Working Paper 92-29
July 1992

División de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe
Fax: (341) 624-9849

POSTERIOR INFERENCE ON LONG-RUN IMPULSE RESPONSES

Gary Koop, Jacek Osiewalski and Mark F.J. Steel*

Abstract

This paper describes a Bayesian analysis of impulse response functions. We show how many common priors imply that posterior densities for impulse responses at long horizons have no moments. Our results suggest that impulse responses should be assessed on the basis of their full posterior densities, and that standard estimates such as posterior means, variances or modes may be very misleading.

Key words:
Persistence of shocks, prior sensitivity, ARIMA models, invertibility, existence of moments.

*Koop, Boston University; Osiewalski, Academy of Economics, Krakow, Poland; Steel, Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. The first author acknowledges the hospitality of the Center for Economic Research, Tilburg University. The second author received support from the Polish Committee for Scientific Research (KBN Grant number 1-0894-91-01) and benefitted from the hospitality of Universidad Carlos III, Madrid. The third author held a research fellowship of the Royal Netherlands Academy of Arts and Sciences (KNAW).

Introduction

To understand the dynamic properties of time series, macroeconomists frequently calculate impulse response functions, which describe the effect of an innovation on future values of a series. In classical analyses, point estimates of impulse responses are usually reported, but Runkle (1987) argues that these can be quite misleading because confidence intervals can be quite wide. With a few exceptions, notably Koop (1992) and DeJong and Whiteman (1991), few papers have analyzed impulse response functions from a Bayesian perspective. This is surprising since Bayesian techniques offer a natural framework for dealing with them.

Frequently, macroeconomists are interested in the properties of impulse responses at long horizons (see for instance, Campbell and Mankiw (1987) and Lee et al. (1992) and the references therein). However, we have demonstrated in a related paper (Koop, Osiewalski and Steel (1992), hereafter KOS) how Bayesian long-run forecasts can be very sensitive to apparently innocuous assumptions made in the prior. Building on our previous research, this paper will demonstrate how such sensitivity can also occur with impulse responses; specifically, we show that impulse responses at some forecast horizons may have density functions which have no finite moments and that standard Bayesian estimators based on squared error loss functions do not exist.

The remainder of the paper is organized as follows. Section 1 presents some theoretical results intended to illustrate how impulse responses at long horizons can be sensitive to seemingly innocent prior assumptions. Section 2 presents an empirical illustration and Section 3 concludes.

Section 1: Theoretical Results

Consider the following ARIMA(p,1,q) model[1]

---

[1]In this paper, we consider only univariate models. But similar results will hold for multivariate models such as VARs. Furthermore, the presence of nonstochastic terms as in Appendix A will not affect the form of the impulse response function, except insofar as it affects the posterior density of the parameters.

$$\Delta y_t=\beta+\alpha_1\Delta y_{t-1}+...+\alpha_p\Delta y_{t-p}+u_t \, , \tag{1}$$

where

$$u_t=\epsilon_t-\eta_1\epsilon_{t-1}-...-\eta_q\epsilon_{t-q}, \quad \epsilon_t \text{ are } i.i.N(0,\sigma_\epsilon^2).$$

y_t is observed for t=1,...,T and we have initial values $y_{(0)}=(y_{-p},...,y_0)'$. We assume that $\beta\in R$, $\alpha=(\alpha_1,..,\alpha_p)'$ is an element of the stationary region, C_p, and $\eta=(\eta_1,..,\eta_q)'$ is an element of the union of the invertibility region and the set of points allowing for a moving average (MA) root of one. That is, $\alpha\in C_p$, and $\eta\in C_q^*$ where $C_k=\{x\in R^k$: all roots, z_j (j=1,..,k), of $1-\Sigma x_j z^j=0$ lie outside the unit circle} and $C_k^*=\{x\in R^k$: all roots, z_j (j=1,..,k), of $1-\Sigma x_j z^j=0$ lie outside or on the unit circle}.

Using the following one-to-one reparameterization of (1): $(\rho\ \gamma_1...\gamma_{p-1})'=G_p(\alpha_1...\alpha_p)'$ and $(\lambda\ \phi_1...\phi_{q-1})'=G_q(\eta_1...\eta_q)'$ where G_i is an ixi matrix taking the form:

$$G_i=\begin{bmatrix}1 & 1 & .. & 1\\ 0 & -1 & .. & -1\\ . & . & ... & .\\ . & . & ... & .\\ 0 & 0 & .. & -1\end{bmatrix},$$

we can write (1) as

$$\begin{aligned}\Delta y_t=\beta+\rho\Delta y_{t-1}+\gamma_1\Delta^2 y_{t-1}+..+\gamma_{p-1}\Delta^2 y_{t-p+1}+\\ \epsilon_t-\lambda\epsilon_{t-1}-(\phi_1\Delta\epsilon_{t-1}+..+\phi_{q-1}\Delta\epsilon_{t-q+1}).\end{aligned} \tag{2}$$

If (1) has an MA unit root of one, then λ=1, and (1) collapses to a trend-stationary specification for y_t:

$$\begin{aligned}y_t=\psi+\mu t+u_t,\\ u_t=\alpha_1 u_{t-1}+..+\alpha_p u_{t-p}+\epsilon_t-\phi_1\epsilon_{t-1}-..-\phi_{q-1}\epsilon_{t-q+1}.\end{aligned} \tag{3}$$

