to pre-tests for unit roots and cointegration prior to the vector autoregressive (VAR) analysis to determine the appropriate transformations that render the data stationary. Incorporating information about the integration and cointegration properties of the data in VAR models reduces the estimation uncertainty and the degree of small-sample bias of impulse response estimates. Such pre-tests, however, suffer from low power in the neighborhood of the null hypothesis. In fact unit root pretests lack robustness to small deviations from unit roots and cointegration as they do not reject the null with probability approaching one for near integrated processes that are parameterized as local-to-unity. For example, Gospodinov, Maynard and Pesavento (2011) demonstrate that arbitrarily small deviations from an exact unit root may produce impulse response estimators that are highly distorted and misleading. The main goal of this article is to study and quantify the role of pre-tests for unit roots and cointegration for VAR specifications and identification schemes for impulse response analysis. We contrast the pre-test approach to the specification of VAR models with the alternative of expressing all VAR variables in levels. The latter practice, which is also common in applied work, is susceptible to problems of small-sample bias, especially when the data are highly persistent. Moreover, it requires appropriate choices of bootstrap or asymptotic methods for inference that are developed to accommodate near-integrated processes. However, the levels specification makes the estimates of structural impulse responses, identified by short-run restrictions, asymptotically robust to the unknown order of integration of the model variables. The relative merits of the pre-test VAR approach and the levels approach for VAR impulse response analysis are not well understood. The objective of this article is to examine this issue and to provide users of VAR models with advice on how to proceed in practice.

The limitations of the pre-test approach to VAR specification have been discussed previously in the literature. Elliott (1998) illustrated the possibly large size distortions of the cointegration methods that arise in systems with near unit roots. It is widely documented that similar distortions tend to characterize the properties of sequential modeling and specification procedures that are based on pre-tests for unit roots. These tests are known to have low power in rejecting the null of a unit root when the data are highly persistent and nearly integrated. Nevertheless, Elliott's

(1998) warning has been ignored by many applied researchers who continue to rely on pre-tests, perhaps because of the perception that the levels specification is not without its own potential drawbacks. This chapter examines how severe the problem documented by Elliott (1998) is for estimators of VAR impulse responses and which of the commonly used approaches to determining the VAR specification is most accurate in practice. We conduct a Monte Carlo study that assesses the mean squared errors (MSEs) of selected impulse response estimates and quantify how dangerous the pre-test strategy can be in practice. The pre-tests include individual unit root tests such as the augmented Dickey-Fuller test with GLS detrending (ADF-GLS) and tests for cointegrating rank, which determine whether the model should be estimated as a VAR in differences, a vector error-correction model (VECM) or a VAR in levels. Our simulation results provide support for the robustness of the level VAR specification to departures from exact unit roots and cointegration and to the possible presence of unmodelled low frequency co-movement among the variables of interest.

Our evidence in favor of the levels specification highlights the importance of valid inference on impulse response estimators when there is uncertainty about the integration and cointegration properties of the data. We provide a brief review of several asymptotic and bootstrap approaches and methods for inference in levels models including local-to-unity asymptotics (Gospodinov, 2004, 2010; Mikusheva, 2012; Pesavento and Rossi, 2006; Phillips, 1998), delta method asymptotics (Sims, Stock, and Watson, 1990), and bootstrap (Kilian, 1998; Inoue and Kilian, 2002). We show that the best approach may differ depending on the properties of the data, the dimensionality of the system, the forecast horizon, the identification scheme, etc.

The remainder of this chapter is structured as follows. Section 2 provides a short description of various representations of multivariate integrated processes. The following section introduces the common trend decomposition for nearly integrated process and discusses some theoretical results for impulse response analysis with nearly integrated (and possibly cointegrated) processes based on short-run and long-run identifying restrictions. Section 4 reviews the theoretical literature on inference for impulse response functions which is robust to the magnitude of the largest roots of the process. Section 5 presents the results of several simulation exercises. Section 6 concludes.

## 2 A Brief Review of Representations of Multivariate Integrated Processes

Let  $y_t$  be an  $I(1)$  multivariate ( $m \times 1$ ) process. Since  $\Delta y_t$  is  $I(0)$  by assumption,  $\Delta y_t$  has a moving average (MA) representation of the form

$$\Delta y_t = C(L)u_t, \quad (1)$$

where  $C(L) = C_0 + C_1L + C_2L^2 + \dots$  is a  $m \times m$  matrix lag polynomial and  $u_t \sim iid(0, \Omega)$ . Then,  $\Delta y_t$  has a long-run covariance matrix  $C(1)\Omega C(1)'$ .

For any  $\beta \neq 0_{m \times r}$  ( $r = 1, \dots, m$ ), the long-run variance of  $\beta' \Delta y_t$  is  $\beta' C(1)\Omega C(1)' \beta$ . If  $\beta$  are  $r$  cointegrating vectors of dimension  $m \times 1$ , then  $\beta' y_t$  is  $I(0)$  and  $\beta' \Delta y_t$  is overdifferenced with a zero long-run variance. Therefore, for  $\beta$  to be cointegrating vectors, we must have

$$\beta' C(1)\Omega C(1)' \beta = 0_{r \times r}.$$

This equality cannot hold if  $C(1)$  has full rank. Thus, cointegration requires  $C(1)$  to have rank less than  $m$  and  $\beta$  must lie in the null space of  $C(1)$ , i.e., so that  $\beta' C(1) = 0_{r \times m}$ .

Assuming that  $C(L)$  is invertible, model (1) has the VAR representation

$$A(L)y_t = u_t, \quad (2)$$

where  $A(L) = (1 - L)C(L)^{-1}$  is a  $p$ -th order lag polynomial. Evaluating  $(1 - L)I_m = A(L)C(L)$  at  $L = 1$  yields

$$A(1)C(1) = 0_{m \times m}.$$

Therefore,  $A(1)$  lies in the null space of  $C(1)$  and, from the previous results for the MA representation, must be a linear combination of the cointegrating vectors  $\beta$ . More precisely,  $A(1) = -\alpha\beta'$ , where  $\alpha$  is an  $m \times r$  matrix of rank  $r$ .

Finally, model (2) admits an VECM representation

$$\Delta y_t = \Pi y_{t-1} + \Gamma(L)\Delta y_t + u_t,$$

where  $\Gamma(L) = \Gamma_1L + \Gamma_2L^2 + \dots + \Gamma_{p-1}L^{p-1}$ ,  $\Gamma_i = -\sum_{j=i+1}^p A_j$  and  $\Pi = -A(1) = \alpha\beta'$ . In this representation,  $\text{rank}(\Pi) = r < m$ , i.e.,  $\Pi$  has a reduced rank if there are  $r$  cointegrating relations. If  $\text{rank}(\Pi) = m$ , all elements of  $y_t$  are  $I(0)$  and the appropriate model is an unrestricted VAR in

levels; whereas if  $\text{rank}(\Pi) = 0$ , all elements of  $y_t$  are  $I(1)$  and not cointegrated, and the appropriate model is VAR in first differences. Practitioners are commonly faced with the decision of which of these three models (VAR in levels, VAR in first differences, or VECM) to estimate, and this decision is often made on the basis of pre-tests for unit root and cointegration.

A second issue faced by practitioners interested in identifying the effect of a structural shock of interest (e.g., monetary policy shocks, oil price shocks, or technological innovations) is the choice of identifying restrictions. In the applied literature, the most commonly used identification strategies are short-run and long-run restrictions. Short-run identification schemes can be characterized as follows. Let  $B_0$  denote an  $m \times m$  invertible matrix with ones on the main diagonal. Pre-multiplying both sides of (2) by  $B_0$  yields the structural VAR (SVAR) model

$$B(L)y_t = \varepsilon_t,$$

where  $B(L) = B_0A(L)$  and  $\varepsilon_t = B_0u_t$  denotes an  $m \times 1$  vector of structural shocks which are assumed to be orthogonal with a diagonal covariance matrix  $\Sigma$ . Hence,

$$E(u_t u_t') \equiv \Omega = B_0^{-1} \Sigma B_0^{-1'}$$

Given that  $\Omega$  is symmetric with  $\frac{m(m+1)}{2}$  elements, in order to identify all of the free parameters of  $B_0$  we need to impose  $\frac{m(m-1)}{2}$  restrictions. One possibility is to restrict  $B_0$  to be lower triangular which is equivalent to a Choleski decomposition of  $\Omega$ . Note that the short-run identifying restrictions do not depend on the specification of the reduced-form VAR model (e.g., Lütkepohl and Reimers, 1992).

Alternatively, one can impose long-run identifying restrictions that render the moving average matrix  $A(1)^{-1}B_0^{-1}$  lower triangular. The use of long-run restrictions is less general in that it requires some model variables to be  $I(1)$  and others to be  $I(0)$ , making this approach particularly susceptible to any misspecification of the integration properties of the individual series. For example, Gospodinov, Maynard and Pesavento (2011) demonstrate that even arbitrary small deviations from the exact unit, when combined with long run identification restrictions, can produce impulse response estimates that are highly distorted.

In Section 4 we study how alternative choices of model specification and identification restrictions can affect the MSE of the structural impulse responses.

### 3 Some Theoretical Results for Near-Integrated Processes

In this chapter, our main interest lies in data generating processes (DGPs) in which the variables in the VAR are highly persistent with roots that are either close to one or equal to one.<sup>1</sup> This is the typical situation a practitioner estimating a VAR with macroeconomic variables would face. Theoretically, these variables are well approximated by near-integrated processes. Thus, to understand the effect of persistent variables on the estimation of the VARs, we start by reviewing some theoretical results for near-intergrated processes.

#### 3.1 Common trend decomposition of near-integrated processes

Let  $y_t$  denote a multivariate  $m \times 1$  process

$$(I_m - \Phi L)y_t = \Psi(L)u_t,$$

where  $\Phi = I_m + C/T$  and  $C$  is a matrix of fixed constants. It is often convenient to define  $C = \text{diag}(c_1, c_2, \dots, c_m)$  in order to rule out  $I(2)$  processes when the diagonal elements are non-zero. We discuss the effect of non-zero off-diagonal elements later. Assume that  $y_0 = 0$ ;  $\Psi(L) = \sum_{i=0}^{\infty} \Psi_i L^i$  with  $\Psi_0 = I_m$ ,  $\sum_{i=0}^{\infty} i|\Psi_i| < \infty$  and  $\sum_{i=0}^{\infty} \Psi_i \neq 0_{m \times m}$ ; and  $u_t$  is a homoskedastic martingale difference sequence with a covariance matrix  $\Omega$  and finite fourth moments. Deterministic terms are assumed away for notational convenience.

Let  $e_t = \Psi(L)u_t$ . Then,  $e_t = [\Psi(1) + (1-L)\Psi^*(L)]u_t$  by the Beveridge-Nelson (BN) decomposition, where  $\Psi(1) = \sum_{i=0}^{\infty} \Psi_i$ ,  $\Psi^*(L) = \sum_{i=1}^{\infty} \Psi_i^* L^{i-1}$  and  $\Psi_i^* = -\sum_{j=i}^{\infty} \Psi_j$ . Denoting  $S_t = \sum_{j=1}^t u_j$  and using recursive substitution and summation by parts,

$$y_t = \sum_{j=1}^t \Phi^{t-j} \Psi(L)u_j = \Psi(1)S_t + v_t + \frac{C}{T} \sum_{j=1}^{t-1} \Psi(1)\Phi^{t-j} S_{j-1} + \frac{C}{T} \sum_{j=1}^{t-1} \Phi^{t-j} v_{j-1}, \quad (3)$$

where  $v_t = \Psi^*(L)u_t$ .

Expression (3) is an algebraic decomposition (or factorization) of the process  $y_t$  that contains the standard BN decomposition for exact unit root processes as a special case ( $C = 0_{m \times m}$ ). While the standard BN decomposition is given by the permanent,  $\Psi(1)S_t$ , and transitory,  $v_t$ , components, the decomposition in (3) contains two additional terms. The fourth term  $\frac{C}{T} \sum_{j=1}^{t-1} \Phi^{t-j} v_{j-1}$  is

---

<sup>1</sup>Persistent variables can also be approximated by fractionally integrated processes. Allowing for this possibility is outside the scope of this chapter, although some of the results in our simulations may generalize to the case in which the DGP contains fractionally intergrated variables.

asymptotically negligible, whereas the term  $\frac{C}{T} \sum_{j=1}^{t-1} \Psi(1) \Phi^{t-j} S_{j-1}$  is not and is of the same order  $O_p(T^{1/2})$  as  $\Psi(1)S_t$ . Therefore, (3) contains two permanent component terms (i.e., two terms of order  $O_p(T^{1/2})$ ). Omitting one of the them (by wrongly assuming an exact unit root, for example) will give rise to a misspecification bias and undermine the size of hypothesis tests.

Consider, for example, the cointegration model (Elliott, 1998)

$$\begin{aligned} y_{1,t} &= \beta y_{2,t} + e_{1,t} \\ y_{2,t} &= (1 + c/T)y_{2,t-1} + e_{2,t}, \end{aligned}$$

which can be cast in a similar representation as the model above. In this case, the restriction  $\beta\Psi(1) = 0$  does not annihilate the permanent component. As a result, reduced rank regressions that impose only this restriction are not valid and would cause misspecification bias and test size distortions as shown by Elliott (1998). Similar misspecification biases are expected to play a role in impulse response analysis based on long-run identifying restrictions on  $\Psi(1)$ , i.e., restrictions imposed on the 'wrong' permanent component when  $C \neq 0_{m \times m}$ .

In brief, wrongly assuming that a multivariate process  $y_t$  has one or more unit roots, when the roots are close but not equal to unity, leads to distortions in reduced rank regressions, as well as in the BN decomposition and impulse response analysis.

### 3.2 Impulse response analysis in VAR models with near unit roots identified by short-run restrictions

Let the data generating process be given by

$$\tilde{A}(L)\tilde{y}_t = \tilde{u}_t \tag{4}$$

or

$$(I_m - \Lambda L) = \Psi(L)\tilde{u}_t,$$

where  $\tilde{A}(L) = I_m - \tilde{A}_1 L - \dots - \tilde{A}_p L^p$  has roots outside or on the unit circle, and  $\Lambda$  can be expressed in terms of its eigenvalues and eigenvectors as  $\Lambda = V^{-1}\Phi V$  with  $\Phi$  as defined above (Pesavento and Rossi, 2006). This model allows for unit roots and possible cointegration and several representations in other parts of the chapter can be regarded as specifications in the rotated variables  $y_t = V\tilde{y}_t$  and  $u_t = V\tilde{u}_t$ .

