The importance of long run structure for impulse response analysis in VAR models

James Mitchell*
National Institute of Economic and Social Research

Abstract

This paper examines the finite sample accuracy of impulse response functions in VAR models when cointegration is present but not imposed in estimation. It finds that in typical applications there can be substantial biases at the short horizon, as well as the long. These results suggest that, in contrast to common practice, before conducting impulse response analysis researchers should analyse the cointegrating properties of the data and restrict the VAR (i.e. estimate a reduced rank VAR) when the data are cointegrated. But there are practical dangers associated with following this advice. When the data are "nearly" cointegrated then not just does the VAR in levels provide misleading inference about long-run impulse responses but so does the reduced rank VAR.

Keywords: VAR model; Cointegration; Impulse Response Functions; Monetary Policy

JEL classification: C32 E17 E24 E32

1 Introduction

Impulse response analysis is used widely in the empirical literature to uncover the dynamic relationship between macroeconomic variables within vector-autoregressive (VAR) models. Impulse responses measure the time profile of the effect of a shock, or impulse, on the (expected) future values of a variable. By imposing specific restrictions on the parameters of the VAR model the shocks can be attributed an economic meaning; for a review see Watson (1994). A popular application has been the identification of monetary policy shocks.

The standard approach to identifying impulse responses imposes restrictions on a VAR model estimated in the (log) levels of the variables. However, if present cointegration imposes restrictions on the VAR. It is known that if these restrictions are not imposed

*Thanks to Joe Byrne and Stephen Hall for helpful comments. Address for correspondence: NIESR, 2 Dean Trench Street, Smith Square, London, SW1P 3HE, U.K. Tel: +44 (0) 207 654 1926. Fax: +44 (0) 207 654 1900. E-Mail: j.mitchell@niesr.ac.uk.

and a nonstationary VAR model in levels is estimated, impulse responses are inconsistent at long horizons. More specifically, the impulse responses are inconsistent at long horizons as the horizon increases with the sample size; this is because the nonstationarity means the true impulse responses do not tend to zero as the horizon increases - the effect of the unit root persists, but the unit roots are estimated with error; see Phillips (1998) for details.[1]

This paper informs applied users that in applications of the type often considered in macroeconomics, with 4-7 variables, there can be substantial biases involved in estimating VAR models in levels if there is in fact cointegration. These biases are important at the short horizon, as well as the long. They can affect economic inference drawn from the estimated impulse response functions; in fact "puzzles" can appear when there are in fact none. These results suggest that, in contrast to common practice, before identifying and estimating any *economic* shock, researchers should analyse the cointegrating properties of the data and restrict the VAR (i.e. estimate a reduced rank VAR, a vector error correction model) when the data are cointegrated. We also find that not knowing the true long run structure and having to estimate it is more accurate than ignoring cointegration and estimating a VAR in levels. Unfortunately, we then illustrate the dangers of following this practical advice and estimating a reduced rank VAR when the data appear cointegrated. When the data are "nearly" cointegrated (i.e. there are "near" unit roots in the VAR model) then not just does the VAR in levels provide misleading inference about long-run impulse responses but so does the reduced rank VAR. Inference about longer horizon impulse response functions therefore remains an important and open research area.

2 Monte-Carlo experiments based on typical applications

We use Monte-Carlo experiments to quantify the biases involved in typical VAR applications to monetary policy if there is cointegration but it is ignored. Rather than using artificial data generating processes (DGPs) without economic interpretation, that make it difficult for applied users to interpret the results, we use as DGPs VAR models of the type typically estimated in applications.[2] We focus on responses to monetary policy shocks.

We consider two representative VAR applications to monetary policy. The first is based on a VAR(6) model of U.S. industrial production (Y), U.S. consumer prices (P), U.K. output (YUK), the U.K. interest rate (RUK), the ratio of nonborrowed reserves to total reserves (NBRX), the three month U.S. Treasury bill rate (RUS) and the real exchange rate against the U.K. pound (E) using 197 monthly observations from the period 1974.1 to 1990.5. A negative *orthogonalized* shock to NBRX is interpreted as a (contractionary)

[1]This result for the impulse response functions can be juxtaposed with the well known result that the estimated coefficients in the VAR model remain consistent even if there are unit roots; indeed they can converge at faster rates.