In the structural model, (3), the deviations from the linear trend then follow a stationary ARMA(p,q-1) process, and the reduced form is:

$$\begin{aligned}y_t=\beta_0+\beta t+\alpha_1 y_{t-1}+..+\alpha_p y_{t-p}+\\ \epsilon_t-\phi_1\epsilon_{t-1}-..-\phi_{q-1}\epsilon_{t-q+1},\end{aligned} \tag{4}$$

where

$$\beta_0=\psi(1-\rho)+\mu\sum_{i=1}^{p} i\alpha_i=\psi(1-\rho)+\mu(\rho-\sum_{i=1}^{p-1}\gamma_i),$$
$$\beta=\mu(1-\rho).$$

Hence, introducing a constant, β, and allowing for an MA root of one in the ARIMA(p,1,q) model keeps open the possibility that trend-stationarity will be present.[2]

To calculate impulse response functions we assume that a shock of size δ has occurred in period i (usually i=T+1). Macroeconomists are typically interested in the effect of a unit shock (δ=1)[3] at time T+1 on the variable in time T+n, defined as:

$$I_{T+n}=y^*_{T+n}-y_{T+n}\ ,$$

where the notation implicitly assumes that the shock occurs in period T+1, and y_{T+n}^* is the observable in period T+n conditional on a shock of size 1 occurring (i.e. adding unity to ϵ_{T+1}).

As we show in Appendix A, the impulse response is given by

$$I_{T+n}=-\sum_{j=0}^{q}\eta_j\sum_{i=1}^{n}b_{i-j}, \qquad (5)$$

where η_0=-1 and b_h=0 for h≤0, b_h=1 for h=1, and $b_h=\alpha_1 b_{h-1}+..+\alpha_p b_{h-p}$ for h≥2. Note that the impulse response is an (n-1)th order polynomial in the α_is, and its posterior distribution depends entirely on the marginal posterior distribution of $\theta=(\alpha'\ \eta')'$.

Under stationarity, the limit of the sequence of impulse responses as the forecast horizon goes to infinity exists and can be written as:

$$I_{T+\infty}=\frac{1-\lambda}{1-\rho}. \qquad (6)$$

Note that the derivation of (6) does not require invertibility,

---

[2]Campbell and Mankiw (1987) work with deviations from the sample mean of Δy and impose β=0 in their model on page 861.

[3]For ARIMA(p,d,q) models impulse response functions are independent of the size of the shock (i.e. doubling δ will merely double the impulse response function). Note also that, instead of unit shocks, we could also consider shocks equal to σ_ϵ, as discussed in Appendix C.

$I_{T+\infty}$ depends only on λ and ρ, and is zero iff $\lambda=1$, regardless of the value for ρ. Since the model with $\lambda=1$ corresponds to the trend-stationary process (3), it needs special treatment (i.e. positive prior probability for this model requires prior mass over 1). Since stationarity implies $\rho<1$ and invertibility implies $\lambda<1$, $I_{T+\infty}$ is strictly positive for any $\theta \in C_p x C_q$. In this paper, we rule out MA roots inside the unit circle to ensure identification. Looking at (6) it is not immediately obvious that this is a harmless assumption. However, in Appendix C we show how impulse responses from the non-invertible model are, in a sense, equivalent to impulse responses from the corresponding invertible model.

As the following result shows, posterior distribution of $I_{T+\infty}$ need not have finite integer moments.

Theorem: Suppose that the conditional prior density of ρ given (η,γ) fulfills the condition $p(\rho|z_{(0)},\eta,\gamma)\geq c>0$ for $\rho\in(d,1)$ for some $d<1$. Then $E[(I_{T+\infty})^j|z,z_{(0)}]$ does not exist for $j\geq 1$. (z and $z_{(0)}$ are the data and the initial conditions, respectively. They are described in more detail in Appendix B.)

Proof: As is shown in Appendix B, the conditional posterior density of ρ given (η,γ) is proportional to the product $p(\rho|z_{(0)},\eta,\gamma)f_{S,1}(\rho|T-2,a,h)$, where a and h are positive functions of (μ,γ) (which can be obtained from the p-variate Student t factor in $p(\rho,\gamma|z,z_{(0)},\rho,\gamma))$, and $f_{S,r}$ indicates the r-variate Student t density. Let $f_{min}=\inf_{\rho\in(d,1)}p(\rho|z,z_{(0)},\eta,\gamma)$ and note that $f_{min}>0$ due to the assumption about the prior. $E[(I_{T+\infty})^j|z,z_{(0)},\eta,\gamma]$ can be written as the sum of two terms, one of which is non-negative and the other is:

$$(1-\lambda)^j\int_d^1(1-\rho)^{-j}p(\rho\,|\,z,z_{(0)},\eta,\gamma)\,d\rho\,. \tag{7}$$

The integral in (7) can be bounded from below by

$$f_{min}\int_d^1(1-\rho)^{-j}d\rho=f_{min}\int_0^{1-d}w^{-j}dw,$$

which diverges to $+\infty$ for $j\geq 1$. Since the conditional moments of $I_{T+\infty}$ do not exist, the unconditional moments also do not exist.□

It is worth emphasizing that the theorem holds for all prior structures for which $p(\rho|z_{(0)},\eta,\gamma)$ does not tend to zero as ρ tends to one. This implies, for example, that the posterior density of $I_{T+\infty}$ will have no moments if the prior is uniform over the stationary region, or is a Normal or Student t density truncated by the stationary region. However, for Beta priors for ρ given $(z_{(0)},\eta,\gamma)$ which have

$$\lim_{\rho\to 1^-} p(\rho|z_{(0)},\eta,\gamma)=0,$$

the existence of posterior moments of $I_{T+\infty}$ is possible. In particular, if the conditional prior for ρ is a Beta density with parameters r and s defined on (-a,1) with a>0, then the conditional jth order moment exists if $r\geq 1$ and $s>j$.