Despite the possible presence of unit roots and cointegration, the estimates of  $\tilde{A}_1, \dots, \tilde{A}_p$  from the levels VAR (4) are root- $T$  consistent and individually normally distributed (Sims, Stock and Watson, 1990). Using these estimates and a short-run (recursive) identification scheme, one can construct  $l$ -period ahead orthogonalized impulse responses  $\Theta_l$ . Provided that the response horizon is fixed with respect to the sample size, the impulse response estimators are, under weak regularity conditions, consistent and asymptotically normally distributed (see Lütkepohl and Reimers, 1992). This explains the robustness of impulse response estimators based on the level specification to possible deviations from exact unit roots and cointegration, as long as the structural parameters and shocks are identified using short-run restrictions. In contrast, inference becomes nonstandard if we rely on long-run restrictions for identification, as discussed below. A similar problem would arise, regardless of the identification, if we allow the response horizon to grow fast enough with respect to the sample size. The latter result casts doubt on impulse response estimates for long horizons when the data are persistent. Indeed, Kilian and Chang (2000) document that the reliability of impulse response estimates from standard levels specifications may deteriorate substantially at longer horizons.

### 3.3 Impulse response analysis in VAR models with near unit roots identified by long-run restrictions

This case is best illustrated in the context of the Blanchard and Quah (1989) model for which  $y_t = (\Delta Y_t, U_t)'$ , where  $Y_t$  is the output (log real GNP) and  $U_t$  is the unemployment rate,  $A(L) = \Psi(L)^{-1}(I_2 - \Phi L)$  and  $B_0 = \begin{bmatrix} 1 & -b_{12}^{(0)} \\ -b_{21}^{(0)} & 1 \end{bmatrix}$ . The structural shocks  $\varepsilon_t = B_0 u_t$  can be interpreted as aggregate supply and demand shocks  $(\varepsilon_t^s, \varepsilon_t^d)'$ . The long-run identifying restriction that the demand shocks  $\varepsilon_t^d$  have no long-run effect on output imposes a lower triangular structure on the moving average matrix  $A(1)^{-1}B_0^{-1}$ . Hence, under this identifying restriction, the matrix of long-run multipliers in the structural model,  $B(1) = B_0 A(1)$ , is also lower triangular.

The robustness of the impulse response analysis depends crucially on the persistence of  $U_t$  and, hence, on the parameterization of the matrix  $\Phi$ . We follow Gospodinov, Maynard and Pesavento (2011) and express  $\Phi$  as  $\Phi = \begin{bmatrix} 1 & \delta \\ 0 & \rho \end{bmatrix}$ , where  $\delta = -\gamma(1 - \rho)$  is the parameter that determines the low frequency co-movement between the variables, and  $\rho = 1 + c/T$  (for  $c < 0$ ) denotes the largest root of the unemployment rate. The non-zero off-diagonal element  $\gamma(1 - \rho)$  allows for the



possibility that a small low frequency component of unemployment affects output growth.

To gain some further insight into this parameterization, let us assume for simplicity that  $\Psi(L) = I_2$  and note that the vector error-correction form of the model is given by

$$\Delta y_t = (\rho - 1) \begin{bmatrix} 0 & -\gamma \\ 0 & 1 \end{bmatrix} y_{t-1} + u_t,$$

which can be written as

$$\Delta y_t = \alpha \beta' y_{t-1} + u_t,$$

where  $\alpha = (\rho - 1) \begin{pmatrix} \gamma \\ 1 \end{pmatrix}$  and  $\beta = (0, 1)'$ . Therefore, the model has a trivial cointegration vector  $(0, 1)'$  and the speed of adjustment  $\alpha$  is a function of  $\gamma$  which carries important information from the levels of the series, provided that  $\rho < 1$ .

With this parameterization of  $\Phi$ , the long-run multiplier matrix is given by

$$B(1) = \begin{bmatrix} \psi_{11}(1) - b_{12}^{(0)} \psi_{21}(1) & (1 - \rho) \left( [\gamma \psi_{11}(1) + \psi_{12}(1)] - b_{12}^{(0)} [\gamma \psi_{21}(1) + \psi_{22}(1)] \right) \\ \psi_{21}(1) - b_{21}^{(0)} \psi_{11}(1) & (1 - \rho) \left( [\gamma \psi_{21}(1) + \psi_{22}(1)] - b_{21}^{(0)} [\gamma \psi_{11}(1) + \psi_{12}(1)] \right) \end{bmatrix},$$

where  $\psi_{ij}(1)$  denotes the corresponding elements of the matrix  $\Psi(1)^{-1}$ . Imposing the long-run restriction that aggregate demand shocks have no permanent effect on output implies that  $b_{12}^{(0)} = [\gamma \psi_{11}(1) + \psi_{12}(1)] / [\gamma \psi_{21}(1) + \psi_{22}(1)]$ . In contrast, imposing a unit root in unemployment ( $\rho = 1$ ) and differencing both output and unemployment renders the restriction  $b_{12}^{(0)} = \psi_{12}(1) / \psi_{22}(1)$ . Thus, the differenced VAR ignores any information contained in the levels of output and unemployment.

The main message to the practitioner is that the differenced VAR specification is not robust to small low frequency co-movements. While the levels specification (in which some of the variables are in first differences and some are in levels) preserves the information on the low-frequency co-movement, the estimates of the impulse response functions are inconsistent and their distribution is fat-tailed in the presence of local-to-unity processes (Gospodinov, 2010). In summary, the applied researchers should exert extreme caution when working with near-integrated variables and the SVAR is identified by long-run restrictions.

### 3.3.1 Unit root pre-test VAR specification

Similarly to the case of the differenced VAR, lack of robustness is expected to characterize the behavior of specifications based on pre-test for a unit root given that this pre-test will select the

differenced specification with probability approaching one when the process is near-integrated. Hence, the unit root pre-test VAR specification will inherit the non-robustness properties of the differenced VAR specification for slowly co-moving, near-integrated variables.

## 4 Robust inference for impulse response functions

Given the low power of most pre-tests for unit root and cointegration, and costs from estimating a VAR that erroneously imposes unit roots, the current literature has moved in the direction of methods for inference that are robust to the possible presence of unit roots. These robust methods are designed for VAR models based on short-run identifying restrictions only, of course, as departures from exact unit roots immediately invalidate the use of long-run identifying restrictions. Consider the model

$$(I_m - \Phi L)y_t = \Psi(L)u_t,$$

where  $\Phi = I_m + C/T$  and  $C = \text{diag}(c_1, c_2, \dots, c_m)$ . Note that  $y_t$  could denote rotated variables to account for the possibility of cointegration with a known cointegration vector (for more details, see Pesavento and Rossi, 2006). Suppose that the interest lies in inference on the impulse responses of  $y_t$  at long response horizons  $l$ ,  $y_{t+l}$ , to the structural shocks  $\varepsilon_t = B_0 u_t$ .

To better capture the uncertainty in estimating the impulse responses at long horizons, it is convenient to parameterize  $l$  as a function of the sample size. More specifically, let  $l = [\delta T]$  for some fixed  $\delta > 0$ . Under this parameterization, we have that  $C^l \rightarrow e^{C\delta}$  as  $T \rightarrow \infty$ , where  $e^{\delta C}$  is a diagonal matrix with  $(e^{\delta c_1}, e^{\delta c_2}, \dots, e^{\delta c_m})$  on the main diagonal (Pesavento and Rossi, 2006; Gospodinov, 2004). Furthermore,

$$\Theta_l \equiv \frac{\partial y_{t+l}}{\partial u_t} = \Phi^l (I_m + \Phi^{-1}\Psi_1 + \Phi^{-1}\Psi_2 + \dots) + o(1) \simeq \Phi^l \Psi(1) \quad (5)$$

and the  $l$ -period impulse response function of the  $j$ -th variable in  $y_t$  to the  $k$ -th structural shock in  $\varepsilon_t$  is given by (Pesavento and Rossi, 2006)

$$\frac{\partial y_{j,t+l}}{\partial \varepsilon_{k,t}} \simeq \boldsymbol{\nu}'_j \Phi^l \Psi(1) B_0^{-1} \boldsymbol{\nu}_k \rightarrow \boldsymbol{\nu}'_j e^{C\delta} \Psi(1) B_0^{-1} \boldsymbol{\nu}_k \quad (6)$$

as  $T \rightarrow \infty$ , where  $\boldsymbol{\nu}_j$  and  $\boldsymbol{\nu}_k$  are the corresponding columns of the identity matrix  $I_m$ . Inference on the impulse response function can then proceed by plugging a consistent estimate of  $\Psi(1)$  and obtaining the confidence interval limits for the localizing constant  $c_j$  of the  $j$ -th variable  $y_{j,t}$ , which

is typically done by inverting a unit root test (see, for example, Stock, 1991). In this sense, this method is a two-step procedure, which involves the construction of a confidence interval for  $c_j$  in the first step and using this confidence interval and the expression (6) to construct a confidence interval for the impulse response function  $\frac{\partial y_{j,t+l}}{\partial \varepsilon_{k,t}}$ .

Gospodinov (2004) proposes a one-step test-inversion method for directly constructing confidence bands for the impulse response function. Let  $\theta = (\text{vec}(\Phi)', \text{vec}(\Psi_1)', \text{vec}(\Psi_2)', \dots)'$ , where in practice  $\Psi_1, \Psi_2, \dots$  are approximated by a finite-order VAR, and  $h_l(\theta) = \Theta_l(\theta) - \Theta_{0,l}$  denotes the constraint that the expression in (5) is equal to a particular value  $\Theta_{0,l}$  under the null hypothesis  $H_0 : \Theta_l(\theta) = \Theta_{0,l}$ . Finally, let  $LR_T = T \text{tr} \left[ \widehat{\Omega}^{-1} \left( \widetilde{\Omega} - \widehat{\Omega} \right) \right]$  denote the likelihood ratio statistic of the hypothesis  $H_0 : h_l(\theta) = 0$ , where  $\widetilde{\Omega}$  and  $\widehat{\Omega}$  denote the variance-covariance matrices of the restricted and the estimated residuals, respectively, from the approximating finite-order VAR model. Then, under the null  $H_0 : h_l(\theta) = 0$ ,  $\Phi = I_m + C/T$  and  $l = [\delta T]$ ,

$$LR_T \xrightarrow{d} \text{tr} \left[ \left( \int_0^1 \mathbf{J}_c(s) d\mathbf{J}_c(s)' \right)' \left( \int_0^1 \mathbf{J}_c(s) \mathbf{J}_c(s)' ds \right)^{-1} \left( \int_0^1 \mathbf{J}_c(s) d\mathbf{J}_c(s)' \right) \right],$$

where  $\mathbf{J}_c(s)$  is an Ornstein-Uhlenbeck vector process. Confidence sets can be obtained by inverting the  $LR_T$  test on a grid of possible values of  $\Theta_{0,l}$  and using projection methods for constructing confidence bands for individual impulse responses.

One drawback of the methods proposed by Gospodinov (2004) and Pesavento and Rossi (2006) is that they provide accurate approximations only for near-integrated processes ( $\Phi = I_m + C/T$ ) and long response horizons ( $l = [\delta T]$ ). While these methods continue to remain asymptotically valid over the other parts of the parameter space and horizon space, they tend to be conservative. One solution to this problem is to adopt an approximation procedure which is uniform over the whole parameter space. Mikusheva (2012) shows that the grid bootstrap of Hansen (1999) adapted to the  $LR$  test inversion method of Gospodinov (2004) delivers the desirable uniform approximation regardless of the persistence of the underlying process. Unfortunately, this approximation method suffers from the curse of dimensionality and can be computationally demanding (or even infeasible) as the dimensionality of the VAR model increases, due to the multi-dimensional grid required to perform the test inversion.

Other practically appealing bootstrap-based methods for inference on impulse response functions are also available. For example, Inoue and Kilian (2002), building on the work of Sims, Stock and Watson (1990), show that the conventional bootstrap (Runkle, 1987) is asymptotically valid for

inference on typical nonlinear combinations of VAR parameters even in the presence of unit roots or near unit roots, provided the VAR model includes more than one lag. However, despite their asymptotic validity in the presence of highly persistent variables, standard bootstrap methods for higher-order VAR models may have poor coverage accuracy for impulse responses in small samples. One source of the unsatisfactory behavior of the conventional bootstrap is that it further exacerbates the bias that characterizes the least-squares estimator of VAR models with highly persistent variables. For this reason, Kilian (1998) develops a two-stage bootstrap method that explicitly estimates and removes the bias in the VAR parameters before approximating the impulse response distribution. However, this bias-correction method is not designed for unit root processes, unless the unit root is imposed in the estimation, and its theoretical validity to date has been established only for stationary VAR models. We consider these bootstrap methods in our simulation section.

## 5 Monte Carlo Simulations

### 5.1 Bivariate DGP with near unit roots, but no near cointegration

Our first simulation experiment follows closely the structure of the models used for quantifying the contribution of aggregate demand (non-technology) and supply (technology) shocks to business cycle fluctuations (Blanchard and Quah, 1989; Galí, 1999, Christiano, Eichenbaum and Vigfusson, 2006; among others). Blanchard and Quah (1989) use output and unemployment to identify the demand and supply shocks while Galí (1999) and Christiano, Eichenbaum and Vigfusson (2006) model the dynamic behavior of labor productivity and hours worked to study the prediction of the real business cycle theory that technology shocks have a positive effect on hours worked. These papers impose a unit root on output or technology, i.e., the first variable (output or labor productivity) in the SVAR is expressed in first differences. The second variable (unemployment or hours worked) is highly persistent which has generated a debate about the appropriate empirical specification of this variable; the question is whether this second variable should be included in levels or in first differences (see Gospodinov, Maynard and Pesavento, 2011). While the two variables are not cointegrated, they appear to be driven by some low-frequency co-movement which is preserved in the levels VAR specification but is eliminated after differencing.

In this simulation exercise, we investigate the robustness of the levels VAR specification (first variable in differenced form and second variable in levels) and a specification based on a pre-test of a

unit root for the second variable to various degrees of persistence and low-frequency co-movement. We employ the ADF-GLS test at 5% significance level to pre-test for a unit root in the second variable. If the test rejects the null, we model the second variable in levels; instead, if the null is not rejected, we model it in first differences. We also follow Gospodinov, Maynard and Pesavento (2011) in allowing for a small low-frequency co-movement when the root of the second variable is strictly less than unity. Finally, we consider both long-run identification (based on the assumption that the shocks to the second variable have no permanent effect on the level of the first variable) and short-run (recursive) identification (based on the assumption that the second variable has no contemporaneous effect on the first variable). More specifically, the data are generated from

$$\begin{bmatrix} \Delta y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} 0 & (1 - \rho) \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \Delta y_{1,t} \\ y_{2,t} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}, \quad (7)$$

where  $\begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \sim iid N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix} \right)$ . The sample size is  $T = 200$  (fifty years of quarterly data that matches our empirical application) and  $\rho \in \{0.92, 0.95, 0.98, 1\}$ . The two variables are demeaned prior to estimation. The number of Monte Carlo replications is 10,000. Note that in what follows, the first variable is always modeled in first differences ( $\Delta y_{1,t}$ ) and the second variable is either in levels  $y_{2,t}$  (we refer to this as levels VAR specification) or in first-differences  $\Delta y_{2,t}$ .