[2]Previous Monte-Carlo work has considered simpler bivariate VARs that are not of direct relevance to applied practitioners; see Naka and Tufte (1997).

monetary policy shock; the variables are ordered (Y,P,YUK,RUK,NBRX,RUS,E).[3] The second application is a VAR(12) model of U.S. industrial production (Y), U.S. consumer prices excluding shelter (P), the commodity price index (PCOM) and the federal funds rate (FF) using 348 monthly observations from the period 1965.1 to 1993.12. An orthogonalized shock to FF is interpreted as a monetary policy shock; the variables are ordered (Y,P,PCOM,FF). These two applications are similar to those of Eichenbaum and Evans (1995) and Bernanke and Gertler (1995), respectively.

For each application we estimate a VAR model with an intercept and a linear time trend.[4] We then test for cointegration using standard techniques; see Johansen (1988). It is well known that these tests lack power. As a result evidence showing that the tests do not reject unit roots and cointegration does not mean that there really are unit roots/cointegration. This unreliability is particularly acute in the presence of roots close to unity. However, commonly in applied work users do act upon the results of unit root/cointegration tests; therefore it is interesting to see what the tests say about the two representative datasets. At a 95% level of significance the maximum eigenvalue statistic points to one cointegrating vector in the first model and one vector in the second model. Use of an alternative test or *a priori* theory may well lead us to find, or expect, a different number of cointegrating vectors. However, our objective here is to examine the sensitivity of impulse response analysis to the assumed cointegrating rank. What is of interest is that standard tests suggest the data are cointegrated.

Cointegrating series are then generated for each application *via* nonparametric residual resampling from the estimated vector error correction model (VECM) that imposes cointegration of the order suggested. The cointegrating vectors are exactly identified following Johansen (1988). Using the simulated series we then estimate the following three models:

1. **VAR**: a VAR with an intercept and trend. The lag length is fixed at the true value.

2. **VECM**: a VECM with an intercept and restricted trend. The lag length, cointegrating rank and cointegrating vectors are fixed at their true values.

3. **RRR**: a VECM with an intercept and restricted trend. The lag length and cointegrating rank are fixed at their true values. The cointegrating vectors are data determined; they are identified and estimated following Johansen (1988). In contrast to VECM in this reduced rank regression (RRR) model the number but not the form of the cointegrating vectors is known.

[3] The chosen ordering affects economic interpretation. Nonrecursive contemporaneous structures have been considered too, see Kim and Roubini (2000), but we do not for computational reasons. Efficient estimation of nonrecursive contemporaneous structures, e.g. by FIML, requires recourse to iterative algorithms and therefore is not so well suited to Monte-Carlo analysis. Moreover, nonrecursive structures will be subject to the same issues raised here for recursive structures. It is of interest to mention an alternative approach to identification. It identifies *economic* shocks by exploiting explicitly the cointegrating properties of the data; see King *et al.* (1991) and Crowder *et al.* (1999). We do not consider such an approach here. We concentrate on identification achieved by imposing restrictions on contemporaneous behaviour, a practice common in applied macroeconomic work.

[4] In the case of cointegration the trend coefficient is restricted; see Pesaran *et al.* (2000).

This process is repeated 10,000 times. At each replication the parameters of the three models are estimated and the impulse responses due to a contractionary monetary policy shock calculated.

The results are shown in Figures 1 and 2. These figures plot various quantiles; they plot the median, 10% and 90% (of the 10,000 replications) impulse responses up to 50 periods out for each of the datasets.[5] The true impulse responses (**TRUE**) from the DGP are also plotted; bias is any difference between TRUE, on the one hand, and VAR, VECM and RRR on the other.