We assume in this paper that α lies in the stationary region. This assumption ensures that all moments exist for the posterior of I_{T+n} for finite n. If we allow for nonstationarity the problem of non-existence of moments is present even for finite horizons. For example, if we use an improper uniform prior that is not truncated to ensure stationarity, then $E[(I_{T+n})^j|z,z_{(0)},\eta]$ does not exist when $n\geq 1+(T-p-1)/j$ $(j=1,2,\ldots)$. A prior that puts more weight on the explosive region, such as Phillips' prior (see Phillips (1991)), would cause impulse responses at even smaller n to have no moments. In fact, for the ARIMA(1,1,0) model, Phillips' prior implies that the posterior of I_{T+n} has no integer moments for $n\geq 2$.[4]

In view of this lack of moments, we do not focus on the calculation of posterior means and variances. Instead, we advocate calculating the whole posterior density of $I_{T+\infty}$; or, for particular values of d,

$$p(I_{T+\infty}\geq d|Data)=p(1-\lambda\geq d(1-\rho)|Data). \quad (8)$$

Alternatively, highest posterior density or posterior coverage intervals can be calculated. All these features are by-products of Monte Carlo analysis of the posterior distribution of $\boldsymbol{\theta}$.

Special Case: ARIMA(1,1,0)

This case corresponds to (1) with $\alpha=\alpha_1=\rho$, $\alpha\in C_1=(-1,1)$, and

[4]The results in this paragraph are straightforward extensions of results given in KOS.

q=0. For this model, $b_h=\alpha^{h-1}$ for h≥1, and (5) reduces to

$$I_{T+n}=\sum_{i=0}^{n-1}\alpha^i=\frac{1-\alpha^n}{1-\alpha},$$

and

$$I_{T+\infty}=\frac{1}{1-\alpha}>\frac{1}{2},$$

which are in full agreement with (6). Note that, for d>0,

$$p(I_{T+\infty}\geq d|Data)=P(\alpha\geq 1-\frac{1}{d}|Data),$$

which gives a simple relationship between the quantiles of α and $I_{T+\infty}$. Note also that, for ARIMA(1,1,0) models, long-run impulse responses are bounded away from zero since $I_{T+\infty}>\frac{1}{2}$. Moving average errors are necessary if more flexibility is to be achieved, a point stressed by Campbell and Mankiw (1987).

If the prior for α is restricted to lie in the stationary region, then, for any finite n, all posterior moments of I_{T+n} exist. However, our theorem implies that the existence of posterior moments of $I_{T+\infty}$ depends crucially on the behavior of the prior for α near 1. It is worth emphasizing that for most macroeconomic examples (using differenced data), 1 will lie far out in the tail of the posterior. Hence the existence of posterior moments for $I_{T+\infty}$ depends on properties of the prior in apparently insignificant areas.

As far as posterior modal values are concerned, the analysis for $I_{T+\infty}$ is easier than for any finite horizon.

<u>Proposition</u>:
Under a uniform prior for α truncated to lie in (-1,1), the posterior of $I_{T+\infty}$ is proportional to an inverted Student t distribution truncated from below at $\frac{1}{2}$. This posterior density is either monotonically decreasing or has a unique mode in $(\frac{1}{2},\infty)$.

<u>Proof</u>:
If $p(\alpha|z_{(0)})=p(\alpha)=\frac{1}{2}I(-1<\alpha<1)$, then it follows from Appendix B that the posterior density of $1-\alpha$ is a truncated Student t density which is nonzero over the interval (0,2). If we denote the degrees of freedom, location and precision of the underlying untruncated Student t density by ν, a, and h respectively, then the posterior density of $I_{T+\infty}=(1-\alpha)^{-1}$ is

$$p(I_{T+\infty}|z_{(0)},z) = I_{T+\infty}^{-2} f_{S,1}(I_{T+\infty}^{-1}|\nu,a,h)\, I(I_{T+\infty} > \frac{1}{2}),$$

where I(.) is the indicator function. To find the mode of this posterior we take the first derivative and set it to zero. Details of this messy calculation are available on request. Ignoring the truncation, we can show that there are two solutions to the first order condition: one positive and one negative.[5] The $I(I_{T+\infty}>\frac{1}{2})$ term rules out the negative solution. It follows that, if the positive solution is greater than ½, it is a unique mode; otherwise $p(I_{T+\infty}|z_{(0)},z)$ is decreasing over $(\frac{1}{2},\infty)$ since the gradient can be shown to be negative after passing through its positive solution.□[6]

A Bayesian econometrician might be tempted to abandon the stationarity region and assume, for analytical convenience, an improper uniform prior on $\alpha \in R$. Although such a move may have little consequence for posterior inference on α itself, it would lead to posterior point mass at infinity for $I_{T+\infty}$, since $I_{T+\infty}=\infty$ for $\alpha \geq 1$.