The simulation results are summarized by plotting the coverage rates of 95% confidence intervals of the estimated impulse responses and the MSEs of the impulse response estimator for horizons 1, 2, ..., 24. The confidence bands for the impulse responses are constructed using the bias-corrected bootstrap method proposed by Kilian (1998a,b). To illustrate the advantages of Kilian's (1998a,b) method over the conventional bootstrap (Runkle, 1987), Figures 1 and 2 report the coverage rates of the 95% confidence intervals of the  $(i, j)$  impulse response ( $i = 1, 2; j = 1, 2$ ), constructed using Kilian's (1998a,b) and Runkle's (1987) bootstrap methods, for the levels VAR specification using long-run identification ( $\rho = 0.95$  and  $0.98$ ). Overall, the coverage rates of Kilian's (1998a,b) bootstrap method are very close to the nominal level while Runkle's (1987) conventional bootstrap method tends to undercover. In what follows, we use Kilian's (1998a,b) bias-corrected bootstrap method to construct confidence bands for the impulse response of interest.

We start with the case of long-run identification and consider both coverage rates and mean squared errors of the (2,1) and (2,2) impulse response functions. These impulse responses trace the

responses of the second variable, for which there is uncertainty about the value of its largest AR root, to both shocks. Recall that the levels VAR specification always models the first variable in first differences and the second variable in levels. Instead, the pre-test VAR specification is based on a sequential procedure. In the first step, the ADF-GLS test is used to determine if the second variable is integrated of order one or zero. If the null hypothesis is not rejected, the second variable is modeled in first differences; otherwise, the second variable enters the model in levels.

Figures 3, 4 and 5 plot the coverage rates and mean squared errors for  $\rho = 0.92, 0.95$  and  $0.98$ . The coverage rates and the mean squared errors of the (2,1) and (2,2) impulse responses from the pre-test model deteriorate when the root  $\rho$  approaches one. This is due to the presence of a low frequency component which is differenced away by the pre-test specification when the null of a unit root is not rejected. In the absence of this low-frequency co-movement, the coverage rates and MSEs of the pre-test specification should improve. Gospodinov, Maynard and Pesavento (2011) illustrate how imposing a unit root in the presence of an arbitrarily small low-frequency co-movement leads to estimated impulse responses that are vastly different from the true ones. In the simulation setup (7), the VAR model in first differences (except for  $\rho = 1$ ) is misspecified and the estimated impulse responses from this model perform worse than the ones from the pre-test specification. In contrast, the levels specification performs well and appears to be robust to the different degrees of persistence and the presence of a low-frequency co-movement. However, as the root approaches unity, the structural parameters are only weakly identified and the estimation uncertainty of the impulse responses is large (see Gospodinov, 2010).

The corresponding short-run identification scheme seems to be immune to some of the problems that arise under the long-run identifying restrictions (Christiano, Eichenbaum and Vigfusson, 2006; Gospodinov, Maynard and Pesavento, 2011). Figures 6, 7, 8 and 9 plot the coverage rates and mean squared errors for  $\rho = 0.92, 0.95, 0.98$  and  $1$  under the short-run identification scheme. Again, the pre-test specification is not robust to deviations from the unit root and exhibits distorted coverage rates and increased MSEs. When  $\rho = 1$ , the differenced model is correct and the pre-test VAR specification dominates the levels VAR. But even in this case, the costs from estimating the levels VAR are not large especially if one employs inference procedures that are asymptotically valid in the case when the root lies on or near the unit circle.

To summarize, both for the long-run and the short-run identification schemes, the levels VAR specification fares well in terms of coverage rates and MSEs. The sampling uncertainty associated

with the estimated impulse response functions tends to be well approximated by the bias-corrected bootstrap procedure of Kilian (1998a,b). The coverage rates and the MSEs of the levels specification continue to be satisfactory even when the largest root is very close to unity ( $\rho = 0.95$ ). The specification based on the pre-test for a unit root exhibits finite-sample distortions and inflated MSEs. Overall, the levels specification emerges as the preferred specification for this simulation experiment.

## 5.2 Bivariate DGP with near unit roots and near cointegration

The DGPs so far did not allow for cointegration among the model variables. In the next simulation experiment, we consider a bivariate system with possible cointegration. More specifically, we simulate data from the following model:

$$\begin{aligned} y_{1,t} &= \beta y_{2,t} + u_{1,t} \\ y_{2,t} &= \rho y_{2,t-1} + u_{2,t} \\ u_{1,t} &= \rho_u u_{1,t-1} + \xi_t, \end{aligned}$$

where  $\beta = 1 - \rho_u$ , and  $\xi_t$  and  $u_{2,t}$  are mutually uncorrelated *iid*  $N(0, 1)$  random variables. The variable  $y_{2,t}$  is generated as a (near-) integrated process with  $\rho \in \{1, 0.98, 0.95, 0.90, 0.85\}$ . The autoregressive parameter  $\rho_u$  determines the persistence of  $y_{1,t}$  and the cointegration properties between  $y_{1,t}$  and  $y_{2,t}$ . In our numerical experiment, we set  $\rho_u$  to 0.4. Note that when  $\rho_u = 1$ ,  $y_{1,t}$  is an exact unit root process and there is no cointegration between the two variables, whereas for  $\rho_u = 0.4$ ,  $u_{1,t}$  is stationary and there is cointegration between the two variables.

Consider first the sequential pre-test procedure. The two series are first subjected to a unit root test. If the null of a unit root is not rejected for both series, we test for cointegration between the two series. Depending on the outcome of the test, the model is estimated as a differenced VAR (if no cointegration is detected) or as a VECM using the estimated cointegrated vector (if cointegration is detected). If the null of a unit root is rejected, the levels VAR is estimated. We contrast this sequential modeling strategy to an unrestricted VAR in levels or first differences regardless of the pre-test results. We also consider the case in which the practitioner only tests for unit roots but no cointegration, and then uses the results of these pre-tests to decide whether to include each series in levels or in first differences. These models are used to estimate impulse responses, identified by short-run (recursive) restrictions, which are compared to the true impulse responses from the

model. Unit root pre-tests are performed using the ADF-GLS test at 5% significance level, with the number of lags chosen by the modified information criterion of Ng and Perron (2001). For the VAR and VECM estimation, we select the number of lags using the BIC (the maximum lag order is set to 4). We run 10,000 replications for simulated samples of 200 observations. The results are presented in Figures 10 to 14.

Figures 10 to 14 show the MSE of the  $(i, j)$  impulse response ( $i = 1, 2; j = 1, 2$ ) for the case in which there is cointegration in the DGP ( $\rho_u = 0.4$ ). Similar to the results in the previous section, the levels VAR performs better in terms of MSE than a VAR in first differences except in the case in which we know that the root is equal to one. Pretesting for a unit root has a similar effect as in the no-cointegration case: pre-testing provides very little gain when the roots are large and it is not better than estimating the impulse response from the levels VAR in most cases. For the (2,1) response, when the root is large, pretesting is actually worse in term of MSE than simply running the VAR in first differences. As expected, when  $\rho$  gets smaller, pre-tests have good power and they are able to correctly suggest to estimate the VAR in levels.

Comparing the results from estimating a VAR in levels ignoring and not ignoring cointegration reveals some interesting results: the levels VAR provides a smaller MSE for the (1,1) and (2,1) impulse responses compared to the specification based on a pre-test for cointegration even if cointegration is actually present ( $\rho = 1$ ). For the (1,2) and (2,2) impulse responses there is a gain from estimating a VECM when indeed there is cointegration ( $\rho = 1$ ). When  $\rho$  is less than one, technically there is no cointegration but we would still expect the variables to behave similarly to the unit root case. For roots equal to 0.98 and smaller, the levels VARs performs better for all four impulse responses. Overall, the results seem to suggest that, in most cases, the levels VAR dominates the other specifications in terms of MSE although this depends on the impulse responses of interest and the magnitude of the largest roots in the model.

### 5.3 A multivariate model used in monetary policy analysis

In the previous section, we considered a bivariate DGP and evaluated the options faced by the practitioner regarding model specification. When only one unit root is suspected, these options include: (a) running a VAR with  $y_{2,t}$  in levels ignoring a possible unit root; (b) imposing a unit root and running a VAR in first differences; or (c) carrying out a pre-test for a unit root in  $y_{2,t}$  and specifying the VAR based on the outcome for the pre-test. Instead, when two unit roots are



suspected, the practitioner would proceed with a sequential procedure that could lead to a VAR model in first-differences, a VECM, a VAR model in levels, or a mixed VAR model with some variables in levels and some in first-differences.

A similar problem is faced by the practitioner when estimating the multivariate VARs that are commonly employed in analyzing the effect of monetary policy shocks. Below, we use the benchmark model of Christiano, Eichenbaum and Evans (1998) – hereafter CEE – to study the effect of exogenous monetary policy shocks on economic activity. In this model, the economy is described by a 7-dimensional VAR(4) given by

$$A(L)y_t = c + u_t, \tag{8}$$

where  $y_t = [Y_t, P_t, PCOM_t, FF_t, TR_t, NBR_t, M_t]'$ ,  $Y_t$  is the log of real GDP,  $P_t$  is the log of the implicit GDP deflator,  $PCOM_t$  is the log of the Conference Board commodity price index,  $FF_t$  is the federal funds rate,  $TR_t$  is the log of total reserves,  $NBR_t$  is the log of nonborrowed reserves,  $M_t$  is the log of M1, and  $u_t \sim iid(0, \Omega)$ .

Motivated by the literature on the transmission of monetary policy shocks, we consider a short-run (recursive) identification scheme based on the lower triangular decomposition of the variance-covariance matrix. As in Bernanke and Blinder (1992) and Sims (1986, 1992), among others, we measure the monetary policy instrument by the federal funds rate.<sup>2</sup> The variables in  $y_t$  are ordered as described above such that  $Y_t$ ,  $P_t$ , and  $PCOM_t$  do not react on impact to monetary policy shocks.

In the subsequent analysis, we treat the 7-dimensional VAR(4) in (8) as the data generating process. We follow the standard practice to estimate the model in levels and with an intercept included. The estimation method is ordinary least squares.

The data are quarterly and, with the exception of the commodity price index, are obtained from the Federal Reserve Bank of St. Louis (FRED) database. The commodity price index corresponds to the close of the Reuters/Jefferies-CRB Index at the end of the quarter.  $FF_t$  equals the average of the monthly federal funds rate over the quarter. The monetary aggregates are measured in millions of dollars at the end of the quarter. The data span the period between 1959:1 and 2006:4. We exclude the data prior to the onset of the Great Recession as well as subsequent data, given that the conduct of monetary policy during this period has been markedly different from earlier years. In

---

<sup>2</sup>The results obtained when shocks to monetary policy are identified with shocks to nonborrowed reserves (e.g., Eichenbaum 1992; Christiano and Eichenbaum 1992) are virtually the same.

particular, nonborrowed reserves took on negative numbers during almost all of 2008 – something that has never been observed in the documented history of this series – and then increased at a very fast pace in the following year. Total reserves, on the other hand, exploded during 2008 and have continued to increase at a very rapid pace. As can be seen in Figure 15, all of the variables appear to exhibit a high degree of persistence.

### 5.3.1 Multivariate model without considering cointegration

In this section, we consider the approach followed by a practitioner who would first pre-test for unit roots in the individual series before estimating the VAR. Then, based on the test results, the practitioner would include the variables either in levels or first-differences. As we mentioned before, we treat the estimated 7-variable VAR(4) model in levels as the data generating process.

To simulate the data, we draw randomly from the vector  $y_t$  to obtain the starting values  $y_1^{(i)}, \dots, y_4^{(i)}$ , and draw with replacement from the estimated vector of residuals  $\hat{u}_t$  to obtain  $u_t^{(i)}$ . We generate 10,000 samples,  $y_t^{(i)}$  for  $i = 1, 2, \dots, 10,000$ , each consisting of a sequence of data of the same length as that used to estimate the 7-variable VAR(4).

Then, for each of the 10,000 trials we pre-test the individual series in the vector  $y_t^{(i)}$  ( $Y_t^{(i)}, P_t^{(i)}, PCOM_t^{(i)}, FF_t^{(i)}, TR_t^{(i)}, NBR_t^{(i)}, M1_t^{(i)}$ ) using an ADF-GLS test at a 5% significance level. A time trend is included in the GLS detrending step for all variables except for  $FF_t^{(i)}$ . As in the previous section, the lag order for the unit root test is chosen by the modified information criterion of Ng and Perron (2001). If the pre-test rejects the null, we keep the simulated variable in levels. Instead, if we fail to reject, we first difference the simulated variable. In all of the 10,000 replications, we estimate a VAR(4) model with each series either in levels or first differences depending on the outcome of the individual ADF-GLS test. Impulse response functions are then computed using the recursive identification scheme (Choleski decomposition) described in the previous section. Finally, we compute the responses to a federal funds shock and report the mean squared errors with respect to the impulse responses from the data generating process.

The solid and dashed lines in Figure 16 represent the mean squared errors of the impulse responses for the levels specification and the ADF-GLS pre-testing strategy, respectively. The numbers on the horizontal axis represent quarters since the monetary policy shock. Note that for the variables that exhibit a large degree of persistence, the mean squared error deteriorates rapidly as the horizon increases. This is particularly the case for the log of the GDP deflator,  $P_t$ . Instead,

for the federal funds rate, the MSE increases a quarter after the shock but it drops in the following quarter staying about the same level in the long-run.

Comparing the results for the levels specification and the DF-GLS pre-test specification provides some interesting insights. First, pre-testing for unit roots has the effect of increasing the MSE in the long-run for all impulse response functions but real GDP. The increase in the MSE, relative to the levels specification, is larger for the monetary variables, which are characterized by lower persistence. As it can be seen from the scale of the mean squared errors, the cost associated with pre-testing for unit roots is not large.<sup>3</sup>

### 5.3.2 Multivariate model with possible cointegration

An alternative model selection strategy faced by a practitioner involves considering the possibility that some of the variables in the system might be cointegrated. In this case, a possible avenue would be to pre-test for unit roots in each of the series in  $y_t$ , and then consider cointegration among the subset of variables that have a unit root. Although this strategy might be employed in practice, the most commonly used approach is to directly tests for the cointegration rank, without pre-testing for unit roots. We thus follow the bulk of the literature and address the issue of cointegration in the full set of variables.

