DISCUSSION PAPER SERIES

No. 4536
CEPR/EABCN No. 7/2004

SMALL SAMPLE CONFIDENCE INTERVALS FOR MULTIVARIATE IMPULSE RESPONSE FUNCTIONS AT LONG HORIZONS

Elena Pesavento and Barbara Rossi

INTERNATIONAL MACROECONOMICS

€ABCN
Euro Area Business Cycle Network
www.eabcn.org

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP4536.asp

ISSN 0265-8003

SMALL SAMPLE CONFIDENCE INTERVALS FOR MULTIVARIATE IMPULSE RESPONSE FUNCTIONS AT LONG HORIZONS

Elena Pesavento, Emory University
Barbara Rossi, Duke University

Discussion Paper No. 4536
September 2004

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 4536

September 2004

ABSTRACT

Small Sample Confidence Intervals for Multivariate Impulse Response Functions at Long Horizons*

Existing methods for constructing confidence bands for multivariate impulse response functions depend on auxiliary assumptions on the order of integration of the variables. Thus, they may have poor coverage at long lead times when variables are highly persistent. Solutions that have been proposed in the literature may be computationally challenging. The goal of this Paper is to propose a simple method for constructing confidence bands for impulse response functions that is not pointwise and that is robust to the presence of highly persistent processes. The method uses alternative approximations based on local-to-unity asymptotic theory and allows the lead time of the impulse response function to be a fixed fraction of the sample size. These devices provide better approximations in small samples. Monte Carlo simulations show that our method tends to have better coverage properties at long horizons than existing methods. We also investigate the properties of the various methods in terms of the length of their confidence bands. Finally, we show, with empirical applications, that our method may provide different economic interpretations of the data. Applications to real GDP and to nominal versus real sources of fluctuations in exchange rates are discussed.

JEL Classification: C12, C32 and F40
Keywords: impulse response functions, local to unity asymptotics, persistence and VARs

Elena Pesavento
Department of Economics
Emory University
Atlanta GA 30322-2240
USA
Email: epesave@emory.edu

Barbara Rossi
Department of Economics
Duke University
305 Social Sciences Building
Durham, NC 27708-0097
USA
Email: brossi@econ.duke.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=159908

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157554

*This Paper is funded by the Euro Area Business Cycle Network (www.eabcn.org). This Network provides a forum for the better understanding of the euro area business cycle, linking academic researchers and researchers in central banks and other policy institutions involved in the empirical analysis of the euro area business cycle. We gratefully acknowledge discussions with F Altissimo, T Bollerslev, G Elliott, R Gallant, L Kilian, D Osborn, A Pagan, A Tarozzi, G Tauchen, M Watson and participants to the University of Houston/Rice, Virginia Tech, the Financial Econometrics and Macroeconomics seminars at Duke, the 2003 NBER-NSF Conference on Time Series Analysis, the 2003 EABCN Conference, the 2004 Conference of the Society for Non-linear Dynamics and Econometrics, and the 2004 Winter Meetings of the Econometric Society. All mistakes are ours.

Submitted 16 September 2004

1. Introduction

Impulse response functions (IRFs) play an important role in describing the impact that shocks have on economic variables and their propagation mechanisms. They are generally obtained from Vector Autoregressions (VAR) and commonly used to analyze the response of current and future values of economic variables to a one standard deviation increase in the current value of the VAR identified shocks. The estimate of the IRFs and their confidence intervals are commonly based on Lütkepohl (1990) asymptotic normal approximation or bootstrap approximations to that distribution (see Kilian (1998a, b)).

Existing methods for constructing IRFs and their confidence intervals, however, depend on auxiliary assumptions on the order of integration of the variables. In fact, they may provide very different results depending on whether the series are assumed to be stationary, exactly integrated or exactly cointegrated. In addition, confidence bands may have poor coverage properties in small samples in the presence of highly persistent variables, as shown in Kilian and Chang (2000). These authors compare finite-sample accuracy and length of commonly used confidence intervals of IRFs calibrated on existing macroeconomic studies and caution applied researchers against inference at horizons bigger than 16 quarters, as inference becomes very unreliable. Unit root pretests do not solve the problem, as the actual coverage of IRF bands obtained after a pretest can be quite different from the nominal one. The problem is empirically relevant, as most macroeconomic time series are highly persistent (see e.g. Stock (1991, 1996)).

The goal of this paper is to propose a method for constructing confidence bands for IRFs that is robust to the presence of highly persistent processes, that theoretically (asymptotically) does not depend on the VAR lag length in the long-run, and that takes into account sampling variability in an appropriate way. We do so by using alternative asymptotic approximations based on local-to-unity asymptotic theory and allowing the lead time of the IRF to be a fixed fraction of the sample size. These devices provide better approximations in small samples. The advantages of our method are that: *(i)* it does not require a researcher to decide whether the process has a unit root or not; *(ii)* it is easy to compute (even in multivariate setups, if the uncertainty is about one root only); *(iii)* the confidence bands asymptotically contain the *whole* true IRF with a pre-specified confidence level (i.e. they are not pointwise); *(iv)* it is robust to the presence of deterministic components; *(v)* it allows also mildly non-stationary processes (e.g. roots about 1.001). Due to the nature of our approximation, our confidence bands are appropriate at long lead times and as long as the process is highly persistent (including a unit root). How long the lead time in practice has to be in order for our approximation to be accurate will be investigated in a Monte Carlo experiment.

The empirical literature commonly estimates VARs either in levels or in first differences, sometimes after unit root pretest procedures. The results are usually sensitive to the order of integration of the economic variables, and ad-hoc robustness checks are routinely performed to verify the empirical conclusions. That is, researchers verify that the same results hold whether one uses a specification in levels or in first differences. However, the latter method will not give any indication of the overall size of the procedure and, as we show, pretests create considerable size distortions in confidence bands for IRFs. This paper not only shows how bad size distortions researchers actually face by applying IRF based on pretests, levels or first differences (see also Ashley and Verbrugge (2001)), but it also constructively indicates when these distortions are more likely. The paper also shows when and by how much our procedure improves over these methods. Our method in general works fine for horizons bigger than or equal to 10. Depending on the persistence, our method performs better than pretest-based IRFs also for smaller horizons. We also discuss a simple way for constructing conservative confidence bands that provides reliable (although conservative) inference at shorter horizons too.

Alternative methods are available in the literature. Andrews (1993) and Andrews and Chen (1994) provide bias-corrected parameter estimates for obtaining IRFs for univariate time series. Their method is an important improvement over normal sampling methods, but the coverage is poor at long lead times and it is computationally demanding. Another available method is Hansen (1999) grid bootstrap method. However, to date there is no extension of the aforementioned methods to deal with multivariate processes. Kilian (1998a) provides a useful, improved bias-corrected bootstrap method that explicitly accounts for the bias and skewness of the small-sample distribution of the IRF estimator. Sims and Zha (1999) propose an alternative Bayesian method that characterizes the shape of the likelihood. However, these methods may not be robust to the presence of highly persistent processes (roots equal to one or mildly explosive) or deterministic terms, and it is important to investigate whether there are alternative methods that can provide better coverage or smaller length of the confidence bands. A recent solution has been proposed by Gospodinov (2002). His method relies on the inversion of a likelihood ratio test in which the constrained estimate exploits a null hypothesis on the value of the IRF at some horizon of interest. His method has the correct size pointwise and confidence bands have small length. Compared to his, the method proposed in this paper is much less computationally intensive and it is not pointwise; on the other hand, we rely explicitly on long horizon asymptotics so our method might have worse size properties at short horizons.

As empirical applications, we revisit the effects of a shock to real GDP and the nominal versus real sources of fluctuations in exchange rates. We find that shocks are more persistent than commonly found in VARs estimated in levels due to the downward bias estimate of the largest root. Thus, in Eichenbaum and Evans (1995), shocks would seem to disappear more quickly than they really do. On the other hand, in some other interesting empirical applications where pretests were used, the processes might have been incorrectly identified as I(1), due to the low power of unit root tests. This ends up over-estimating the long run effects of the shocks. This would happen, for example, in Lastrapes (1992) and Clarida and Gali (1994).

2. Motivation and Preview of the Results

Consider a researcher interested in analyzing whether monetary shocks have an effect on the real exchange rate. An answer to this question is relevant, as it would provide an empirical contribution to the longstanding debate on flexible versus sticky price models of exchange rate determination. This question has been analyzed by Christiano et al. (1995) and Rogers (1999), among others. Researchers working on this topic typically run VARs to estimate IRFs, and have to face the choice of whether to use variables in levels or in first differences. While Christiano et al. (1995) use a VAR in levels, it is also common practice to rely on unit root pretests, as in Rogers (1999). However, neither a VAR in levels nor pretest based VARs are guaranteed to provide accurate IRF confidence bands in finite samples.

To document this problem, we show Monte Carlo simulations for a bivariate VAR where one variable has a root that is close to unity. The experiment, explained more in details in Section 5, is representative of the practical situation outlined above, where the researcher needs to include a "key" variable, in that case the real exchange rate, but neither theory nor unit root pretests provide conclusive evidence on whether this variable has a unit root or not.[1]

Figure 4(a) shows that IRFs based on both VARs in levels and unit root pretest-based VARs provide unreliable inference. The figure shows one minus the actual coverage rate of the various methods used to construct IRFs confidence bands. The nominal (desired) coverage rate is 0.90, so that a method performs well when its line is around 0.10. These lines are reported as a function of δ, a parameter that plays an important role in this paper. δ denotes the horizon of the IRF as a fraction of the total sample size. For example, in a sample of 100 monthly observations, a horizon of 12 months would correspond to $\delta = 0.12$.[2] The upper panel in Figure 4(a) shows that when the root is exactly one, estimation in levels (the line with diamonds) produces IRFs bands that undercover. In this case, a pretest (the line with circles) would be a much better choice, as it is approximately around 0.10 at any δ. However, when the root is close to unity, say 0.97, the bottom panel in Figure 4(a) shows that the situation is completely different. Unit root pretest lack power and induce the

[1] In some cases there may be more than one variable for which we can question the presence of an exact unit root. The results section 5 show that our method is robust to misspecification of the largest root of other variables in the VAR.

[2] We chose to report δ rather than the horizon because our analysis focuses on small samples, so that the horizon per se is less important than its ratio to the available sample size. This happens because, as explained later, the sample size determines the degree of imprecision of the estimate of the unit root, and the horizon determines how much this imprecision is blown up. See Rossi (2001a) for more details.

researcher to estimate a misspecified model, that is the model in first differences, which causes the coverage to be extremely poor. On the other hand, estimation in levels is now better for relatively large horizons, $\delta \geq 0.10$. Thus, the applied researcher faces the problem of choosing between levels and pretest-based methods, knowing that neither method is superior to the other, and that their relative performance crucially depends on the unknown root.

To our knowledge, there is no method that successfully provides a solution to this problem, although some have been proposed and will be discussed in the rest of the paper. To overcome this problem, Rogers (1999) estimates a VAR with the real exchange rate both in levels and in first differences. While this is a clever way of assessing the robustness of the results, this solution is not satisfactory, as nothing guarantees that the overall coverage will be correct.

In this paper we instead propose methods that have the correct coverage at medium to long horizons, where both the level and the pretest based IRFs have problematic coverage. The dotted line in Figure 4(a) shows one such method. It is clear that the method has exactly the correct coverage at medium to long horizons ($\delta \geq 0.10$), and its coverage properties are robust to whether there is an exact unit root or not, that is exactly in that region where the usual pretest methods lead to unreliable inference. In the next section we present these methods, and discuss how to implement them. More details and Monte Carlo results are available in Section 5.