Special Case: AR(1) plus trend

As demonstrated, the case of AR(1) deviations from a linear trend may be treated as the limit of an ARIMA(1,1,1) model as the MA parameter, $\eta=\eta_1=\lambda$, approaches one from the left. If we do not impose stationarity, but allow $\alpha=\alpha_1=\rho$ to be an element of $(-1,1+d)$ for $d>0$, then $I_{T+n}=\alpha^{n-1}$. This implies $I_{T+\infty}=0$ if $\alpha \in (-1,1)$, $I_{T+\infty}=1$ if $\alpha=1$, and $I_{T+\infty}=\infty$ if $\alpha>1$. Under any continuous prior for α, the posterior for I_{T+n} is also continuous for finite n; but $I_{T+\infty}$ has two point masses:

---

[5]This is just an illustration of Theorem 2.3 in Lehmann and Schaffer (1988) which states that, if the underlying distribution has both tails lighter than Cauchy, then the corresponding inverted distribution has at least one positive and one negative mode.

[6]The same reasoning could be applied to the conditional posterior of $I_{T+\infty}$ given (η,γ) in the general ARIMA(p,1,q) model. However, unimodality of the posterior may be lost when η and γ are integrated out.

$$p(I_{T+\infty}=0\,|\,Data)=P(|\alpha|<1\,|\,Data),$$
$$p(I_{T+\infty}=\infty\,|\,Data)=p(\alpha>1\,|\,Data)=p(\alpha\geq 1\,|\,Data).$$

If we put some prior mass at the point $\alpha=1$, then there is some posterior mass at this point, and the posterior of $I_{T+\infty}$ has one more mass point:

$$p(I_{T+\infty}=1\,|\,Data)=p(\alpha=1\,|\,Data).$$

Being simple, AR(1) deviations from a linear trend model can offer only crude answers to persistence questions at long forecast horizons. If stationarity is assumed, the answer is presupposed (no persistence at all). This is true for stationary ARMA deviations about a linear trend, and is the motivation for working with differenced data (i.e. starting with an ARIMA(p,1,q) model).

Section 2: Empirical Illustration

In this section we illustrate some of our theoretical findings using quarterly, seasonally adjusted real U.S. GNP from 1947:4 to 1989:4.[7] In order to illustrate our techniques, we focus on two specifications: ARIMA(1,1,0) and the ARIMA(1,1,3). We choose the ARIMA(1,1,0) since it is computationally easy (it has no MA component) and receives a good deal of support from the data. The ARIMA(1,1,3) poses more computational problems[8] and the likelihood exhibits some interesting properties. By including the ARIMA(1,1,0) and the ARIMA(1,1,3) models we are covering both a "well-behaved" case and a "poorly-behaved" case. In all cases, we use a flat prior on θ over the stationarity and invertibility regions; hence our Theorem implies that no posterior moments exist for $I_{T+\infty}$.

[7]We use Citibase data series GNP82. Our data only differs from that used Campbell and Mankiw (1987) and DeJong and Whiteman (1991) in that we include data for 1986 through 1989 which was unavailable to the previous authors.

[8]We use Kalman filtering techniques to deal efficiently with the moving average covariance matrix described in Appendix B. Hence, we follow Campbell and Mankiw and use exact methods for dealing with the moving average component.

For the ARIMA(p,1,3) models, Campbell and Mankiw find that the likelihood function is bimodal, and that there is evidence for a moving average root of unity. The authors also find that results for these models differ radically from any of the others they consider. DeJong and Whiteman criticize Campbell and Mankiw for ignoring the evidence from the ARIMA(p,1,3) models and argue that the evidence in favor of shocks being transitory is much stronger if these models are included. DeJong and Whiteman's priors differ from ours in that they are either flat over $\boldsymbol{\theta}$ or flat over the impulse responses themselves. However, the authors do not impose invertibility since they obtain negative lower bounds for 95% coverage intervals for the ARIMA(1,1,3) case. Furthermore, the second prior rule they adopt implies a different prior on θ for each forecast horizon, n.

Table 1 and Figure 1 present results for the ARIMA(1,1,0) model. Table 1 presents moment-based quantities, while Figure 1 plots the posterior for $I_{T+\infty}$. Posterior means and standard deviations for this simple case are similar to those given in Campbell and Mankiw's Table IV. Also, their maximum likelihood estimate of I_{T+81} is 1.571 which is very close to our posterior mode of $I_{T+\infty}$. However, there is some evidence of excess skewness and kurtosis. Figure 1 indicates these departures from Normality. Finally, as expected from our Proposition, the posterior density of $I_{T+\infty}$ is unimodal.

Given our prior, it is important to note that no posterior moments exist for $I_{T+\infty}$ but that changing the prior so that ρ no longer comes arbitrarily close to one, leads to all posterior moments existing. We reran our Monte Carlo integration program using a flat prior for ρ on [-.9999,.9999] and found that the moments for $I_{T+\infty}$ were identical to those of I_{T+81} to 3 decimal places. This finding indicates that care should be taken not only when selecting priors, but when using Monte Carlo integration as well. Since the Monte Carlo procedure cannot tell the difference between intervals (-1,1) and [-.9999,.9999] with a finite number of draws, a naive application of the procedure could lead a researcher to report moments of $I_{T+\infty}$ when no such moments exist.