The experiments suggest two conclusions.

First, the median estimates from VECM are, in general, more accurate than VAR. Differences between VAR and VECM can be both economically and statistically significant. For example, consider P's response in Figure 1 and 2. Unsurprisingly as the dominant root is less than unity, the VAR suggests that in the long run the monetary shock has no effect on prices. By contrast the truth is that prices are raised; there is a *price puzzle*; see Sims (1992). Estimating a VECM yields correct inference; even the 10% quantile is greater than zero. Furthermore, it is not just inference about long run behaviour that is affected. After just 5 periods in Figure 1, and 20 periods in Figure 2, the VAR falsely estimates that prices begin to fall. By contrast imposing cointegration we correctly see that prices rise persistently after the monetary shock. For a given set of restrictions on contemporaneous behaviour, whether the effect of a monetary policy shock on domestic prices is consistent with priors depends on the assumed long run of the VAR model. Further evidence of the sensitivity of inference to the assumed long run structure is seen in Y's response in Figure 2. In contrast to the truth (and VECM and RRR), the VAR suggests a (familiar) pronounced hump-shaped response of output to the shock. Note also that in many cases the 10% and 90% quantiles for VAR, unlike VECM, do not even 'cover' the true response.

Secondly, it is better to know just the number of cointegrating vectors and have to estimate their form than assume (incorrectly) that the data are not cointegrated; RRR is very accurate and is closer to TRUE than VAR. Knowing the true cointegrating rank leads to more accurate inference. However, the manner in which the cointegrating vectors are identified can be important. Structural knowledge improves accuracy - not knowing the true long run structure and having to identify and estimate it at each replication leads to increased uncertainty; see in particular Y's response in Figure 2 where the 10% and 90% impulse response bands are noticeably wider for RRR than VECM.

However, it is dangerous to interpret these findings as offering practical advice that one should estimate a reduced rank VAR when the data appear cointegrated. It is has been shown theoretically that when the data are "nearly" cointegrated (i.e. there are "near" unit roots in the VAR model) then not just does the VAR in levels provide misleading inference about long-run impulse responses but so does the reduced rank VAR; see Phillips

[5]There are occasions when the estimated roots of the VAR are greater than unity. In these cases the impulse responses blow up. This explains the finding that impulse response bands can be much wider for VARs than for VECMs; see Naka and Tufte (1997) and Phillips (1998). However, focusing on the 10% and 90% quantiles we are able to abstract from the explosive cases.

(1998).

These dangers are again illustrated empirically by re-visiting the applications above and undertaking one more set of Monte-Carlo experiments. For simplicity attention is now restricted both to the second application considered above and to VAR(1) models with error correction representations of the form $\Delta \mathbf{y}_t = \mathbf{\Pi} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$, with an intercept also included. In contrast to the experiments above, the rank of the coefficient matrix in the VECM, $\mathbf{\Pi}$, is not imposed to be deficient. Rather than imposing zero eigenvalues (given that there are four variables in this VAR application and one cointegrating vector was imposed, there were three zero eigenvalues) we consider the following three sets of eigenvalues for $\mathbf{\Pi}$:

1. **Case 1**: -0.1, -0.005, -0.005 and -0.005. This is the case closest to the imposition of one cointegrating vector; 0.005 is subtracted from the three zero eigenvalues considered previously.

2. **Case 2**: -0.1, -0.01, -0.01 and -0.01.

3. **Case 3**: -0.1, -0.05, -0.05 and -0.05. This is the most stationary case.

Note that all the eigenvalues are negative to ensure the eigenvalues of $(\mathbf{\Pi} + \mathbf{I})$, i.e. the cofficient matrix in the corresponding VAR(1) model, are less than unity. A standard normal random matrix of eigenvectors, $\boldsymbol{E}$, is then constructed; the eigenvectors are almost surely linearly independent. These are transformed into an orthonormal basis, $\tilde{\boldsymbol{E}}$, using the Gram-Schmidt process. The coefficient matrix in the VECM is then given by $\tilde{\boldsymbol{E}} \boldsymbol{\Lambda} \tilde{\boldsymbol{E}}'$ where $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of the required coefficient matrix. Using this matrix and random normal disturbances generated by the GAUSS random number generator, nearly cointegrating series are then generated *via* nonparametric residual resampling from the VECM with a covariance matrix equal to that as estimated by a VECM(1) with one cointegrating vector imposed.