3. The Model

Let the data generating process (hereafter DGP) be:

$$(I - \Phi L)\, w_t = u_t \tag{1}$$

w_t is a $(m \times 1)$ vector of variables, u_t is a $(m \times 1)$ stationary and ergodic moving average sequence:[3]

$$u_t = \Theta(L)\epsilon_t \tag{2}$$

where ϵ_t is a martingale difference sequence with covariance Σ, $\Theta(L) \equiv \sum_{i=0}^{\infty} \Theta_i L^i$, $\Theta_0 = I$, I is the $(m \times m)$ identity matrix and $\Omega^{1/2} \equiv \Theta(1)\Sigma^{1/2}$ is invertible. We will denote by η_t the structural shocks:

$$\eta_t = A_0 \epsilon_t \tag{3}$$

Without loss of generality, we can express Φ in terms of its eigenvalues and eigenvectors: $\Phi = V^{-1}\Lambda V$. We will assume that $V w_t$ is a persistent (close to I(1)) process

[3] w_t may contain deterministic components (constants and time trends) but they are irrelevant, as the IRFs are defined as deviations from the deterministic components (see Phillips (1998)).

and we rule out processes that behave like I(2) – or "almost" I(2) – in small samples. Thus, Λ is diagonal with the largest roots of the process on the main diagonal whereas the eigenvectors V describe possible cointegrating vectors.[4] To improve the asymptotic approximation to highly persistent processes in small samples we use multivariate local-to-unity asymptotic theory. That is, we model the real part of the (distinct) largest roots of the VAR, Λ, as local-to-unity:

$$\Lambda = I + \frac{1}{T}\mathbf{C} \tag{4}$$

where $\mathbf{C} = diag(c_1, c_2, ...c_m)$. By allowing these components to be either negative or positive, our approximation encompasses both close to unit root and mildly explosive processes.

To obtain better asymptotic approximations to IRFs in small samples, we also assume that the lead time of the IRF is a fixed fraction of the sample size:

$$\frac{h}{T} \underset{T\to\infty}{\to} \delta \tag{5}$$

Considering the two assumptions (4) and (5) together, we have that:

$$\Lambda^h \underset{T\to\infty}{\to} e^{\mathbf{C}\delta}$$

where $e^{\mathbf{C}\delta}$ denotes a diagonal matrix with $(e^{c_1\delta}, e^{c_2\delta}, ...e^{c_m\delta})$ on the main diagonal.

When the variables are highly persistent, from (1), the DGP for the rotated variables $y_t = Vw_t$ can be approximated as:[5]

$$y_{t+h} = \sum_{i=0}^{h-1} \Lambda^i \Theta\left(I\right) A_0 \eta_{t+h-i} + \Lambda^h y_t + o_p(T^{1/2}) \tag{6}$$

where we used (3). IRFs are easily obtained by using (6) after the shocks are identified. The IRF of the effect of a unitary $j-th$ structural shock on the $k-th$ variable in y_{t+h} will be:

$$\frac{\partial y_{t+h}^{(k)}}{\partial \eta_t^{(j)}} = \mathbf{i}_k' \Lambda^h \Theta\left(I\right) A_0 \mathbf{i}_j + o_p(T^{1/2}) \Rightarrow \mathbf{i}_k' e^{\mathbf{C}\delta} \Theta\left(I\right) A_0 \mathbf{i}_j \tag{7}$$

where $\mathbf{i}_s$ denotes the $s-th$ column of the $m \times m$ identity matrix and "$\Rightarrow$" denotes weak convergence. Note that the IRF depends on the largest roots, described by $\mathbf{C}$,

[4] We assume throughout that the cointegrating vectors are known (as in Elliott, Jansson and Pesavento (2001)). The general model (1) nests some popular models as particular cases (see the Appendix).

[5] For a proof of (6), see Rossi (2001b). Note that (6) is a device for obtaining better asymptotic approximations in small samples. Assumptions (4) and (5) govern its validity.

and on the cumulated short run dynamics, described by $\Theta(I)$. At long horizons, the uncertainty associated with the short run parameters is of smaller order than the uncertainty associated with the largest root, and we can ignore the uncertainty in the estimation of $\Theta(I)$ by simply replacing it with a consistent estimator. Although $\mathbf{C}$ cannot be consistently estimated, methods to construct valid confidence intervals for $\mathbf{C}$ are available. Confidence intervals for the IRF at long horizons can then be constructed from confidence intervals for $\mathbf{C}$. More details on how to implement this method in practice are provided in Sections 4 and 5.

As usual, different types of identification will result in different IRFs. The long-run identification (proposed by Blanchard and Quah (1989)) imposes a triangular structure to $\Theta(I)\,A_0$. The Wold ordering identification (Sims (1980)) imposes constraints on A_0. We will be agnostic about the identification procedure and will take it as given, as our goal is to provide a method for constructing IRF bands that have correct coverage, not to propose new methods to identify shocks.

4. Univariate impulse response functions with a possible unit root

We perform a simple Monte Carlo experiment to compare the coverage of the method proposed in this paper with those existing in the literature. The DGP is the following:

$$\prod_{j=1}^{p}(1-\lambda_j L)\,y_t=\epsilon_t \tag{8}$$

where λ_j are the possible roots of the process and $\lambda_1 \equiv \rho = 1 + c/T$. Nominal coverage is 0.90, $T = 100$, and the number of Monte Carlo replications is 5,000.

Our method is implemented as follows. First, we construct a median unbiased confidence interval for "c" by using methods that are available in the literature (in this case, Stock (1991)). Denote this confidence interval by $[c_0, c_1]$. With $m = 1$, (7) becomes $e^{c\delta}\Theta(1)$, where $\Theta(1)$ from (8) is equal to $[\prod_{j=2}^{p}(1-\lambda_j)]^{-1}$. Note that $\Theta(I)$ can be consistently estimated by an ADF regression (see Rossi (2001a)). The confidence interval for the IRF at horizon $h = [\delta T]$ then becomes:

$$[e^{c_0\delta}\widehat{\Theta}(1)\,,\ e^{c_1\delta}\widehat{\Theta}(1)] \tag{9}$$

The methods that we will consider are:

- the method proposed in this paper, labeled "*Big h*";
- the method proposed by Andrews and Chen (1994), where we obtain Approximately Median Unbiased estimates of ADF coefficients and use them to simulate quantiles of IRFs, labeled "*AMU*";
- Wright (2000) method;
- the standard (Runkle) bootstrap method;

- Kilian (1998a) "bootstrap after bootstrap" method.[6]

The method proposed in this paper is not a Bonferroni method and has the advantage of not relying on pretests, which may introduce pretest bias (see Cavanagh, Elliott and Stock (1995)). To show the consequences of pretest biases, we also report the coverage of confidence bands based on unit root pretests. They are constructed as follows: if the ADF test rejects a unit root then the VAR is estimated in levels, otherwise the VAR is estimated in first differences.

Figures 1 and 2 show *one minus* the empirical Monte Carlo coverage of the different methods for different DGP's as a function of δ. Its nominal value is 0.10. We consider a variety of AR(2) representative processes. In addition, Table 1 reports additional details on the percentage of samples in which the true value of the IRF lays below and above the estimated confidence bands (ideally, these percentages should be 0.05). Figure 1 shows that, for most methods, coverage is close to nominal when there is no serial correlation other than the largest root. However, things change in the presence of additional serial correlation. Here below we suummarize the main findings.

Our method has coverage close to nominal in most situations. On the other hand, the Andrews and Chen method quickly worsens as soon as the amount of serial correlation (described by λ_2) becomes non-negligible, and the performance is usually worse the bigger the horizon of prediction. This is not surprising: Andrews and Chen (1994) reported that the IRFs based on their method were median biased and the median bias increased with the horizon of prediction (see their Table 2, last DGP). The method proposed by Wright (2000) is conservative, the reason being that it simulates the DGP over the confidence interval for "c" in order to get the IRFs and then takes its maximum and minimum values[7]. We don't have to simulate the IRFs. In

[6] We focus on methods that are computationally easy. Also, we do not consider Sims and Zha (1999), nor other Bayesian methods. A thorough investigation of Sims and Zha method is provided in Kilian and Chang (2000). These authors show that Sims and Zha's intervals, by classic criteria, tend to have comparable accuracy to the bias-corrected bootstrap, which we already consider.

An alternative to Andrews and Chen (1994) is Hansen (1999) grid bootstrap. Gospodinov (2002) and Rossi (2001a) already analyzed its small sample properties for confidence intervals for half-lives and IRFs, so we do not include this method.

When implementing Kilian (1998), we use 100 bootstrap replications and Kilian's "bootstrap after bootstrap" short-cut rather than the "bootstrap plus double bootstrap".

We implement Andrews and Chen (1994) confidence intervals based on the appropriate quantiles unbiased estimates of the parameters. In an appendix available upon request we also provide coverage for IRF bands based on quantiles of the simulated IRFs constructed from unbiased estimates. We are grateful to C. Murray for providing the codes for Andrews and Chen for the latter simulations. The coverage properties based on the latter seem to be worse than those based on the former, and this is the reason why we report the former.

[7] The results for the method proposed by Wright are obtained by calculating the gradient by using analytical derivatives. This method makes the Monte Carlo simulations faster. We also experimented confidence sets based on Monte Carlo simulations with random normal draws on the distribution of the root other than the largest one. The latter had a slightly better performance, i.e. coverage

addition, Wright imposes a lower bound on "c" (zero), whereas we do not necessarily impose that. Note that the bootstraps (both Runkle's and Kilian's) appear to work quite well overall. However, Table 1 shows that the two-sided confidence interval is highly non-symmetric and, thus, it might be unappealing if the objective is to obtain median unbiased confidence bands (note that it is not possible to control which side will never reject). Our method, instead, is median unbiased at all horizons. Furthermore, it is well known that the performance of the bootstrap worsens if the DGP includes deterministic components, especially time trends. We emphasize that our method is robust to that, as confidence intervals for c constructed using Stock (1991) method are invariant to deterministic component, and our method inherits that property as well. Also, our method is an improvement over Kilian's as his method is not designed to work for roots equal to one or mildly explosive (i.e. his method relies on bias estimates, the validity of which so far has only been established for the weakly stationary case).

Figure 2 shows that estimating regressions in first differences leads to big size distortions (see also Ashley and Verbrugge (2001)). One might hope that the use of pretests to decide whether to use a specification in levels or first differences would alleviate this problem. The Figure also shows how bad confidence bands based on pretests can be. The pretest results that we report are obtained by a first stage unit root test with size 0.05 and a second stage two sided confidence interval with coverage 0.95. If the two stages of the pretest were independent, the final coverage should be roughly 0.90.[8] Finally, note from Table 1 that confidence intervals for IRFs of VARs in levels are not median unbiased.

INSERT FIGURES 1 AND 2, AND TABLES 1 AND 2

The above evidence leads to the natural question: how "big" does the horizon have to be in order to be considered "long", in which case commonly used methods fail and our procedure improves upon VARs in levels, first differences or pretests?

was close to nominal at sufficiently long horizons (0.6,0.8), although it was conservative at smaller horizons. However, this variant was extremely computationally intensive, as it required a Monte Carlo inside a Monte Carlo, so we preferred the former in our simulations. For a description of the two methods see Hamilton (1994), pag. 336-337.