Thus, consider the case where  $y_t$  might have a VECM representation:

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t.$$

For instance, a number of possible cointegrating relationships in a model of monetary transmission is explored by Juselius (1998a, b). One possibility is that equilibrium in the money market is attained via stationarity of the liquidity ratio  $M_t - P_t - Y_t$ , so that a cointegration relationship among these three variables exists. In addition, inflation and the nominal interest rate could be cointegrated given the stationarity of the real interest rate. Moreover, the IS relationship would suggest that the trend-adjusted real GDP and inflation (or the interest rate) are cointegrated. Finally, Strongin (1995) argues that, initially, monetary policy shocks lead only to changes in the composition of total reserves between borrowed and nonborrowed reserves. Thus, one could conjecture that changes in

---

<sup>3</sup>Simulation results (not reported here) suggest that using nonborrowed reserves to identify the monetary policy shock does not alter our conclusions. In fact, when the data are subject to an ADF-GLS pre-test specification, the federal funds rate model and the nonborrowed reserves model exhibit mean squared errors of similar magnitudes for the main variables of interest (i.e.,  $Y_t$  and  $P_t$ ).

the ratio of nonborrowed to total reserves are only short-lived and that a long-run relationship exists so that  $TR_t$  and  $NBR_t$  are cointegrated.

In this section, we evaluate the effect of Johansen's method (Johansen, 1988), used to pre-test for the cointegration rank, on the MSE of the impulse response functions. In order to do this, we simulate 10,000 samples of the data generating process as described in the previous section. For each of these samples, we follow a sequential procedure to determine the cointegrating rank of the 7-variable system (see Lütkepohl 2005, and the references therein). That is, using the trace statistic, we test the sequence of null hypotheses

$$H_0 : \text{rank}(\Pi) = 0; H_0 : \text{rank}(\Pi) = 1, \dots; H_0 : \text{rank}(\Pi) = 6 \quad (9)$$

and stop the test procedure when the null hypothesis cannot be rejected for the first time. The cointegrating rank is selected to be the value when the test is stopped. That is, if we cannot reject the null that  $\text{rank}(\Pi) = 0$ , the cointegrating rank is taken to be 0 and we estimate a model in first differences. If the null is rejected, we proceed to test  $H_0 : \text{rank}(\Pi) = 1$ . This continues until we are not able to reject the null. If the first time we reject the null is for  $0 < \text{rank}(\Pi) = r < 7$ , then the analysis proceeds with a cointegrating rank of  $r$  and a VECM is estimated. Finally, if we cannot reject the null that  $\text{rank}(\Pi) = 7$ , then we estimate a VAR in levels. We use the trace statistic at 5% significance level to test the sequence of null hypothesis describe in (9). The critical values for the trace test are taken from MacKinnon, Haug and Michelis (1999).

Before we proceed to describe the effects of pre-testing on the impulse response functions, it is worth mentioning that in 55% percent of the simulations, the trace statistic leads us to conclude that the rank  $\text{rank}(\Pi) = 7$ . Hence, in more than a half of the simulations we use a VAR in levels when computing the impulse response functions. This result is not surprising because our data generating process is given by the estimated VAR in levels (equation (8)).

The comparison of the dashed and the dotted lines in Figure 16 suggests some differences in the MSEs for the impulse responses when we use the ADF-GLS pre-test specification and the Johansen pre-test specification. Only slight differences are observed in the MSEs on impact and in the short-run. As the horizon increases, the MSE for the impulse responses of  $Y_t$ ,  $P_t$ ,  $TR_t$ , and  $NBR_t$  are larger under the Johansen pre-test strategy than if we pre-test for unit roots.

The cost associated with pre-testing for cointegrating rank, as measured by the mean squared errors, is not large. This finding is consistent with our previous findings based on short-run (re-

cursive) identification scheme. Nevertheless, in the long-run, pre-testing for cointegration results in larger MSEs for all impulse responses –except for the log of the GDP deflator– than the levels specification.

## 6 Concluding Remarks and Practical Recommendations

The practical relevance of structural VARs has been recently questioned (Cooley and Dwyer, 1998; Chari, Kehoe and McGratten, 2008; among others) with criticisms targeting their theoretical underpinnings, identification assumptions and statistical methodology. Despite these criticisms, the SVARs prove to be an indispensable tool for policy analysis and evaluation of dynamic economic models. However, their robustness to uncertainty about the magnitude of the largest roots in the system and possible co-movements between the variables has not been fully investigated. For example, the costs from erroneously imposing restrictions (by differencing the data and/or incorporating cointegrating relationships) on the models or from estimating unrestricted VAR models in levels (when the true model is a VECM) have not been quantified for the purpose of impulse response analysis under different identification schemes. In this chapter, we evaluated the robustness of alternative VAR specifications to deviations from exact unit roots and cointegration.

The main findings and practical recommendations that emerge from our results are the following. First, under the long-run identification scheme, specification strategies based on pre-tests and restricted VAR models are not robust to uncertainty about the largest roots of the process and may lead to highly distorted inference. On the other hand, impulse response estimates from VAR models using long-run restrictions are inconsistent when the non-unit root variables are expressed in levels, but have near-unit roots. Thus, applied researchers should exert caution in using models based on long-run restrictions. Nevertheless, we showed that in practice, the VAR specification with the non-unit root variable in levels is more robust to possible low-frequency co-movements and departures from exact unit roots.

Second, under the short-run identification scheme, the restricted (based on pre-tests for unit roots and cointegration) and unrestricted VAR specifications do not exhibit substantial differences in their computed impulse responses. In other words, the costs from imposing restrictions on the variables of the model (first differences or cointegration), when these restrictions do not hold exactly in the true model, do not tend to be too large when the structural impulse responses are

identified through short-run restrictions. However, the unrestricted VAR in levels appears to be the most robust specification when there is uncertainty about the magnitude of the largest roots and the co-movement between the variables. We conclude that estimating VAR models in levels and identifying the structural impulse responses through short-run restrictions emerges as the most reliable strategy for applied work.

## References

- [1] Bernanke, B. S., and A. S. Blinder (1992), “The federal funds rate and the channels of monetary transmission,” *American Economic Review* **82**, 901–921.
- [2] Blanchard, O., and D. Quah (1989), “The dynamic effects of aggregate demand and supply disturbances,” *American Economic Review* **79**, 655–673.
- [3] Chari, V. V., P. J. Kehoe and E. R. McGrattan (2008), “Are structural VARs with long-run restrictions useful in developing business cycle theory?,” *Journal of Monetary Economics* **55**, 1337–1352.
- [4] Christiano, L. J., and M. Eichenbaum (1992), “Identification and the liquidity effect of a monetary policy shock,” in *Political Economy, Growth and Business Cycles*, eds., A. Cukierman, Z. Hercowitz, and L. Leiderman, Cambridge and London: MIT Press, 335–370.
- [5] Christiano, L. J., M. Eichenbaum and C. Evans (1999), “Monetary policy shocks: What have we learned and to what end?,” in *Handbook of Macroeconomics*, Vol. 1A, eds., M. Woodford and J. Taylor, Amsterdam; New York and Oxford: Elsevier Science, North-Holland.
- [6] Christiano, L. J., M. Eichenbaum, and R. Vigfusson (2006), “Assessing structural VARs,” In: D. Acemoglu, K. Rogoff and M. Woodford, eds., *NBER Macroeconomics Annual* Cambridge, MA: MIT Press.
- [7] Cooley, T., and M. Dwyer (1998), “Business cycle analysis without much theory: A look at structural VARs,” *Journal of Econometrics* **83**, 57–88.
- [8] Eichenbaum, M. (1992), “Comment on interpreting the macroeconomic time series facts: The effects of monetary policy,” *European Economic Review* **36**, 1001–1011.
- [9] Elliott, G. (1998), “On the robustness of cointegration methods when regressors almost have unit roots,” *Econometrica* **66**, 149–158.
- [10] Gali, J. (1999), “Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?” *American Economic Review* **89**, 249–271.
- [11] Gospodinov, N. (2004), “Asymptotic confidence intervals for impulse responses of near-integrated processes,” *Econometrics Journal* **7**, 505–527.

- [12] Gospodinov, N. (2010), “Inference in nearly nonstationary SVAR models with long-run identifying restrictions,” *Journal of Business and Economic Statistics* **28**, 1–12.
- [13] Gospodinov, N., A. Maynard, and E. Pesavento (2011), “Sensitivity of impulse responses to small low frequency co-movements: Reconciling the evidence on the effects of technology shocks,” *Journal of Business and Economic Statistics* **29**, 455–467.
- [14] Hansen, B. E. (1999), “The grid bootstrap and the autoregressive model,” *Review of Economics and Statistics* **81**, 594–607.
- [15] Hamilton, J. D., and A. M. Herrera (2004), “Oil shocks and aggregate macroeconomic behavior: The role of monetary policy,” *Journal of Money, Credit, and Banking* **36**, 265–286.
- [16] Inoue, A., and L. Kilian (2002), “Bootstrapping autoregressive processes with possible unit roots,” *Econometrica* **70**, 377–391.
- [17] Johansen, S. (1988), “Statistical analysis of cointegration vectors,” *Journal of Economic Dynamics and Control* **12**, 231–254.
- [18] Juselius, K. (1998a), “A structured VAR for Denmark under changing monetary regimes,” *Journal of Business and Economic Statistics* **16**, 400–411.
- [19] Juselius, K. (1998b), “Changing monetary transmission mechanisms within the EU,” *Empirical Economics* **23**, 455–481.
- [20] Kilian, L. (1998a), “Small-sample confidence intervals for impulse response functions,” *Review of Economics and Statistics* **80**, 218–230.
- [21] Kilian, L. (1998b), “Confidence intervals for impulse responses under departures from normality,” *Econometric Reviews* **17**, 1–29.
- [22] Kilian, L. and P. L. Chang (2000), “How accurate are confidence intervals for impulse responses in large VAR models?,” *Economics Letters* **69**, 299–307.
- [23] Lütkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin Heidelberg.
- [24] Lütkepohl, H., and H. E. Reimers (1992), “Impulse response analysis of cointegrated systems,” *Journal of Economic Dynamics and Control* **16**, 53–78.



- [25] MacKinnon, J. G., A. A. Haug, and L. Michelis (1999), “Numerical distribution functions of likelihood ratio tests for cointegration,” *Journal of Applied Econometrics* **14**, 563–577.
- [26] Mikusheva, A. (2012), “One-dimensional inferences in autoregressive models in a potential presence of a unit root,” *Econometrica* **80**, 173–212.
- [27] Ng, S. and P. Perron (2001), “Lag length selection and the construction of unit root tests with good size and power,” *Econometrica* **69**, 1519–1554.
- [28] Pesavento, E., and B. Rossi (2006), “Small sample confidence intervals for multivariate impulse response functions at long horizons,” *Journal of Applied Econometrics* **21**, 1135–1155.
- [29] Phillips, P. C. B. (1998), “Impulse response and forecast error variance asymptotics in non-stationary VARs,” *Journal of Econometrics* **83**, 21–56.
- [30] Runkle, D. E. (1987), “Vector autoregressions and reality,” *Journal of Business and Economic Statistics* **5**, 437–442.
- [31] Sims, C. A. (1986), “Are forecasting models usable for policy analysis?,” *Federal Reserve Bank of Minneapolis Quarterly Review* **10**, 2–16.
- [32] Sims, C. A. (1992), “Interpreting the macroeconomic time series facts: The effects of monetary policy,” *European Economic Review*, **36**, 975–1000.
- [33] Sims, C. A., J. H. Stock, and M. W. Watson (1990), “Inference in linear time series with some unit roots,” *Econometrica* **58**, 113–144.
- [34] Stock, J. H. (1991), “Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series,” *Journal of Monetary Economics* **28**, 435–459.
- [35] Strongin, S. (1995), “The identification of monetary policy disturbances explaining the liquidity puzzle,” *Journal of Monetary Economics* **35**, 463–497.

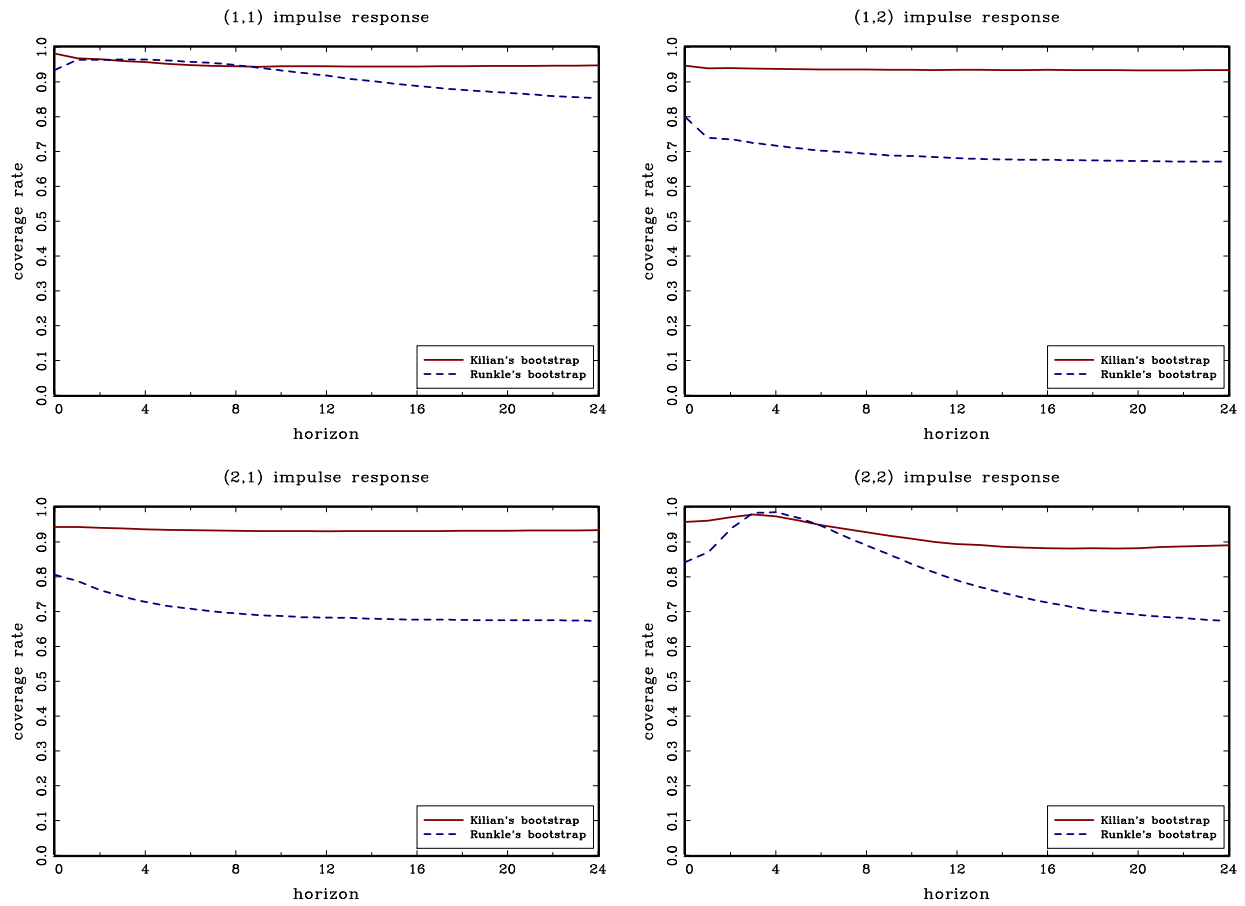


Figure 1: Coverage rates of 95% confidence intervals for the  $(i, j)$  impulse response (response of the  $i$ -th ( $i = 1, 2$ ) variable to the  $j$ -th ( $j = 1, 2$ ) shock) constructed by Kilian's (1998a,b) bias-corrected bootstrap and Runkle's (1987) conventional bootstrap methods. The data are generated from model (7) with  $\rho = 0.95$ . The impulse responses are identified using the long-run identification scheme.