Using the simulated series for each case we then estimate: (i) a VAR with an intercept and its lag length fixed at the true value; (ii) a RRR, i.e. a VECM with an intercept where the lag length and cointegrating rank are fixed at their true values but the cointegrating vectors are data determined. This process is repeated 10,000 times. At each replication the parameters of the two models are estimated and the impulse responses due to a contractionary monetary policy shock calculated.

The results are shown in Figure 3. These figures focus on the median (of the 10,000 replications) impulse responses up to 50 periods out for each of the three cases. The true impulse responses (**TRUE**) from the DGP are also plotted; bias is any difference between TRUE, on the one hand, and VAR and RRR on the other.

As we should expect, the experiments clearly show that as we move from case 1 to case 3, that is as the data become increasingly stationary, the long horizon biases associated with estimation of a VAR model become less acute while those associated with the estimation of a reduced rank VAR increase. Conversely as we move from case 3 to case 1, the closer the data become to being cointegrated, the greater the biases at long horizons associated with estimation of a VAR model and the lesser the biases associated with estimation of a reduced rank VAR. But there is a point corresponding to "near" cointegration, here seen for case 2, when there are substantial biases at the long horizon involved with estimation both of a VAR and a reduced rank VAR. While the data may appear cointegrated according to standard tests, if there is "near" cointegration estimation of a reduced rank VAR, as well as a VAR in levels, delivers misleading inference about longer horizon impulse responses. Reliable inference can be conducted for short-horizon impulse response functions only using either levels or reduced rank VAR models. How one can in practice conduct reliable and robust inference about longer horizon impulse responses remains an open research area.

References

[1] Bernanke, B.S. and Gertler, M. (1995). 'Inside the black box: the credit channel of monetary policy transmission', *Journal of Economic Perspectives*, Vol. 9, pp. 27-48.

[2] Crowder, W.J., D.L. Hoffman and R.H. Rasche (1999), "Identification, long run relations and fundamental innovations in a simple cointegrated system", *Review of Economics and Statistics*, 81(1), 109-121.

[3] Johansen, S. (1988). 'Statistical analysis of cointegrating vectors', *Journal of Economic Dynamics and Control*, Vol. 12, pp. 231-54.

[4] Kim, S. and Roubini, N. (2000). 'Exchange rate anomalies in the industrial countries: a solution with a structural VAR approach', *Journal of Monetary Economics*, Vol. 45, pp. 561-586.

[5] King, R., C.I. Plosser, J.H. Stock and M.W. Watson (1991), "Stochastic trends and economic fluctuations", *American Economic Review*, 81, 819-40.

[6] Naka, A. and Tufte, D. (1997). 'Examining impulse response functions in cointegrated systems', *Applied Economics*, Vol. 29, pp. 1593-1603.

[7] Pesaran, M.H., Shin, Y. and Smith, R.J. (2000). 'Structural analysis of vector autoregressive models with exogenous $I(1)$ variables', *Journal of Econometrics*, Vol. 97, pp. 293-343.

[8] Phillips, P.C.B. (1998). 'Impulse response and forecast error variance asymptotics in nonstationary VARs', *Journal of Econometrics*, Vol. 83, pp. 21-56.

[9] Sims, C.A. (1992). 'Interpreting the macroeconomic time series facts: the effects of monetary policy', *European Economic Review*, Vol. 36, pp. 975-1000.

[10] Watson, M.W. (1994). 'Vector autoregressions and cointegration'. In *Handbook of Econometrics, Volume IV* (eds. Engle, R.F. and McFadden, D.L.), Elsevier Science B.V.