[8] We do not report results for Bonferroni confidence intervals, as they are not used in practice for this kind of problems and they would be conservative anyway. We are aware of the fact that the total coverage of the two stage procedure is only roughly 0.90, but, again, in practice nobody really corrects for that anyway. Even if the researcher does a pretest, the size distortions persist. We should note that most researchers, in practice, run the pretest but don't take that into account when choosing the coverage of the confidence interval for the IRF (i.e. if the desired coverage is 0.10, the nominal coverage is chosen to be 0.90 independently of the size of the pretest). We chose not to report this, as, in practice, it would make coverage properties even worse.

First, note that our procedure works quite well for delta bigger than or equal to 0.10, which for a sample of 100 observations corresponds to an horizon equal to 10. This provides a "rule-of-thumb" for evaluating the expected coverage of existing methods relative to ours. Also, even for horizons as small as 6, our procedure is at least as good as the pretest-based procedure, unless there is a large second root (see Figure 2a, upper panels). The point at which our procedure starts to improve relative to the normal-based approximation depends somehow on the underlying serial correlation in the data, but not on the frequency of the data, so the rule-of-thumb will be the same for monthly or quarterly series. For example, consider a case in which we are interested in the IRF at an horizon of 2 years when we have a sample of data spanning 25 years. With quarterly data this situation corresponds to $\delta = 0.08$ (8 quarters over $T = 100$). With monthly data, this situation corresponds to a sample of 300 observations but the same δ (24 months over 300). The "rule-of-thumb" will be the same in both cases as the span of data is the same and, in fact, the power of the unit root test used to construct the confidence interval for c is the same (Ng (1995), Perron (1991)). For medium to long horizon IRFs, the relevant parameter is the magnitude of the horizon with respect to the sample size. A horizon of one year, corresponds to $\delta = 0.04$ with 100 quarterly data spanning 25 years and to $\delta = 0.12$ with 100 monthly data spanning a little over 8 years. Although the horizon is the same, we have much more information about the long run in 25 years of data than in 8 years of data. The "rule-of-thumb" tells us that our method works well relatively to standard methods for situations in which the practitioner has data that cover a limited span of years and he is interested in IRFs at the medium to long horizon.

Note that our method is *not pointwise*. In fact, it is based on a monotone transformation of the confidence interval for the largest root of the process, so that if the confidence interval for the root contains the true root with a pre-specified probability, the confidence interval for the whole IRF will contain the true IRF trajectory with a pre-specified probability. Table 2 reports one minus the coverage probability of the whole IRF from 10 periods onwards for the methods discussed above. Our method works pretty well.

As a robustness check we investigate the sensitivity of our method to the choice of the lag length and the presence of deterministic terms. Figure 3 shows one minus the actual coverage of IRFs confidence bands obtained by using the different lag length selection procedures. The lag length procedures are AIC, BIC and Ng and Perron (2001) "MAIC" based on GLS de-trending.[9] For comparison, we also report results based on the true lag length (labeled "True lag"). The plots are shown as a function of δ, and for a representative DGP.[10] The Figure shows that the use of

[9] We focus on AIC, BIC and MAIC for the following reasons. AIC and BIC are widely used. In addition, Ivanov and Kilian (2003) show that BIC performs quite well at long horizons and small samples relative to other methods. Ng and Perron (2001) suggest that MAIC performs better than other criteria when the persistence in the data is high.

[10] In simulations not reported, but available upon request, we show that this result holds in general

different methods to select the lag length does not make much difference in terms of the empirical coverage. This result is due to the fact that the estimation of the lag length matters for only two reasons. First, it enters in the estimation of $\Theta(L)$, but the uncertainty over that is negligible relative to the uncertainty in the roots. Second, it affects the confidence interval for c through the order of the ADF regression, but this turns out not to matter too much in terms of coverage. Table 3(a) shows the upper and lower size of the tests for the case in which deterministic terms are included in the DGP. The results indicate that the method is robust to the presence of a deterministic time trend and that the "rule-of-thumb" for evaluating the expected coverage of existing methods relative to ours is the same.

INSERT FIGURE 3, AND TABLE 3

We also compare the length of the confidence bands of the different methods we considered above. In general, our method would perform much worse than the other methods in terms of length. The reason is that we allow for mildly explosive processes, for which the upper confidence band is very wide, whereas the other methods usually impose an upper bound equal to one on the largest root. To compare these methods with ours, we therefore also impose an upper bound in our method. Results are reported in Table 3(b). The Table shows that the length of our method compares favorably to that of the other methods, especially when the process is highly persistent. Note, however, that in general imposing an upper bound may generate size distortions. The closer the root is to unity, the bigger the size distortions will be. As explained before, our method avoids this problem, and will then be more robust to possibly mildly non-stationary processes, whereas the methods proposed in the literature impose stationarity and will not work if the true root is, say, 1.001.

To conclude this section, we provide an empirical application to the analysis of output fluctuations. The question of whether output fluctuations are transitory has attracted much attention in the literature since the early work by Campbell and Mankiw (1987). While the persistence of the log of the real U.S. GDP process has been already analyzed by using local-to-unity methods in Stock (1994), nobody to date has analyzed its implications in terms of the effects of shocks at long horizons. Rather than reporting confidence bands for IRFs, we focus on lower bounds of the confidence intervals for the persistence of the shock. These are measured by how many periods it takes the shock to be back to $\gamma - percent$ of the initial value ($\gamma = 0.5$

(unless the process is very close to stationarity, in which case the MAIC criterion may lead to undercoverage). Also, the DGP is AR; if the DGP were MA, AIC may perform worse; we thank Y. Chang for this point.

corresponds to the half-life of the shock, analyzed in Rossi (2001a)).[11] Using Kilian's-style mean bias corrections for the slope parameters, it takes at least 7 periods to halve the effect of a shock ($\gamma = 0.5$) and 15 to reduce it to 25% ($\gamma = 0.25$). Our method shows that it takes at least 20 quarters to halve it, and 40 quarters to reduce it to 25%. It is clear that the method proposed in this paper, which is robust to mildly explosive as well as highly persistent processes, uncovers that the shocks to real GDP are more persistent than it is commonly found in the literature. Note that the other methods under-estimate the protracted effects of a shock to real GDP.

5. Multivariate impulse response functions with a possible unit root

In this section we compare the coverage and the length of confidence bands for IRFs in multivariate models with a root local to unity. For ease of exposition we focus on a bivariate VAR, where one variable has an exact unit root while the other has a large root that is close to one.[12] This corresponds to the common situation in which the researcher "knows" that one variable is I(1) but is unsure about whether the other variable is stationary or not (see for example Rogers (1999))[13].

The DGP is:

$$\begin{aligned} w_{1t} &= \mu + \rho w_{1t-1} + u_{1t} \\ w_{2t} &= w_{2t-1} + u_{2t} \end{aligned} \tag{10}$$

where $\rho = 1 + \frac{c}{T}$. The dynamics of the model is determined by $\Psi(L) u_t = \varepsilon_t$, where $\Psi(L) = I - \Psi L$, and $\Psi = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.3 \end{bmatrix}$.[14] Let $\Phi = \begin{bmatrix} 1 + \frac{c}{T} & 0 \\ 0 & 1 \end{bmatrix}$ to see that this simple model corresponds to equation (1). The elements of the spectral density at frequency zero of $u_{t,}$, $\Omega = \Psi(1)^{-1} \Sigma \Psi(1)^{-1'}$, are denoted as $\omega_{ij}, i, j = 1, 2$. We assume that $\Sigma^{1/2}$ is lower triangular so that shocks to w_{2t} do not contemporaneously affect w_{1t}. Once the reduced form VAR is estimated, structural IRFs are computed by using a causal ordering identification scheme.

[11] Data are obtained from the Bureau of Economic Analysis, U.S. Department of Commerce, and are seasonally adjusted, in billions of fixed 1996 dollars, quarterly data from 1967:1 to 2002:4, mnemonic "GDP96". We allow for a deterministic time trend in the estimation. The lag length is determined by Ng and Perron (2001) MAIC criterion, and it is equal to one. IRFs are obtained by using Cholesky decompositions, so we measure the consequences of shocks measured in $\sqrt{var(\epsilon_t)}$ units, where $A(L) y_t = u_t$ is the estimated VAR and $u_t = A_0 \epsilon_t$; the variance is a conditional one (i.e. conditioned on (some) variables dated t and (all) variables dated earlier).

[12] All the results in this section easily extend to a multivariate setting with $m-1$ exact unit roots and one local to unity root. While we focus on linear VARs, the same conclusions would hold for non-linear models up to a first order approximation.

[13] We show later that our method is robust to misspecification in the largest root of w_{2t}.

[14] For ease of exposition we consider Ψ diagonal. All the results in this section, with the exception of equation (11) and (13) generalize to the case of Ψ full.

We obtain our IRFs in two steps. First, we construct a confidence interval for c; second, we use the simple closed-form formula for the IRFs at long horizons as a monotone increasing function of c. Given the confidence interval for c, confidence bands for the IRFs at long horizons are obtained from equation (7).

Regarding the first step, we construct confidence intervals for c by inverting the acceptance region of various tests for a unit root in w_{1t}: (i) the Augmented Dickey Fuller (ADF) test; (ii) Elliott, Rothemberg and Stock (1996, ERS thereafter); (iii) Elliott and Jansson (2000, EJ thereafter). These tests differ in their power properties, and the most powerful tests will usualy lead to smaller confidence bands (often at the cost of additional computations). ERS improves on ADF by optimally detrending the data, and EJ improves on ADF by exploiting information on the stationary covariate Δw_{2t}. An important nuisance parameter that affects the power of the EJ test is $R^2 \equiv \omega_{11}^{-1}\omega_{12}\omega_{22}^{-1}\omega_{21}$, the square of the frequency zero correlation between the innovation in Δw_{2t} and the innovation in the quasi difference of w_{1t}.

For this example, under assumptions (4) and (5), the structural IRFs at long horizons are (up to an irrelevant $o_p(\sqrt{T})$):

$$\frac{\partial w_{t+h}}{\partial \eta_t} = \Lambda^h \Omega^{1/2} \Rightarrow \begin{bmatrix} e^{c\delta} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_{11}^{1/2} & 0 \\ \omega_{21}\omega_{11}^{-1/2} & \omega_{2.1}^{1/2} \end{bmatrix} \tag{11}$$

where $\omega_{2.1} \equiv \omega_{22} - \omega_{21}\omega_{11}^{-1}\omega_{12} = \omega_{22}\left(1 - R^2\right)$. Thus:

$$\frac{\partial w_{1t+h}}{\partial \eta_{1t}} \Rightarrow \omega_{11}^{1/2} e^{c\delta}, \quad \frac{\partial w_{1t+h}}{\partial \eta_{2t}} \Rightarrow 0 \tag{12}$$

$$\frac{\partial w_{2t+h}}{\partial \eta_{1t}} \Rightarrow \omega_{21}\omega_{11}^{-1/2} = \omega_{22}^{1/2}\left(R^2\right)^{1/2}, \quad \frac{\partial w_{2t+h}}{\partial \eta_{2t}} \Rightarrow \omega_{2.1}^{1/2} \tag{13}$$

The confidence intervals for the IRF are then estimated as in (9) over the confidence interval for c. In this simple model, since Ψ is diagonal, $\Omega^{1/2}$ is lower triangular, and the structural shocks could also be identified imposing this long run restriction.