Table 2 and Figure 2 plot results for the ARIMA(1,1,3) model. Table 2 indicates a great deal of excess skewness and kurtosis, proving that merely reporting means and standard deviations may be highly misleading. Figure 2 indicates that at an infinite forecast horizon, the posterior of $I_{T+\infty}$ is very non-Normal, viz. multimodal and very skewed with fat tails. One of the modes of our posterior of $I_{T+\infty}$ is close to Campbell and Mankiw's maximum likelihood estimate of I_{T+81} (.026), while another mode is around .85.

When commenting on the ARIMA(p,1,3) models, Campbell and Mankiw (1987, p. 866) remark: "These models have a second peak in the likelihood function with no moving average root and almost the same likelihood... The impulse response function for the second peak is similar to that for lower-order models." The Bayesian approach takes <u>all</u> the evidence in the likelihood function into account and, thus, unlike Campbell and Mankiw, we find that results for the ARIMA(1,1,3) and the ARIMA(1,1,0) models given in Tables 1 and 2 are not very different from each other. In addition, formally treating model uncertainty by mixing over models as in DeJong and Whiteman (1991) could serve to strengthen the evidence in favor of persistence of shocks.

Note that our results for the ARIMA(1,1,3) model also differ substantially from those reported in DeJong and Whiteman, albeit for entirely different reasons. Using our results, a 95% coverage interval for $I_{T+\infty}$ is approximately (0,3.0), whereas DeJong and Whiteman find (-4.5,2.5) for I_{T+81}. Comparing their Figure 1 with our Figure 2 indicates that the difference in our results is due to the fact that DeJong and Whiteman do not impose invertibility. As we show in Appendix C, one cannot compare impulse responses to unit shocks across invertible and non-invertible models.

Although the empirical evidence from both models considered clearly points towards persistence of shocks at an infinite horizon[9], the degree of persistence differs. Whereas $p(I_{T+\infty}\geq 1|\text{Data})$ is 1.0 for the ARIMA(1,1,0) case, it is only .43 for the ARIMA(1,1,3) case.

---

[9]Of course, we impose $I_{T+\infty}>\frac{1}{2}$ in the ARIMA(1,1,0) model, a restriction which is relaxed to $I_{T+\infty}\geq 0$ whenever MA terms are included in the model.

Section 3: Conclusions

In this paper we have used a Bayesian approach to analyze impulse response functions. For ARIMA(p,1,q) models, we derive a simple expression for $I_{T+\infty}$, i.e. the effect of a unit shock in the underlying white noise process on the level of the observable at an infinite horizon. We provide a condition sufficient to preclude the existence of posterior moments of $I_{T+\infty}$, and note that this condition relates to seemingly innocuous properties of the prior. Given the MA parameters and the last p-1 AR parameters, we prove unimodality of the posterior of $I_{T+\infty}$, which is apparently lost upon integrating out the remaining parameters. On the basis of these theoretical results, which we illustrate in an empirical example, we stress that the entire posterior density of $I_{T+\infty}$ should be considered.

We also provide a theoretical justification for imposing invertibility. The relevance of this justification is shown by comparing our empirical results with those obtained by DeJong and Whiteman (1991). We base our empirical example on Campbell and Mankiw (1987), and discuss the consequences of adopting the Bayesian paradigm as opposed to the maximum likelihood approach adopted by Campbell and Mankiw.

In conclusion, and in contrast to the above studies, we find fairly consistent evidence of persistence of shocks for real U.S. GNP.

References

Campbell, J. and Mankiw, N. (1987). "Are Output Fluctuations Transitory?", *Quarterly Journal of Economics*, 102, 857-880.

DeJong, D. and Whiteman, C. (1991). "Are Output Fluctuations Transitory?: A Bayesian Perspective", manuscript.

Koop, G. (1992). "Aggregate Shocks and Macroeconomic Fluctuations: A Bayesian Approach", manuscript.

Koop, G., Osiewalski, J. and Steel, M.F.J. (1992). "Bayesian Long-Run Prediction in Time Series Models", Working Paper 9210, Universidad Carlos III, Madrid.

Lee, K.C., Pesaran, M.H. and Pierse, R.G. (1992). "Persistence of Shocks and their Sources in a Multisectoral Model of U.K. Output-Growth", *Economic Journal*, 102, 342-356.

Lehmann, E. and Popper Schaffer, J. (1988). "Inverted Distributions", *The American Statistician*, 42, 191-194.

Phillips, P. (1991). "To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends", *Journal of Applied Econometrics*, 6, 333-364.

Runkle, D. (1987). "Vector Autoregressions and Reality" (with discussion), *Journal of Business and Economic Statistics*, 5, 437-454.