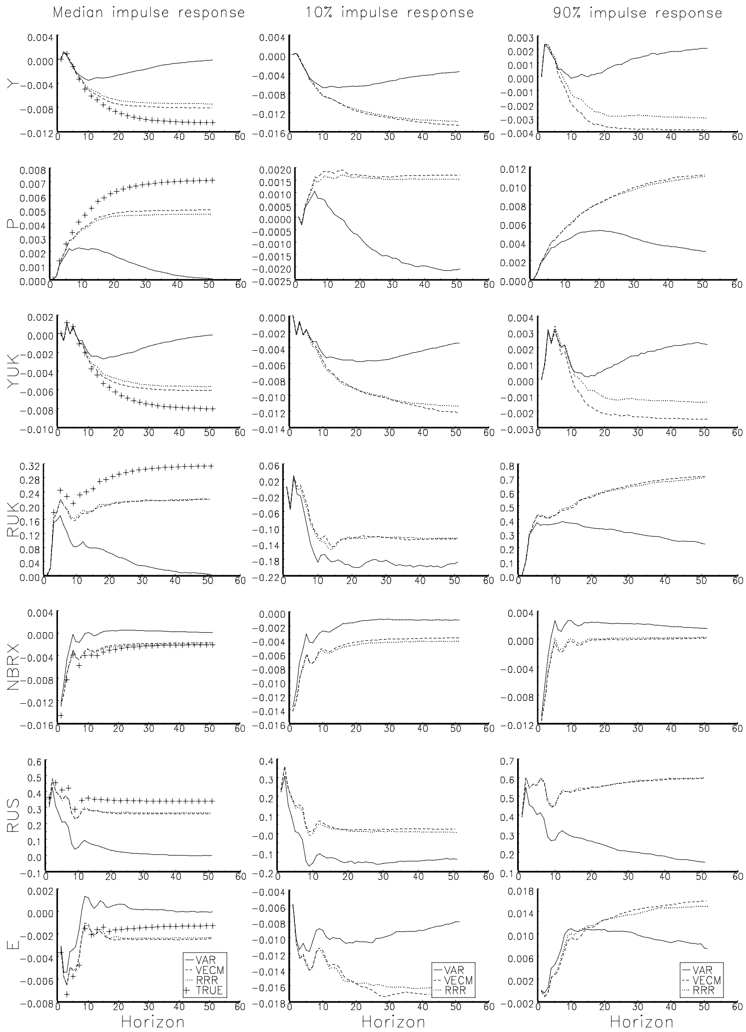

Figure 1: Model 1 - impulse responses due to a standard deviation contractionary monetary policy shock