We proceed by comparing our method with those commonly used in the existing literature in a Monte Carlo experiment. IRFs are commonly computed from VARs estimated either in first differences or in levels. Usually, the decision between a VAR in levels or in first differences is based on a unit-root pretest on w_{1t}. If the pretest fails to reject a unit root at a 5% level, then the VAR is estimated with w_{1t} in first differences, otherwise w_{1t} is used in levels. We report both cases in which the ADF and the ERS tests are used in the pretest. Since pretesting is known not to work well (Cavanagh, Elliott and Stock (1995)), we also report results for a VAR in levels without pretesting. [15]

[15] Given that we are pre-testing y_t for a unit root, if the two stages were independent, the probability

We simulated model (10) with $T = 100$ over 5000 Monte Carlo replications. The VAR confidence intervals are approximated with 500 replications. $R^2 = 0.5$. The 95% confidence intervals of the VAR IRFs are computed by Monte Carlo simulations under a normality assumption (cfr. Hamilton (1994), page 337).

Figure 4 shows the results. It displays one minus the coverage rate of various confidence bands for IRFs for $\rho = 1$, 0.97, 0.90, 0.80. Panels (a) to (b) show that, like in the univariate case, for short horizons our method has coverage that is less than the nominal 90%, while confidence intervals from a VAR have less size distortions. However, for values of δ around 0.10 and above, the confidence intervals computed with the normal approximation almost never contain the true IRF.

For $\rho = 1$, the results are close to the nominal level (0.10) for all methods, except for the confidence intervals computed from the VAR in level, for which it is around 40%. This result reflects the bias in the estimation of ρ from a regression in levels. Our method performs very well for various values of ρ: all three variants (ADF, ERS, EJ) have coverage rates that are close to the nominal level across different values of ρ. On the other hand, as ρ moves away from unity, confidence intervals computed from the VAR in first differences start to behave poorly, with coverage that approaches zero as the horizon increases. In fact, for large enough (though less than one) values of ρ, both pre-tests have low power to reject the hypothesis of a unit root, and they choose a VAR in first difference most of the times. As ρ moves further away from unity, the pretests are able to reject the hypothesis of a unit root more often and the coverage rate of the confidence bands computed from the VAR improves. As ρ becomes very small (say $\rho = .80$), (4) is no longer a good approximation and the coverage rate of our method starts to worsen: ADF and ERS have a coverage rate around 60% while EJ has a coverage rate around 70%. Finally, note from Figure 4, that using different pretests results in significantly different coverage properties. The better coverage rate of ERS relative to ADF reflects the higher power of ERS test against alternatives that are close to one.

As expected, the higher power of EJ test is reflected in the shorter median confidence interval length in Table 4. For values of ρ that are close to one, the inversion of EJ test produces confidence intervals that can be more than half the length of confidence intervals obtained by inverting ADF. At the same time, confidence intervals constructed inverting ERS and EJ are not symmetric (results are not reported). The smaller length of the interval comes at the cost of not having a median unbiased confidence interval. The interval length for VAR is also very small but, as we just noted, having short intervals is irrelevant if the probability of including the true IRF

---

that the confidence interval contain the true IRF when the null of $c = 0$ is true would be $(1 - 0.5)^2$. Using 95% confidence intervals for the IRFs allows us to do a fair comparison of the empirical coverage rates of the different methods. However, the two stages are not independent, because of the correlation between the residuals (as in Cavanagh, Elliott and Stock (1995)), so the pre-test coverage is not 0.90, even asymptotically. The $\bar{c}$ used here is -7, as ERS and EJ recommend.

is zero.

To check the robustness of our method to some of the assumptions in model, we compute the coverage rates for all methods when model (10) is estimated even though the second root is not exactly equal to one. Figure 5 shows that, when the second root is close to one (0.99), our method still produces confidence intervals with the desired coverage. Confidence intervals from VAR either in levels or in differences have still bad coverage at long horizons. As the second root moves away from one, ERS and ADF still performs well, while EJ shows significant size distortions. The robust behavior of ERS and ADF is due to the fact that confidence bands obtained inverting ERS and ADF use information only on w_{1t}. Since we are assuming that Φ is diagonal (we are ruling out processes that are I(2) when $\rho = 1$, or "close to I(2)" when ρ is local to unity) the misspecification in the estimation of the second row of $\Theta(I)$ does not affect the IRF for w_{1t}. On the other hand, with the inversion of EJ test, we are constructing confidence intervals for c using the information contained in Δw_{2t}. Since in this case w_{2t} is not really I(1), we are using incorrect information in calculating the test statistic. Overall, our method performs well even if the other variables in the model are not exactly unit root.

INSERT FIGURES 4, 5 AND TABLE 4

It is possible to generalize the method and analyze the coverage of multivariate confidence sets for IRFs when there is more than one root local to unity. We explore this possibility in a bivariate VAR with two roots local to unity. Note, however, that the method proposed in this paper can theoretically be applied to m local-to-unity VARs, $m > 2$ (although it would be cumbersome in practice).

The confidence set is obtained as follows. First, we construct a one-sided joint confidence set for C with confidence level 0.95 by inverting Stock and Watson (1988) test for common trends for the smallest eigenvalue of Φ. The test statistic for common trends is based on filtering the VAR to take into account the existence of serial correlation (see Stock and Watson (1988)).[16] Then, we project that set onto the parameters of interest, the IRFs, which are monotone functions of C, by using (7).

[16] Consider Stock and Watson (1988) test for k=2 versus m=1<k common trends (i.e. 2 vs. 1 unit root). The system will have 2 unit roots if we do not reject k=2 common trends. So we take the second largest root (the smallest root in our bivariate example, call it λ_2) and test whether it is equal to one. If we accept the null hypothesis then we have two unit roots and if we reject (so it is less than one) we may have either one or zero unit roots. One does not reject if $(|\lambda_2| - 1)T$ is small (since $(|\lambda_2| - 1)T$ is negative, one does not reject when it is bigger than the critical value). The statistic based on filtering the data performs better for VARs than the version obtained by correcting the estimate of Φ for serial correlation, as noticed in Stock and Watson (1988). The confidence set is based on a simulated grid with step 0.05, smoothed using quantile methods over a sparser grid with steps 0.25. See also Stock and Watson (1996) and Rossi (2001b).

The induced confidence set for the IRFs is the convex hull of the projected set. By construction, any point on the IRFs should be contained within its confidence region with probability 0.95.

Table 5 shows one minus the empirical coverage of the joint confidence set for C (labeled "C") and for the IRF (labeled "IRF"). Nominal coverage is 95%. The empirical coverage is defined as the fraction of times in the Monte Carlo in which the true values of C or the IRF are inside the confidence set. The DGP is a bivariate VAR with two lags: $(I - \Psi L)(I - \Phi L) w_t = A_0 \epsilon_t$ where $A_0 = \begin{pmatrix} 1 & 0.4 \\ 0 & 1 \end{pmatrix}$, $\Psi = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix}$, $\Phi = \begin{pmatrix} \phi_{11} & 0 \\ 0 & \phi_{22} \end{pmatrix}$. The column labeled "IRF" reports the joint confidence set for shock 1 to variable 1 and shock 2 to variable 2. As the Table shows, our method results in confidence sets whose coverage is sometimes too small, but the performance improves as the horizon increases. Note that coverage worsens very quickly when the process approaches stationarity, much faster than in the univariate case. The poor performance that we observe for large values of C is due to the fact that we impose the upper bound in the one-sided confidence set.

Unfortunately, the behavior of the confidence bands seems quite erratic and more efficient methods for conducting simulations are needed in the multivariate context with more than one root close to unity.

INSERT TABLE 5

6. A method robust at short horizons

While the main method proposed in this paper applies to long horizons, namely horizons bigger than 6 or 10, it is also possible to refine it in order to make it robust also to short horizons. We discuss one such method in this section. The method is slightly more complicated than the method proposed above, but still easy and very fast to use. However, it builds on Bonferroni methods, so it will be conservative, especially at short horizons. It will also not be median unbiased, again especially at shorter horizons.

The method is as follows:

(a) first, we compute 90% confidence intervals for c by inverting an ADF test. Using the confidence interval for c, (c_L, c_U), we compute $e^{c_L \delta}$ and $e^{c_U \delta}$;

(b) second, we estimate a VAR in first differences and construct a 90% confidence interval for Θ_i by using the same Monte Carlo approximations to the delta method described in section 4;

(c) for each limit of the confidence interval for Θ_i at each horizon, we compute $e^{c_L \delta}\ \Theta_i$ and $e^{c_U \delta}\ \Theta_i$;

(d) the overall confidence interval is then obtained as: $(\min_i e^{C_L \delta}\ \Theta_i; \max_i e^{C_U \delta}\ \Theta_i)$;

By the Bonferroni inequality, the confidence interval should have a coverage of at least 90% at each horizon h. Note that, by construction, this method is now pointwise. Wright (2000) also proposes a method based on Bonferroni inequalities. The difference with his method is that we do not re-estimate Θ_i associated to different values of c in (a).

We perform a simple Monte Carlo experiment to evaluate the performance of this method relative to the one proposed in the main part of the paper. The DGP is the same as described in (10), where $\rho = 0.97$, $T = 100$. The nominal coverage is 0.90. Table 6 shows the results. The new method is labeled "Small h". For comparison, we also report the same probabilities based on our "Big h" method. Subscripts "L" and "R" are used to denote the empirical rejection probabilities of the true IRF laying respectively on the left and on the right of the proposed confidence interval, which should ideally be 0.05. Both methods are based on confidence intervals obtained by inverting a simple ADF regression test for unit root.

Note that the "Small h" method improves the empirical rejection probabilities at short horizons relative to the "Big h" method. However, it is not median unbiased at short horizons, and slightly conservative. Nevertheless, its overall coverage properties are quite good at every horizon.

INSERT TABLE 6

7. An empirical application to exchange rate dynamics

As an empirical application, we analyze the nominal versus real sources of fluctuations in real and nominal exchange rates. There is a large literature on this topic. Eichenbaum and Evans (1995) influential paper use an identification *à* la Sims; Clarida and Gali (1994), Lastrapes (1992) and Rogers (1999), among others, impose a long-run identification. Different types of identification methods will result in different restrictions on $\Theta(I)\,A_0$ in (7) and therefore will result in different IRF. Although our method works regardless of the type of indentification method used, here we will focus on a short run identification based on Wold ordering.[17]

We focus on Eichenbaum and Evans (1995) influential paper, which has been used for illustrative purposes before (e.g. Kilian (1998b)). We use the ratio of the log of non-borrowed reserves to the log of total reserves as a measure of the policy

[17] Traditionally, Wold ordering identification has been imposed on VARs in levels whereas long-run identifications have been imposed on VARs in first differences. It is unclear how much of the differences in the results is due to the different identification procedures and how much is due to the maintained assumption on the order of integration of the variables (which is different in the two cases). By using our approach, we would be able to distinguish how much of these differences are really due to the different economic identifications.

instrument. The five-variables VAR also includes U.S. industrial production, U.S. CPI, a measure of the difference between U.S. and foreign short-term rates and the real exchange rates.

Data for the U.S. (industrial production, 3 months T-bill rates, Total Reserves and Non Borrowed Reserves with extended credit) and bilateral monthly nominal exchange rates are from the Federal Reserve Database. Data for Industrial Production for each foreign country and CPI for all countries including the U.S. are from the IFS database (line 66 and line 64 respectively). Short term money markets rates for each foreign country are also from the IFS database. The starting date for the sample is 1973:1 with different ending dates for each country (2001:12 for Germany, 2002:9 for Japan and the United Kingdom and 2001:5 for Italy). All variables are in logarithms except for the interest rates. Here, an increase in the real exchange rates represents a depreciation of the U.S. real exchange rate. The structural IRFs are calculated using the Wold ordering: output, prices, reserves, interest rate differential and real exchange rate. An exogenous contractionary monetary shock is identified as the component of a negative innovation in $NBRX$ that is orthogonal to prices and output.