Appendix A: Calculation of Impulse Responses

The ARIMA(p,1,q) model with a deterministic component ($x_t'\beta$) can be written as the following vector process:

$$\omega_t=\mu_t+A\omega_{t-1}+R\xi_t, \qquad \textbf{(A.1)}$$

where

$$\omega_t=\begin{bmatrix}\Delta y_t\\ .\\ .\\ .\\ \Delta y_{t-m+1}\end{bmatrix},\quad \mu_t=\begin{bmatrix}x_t'\beta\\ 0\\ .\\ .\\ 0\end{bmatrix},\quad \xi_t=\begin{bmatrix}\epsilon_t\\ .\\ .\\ .\\ \epsilon_{t-m+1}\end{bmatrix},$$

$$A=\left[\begin{array}{c|c}\alpha_1\ldots\alpha_{m-1} & \alpha_m\\ \hline I_{m-1} & 0\end{array}\right],\quad R=\left[\begin{array}{c}\eta'\\ \hline 0\end{array}\right],$$

$$\eta=-\begin{bmatrix}\eta_0\\ \eta_1\\ .\\ .\\ \eta_{m-1}\end{bmatrix},$$

$m=\max\{p,q+1\}$, $\epsilon_t\sim i.i.N(0,\sigma_\epsilon^2)$, $\alpha_i=0$ for $i>p$, $\eta_i=0$ for $i>q$, and $\eta_0=-1$. Note that

$$\omega_{T+n}=A^n\omega_T+\sum_{i=1}^{n}A^{n-i}(\mu_{T+i}+R\xi_{T+i}),\quad (n=1,2,\ldots). \qquad \textbf{(A.2)}$$

Let e_i (i=1,..,m) denote the mx1 vector with 1 in the ith position and zeros elsewhere, and define

$$\begin{aligned}\xi^*_{T+i}&=\xi_{T+i}+e_i,\quad (i=1,\ldots,m),\\ &=\xi_{T+i},\quad (elsewhere),\end{aligned}$$

and

$$\omega^*_{T+n}=A^n\omega_T+\sum_{i=1}^{n}A^{n-i}(\mu_{T+i}+R\xi^*_{T+i}). \qquad \textbf{(A.3)}$$

By defining s=min{m,n} and observing that $Re_i=-\eta_{i-1}e_1$, we can write

$$\omega^*_{T+n}-\omega_{T+n}=\sum_{i=1}^{s} A^{n-i}Re_i=-\sum_{i=1}^{s} \eta_{i-1}A^{n-i}e_1 .$$

A simple inductive reasoning shows that the first column of A^{j-1}, $A^{j-1}e_1$, $(j=1,2,...)$, takes the form

$$A^{j-1}e_1=(b_j \; b_{j-1} \ldots b_{j-m+1})', \tag{A.4}$$

where $b_h=0$ for $h\leq 0$, $b_1=1$, and $b_h=\alpha_1 b_{h-1}+..+\alpha_p b_{h-p}$ for $h\geq 2$. Thus

$$\begin{aligned}\Delta y^*_{T+n}-\Delta y_{T+n}&=e_1'\omega^*_{T+n}-e_1'\omega_{T+n}\\ &=-\sum_{i=1}^{s} \eta_{i-1}e_1'A^{n-i}e_1=-\sum_{j=0}^{q} \eta_j b_{n-j},\end{aligned} \tag{A.5}$$

and, since $y^*_T=y_T$,

$$I_{T+n}=y^*_{T+n}-y_{T+n}=\sum_{i=1}^{n} (\Delta y^*_{T+i}-\Delta y_{T+i}) =-\sum_{j=0}^{q} \eta_j\sum_{i=1}^{n} b_{i-j} . \tag{A.6}$$

Equation (A.6) shows how the impulse response at forecast horizon n can easily be calculated recursively. The existence of $I_{T+\infty}$ (the limit of I_{T+n} as n approaches ∞) can be shown using the properties of A since, using (A.4), we can write (A.5) and (A.6) in the form

$$\Delta y^*_{T+n}-\Delta y_{T+n}=\eta'A^{n-1}e_1,$$

and

$$I_{T+n}=\sum_{i=1}^{n} \eta'A^{i-1}e_1=\eta'(\sum_{i=0}^{n-1} A^{i}) e_1 .$$

It is well known that ΣA^i converges iff all eigenvalues of A are less than one in absolute value, which is implied by stationarity. Therefore, under stationarity,

$$I_{T+\infty}=\lim_{n\to\infty}I_{T+n}=-\sum_{j=0}^{q} \eta_j\sum_{i=1}^{\infty} b_i \tag{A.7}$$

is finite.

Note also that

$$\sum_{i=1}^{\infty} b_i=1+\sum_{i=2}^{\infty} b_i=1+\alpha_1\sum_{i=1}^{\infty} b_i+\ldots+\alpha_p\sum_{i=1}^{\infty} b_i,$$

and thus

$$\sum_{i=1}^{\infty} b_i = \frac{1}{1-\rho},$$

where we have made use of the fact that, under stationarity, $\rho=\alpha_1+\ldots+\alpha_p<1$. Finally,

$$I_{T+\infty} = \frac{1-\lambda}{1-\rho},$$

where $\lambda=-(\eta_0+\ldots+\eta_q)=1-(\eta_1+\ldots+\eta_q)$.