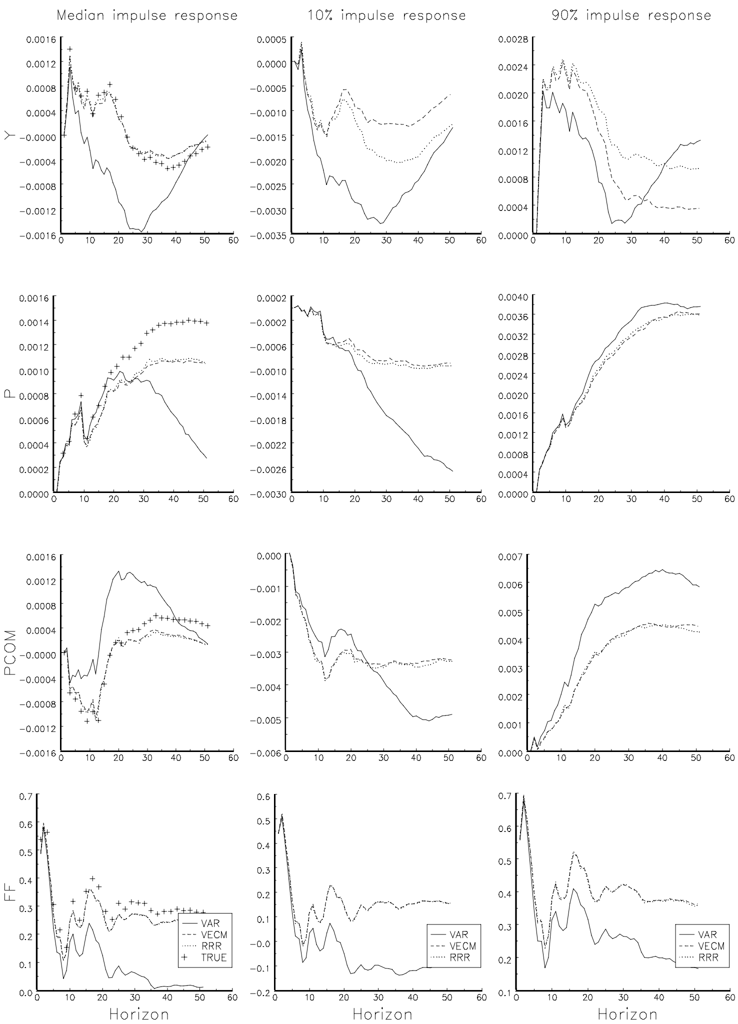

Figure 2: Model 2 - impulse responses due to a standard deviation contractionary monetary policy shock

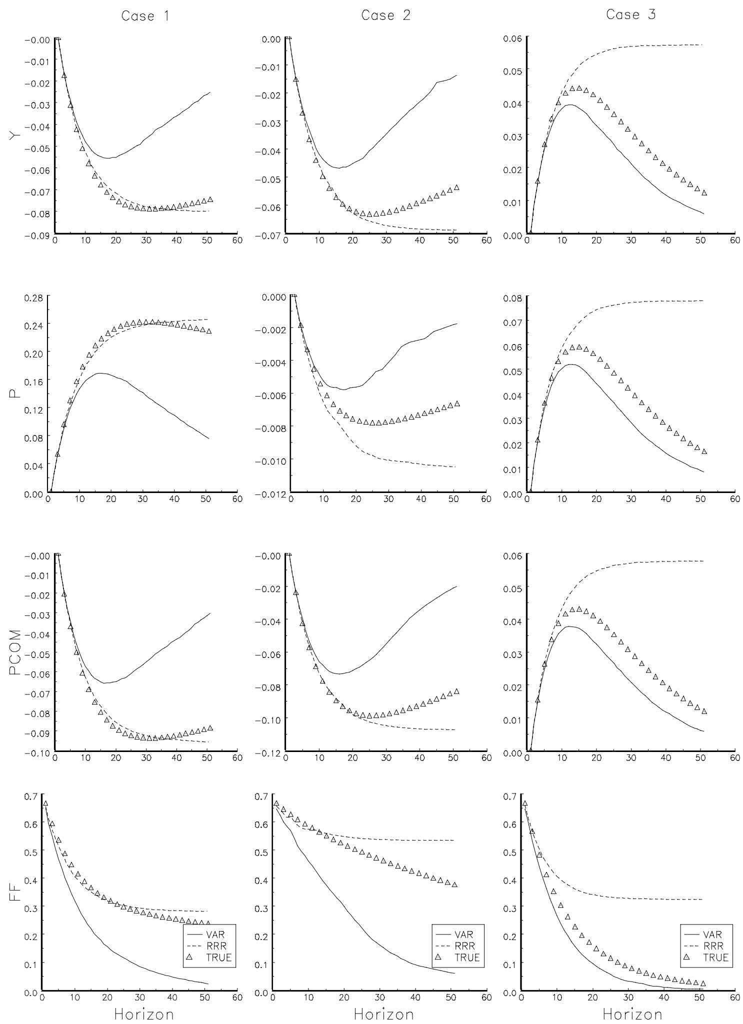

Figure 3: "Near" cointegration: impulse responses due to a standard deviation contractionary monetary policy shock