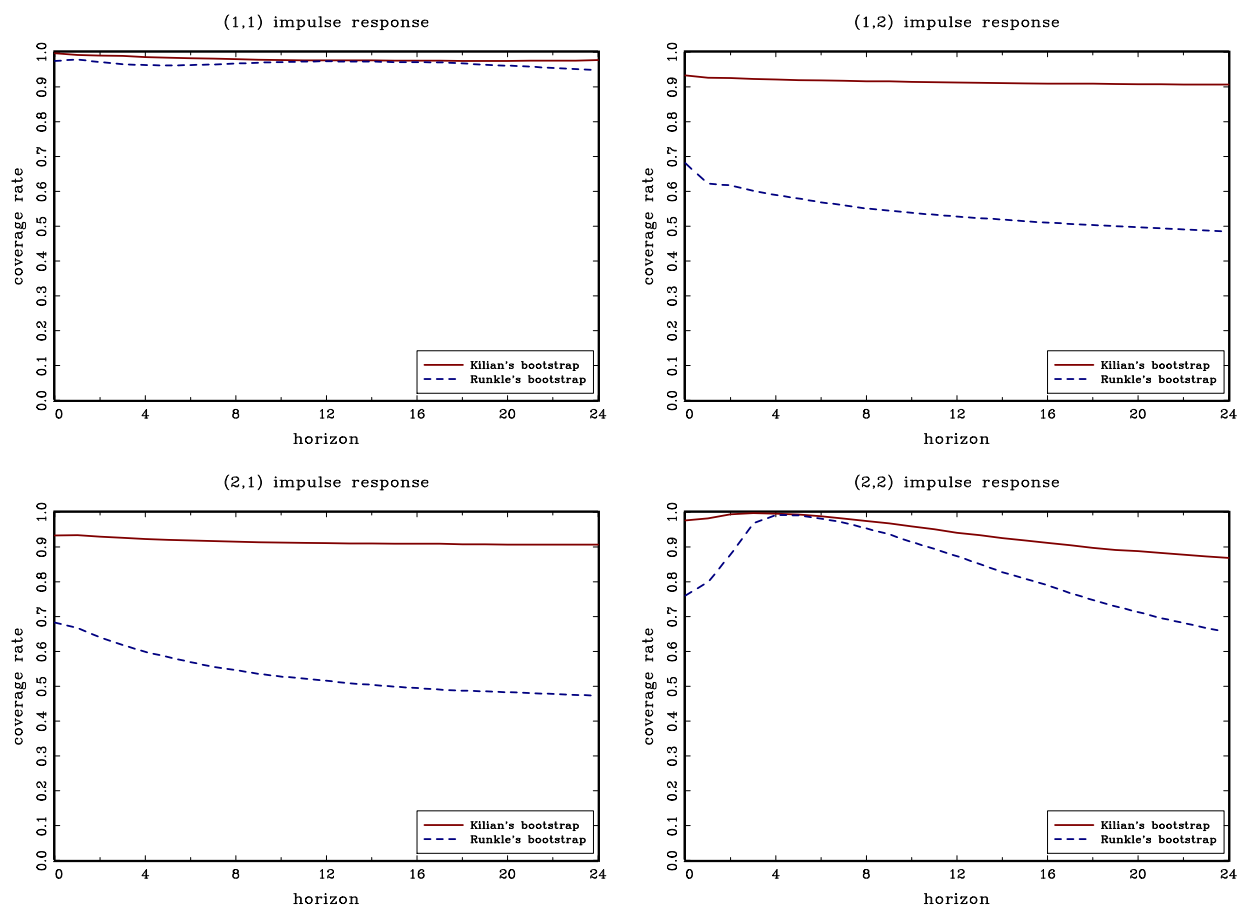


Figure 2: Coverage rates of 95% confidence intervals for the  $(i, j)$  impulse response (response of the  $i$ -th ( $i = 1, 2$ ) variable to the  $j$ -th ( $j = 1, 2$ ) shock) constructed by Kilian's (1998a,b) bias-corrected bootstrap and Runkle's (1989) conventional bootstrap methods. The data are generated from model (7) with  $\rho = 0.98$ . The impulse responses are identified using the long-run identification scheme.

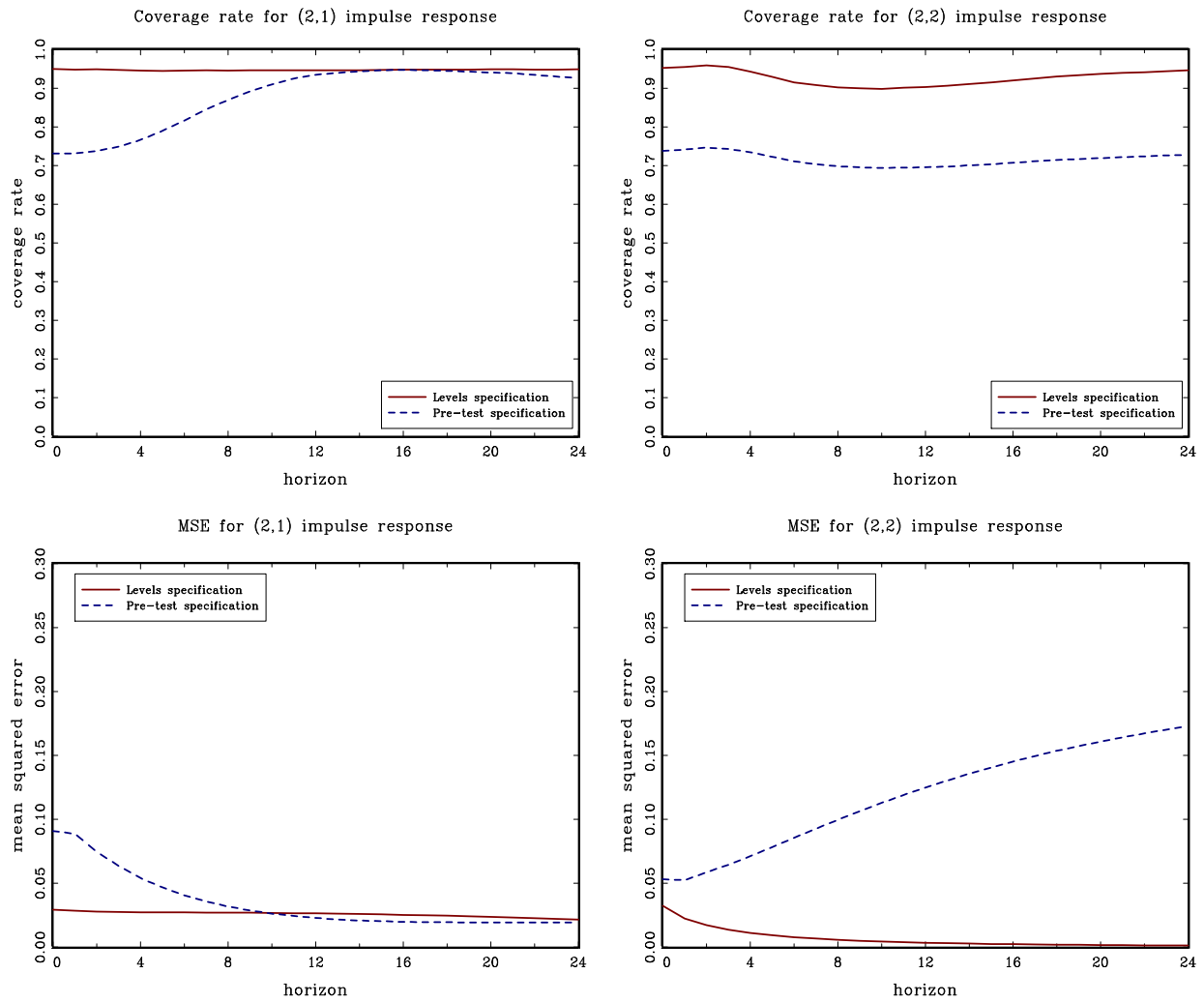


Figure 3: Coverage rates of 95% bias-corrected bootstrap confidence intervals and mean squared errors for the (2,1) and (2,2) impulse responses for the levels (solid line) and pre-test (dashed line) VAR specifications. The data are generated from model (7) with  $\rho = 0.92$ . The impulse responses are identified using the long-run identification scheme.

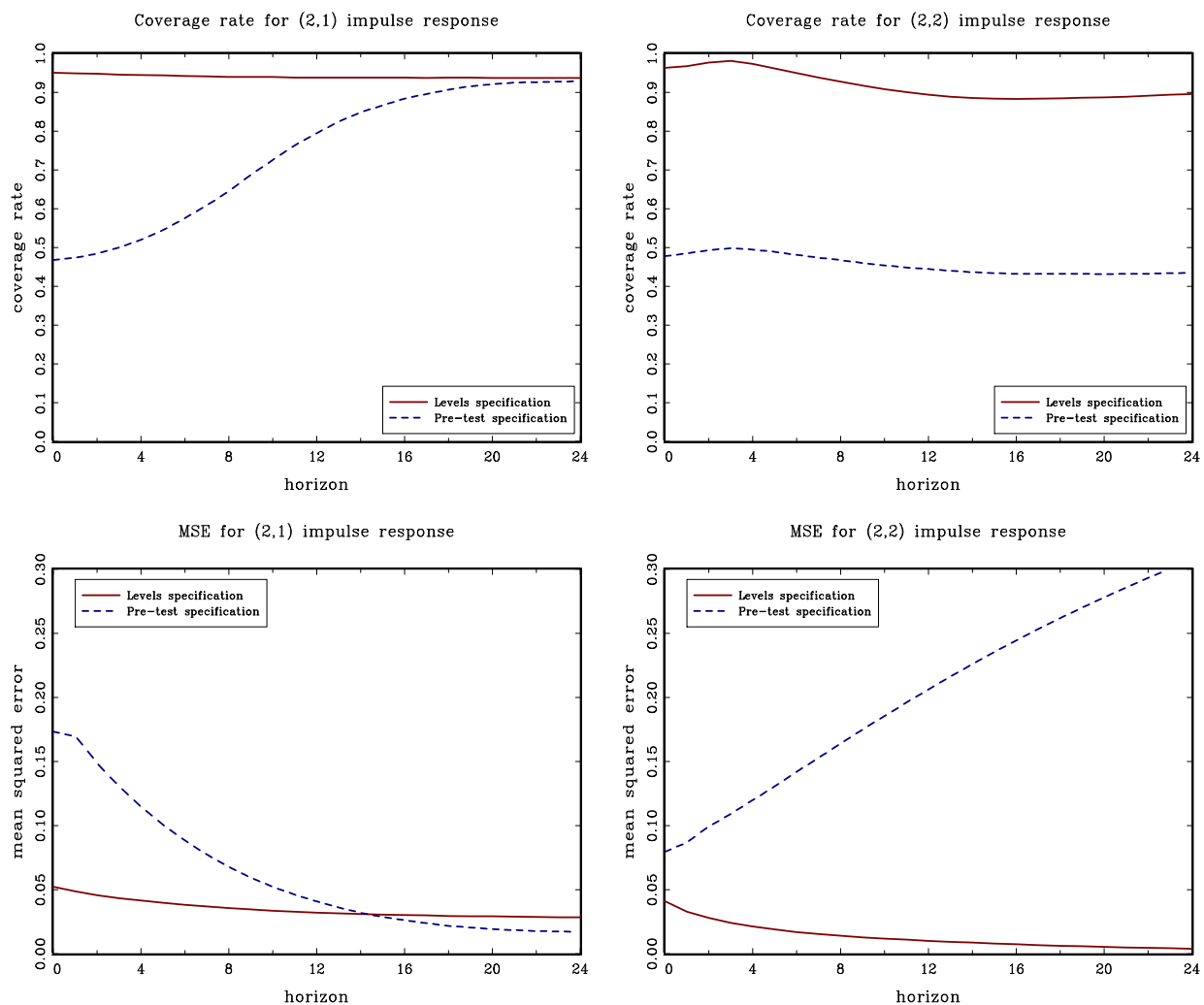


Figure 4: Coverage rates of 95% bias-corrected bootstrap confidence intervals and mean squared errors for the (2,1) and (2,2) impulse responses for the levels (solid line) and pre-test (dashed line) VAR specifications. The data are generated from model (7) with  $\rho = 0.95$ . The impulse responses are identified using the long-run identification scheme.