Appendix B: Marginal Posterior of $\theta=(\alpha',\eta')'$

Let

$$z_t = x_t'\beta + \alpha_1 z_{t-1} + \ldots + \alpha_p z_{t-p} + u_t,$$
$$u_t = \epsilon_t - \eta_1\epsilon_{t-1} - \ldots - \eta_q\epsilon_{t-q},$$
$$\epsilon_t \sim i.i.N(0,\sigma_\epsilon^2),$$

or, in matrix notation

$$z = Z_-\alpha + X\beta + u, \quad u \sim N(0,\sigma_\epsilon^2 V_\eta),$$

where $z=(z_1\ldots z_T)'$, $X=(x_1\ldots x_T)'$, Z_- groups the appropriate lagged values of z_t, $\alpha\in C_p$, $\eta\in C_q^*$, $\beta\in R^k$, and V_η is a TxT matrix with elements given by

$$v_{ij} = \begin{bmatrix} 1+\eta_1^2+\ldots+\eta_q^2, \quad if\ i=j \\ -\eta_\tau+\eta_1\eta_{\tau+1}+\ldots+\eta_{q-\tau}\eta_q, \quad \tau=|i-j|=1,\ldots,q \\ 0, \quad if|i-j|>q. \end{bmatrix}$$

The model given in (1) in the body of the text has $z_t=\Delta y_t$, k=1, and X is a column vector of ones.

Under the prior structure

$$p(\beta,\theta,\sigma_\epsilon^2|z_{(0)}) = p(\beta)p(\sigma_\epsilon^2)p(\theta|z_{(0)}) \propto p(\theta|z_{(0)})\sigma_\epsilon^{-2},$$

where $\beta\in R^k$, $\sigma_\epsilon^2\in R_+$, $\theta\in C_p x C_q^*$, and $z_{(0)}=(z_{1-p}\ldots z_0)'$, the marginal posterior density of θ is

$$p(\theta|z_{(0)},z) \propto (|V_\eta||X'V_\eta^{-1}X|)^{-\frac{1}{2}} p(\theta|z_{(0)})[(z-Z_-\alpha)'M_\eta(z-Z_-\alpha)]^{-\frac{T-k}{2}},$$

where

$$M_\eta = V_\eta^{-1} - V_\eta^{-1}X(X'V_\eta^{-1}X)^{-1}X'V_\eta^{-1};$$

see KOS for details. In particular, the conditional density of α given η is

$$p(\alpha \mid z, z_{(0)}, \eta) = K_\eta^{-1} p(\alpha \mid z_{(0)}, \eta)$$
$$f_{S,p}(\alpha \mid T-k-p, \hat{\alpha}_\eta, \frac{T-k-p}{s_\eta^2} Z_-' M_\eta Z_-),$$

where,

$$\hat{\alpha}_\eta = (Z_-' M_\eta Z_-)^{-1} Z_-' M_\eta z,$$
$$s_\eta^2 = (z - Z_- \hat{\alpha}_\eta)' M_\eta (z - Z_- \hat{\alpha}_\eta),$$

and K_η is a normalizing constant. Note that this density is nonzero over C_p only.

Consider the linear transformation $g: (\alpha_1, \ldots, \alpha_p) \rightarrow (\rho, \gamma_1, \ldots, \gamma_{p-1}) = (\rho \quad \gamma')$. The conditional posterior for these transformed parameters is

$$p(\rho, \gamma \mid z, z_{(0)}, \eta) = K_\eta^{-1} p(\rho, \gamma \mid z_{(0)}, \eta)$$
$$f_{S,p}(\begin{bmatrix}\rho \\ \gamma\end{bmatrix} \mid T-k-p, G_p \hat{\alpha}_\eta, \frac{T-k-p}{s_\eta^2} G_p^{-1} Z_-' M_\eta Z_- G_p^{-1}),$$

where G_p is the transformation matrix corresponding to g. The density is nonzero over $g(C_p)$. Now it is clear that the conditional density for ρ given $(z, z_{(0)}, \eta, \gamma)$ is proportional to the product of some univariate Student t factor, and the conditional prior density for ρ, which imposes a truncation in the right tail at $\rho=1$.

Appendix C: Impulse Responses and Invertibility

In this Appendix we illustrate, for the MA(1) case, the equivalence of impulse responses from invertible and noninvertible models.

Consider a stationary stochastic process, u_t $(t=0,\pm1,\pm2,\ldots)$, with zero mean, finite variance, σ^2, and autocorrelation function: $\rho_0=1$, $0<|\rho_1|\leq\frac{1}{2}$, $\rho_\tau=0$, $\tau\geq2$. This process has two MA(1) representations:

$$u_t = \epsilon_t - \eta\epsilon_{t-1}, \quad \epsilon_t \sim i.i.N(0, \sigma_\epsilon^2), \tag{C.1}$$

and

$$u_t = \xi_t - \theta\xi_{t-1}, \quad \xi_t \sim i.i.N(0, \sigma_\xi^2), \tag{C.2}$$

where

$$\sigma^2 = \sigma_\epsilon^2(1+\eta^2) = \sigma_\xi^2(1+\theta^2), \tag{C.3}$$

and

$$\rho_1 = -\frac{\eta}{1+\eta^2} = -\frac{\theta}{1+\theta^2}. \qquad \text{(C.4)}$$

Note that (C.4) is satisfied iff either $\theta=\eta$ or $\theta=1/\eta$, and therefore (C.1) and (C.2) coincide for $\theta=\eta=1$ or $\theta=\eta=-1$, but are different when $0<|\eta|<1$ and $\theta=1/\eta$.

Multiplying (C.4) and (C.3) gives

$$\sigma^2\rho_1 = -\sigma_\epsilon^2\eta = -\sigma_\xi^2\theta,$$

and thus, for $\theta=1/\eta$, we obtain

$$\sigma_\xi = |\eta|\sigma_\epsilon. \qquad \text{(C.5)}$$

Hence, if (C.1) is invertible, then the standard deviation of ξ_t in (C.2) is smaller than the standard deviation of ϵ_t in (C.1).