Table 7 provides some descriptive statistics and results of various tests for unit roots on the real exchange rate. The evidence suggests that the latter is a highly persistent variable and, for most countries, we cannot reject a unit root. Figure 6 compares various confidence intervals for the response of the real exchange rate to a contractionary monetary policy shock for different countries. The confidence bands from a VAR with all variables in levels (solid line with diamonds) and with all variables in first differences (dotted line with stars) are computed by simulation. The two outer solid lines are the confidence bands computed with our method inverting ERS test for a unit root on the real exchange rate and the middle solid line is the median unbiased estimate of the IRF.[18] As in Eichenbaum and Evans (1995) we use 6 lags in the VAR estimation.

Some interesting results emerge from Figure 6. Note that the confidence bands for VARs in first differences show more persistent effects than those based on VARs in levels, as intuition may suggest. In fact, the former may remain bounded away from zero (as for Japan and the U.K.), showing that the effects of the shock are much more persistent and may never disappear, even in the long run. On the other hand, the confidence bands for the VAR in levels include zero after a few quarters. Our method is somewhere in the middle, but in general it shows that shocks are more persistent than a VAR in levels would predict. As in Eichenbaum and Evans (1995), a contractionary shock to U.S. monetary policy leads to a persistent appreciation of the real exchange rates for Germany, Japan, and UK. For Italy, both the VAR in first

[18] Although inverting ERS test produce confidence bands that are less precise than the ones obtained by inverting EJ test, the simulations in the previous section show that EJ is also less robust to the true DGP of the other variables.

differences and in levels suggest an initial appreciation that dies out in the long run, while the inversion of ERS suggests a small, but persistent, depreciation of the real exchange rate. Although the approximation used in our method is not too informative about the behavior of the real exchange rate at short horizons, in our results the dollar does not depreciate after the initial appreciation. Our confidence intervals suggest a more persistent real exchange rate than what a VAR in levels would suggest, with a dollar not depreciating even after a long period of time for most currencies. In fact, for Germany, Japan and UK, the dollar is still not depreciating even after 4 years (48 months). As in the previous example, our method suggest a slightly less persistent response of the real exchange than that estimated with a VAR in first difference.

INSERT FIGURE 6 AND TABLE 7

8. Conclusions

Whether shocks have long run effects on economic variables, and how persistent these effects are, is a highly debated issue. Results are sensitive to the order of integration of the variables and ad-hoc robustness checks that lack theoretical justifications are routinely performed to verify the empirical conclusions.

We propose a simple, yet rigorous, method to estimate the long run effects of the shocks and the uncertainty around these in a coherent way. The method has the advantage of controlling coverage over the *whole* IRF trajectory at long horizons (i.e. it is not pointwise). With this method, researchers do not have to take a stand on whether the process is I(1) or I(0) before doing inference. Thus, the method provides a feasible alternative to unit root pre-tests, which we show imply considerable coverage distortions.

By using this method, we find that conventional methods that rely on estimation in levels tend to underestimate the persistence of shocks in small samples, and, in accordance with recent evidence by Rogers (1999), that monetary shocks have effects that are more persistent than previously thought.

9. References

Andrews, Donald, "Exact median-unbiased estimation of first order autoregressive/unit root models", *Econometrica* 61(1), January 1993, pp. 139-165.

Andrews, Donald, and Hong-Yuan Chen, "Approximately Median-Unbiased Estimation of Autoregressive Models", *Journal of Business and Economic Statistics* 12(2), April 1994, pp. 187-204.

Ashley, Richard, and Randal Verbrugge, "To difference or not to difference: a Monte Carlo investigation of inference in vector autoregressive models", VPI Economics Department Working Paper E99-15.

Blanchard, Olivier and Danny Quah, "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review* 79(4), September 1989, pp. 655-673.

Campbell, John Y. and Gregory Mankiw, "Are Output Fluctuations Transitory?", *Quarterly Journal of Economics 102*(4)*,* November 1987, pp. 857-880.

Cavanagh, Christopher L., Graham Elliott, and James H. Stock, "Inference in Models with nearly Integrated Regressors", *Econometric Theory* 11, 1995, pp. 1131-47.

Clarida, R., and Jordi Gali, "Sources of real exchange rate fluctuations: how important are nominal shocks?", *Carnegie-Rochester Conference Series on Public Policy* 41, 1994, pp. 1-56.

Eichenbaum, Martin, and Charles Evans, "Some empirical evidence on the effects of monetary policy shocks on exchange rates", *Quarterly Journal of Economics* 110, 1995, pp. 975-1010.

Elliott, Graham and Michael Jansson: "Testing for Unit Roots with Stationary Covariates", UCSD Discussion Paper 2000-06, *Journal of Econometrics*, Forthcoming.

Elliott, Graham and James H. Stock: "Confidence Intervals for Autoregressive Coefficients Near One ", *Journal of Econometrics*, 103, 2001, pp. 155-181.

Elliott, Graham, Michael Jansson, and Elena Pesavento, "Optimal Power for Testing Potential Cointegrating Vectors with Known Parameters for Nonstationarity", *mimeo*, July 2002.

Elliott, Graham, Thomas J. Rothenberg and James Stock: “Efficient Tests for an Autoregressive Unit Root”, *Econometrica*, 64, 1996, 813-836.

Faust, Jon, and Eric Leeper, “When do long-run identifying restrictions give reliable results?”, *Journal of Business and Economic Statistics* 15, 1997, pp. 345-353.

Gospodinov, Nikolay, “Asymptotic confidence intervals for impulse responses of near-integrated processes: An application to Purchasing Power Parity”, *mimeo*, Concordia University, October 2002.

Hamilton, James D., *Time Series Analysis*, Princeton University Press, 1994.

Hansen, Bruce, “Bootstrapping the autoregressive model”, *Review of Economics and Statistics 81(4)*, November 1999, pp.594-607.

Ivanov, Ventzislav, and Lutz Kilian, “A Practitioner’s Guide to Lag-Order Selection for VAR Impulse Response Analysis”, *mimeo,* 2003, University of Michigan.

Kilian, Lutz, “Small-sample Confidence Intervals for Impulse Response Functions”, *Review of economics and statistics* 80, 1998(a), pp.218-230.

Kilian, Lutz, “Confidence intervals for impulse responses under departures from normality”, *Econometric Reviews 17*, 1998(b), pp. 1-29.

Kilian, Lutz, and Paolo-Li Chang, “How accurate are confidence intervals for impulse responses in large VAR models?”, *Economics Letters* 69, 2000, pp. 299-307.

Kilian, Lutz, “Impulse response analysis in vector autoregressions with unknown lag order”, *Journal of Forecasting* 20(3) 2001, pp. 161-179.

Lastrapes, William D., “Sources of Fluctuations in Real and Nominal Exchange Rates, *Review of Economics and Statistics*, 1992, pp. 530-539.

Lütkepohl, Helmut, “Asymptotic Distribution of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models”, *Review of Economics and Statistics 72*, 1990, pp. 116-125.

Ng, Serena, “Testing for Unit Roots in Flow Data Sampled at Different Frequencies”, *Economic Letters* 47, 1995, pp. 237-242.

Ng, Serena and Pierre Perron, “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power”, *Econometrica 69(6),* November 2001, pp. 1519-54.

Perron, P. (1991): " Test Consistency with Varying Sampling Frequency", *Econometric Theory*, 7(3), 341-68.

Phillips, Peter C. B., "Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VARs", *Journal of Econometrics* 83(1-2), 1998, pp. 21-56.

Rogers, John H., "Monetary Shocks and Real Exchange Rates", *Journal of International Economics* 49, 1999, pp. 269-288.

Rogers, John H., "Monetary Shocks and Real Exchange Rates", *International Finance Discussion Paper* #612, Federal Reserve Board, Washington D.C., 1998.

Rossi, Barbara, "Confidence Intervals for Half-Life Deviations From PPP", Duke Working Paper 02-08, 2001(a).

Rossi, Barbara, "Testing Long-Horizon Predictive Ability With High Persistence, and the Meese-Rogoff Puzzle", Duke Working Paper 02-10, 2001(b).

Sims, Christopher A., "Macroeconomics and Reality", *Econometrica* 48, 1980, pp. 1-48.

Sims, Christopher A., and Tao Zha, "Error Bands for Impulse Responses", *Econometrica,* 67(5), September 1999, pp. 1113-1155.

Stock, James H., "Unit Roots, Structural Breaks and Trends", in Engle, R., and Mc Fadden, D., eds., *Handbook of Econometrics*, vol. IV, North Holland, 1994.

Stock, James H., "VAR, Error Correction and Pretest Forecasts at Long Horizons", *Oxford Bulletin of Economics and Statistics*, 58(4), 1996.

Stock, James H., "Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series", *Journal of Monetary Economics*, 28 (1991), pp. 435-459.

Stock, James H., and Mark W. Watson, "Testing for common trends", *Journal of the American Statistical Association*, 83(404), December 1988, pp. 1097-1107.

Stock, James H., and Mark W. Watson, "Confidence sets in regressions with highly serially correlated regressors", *mimeo*, December 1996.

Wright, Jonathan, "Confidence intervals for univariate impulse responses with a near unit root", *Journal of Business and Economic Statistics* 18(3), July 2000.

10. Appendix

The general model (1) nests some popular models. Let $y_t = [y_{1t}', y_{2t}]$, where y_{2t} represents the cointegrating vector.

Case 1: When the researcher knows the cointegration vector a-priori and y_{1t} is known to be I(1), then

$$\Lambda = \begin{bmatrix} I_{m\times m} & 0_{m\times 1} \\ 0_{1\times m} & 1+\frac{c}{T} \end{bmatrix}; \qquad V = \begin{bmatrix} I_{m\times m} & 0_{m\times 1} \\ -\gamma' & 1 \end{bmatrix} \tag{14}$$

This is the DGP considered in Elliott, Jansson and Pesavento (2002). In this case it is possible to use our method to invert their point optimal invariant test to obtain IRFs bands.

Case 2: When y_{1t} is local to unity and y_{2t} is stationary, we obtain the framework used in Elliott (1998). This is equivalent to our DGP where $1+\frac{C}{T}$ with $C = diag(c_1, c_2, ...c_m)$. corresponds to the m local to unity roots for y_{1t}, while λ_1 is the largest but stationary root for the cointegration vector $y_{2t} = w_{2t} - \gamma' w_{1t}$ which can be zero in the trivial case in which $\Theta(L) = I$ and y_{2t} is a martingale difference sequence. In this case:[19]

$$\Lambda = \begin{bmatrix} I+\frac{C}{T} & 0_{m\times 1} \\ 0_{1\times m} & \lambda_1 \end{bmatrix}; \qquad V = \begin{bmatrix} I_{m\times m} & 0_{m\times 1} \\ -\gamma' & 1 \end{bmatrix} \tag{15}$$

Case 3: When the researcher doesn't know whether both y_{1t} or y_{2t} have a unit root (i.e. there is no cointegration) the model becomes (see Rossi (2001)):

$$\Lambda = \begin{bmatrix} I+\frac{C}{T} & 0_{m\times 1} \\ 0_{1\times m} & 1+\frac{c_{m+1}}{T} \end{bmatrix}; \qquad V = \begin{bmatrix} I_{m\times m} & 0_{m\times 1} \\ -\gamma' & 1 \end{bmatrix} \tag{16}$$

Case 4: If the researcher knows that there is no cointegration, but he is not sure whether *one* variable has a unit root or not, while he knows that the other m variables are stationary, we can write the model in the framework of Elliott and Jansson (2000):

$$\Lambda = \begin{bmatrix} \lambda_{m\times m} & 0_{m\times 1} \\ 0_{1\times m} & 1+\frac{c}{T} \end{bmatrix}; \qquad V = \begin{bmatrix} I_{m\times m} & 0_{m\times 1} \\ 0 & 1 \end{bmatrix} \tag{17}$$

where $\lambda = diag\,(\lambda_{11,}\lambda_{12}, ...\lambda_{1m})$ are the m largest (but stationary) roots of the stationary variables y_{1t}. In this case, it is possible to invert a point optimal test to compute IRFs for the variable that is local to unity.