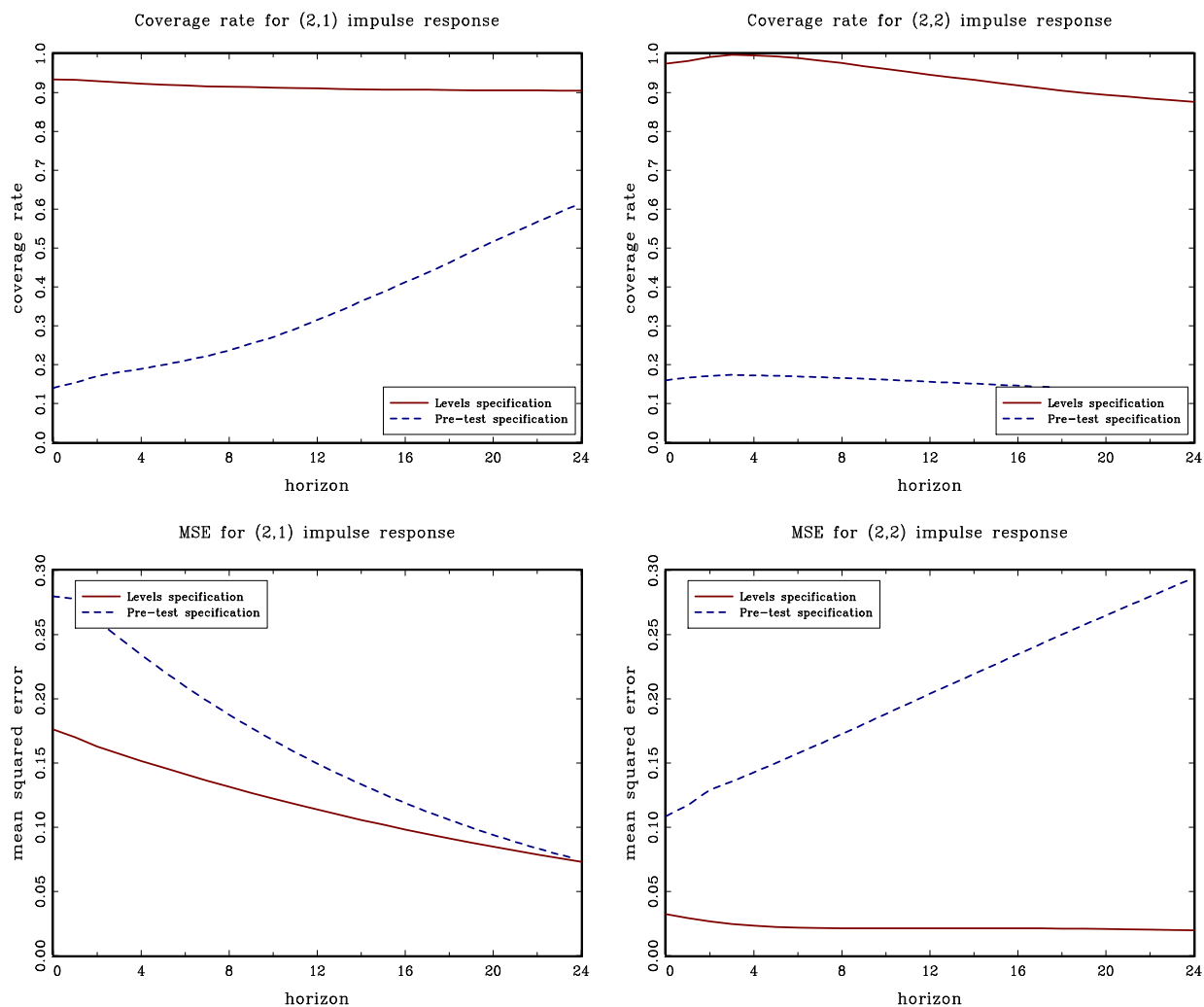


Figure 5: Coverage rates of 95% bias-corrected bootstrap confidence intervals and mean squared errors for the (2,1) and (2,2) impulse responses for the levels (solid line) and pre-test (dashed line) VAR specifications. The data are generated from model (7) with  $\rho = 0.98$ . The impulse responses are identified using the long-run identification scheme.

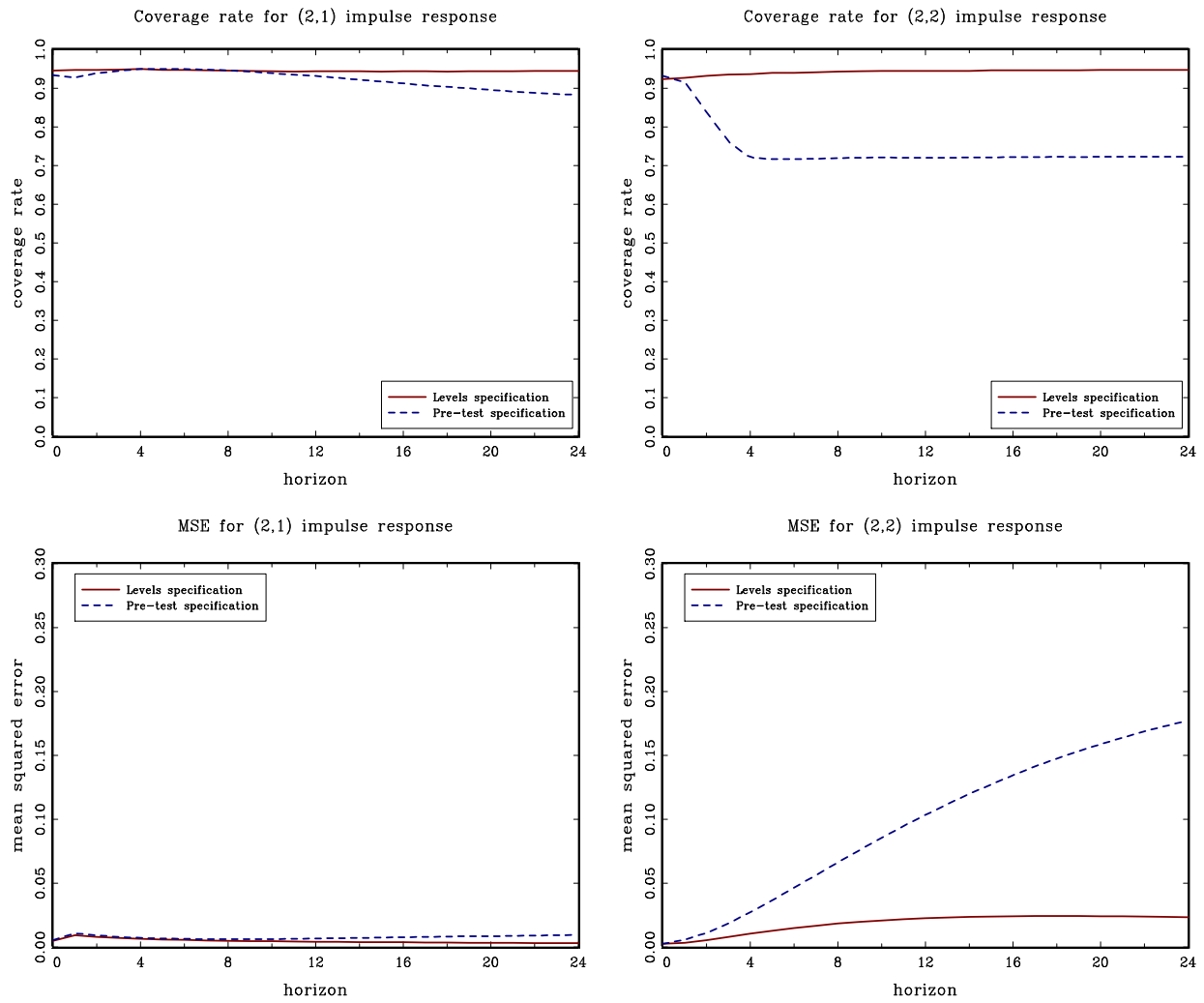


Figure 6: Coverage rates of 95% bias-corrected bootstrap confidence intervals and mean squared errors for the (2,1) and (2,2) impulse responses for the levels (solid line) and pre-test (dashed line) VAR specifications. The data are generated from model (7) with  $\rho = 0.92$ . The impulse responses are identified using the short-run (recursive) identification scheme.

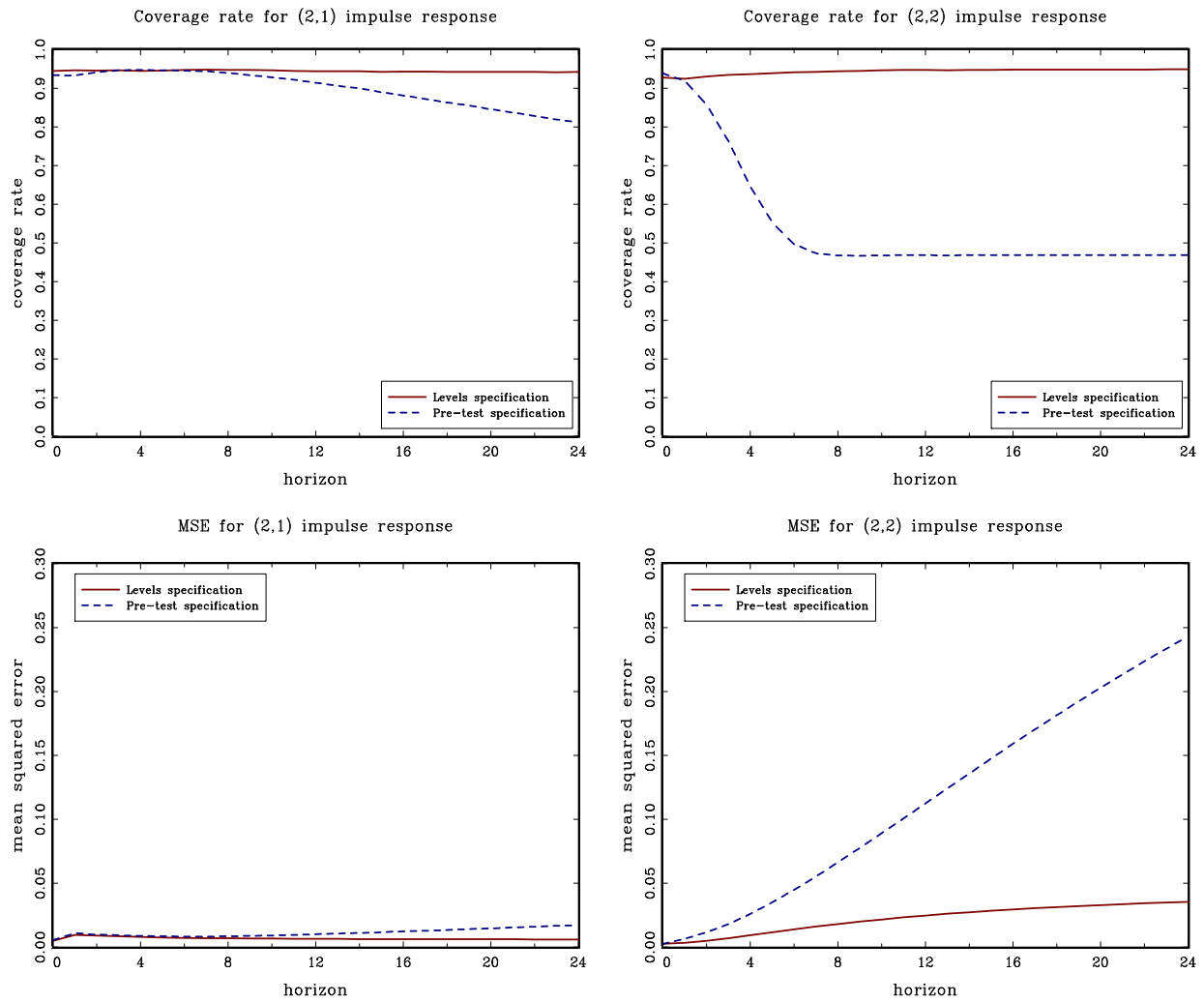


Figure 7: Coverage rates of 95% bias-corrected bootstrap confidence intervals and mean squared errors for the (2,1) and (2,2) impulse responses for the levels (solid line) and pre-test (dashed line) VAR specifications. The data are generated from model (7) with  $\rho = 0.95$ . The impulse responses are identified using the short-run (recursive) identification scheme.



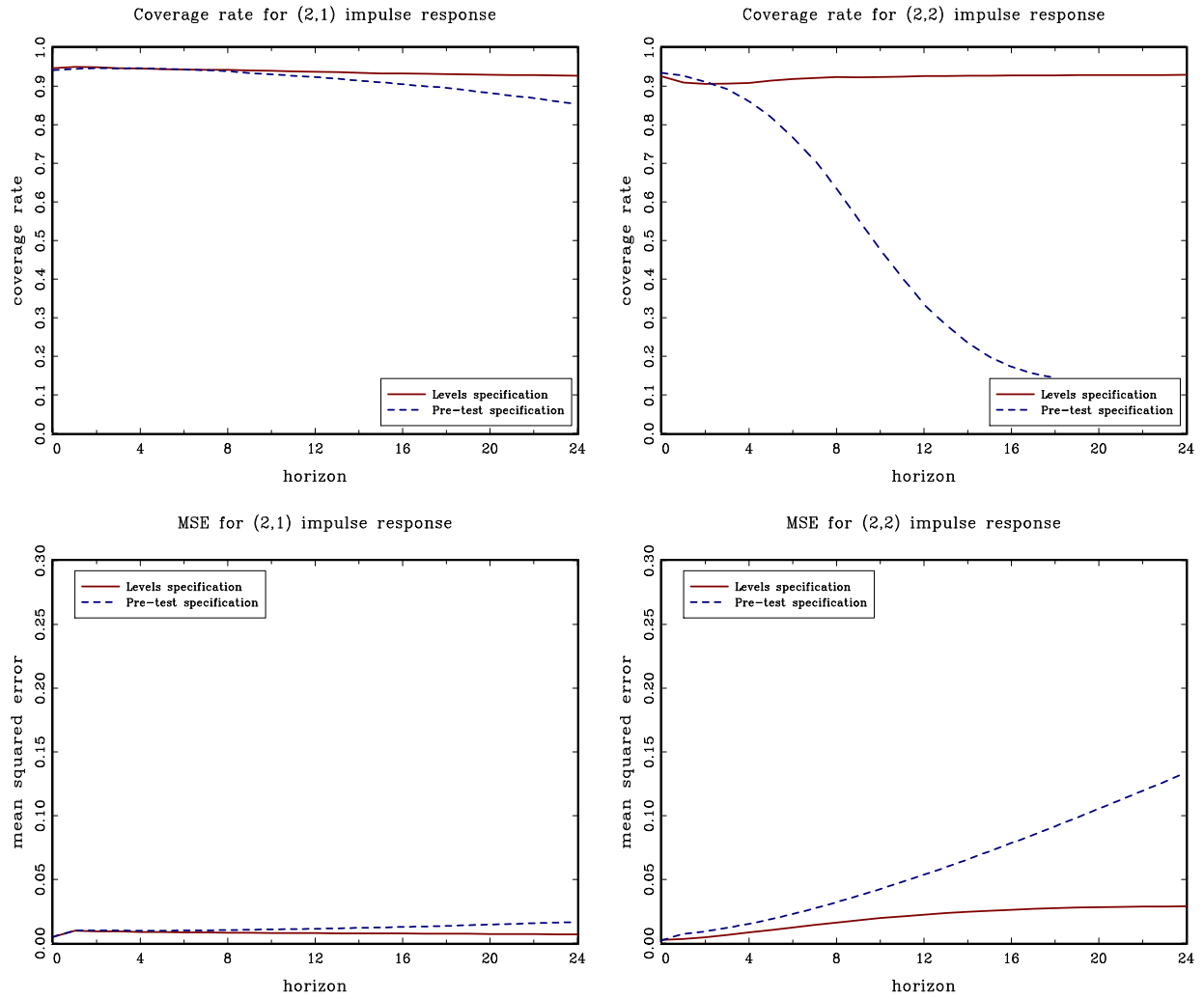


Figure 8: Coverage rates of 95% bias-corrected bootstrap confidence intervals and mean squared errors for the (2,1) and (2,2) impulse responses for the levels (solid line) and pre-test (dashed line) VAR specifications. The data are generated from model (7) with  $\rho = 0.98$ . The impulse responses are identified using the short-run (recursive) identification scheme.