Since the invertible representation, (C.1), involves a different white noise process than the noninvertible representation, (C.2), adding 1 to ξ_t implies a very different magnitude of shock than adding 1 to ϵ_t. Impulse responses calculated for (C.1) and (C.2) using (6) will equal $1-\eta$ and $1-\theta$, respectively. However, these impulse responses are not comparable since they are based on different sized shocks. If we consider a shock of one standard deviation to (C.1) and (C.2), we obtain

$$I_{T+n}(\sigma_\epsilon) = \sigma_\epsilon(1-\eta) = \frac{1}{\eta}\sigma_\xi - \sigma_\xi = -\sigma_\xi(1-\theta) = I_{NI,T+n}(-\sigma_\xi),$$

if $0<\eta<1$, and

$$I_{T+n}(\sigma_\epsilon) = \sigma_\epsilon(1-\eta) = -\frac{1}{\eta}\sigma_\xi + \sigma_\xi = \sigma_\xi(1-\theta) = I_{NI,T+n}(\sigma_\xi),$$

if $-1<\eta<0$. The notation in the previous equations makes clear that we are considering shocks of one standard deviation, not shocks of size one as in the body of the text, and $I_{NI,T+n}(\cdot)$ denotes the impulse response calculated using the non-invertible model.

Thus, a shock of one standard deviation added to ϵ_t in (C.1) has the same consequences as a shock of minus (if $\eta,\theta>0$), or plus (if $\eta,\theta<0$) one standard deviation added to ξ_t in (C.2).

Table 1: Posterior Moments of Impulse Responses for ARIMA(1,1,0)

n-1***	Mean	St. Dev.	Skewness*	Kurtosis**
1	1.370	.072	.022	.059
2	1.511	.126	.073	.222
4	1.591	.174	.148	.484
8	1.607	.192	.182	.609
16	1.608	.193	.185	.620
20	1.608	.193	.185	.620
40	1.608	.193	.185	.620
80	1.608	.193	.185	.620
∞	∞	∞	∞	∞

*Skewness is measured as $E(x^3)-E(x)^3$.
**Kurtosis is measured as $E(x^4)-E(x)^4$.
*** Note that our labelling convention for forecast horizons differs by one from that used in Campbell and Mankiw (1987) (i.e. in our paper $I_{T+1}=1$ by definition).

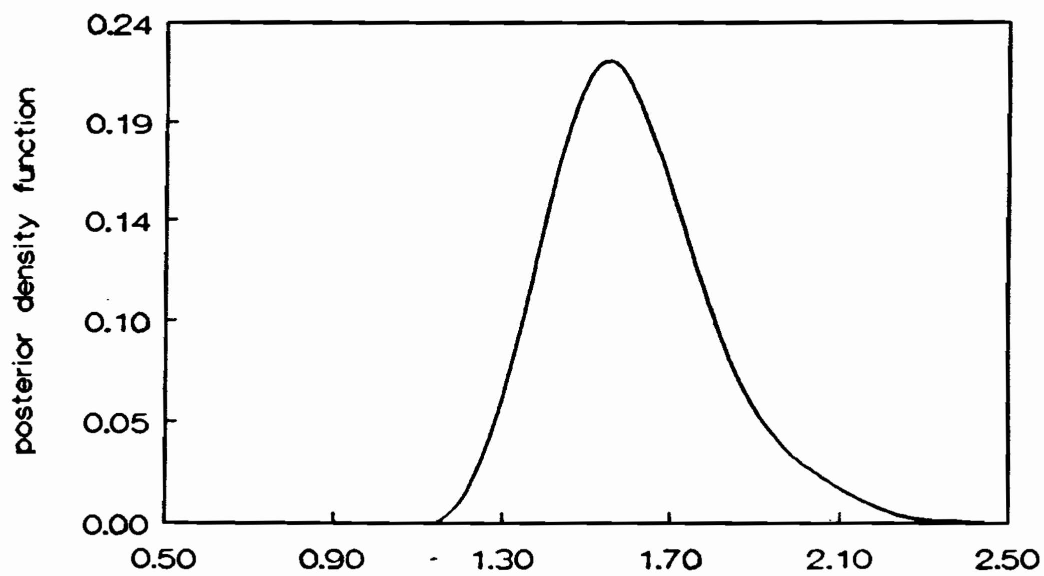
Table 2: Posterior Moments of Impulse Responses for ARIMA(1,1,3)

n-1	Mean	St. Dev.	Skewness*	Kurtosis**
1	.796	.519	.725	1.498
2	1.082	.506	.902	2.314
4	1.101	.470	.785	2.010
8	1.128	.430	.658	1.680
16	1.153	.420	.622	1.589
20	1.157	.429	.651	1.678
40	1.158	.487	.884	2.472
80	1.160	.574	1.445	5.658
∞	∞	∞	∞	∞

*Skewness is measured as $E(x^3)-E(x)^3$.
**Kurtosis is measured as $E(x^4)-E(x)^4$.

IMPULSE RESPONSE AT INFINITY

ARIMA(1,1,0)



IMPULSE RESPONSE AT INFINITY

ARIMA(1,1,3)