[19] Note that the quality of the approximation used in this paper depends on whether λ_1 is close to unity. If λ_1 is very small, the asymptotic approximations used in this paper will not be accurate. The same comment holds for case 4 below.

Table 1(a): Comparison of coverage of IRF confidence bands

			Big h		AMU		Wright		Boot		Kilian	
ρ	λ_2	δ	low	up	low	up	low	up	low	up	low	up
0.99	0	.05	0.06	0.06	0.12	0.41	0.00	0.06	0.01	0.17	0.02	0.11
		0.1	0.05	0.05	0.07	0.27	0.00	0.06	0.00	0.22	0.01	0.13
		0.2	0.05	0.05	0.07	0.10	0.01	0.05	0.00	0.23	0.01	0.13
		0.3	0.05	0.05	0.07	0.07	0.03	0.04	0.00	0.23	0.01	0.13
		0.5	0.05	0.05	0.07	0.07	0.05	0.05	0.00	0.23	0.01	0.13
	0.4	.05	0.07	0.15	0.16	0.46	0.01	0.04	0.01	0.14	0.03	0.08
		0.1	0.05	0.07	0.05	0.38	0.00	0.06	0.00	0.20	0.01	0.10
		0.2	0.05	0.06	0.02	0.19	0.00	0.06	0.00	0.23	0.00	0.11
		0.3	0.06	0.06	0.02	0.08	0.00	0.05	0.00	0.22	0.01	0.12
		0.5	0.06	0.06	0.02	0.03	0.03	0.04	0.00	0.22	0.01	0.12
	0.8	.05	0.32	0.09	0.47	0.32	0.04	0.01	0.00	0.18	0.02	0.09
		0.1	0.06	0.18	0.14	0.49	0.01	0.06	0.00	0.22	0.02	0.09
		0.2	0.04	0.09	0.01	0.47	0.00	0.09	0.00	0.26	0.01	0.10
		0.3	0.05	0.07	0.00	0.41	0.00	0.08	0.00	0.26	0.01	0.11
		0.5	0.06	0.06	0.02	0.25	0.00	0.05	0.00	0.22	0.01	0.11
0.95	0	.05	0.04	0.06	0.04	0.11	0.00	0.04	0.01	0.14	0.02	0.11
		0.1	0.04	0.06	0.03	0.05	0.02	0.05	0.00	0.16	0.01	0.12
		0.2	0.05	0.06	0.04	0.04	0.04	0.05	0.00	0.15	0.01	0.12
		0.3	0.05	0.06	0.04	0.04	0.05	0.05	0.00	0.15	0.01	0.12
		0.5	0.05	0.06	0.05	0.04	0.06	0.06	0.00	0.15	0.01	0.12
	0.4	.05	0.03	0.08	0.05	0.24	0.00	0.04	0.01	0.13	0.03	0.09
		0.1	0.04	0.06	0.01	0.04	0.00	0.03	0.00	0.16	0.01	0.11
		0.2	0.05	0.06	0.01	0.02	0.03	0.04	0.00	0.15	0.01	0.11
		0.3	0.06	0.06	0.02	0.02	0.05	0.04	0.00	0.15	0.01	0.11
		0.5	0.06	0.06	0.02	0.02	0.06	0.05	0.00	0.14	0.01	0.10
	0.8	.05	0.18	0.09	0.25	0.29	0.01	0.02	0.00	0.15	0.03	0.08
		0.1	0.03	0.12	0.02	0.29	0.00	0.04	0.00	0.19	0.02	0.09
		0.2	0.04	0.09	0.00	0.02	0.00	0.02	0.00	0.18	0.01	0.11
		0.3	0.05	0.08	0.00	0.00	0.01	0.02	0.00	0.15	0.01	0.12
		0.5	0.07	0.07	0.00	0.00	0.04	0.02	0.00	0.11	0.01	0.08
1	0	.05	0.05	0.09	0.16	0.56	0.01	0.07	0.00	0.16	0.02	0.10
		0.1	0.05	0.06	0.10	0.58	0.00	0.14	0.00	0.23	0.01	0.12
		0.2	0.05	0.06	0.06	0.62	0.00	0.22	0.00	0.26	0.01	0.13
		0.3	0.06	0.06	0.05	0.66	0.00	0.27	0.00	0.26	0.01	0.13
	0.8	.05	0.35	0.10	0.59	0.31	0.06	0.01	0.00	0.18	0.02	0.09
		0.1	0.08	0.21	0.31	0.52	0.01	0.05	0.00	0.22	0.02	0.10
		0.2	0.04	0.12	0.08	0.61	0.00	0.10	0.00	0.28	0.02	0.10
		0.3	0.04	0.08	0.03	0.67	0.00	0.14	0.00	0.30	0.01	0.11

Table 1(b): Comparison of coverage of IRF confidence bands

λ_2	δ	Level		Pretest		Diff		Level		Pretest		Diff	
		low	up	low	up	low	up	low	up	low	up	low	up
				$\rho =$	0.99					$\rho =$	0.97		
0	.05	0.02	0.13	0.04	0.06	0.10	0.03	0.01	0.13	0.11	0.06	0.29	0.00
	0.1	0.02	0.12	0.10	0.06	0.22	0.01	0.02	0.11	0.45	0.05	0.90	0.00
	0.2	0.02	0.11	0.34	0.05	0.63	0.00	0.02	0.10	0.55	0.05	1.00	0.00
	0.3	0.02	0.11	0.59	0.05	0.95	0.00	0.02	0.10	0.55	0.05	1.00	0.00
	0.5	0.02	0.11	0.65	0.05	1.00	0.00	0.02	0.10	0.55	0.05	1.00	0.00
0.4	.05	0.02	0.12	0.03	0.05	0.06	0.04	0.02	0.13	0.05	0.06	0.12	0.02
	0.1	0.02	0.13	0.05	0.06	0.11	0.02	0.01	0.13	0.19	0.06	0.46	0.00
	0.2	0.02	0.12	0.15	0.05	0.32	0.01	0.01	0.11	0.54	0.06	0.99	0.00
	0.3	0.02	0.11	0.33	0.05	0.61	0.00	0.01	0.10	0.55	0.06	1.00	0.00
	0.5	0.02	0.11	0.62	0.05	0.97	0.00	0.02	0.10	0.55	0.05	1.00	0.00
0.8	.05	0.11	0.05	0.09	0.02	0.14	0.03	0.08	0.05	0.08	0.02	0.15	0.02
	0.1	0.03	0.12	0.03	0.05	0.06	0.06	0.02	0.12	0.04	0.05	0.10	0.03
	0.2	0.02	0.14	0.04	0.06	0.09	0.04	0.01	0.11	0.14	0.06	0.36	0.01
	0.3	0.02	0.14	0.07	0.06	0.17	0.02	0.01	0.10	0.41	0.05	0.81	0.00
	0.5	0.02	0.12	0.22	0.05	0.45	0.01	0.01	0.08	0.56	0.04	1.00	0.00
				$\rho =$	0.95					$\rho =$	1		
0	.05	0.01	0.13	0.21	0.07	0.62	0.00	0.02	0.13	0.03	0.07	0.06	0.06
	0.1	0.01	0.10	0.43	0.06	1.00	0.00	0.02	0.12	0.03	0.07	0.06	0.06
	0.2	0.02	0.09	0.43	0.05	1.00	0.00	0.02	0.11	0.02	0.07	0.06	0.06
	0.3	0.02	0.09	0.44	0.05	1.00	0.00	0.02	0.11	0.02	0.07	0.06	0.06
	0.5	0.03	0.09	0.44	0.05	1.00	0.00	0.02	0.11	0.02	0.07	0.06	0.06
0.4	.05	0.01	0.13	0.05	0.07	0.21	0.01	0.03	0.12	0.03	0.05	0.05	0.06
	0.1	0.01	0.11	0.34	0.06	0.88	0.00	0.02	0.14	0.02	0.07	0.05	0.06
	0.2	0.02	0.10	0.43	0.05	1.00	0.00	0.02	0.14	0.02	0.07	0.05	0.06
	0.3	0.02	0.10	0.44	0.05	1.00	0.00	0.02	0.14	0.02	0.08	0.05	0.06
	0.5	0.02	0.09	0.44	0.05	1.00	0.00	0.02	0.15	0.02	0.08	0.05	0.06
0.8	.05	0.07	0.06	0.06	0.02	0.17	0.02	0.11	0.04	0.10	0.02	0.14	0.03
	0.1	0.02	0.12	0.05	0.06	0.18	0.02	0.02	0.13	0.03	0.06	0.04	0.07
	0.2	0.01	0.12	0.31	0.06	0.78	0.00	0.02	0.17	0.01	0.09	0.03	0.08
	0.3	0.02	0.10	0.47	0.05	1.00	0.00	0.01	0.16	0.01	0.09	0.03	0.08
	0.5	0.02	0.08	0.47	0.04	1.00	0.00	0.02	0.14	0.01	0.09	0.03	0.08

Table 2: Comparison of coverage of whole IRF trajectory

ρ	λ_2	Big h_L	Big h_U	AMU_L	AMU_U	$Wright_L$	$Wright_U$	$Boot_L$	$Boot_U$
1	0	0.06	0.06	0.11	0.71	0.00	0.42	0.00	0.29
	0.4	0.07	0.09	0.10	0.73	0.01	0.34	0.00	0.31
	0.8	0.09	0.22	0.31	0.81	0.01	0.21	0.00	0.37
0.99	0	0.07	0.06	0.10	0.28	0.05	0.09	0.00	0.26
	0.4	0.08	0.07	0.06	0.39	0.03	0.10	0.00	0.29
	0.8	0.09	0.17	0.14	0.50	0.00	0.10	0.00	0.33
0.95	0	0.06	0.06	0.05	0.05	0.06	0.06	0.00	0.19
	0.4	0.07	0.06	0.03	0.03	0.06	0.05	0.00	0.19
	0.8	0.08	0.12	0.02	0.28	0.05	0.05	0.00	0.27
		$Kilian_L$	$Kilian_U$	$Level_L$	$Level_U$	$Pretest_L$	$Pretest_U$	$Diff_L$	$Diff_U$
1	0	0.01	0.15	0.04	0.15	0.05	0.06	0.02	0.07
	0.4	0.02	0.14	0.04	0.17	0.04	0.07	0.02	0.08
	0.8	0.03	0.16	0.04	0.21	0.05	0.08	0.03	0.09
0.99	0	0.01	0.15	0.03	0.13	1.00	0.01	0.64	0.06
	0.4	0.02	0.15	0.04	0.17	0.97	0.02	0.61	0.07
	0.8	0.03	0.15	0.04	0.19	0.43	0.05	0.23	0.08
0.95	0	0.01	0.15	0.03	0.12	1.00	0.00	0.42	0.06
	0.4	0.02	0.14	0.03	0.13	1.00	0.00	0.44	0.07
	0.8	0.03	0.15	0.05	0.18	1.00	0.02	0.49	0.09