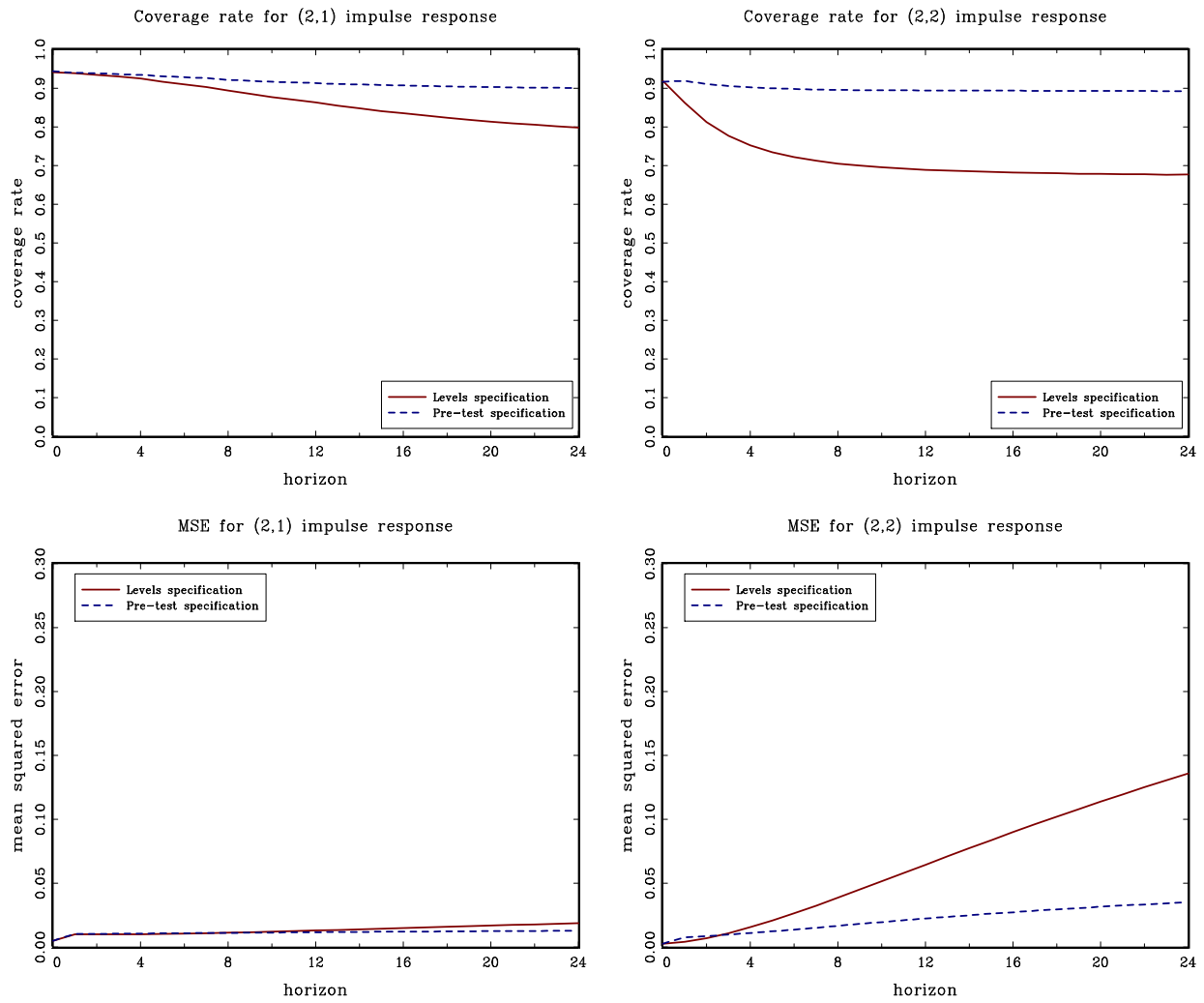


Figure 9: Coverage rates of 95% bias-corrected bootstrap confidence intervals and mean squared errors for the (2,1) and (2,2) impulse responses for the levels (solid line) and pre-test (dashed line) VAR specifications. The data are generated from model (7) with  $\rho = 1$ . The impulse responses are identified using the short-run (recursive) identification scheme.

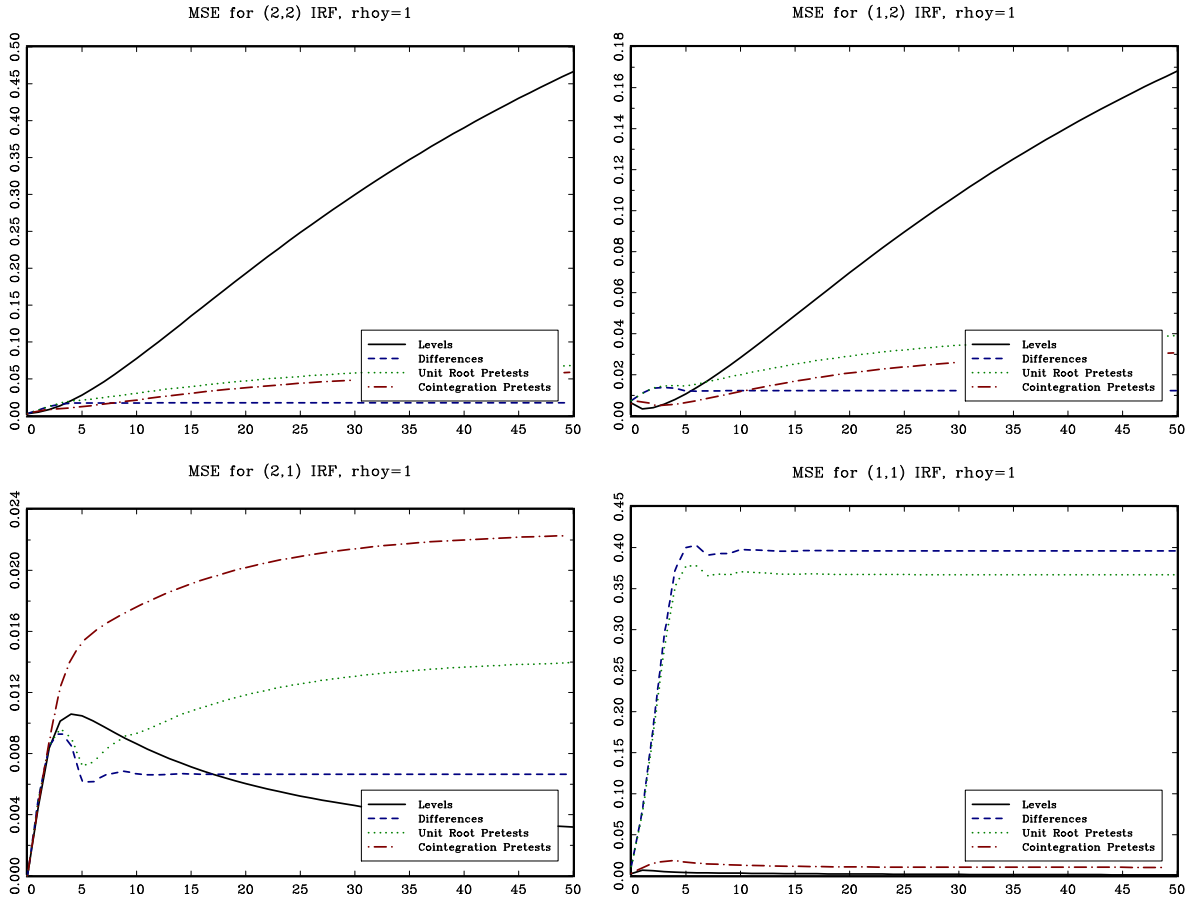


Figure 10: MSEs for  $\rho = 1$ .

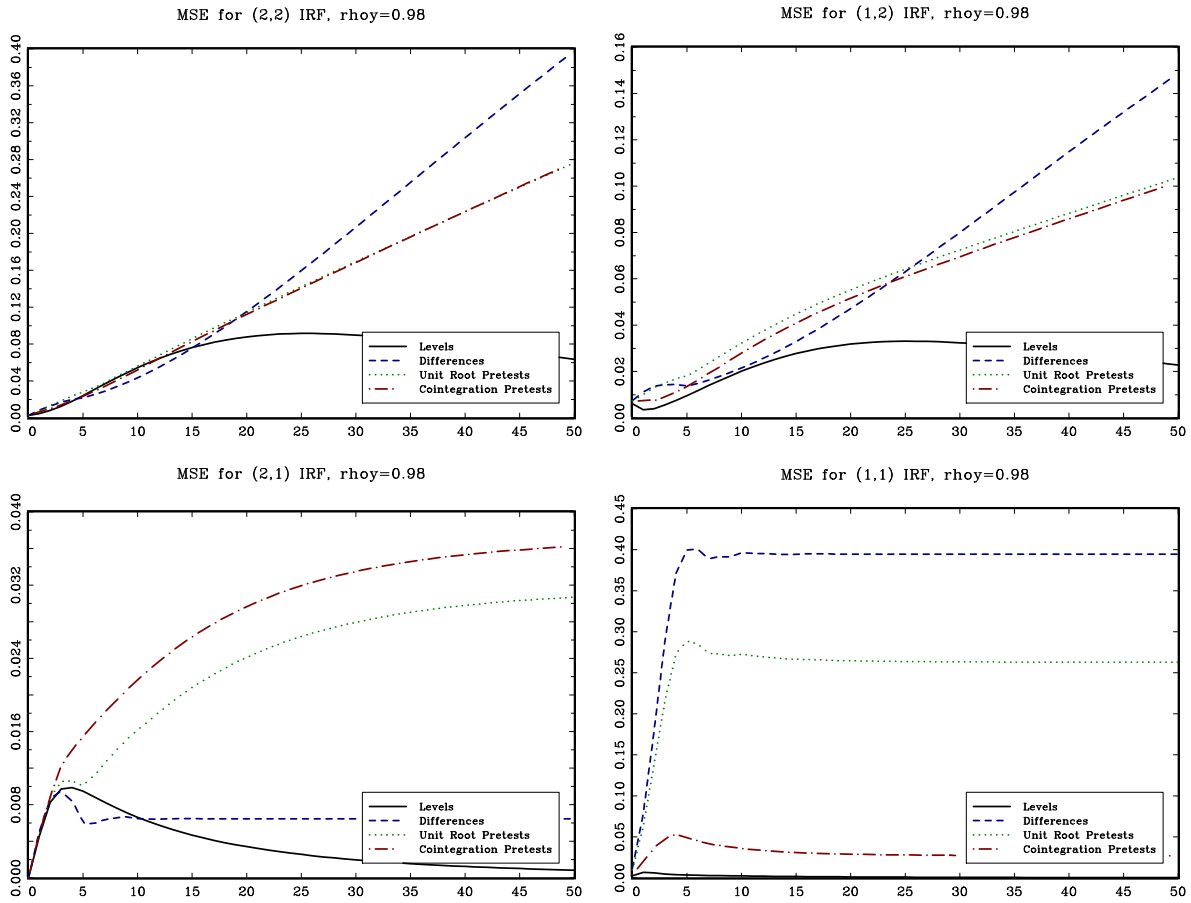


Figure 11: MSEs for  $\rho = 0.98$ .

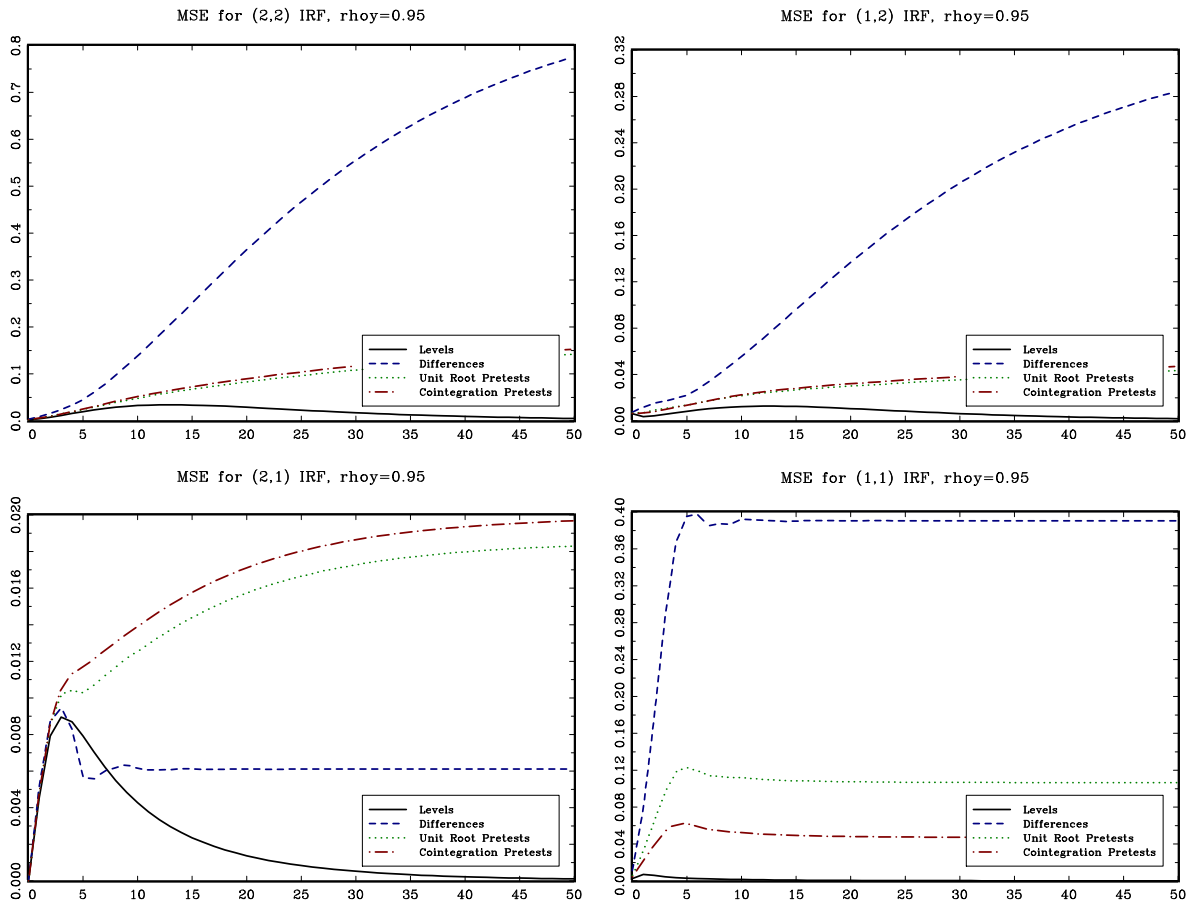


Figure 12: MSEs for  $\rho = 0.95$ .

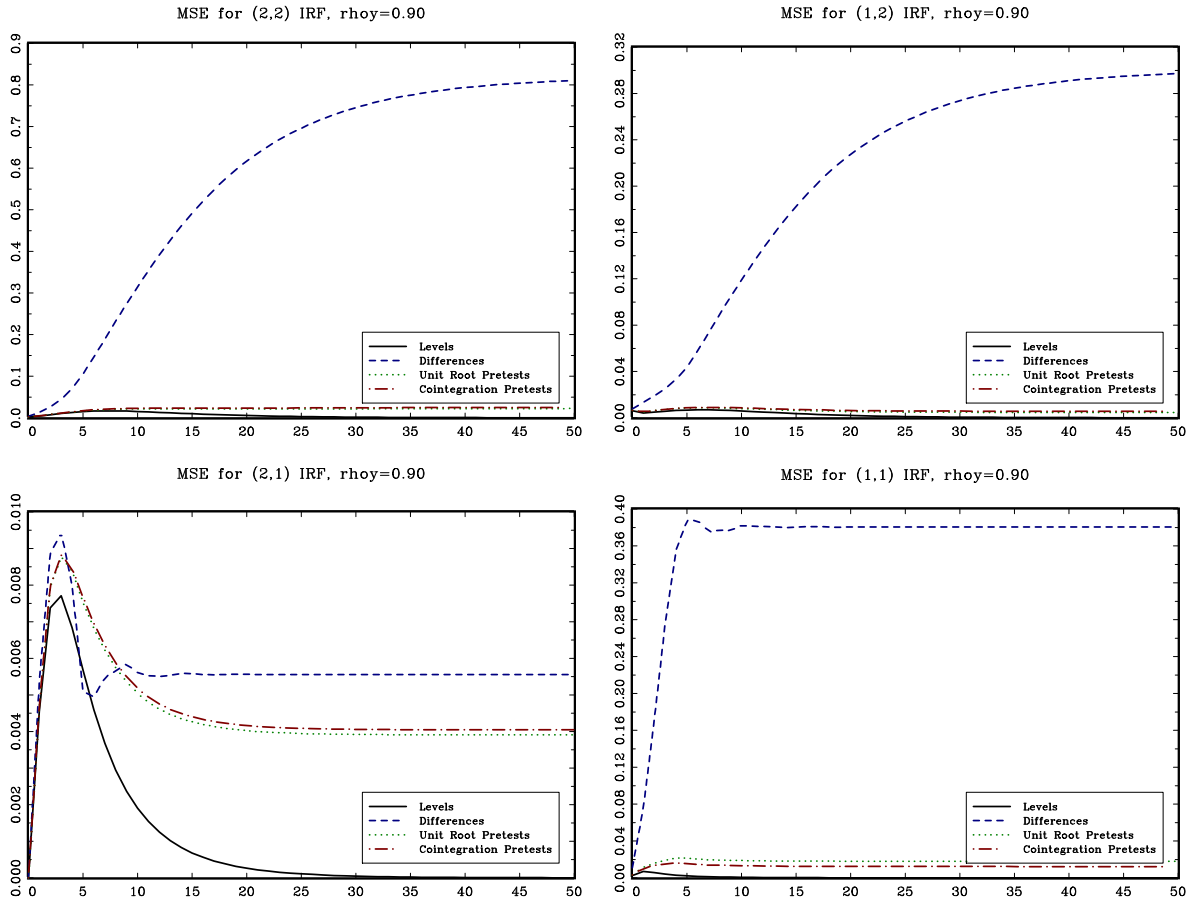


Figure 13: MSEs for  $\rho = 0.90$ .

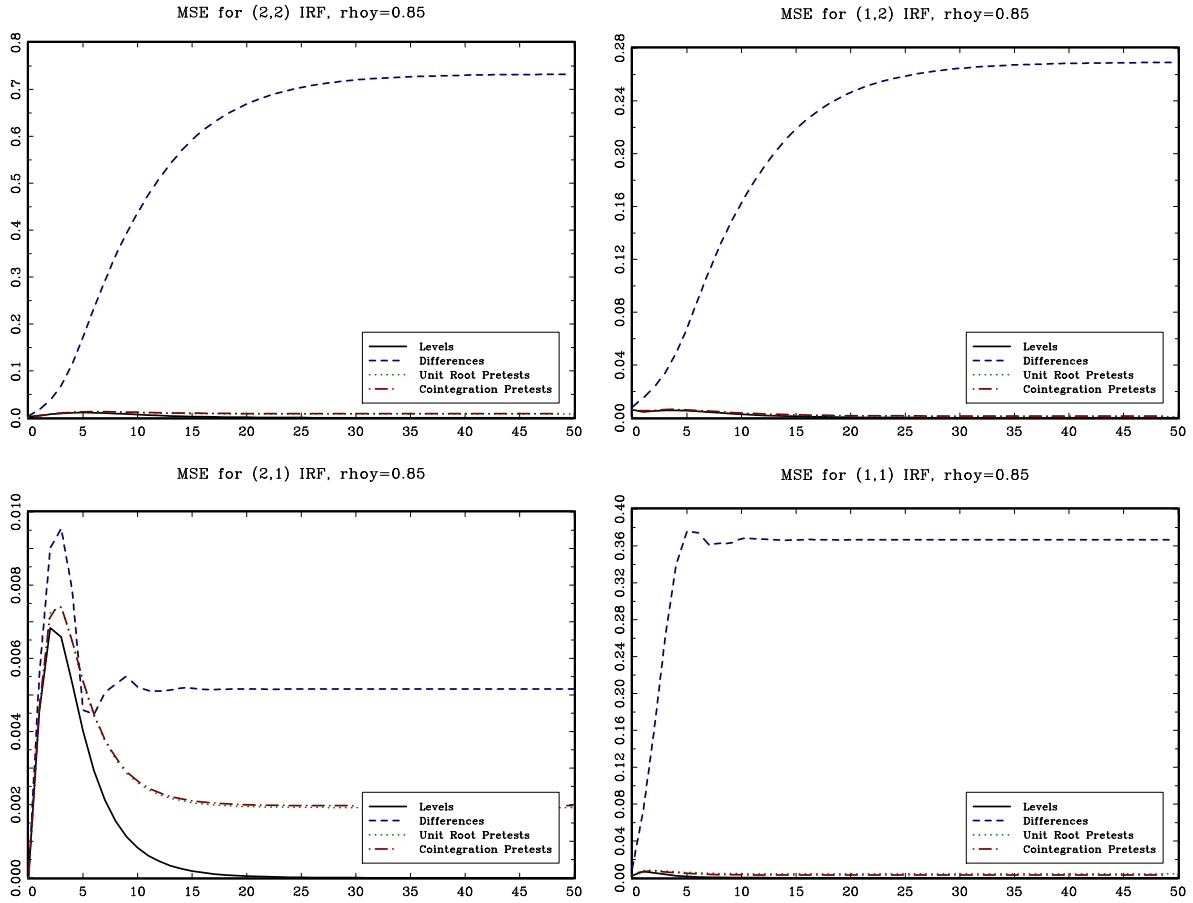


Figure 1: Figure 14: MSEs for  $\rho = 0.85$ .

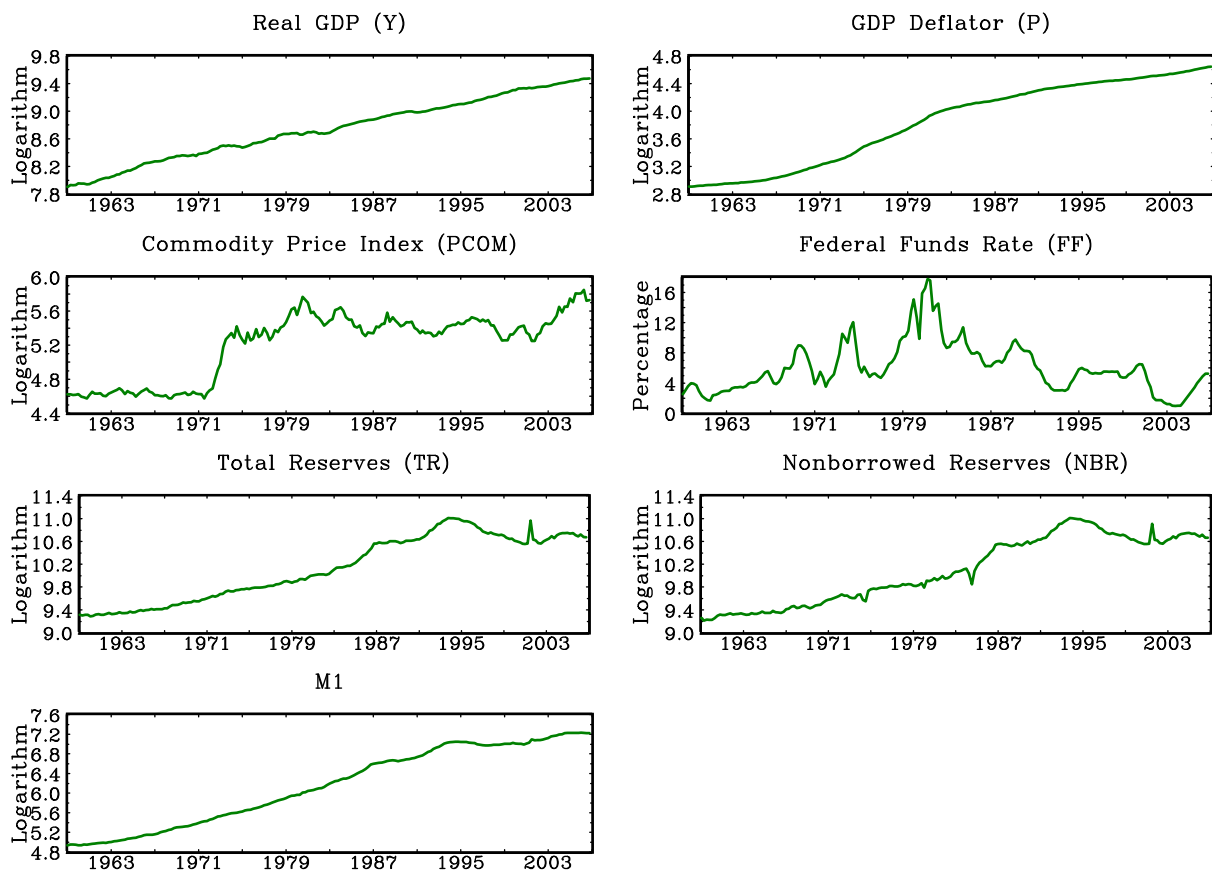


Figure 15: Data used in the Monetary Policy VAR.



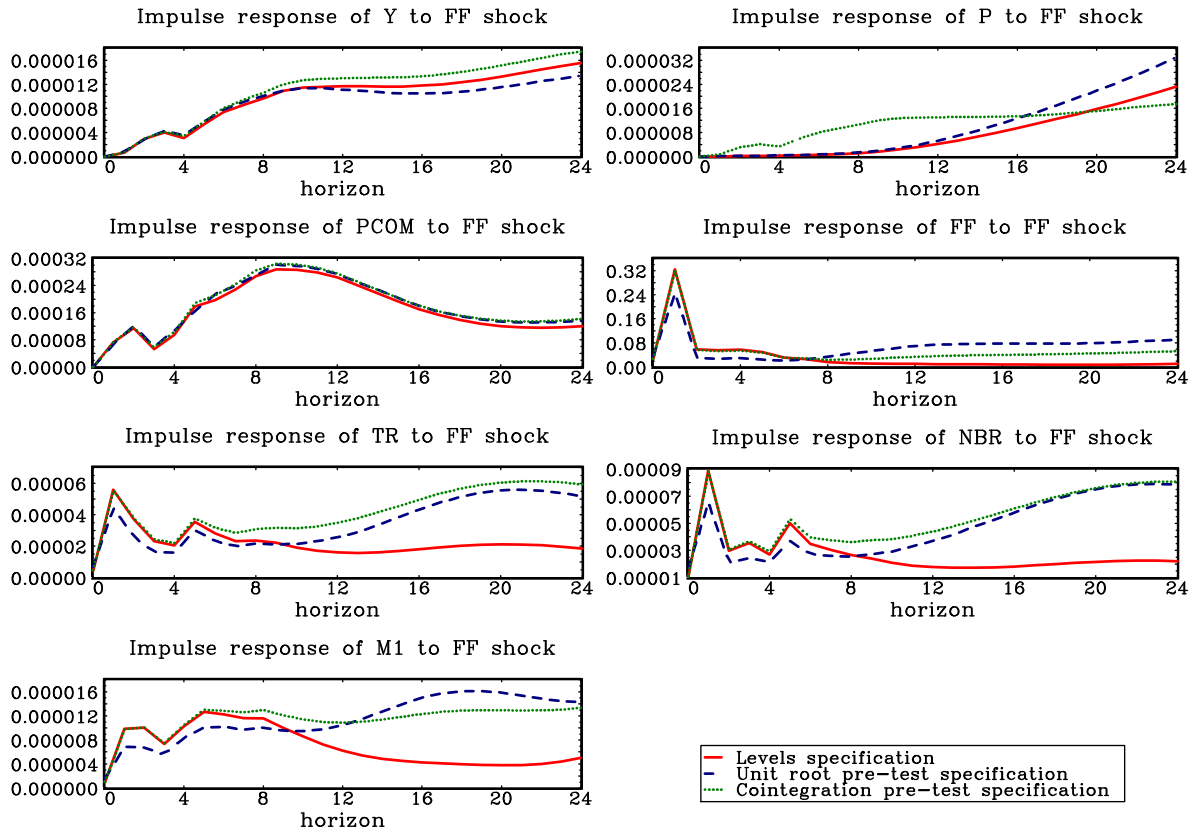


Figure 16: MSEs for the impulse responses to a federal funds rate shock in CEE model in (8). The data are generated as described in Section 5.3. The impulse responses are identified using short-run (recursive) restrictions.