Table 3(a): Robustness to a deterministic time trend

λ_2	δ	$\rho =$	0.99	$\rho =$	0.97	$\rho =$	0.95	$\rho =$	0.92
0	0.05	0.06	0.06	0.07	0.03	0.06	0.03	0.05	0.03
	0.1	0.07	0.04	0.07	0.04	0.06	0.04	0.07	0.04
	0.15	0.07	0.05	0.07	0.05	0.07	0.04	0.07	0.04
	0.2	0.07	0.05	0.07	0.05	0.07	0.05	0.08	0.04
	0.3	0.07	0.05	0.07	0.05	0.07	0.05	0.08	0.04
	0.4	0.07	0.05	0.07	0.05	0.07	0.05	0.08	0.04
0.2	0.05	0.05	0.08	0.06	0.05	0.05	0.03	0.05	0.04
	0.1	0.05	0.04	0.06	0.05	0.06	0.04	0.06	0.04
	0.15	0.06	0.04	0.06	0.05	0.06	0.04	0.07	0.04
	0.2	0.06	0.05	0.06	0.05	0.07	0.04	0.07	0.04
	0.3	0.06	0.05	0.06	0.05	0.07	0.04	0.08	0.04
	0.4	0.06	0.05	0.06	0.05	0.07	0.04	0.08	0.04
0.6	0.05	0.07	0.16	0.07	0.11	0.05	0.09	0.04	0.07
	0.1	0.05	0.08	0.05	0.05	0.05	0.05	0.04	0.06
	0.15	0.06	0.05	0.06	0.05	0.05	0.05	0.05	0.05
	0.2	0.07	0.05	0.06	0.05	0.06	0.05	0.06	0.05
	0.3	0.07	0.05	0.06	0.05	0.07	0.05	0.07	0.05
	0.4	0.07	0.05	0.07	0.05	0.07	0.05	0.08	0.05

Table 3(b): Comparison of length of confidence intervals

ρ	λ_2	δ	Big h_B	AMU	Wright	Boot	Kilian	Level	Pretest	Diff
0.99	0	0.1	0.58	0.39	0.81	0.68	0.75	0.63	0.54	0.35
		0.2	0.75	0.56	0.90	0.92	1.19	0.97	0.66	0.35
		0.3	0.82	0.65	0.93	1.04	1.61	1.29	0.77	0.35
		0.5	0.87	0.73	0.96	1.27	2.75	2.08	1.06	0.35
	0.4	0.1	0.95	0.70	1.83	1.22	1.39	1.12	1.20	0.89
		0.2	1.23	1.06	1.87	1.60	2.20	1.63	1.40	0.90
		0.3	1.34	1.22	1.85	1.79	3.01	2.14	1.58	0.90
		0.5	1.42	1.32	1.79	2.16	5.43	3.37	2.08	0.90
	0.8	0.1	2.64	1.20	4.96	2.94	3.43	3.21	3.99	3.27
		0.2	3.39	2.91	7.56	4.87	6.97	5.16	6.49	5.05
		0.3	3.68	—	8.41	5.66	10.10	6.73	8.02	5.83
		0.5	3.88	—	8.55	6.50	18.20	10.15	10.28	6.44
0.97	0	0.1	0.64	0.52	0.78	0.65	0.71	0.66	0.58	0.34
		0.2	0.81	0.69	0.89	0.76	0.98	0.92	0.69	0.34
		0.3	0.87	0.75	0.92	0.76	1.17	1.10	0.78	0.34
		0.5	0.90	0.77	0.93	0.76	1.64	1.50	1.01	0.34
	0.4	0.1	1.04	0.93	1.54	1.16	1.30	1.12	1.20	0.86
		0.2	1.31	1.26	1.64	1.32	1.78	1.53	1.38	0.87
		0.3	1.41	1.35	1.66	1.32	2.14	1.81	1.50	0.87
		0.5	1.47	1.37	1.64	1.32	3.13	2.48	1.85	0.87
	0.8	0.1	2.74	1.88	4.52	2.76	3.20	2.98	3.72	3.10
		0.2	3.46	3.97	6.21	4.15	5.84	4.58	5.72	4.55
		0.3	3.72	4.08	6.47	4.31	7.52	5.65	6.77	5.12
0.95	0	0.1	0.69	0.59	0.80	0.61	0.67	0.65	0.61	0.33
		0.2	0.84	0.71	0.89	0.63	0.80	0.81	0.69	0.33
		0.3	0.88	0.72	0.91	0.58	0.86	0.88	0.73	0.33
		0.5	0.89	0.69	0.91	0.49	1.02	1.05	0.87	0.33
	0.4	0.1	1.10	1.05	1.47	1.08	1.21	1.10	1.19	0.83
		0.2	1.34	1.31	1.57	1.09	1.45	1.35	1.31	0.83
		0.3	1.41	1.34	1.57	0.99	1.56	1.45	1.37	0.83
		0.5	1.44	1.30	1.55	0.85	1.93	1.73	1.59	0.83
	0.8	0.1	2.75	2.16	4.19	2.57	2.98	2.78	3.47	2.94
		0.2	3.35	3.89	5.42	3.53	4.86	4.06	5.08	4.15
		0.3	3.52	3.75	5.48	3.36	5.58	4.65	5.69	4.57
		0.5	3.61	3.39	5.07	2.82	6.95	5.22	6.20	4.81

Table 4: Median confidence interval length

δ	ADF	ERS	EJ	Pret_ADF	Pret_ERS	Level
		$c = 0$				
0.05	0.487	0.362	0.302	0.5174	0.520	0.531
0.1	0.855	0.698	0.583	0.523	0.529	0.705
0.15	1.172	1.038	0.854	0.520	0.527	0.862
0.2	1.499	1.422	1.143	0.519	0.526	1.007
0.3	2.300	2.351	1.777	0.518	0.525	1.321
		$c = -3$				
0.05	0.544	0.443	0.333	0.513	0.536	0.539
0.1	0.904	0.725	0.562	0.514	0.546	0.726
0.15	1.198	0.966	0.699	0.508	0.533	0.854
0.2	1.484	1.187	0.796	0.505	0.525	0.961
0.3	2.133	1.558	0.924	0.505	0.522	1.194
		$c = -10$				
0.05	0.570	0.472	0.442	0.488	0.578	0.509
0.1	0.856	0.608	0.562	0.466	0.614	0.607
0.15	1.048	0.640	0.570	0.441	0.542	0.592
0.2	1.190	0.615	0.544	0.432	0.503	0.573
0.3	1.417	0.527	0.452	0.430	0.483	0.595
		$c = -20$				
0.05	0.464	0.453	0.457	0.499	0.537	0.450
0.1	0.431	0.415	0.542	0.381	0.438	0.372
0.15	0.341	0.321	0.524	0.301	0.313	0.263
0.2	0.255	0.237	0.478	0.248	0.254	0.208
0.3	0.139	0.126	0.372	0.189	0.192	0.155

Table 5: Multivariate model with two local-to-unity roots

	ψ_{11}=0	ψ_{12}=0	ψ_{21}=0	ψ_{22}=0	ψ_{11}=0.2	ψ_{12}=0	ψ_{21}=0	ψ_{22}=0.3
δ	ϕ_{11}	ϕ_{22}	C	IRF	ϕ_{11}	ϕ_{22}	C	IRF
0.1	0.98	0.97	0.079	0.268	0.98	0.97	0.092	0.372
0.2				0.062				0.154
0.3				0.012				0.04
0.4				0.002				0.006
0.5				0				0.001
0.1		0.96	0.067	0.219		0.96	0.101	0.324
0.2				0.036				0.103
0.3				0.025				0.032
0.4				0.026				0.02
0.5				0.026				0.019
0.1		0.95	0.123	0.207		0.95	0.16	0.292
0.2				0.076				0.099
0.3				0.059				0.048
0.4				0.055				0.042
0.5				0.057				0.042
0.1	0.97	0.96	0.132	0.246	0.97	0.96	0.119	0.326
0.2				0.124				0.152
0.3				0.077				0.097
0.4				0.062				0.065
0.5				0.06				0.059
0.1		0.95	0.155	0.154		0.95	0.155	0.263
0.2				0.072				0.078
0.3				0.069				0.072
0.4				0.072				0.075
0.5				0.072				0.077
0.1	0.96	0.95	0.181	0.285	0.96	0.95	0.163	0.308
0.2				0.174				0.164
0.3				0.142				0.122
0.4				0.128				0.105
0.5				0.12				0.098

Table 6: Short and long horizon comparison

δ	Small h_L	Small h_R	Small h	Big h_L	Big h_R	Big h
0.01	0.016	0.08	0.096	0.244	0.388	0.633
0.02	0.003	0.057	0.06	0.101	0.342	0.443
0.03	0.001	0.041	0.042	0.068	0.238	0.306
0.04	0.001	0.037	0.038	0.057	0.164	0.221
0.05	0.002	0.035	0.036	0.049	0.114	0.163
0.06	0.002	0.036	0.039	0.048	0.086	0.135
0.07	0.004	0.038	0.042	0.047	0.072	0.119
0.08	0.005	0.041	0.046	0.046	0.067	0.113
0.09	0.007	0.042	0.049	0.046	0.064	0.11
0.10	0.01	0.043	0.053	0.047	0.063	0.11
0.15	0.02	0.049	0.069	0.049	0.062	0.111
0.2	0.028	0.051	0.079	0.05	0.061	0.111
0.25	0.033	0.053	0.086	0.051	0.061	0.111
0.3	0.037	0.054	0.092	0.052	0.06	0.112

Table 7: Unit root tests for real exchange rates (short-run identification example)

	GER	JAP	UK	ITA
ADF†	-1.542	-1.850	-2.470	-2.310
ERS‡	5.611	27.102	4.755	5.820
EJ	39.389	34.303	37.034	81.453
$\hat{R}^2$	0.134	0.202	0.192	0.126
EJ 5% c.v.	3.454	3.545	3.530	3.444
N. lags	6	6	6	6

(†) The 5% and 2.5% critical values for ADF are, respectively, -2.890 and -3.170.

(‡) The 5% and 2.5% critical values for ERS are, respectively, 3.11 and 2.47.

Figure 1(*a*) One minus coverage rates of various methods for AR(2) DGPs.

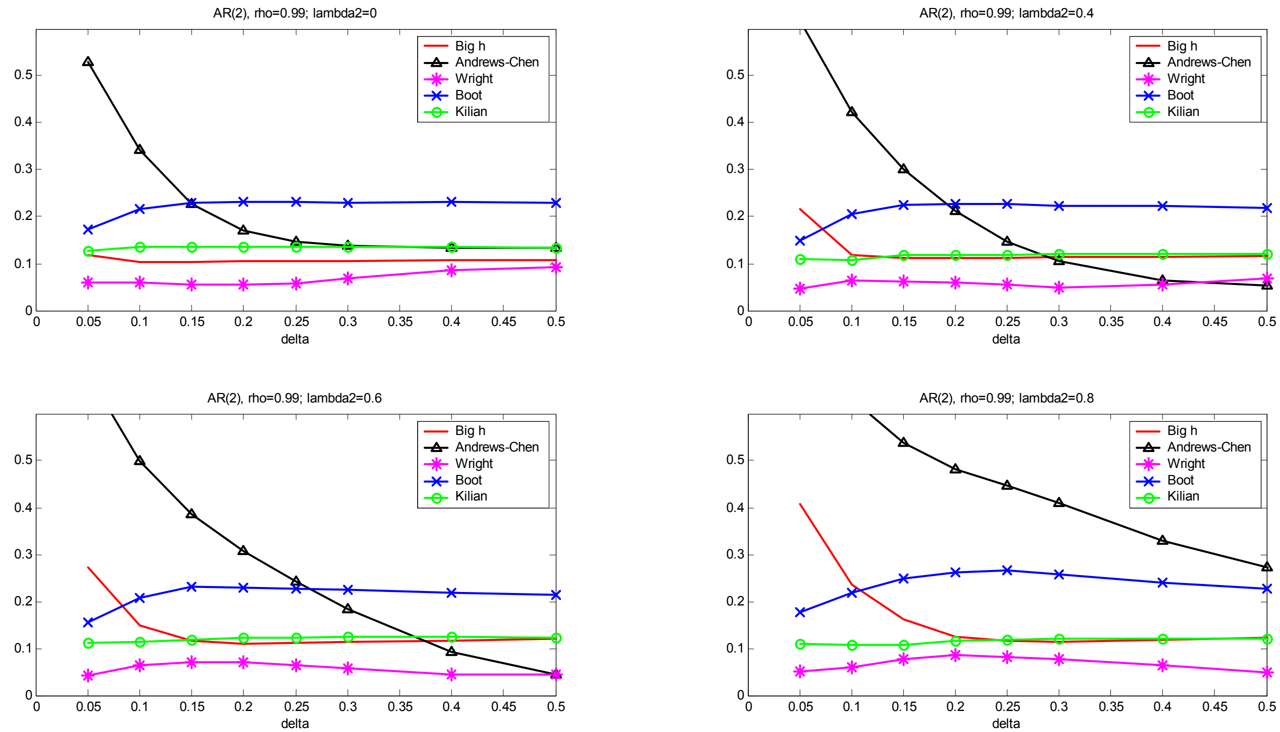


Figure 1(*b*) One minus coverage rates of various methods for AR(2) DGPs.

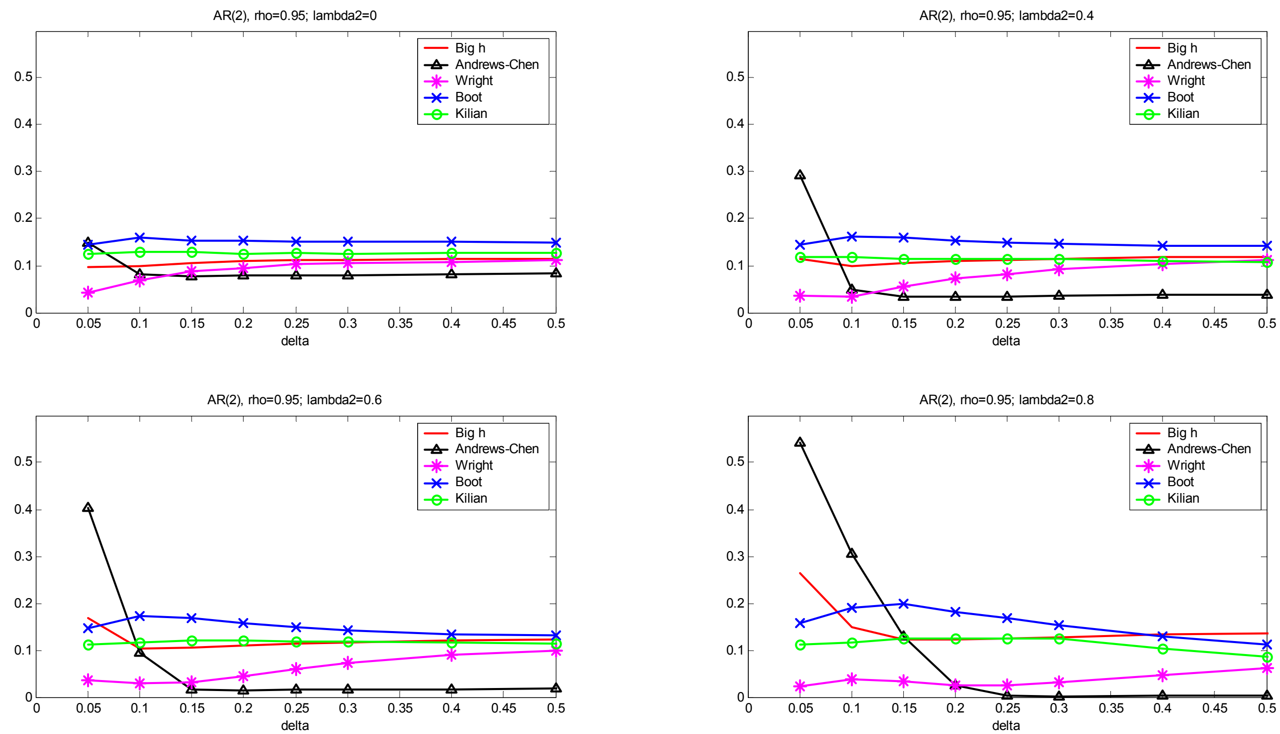


Figure 2(*a*) One minus coverage rates of various methods for AR(2) DGPs.

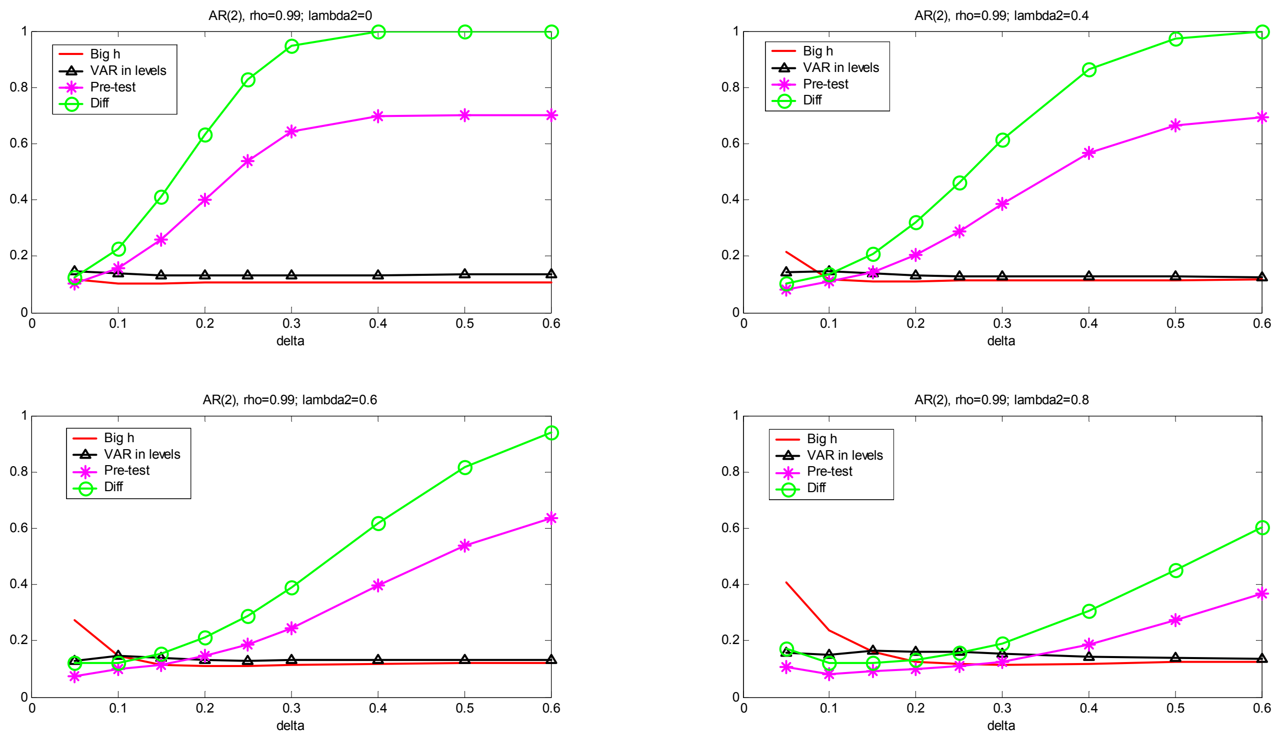


Figure 2(*b*) One minus coverage rates of various methods for AR(2) DGPs.

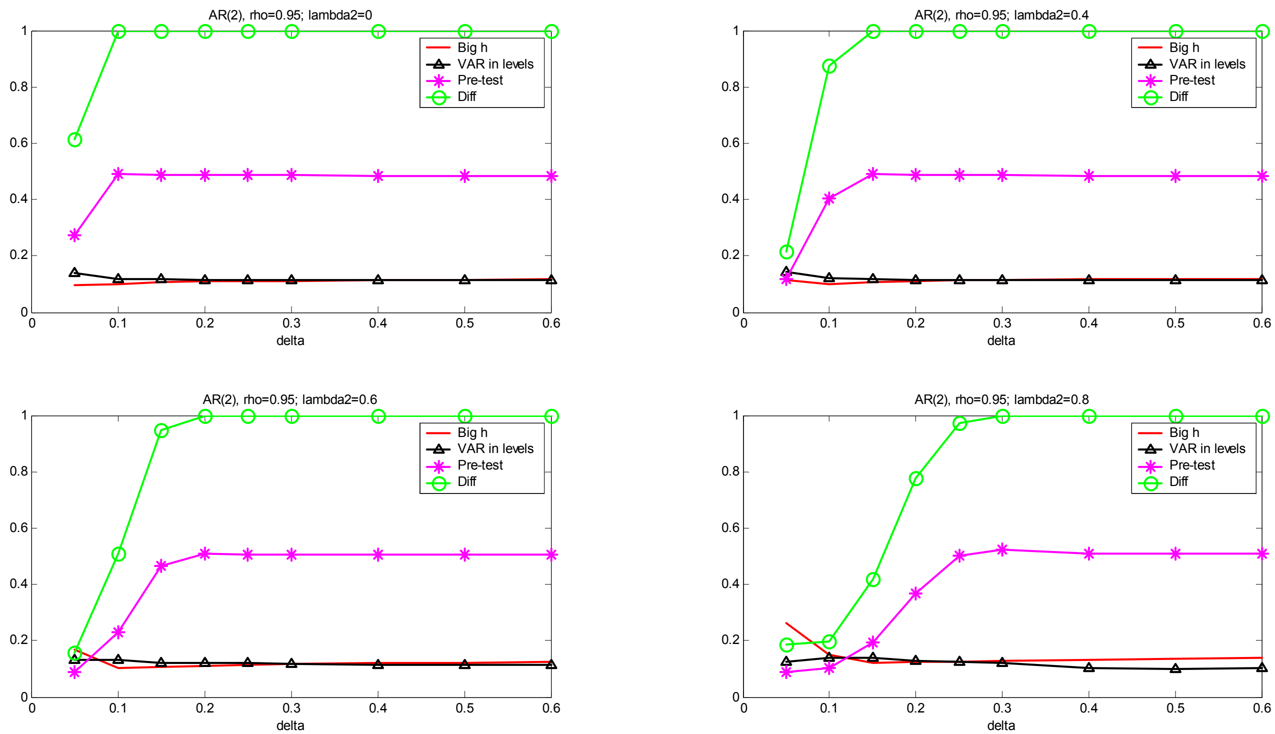
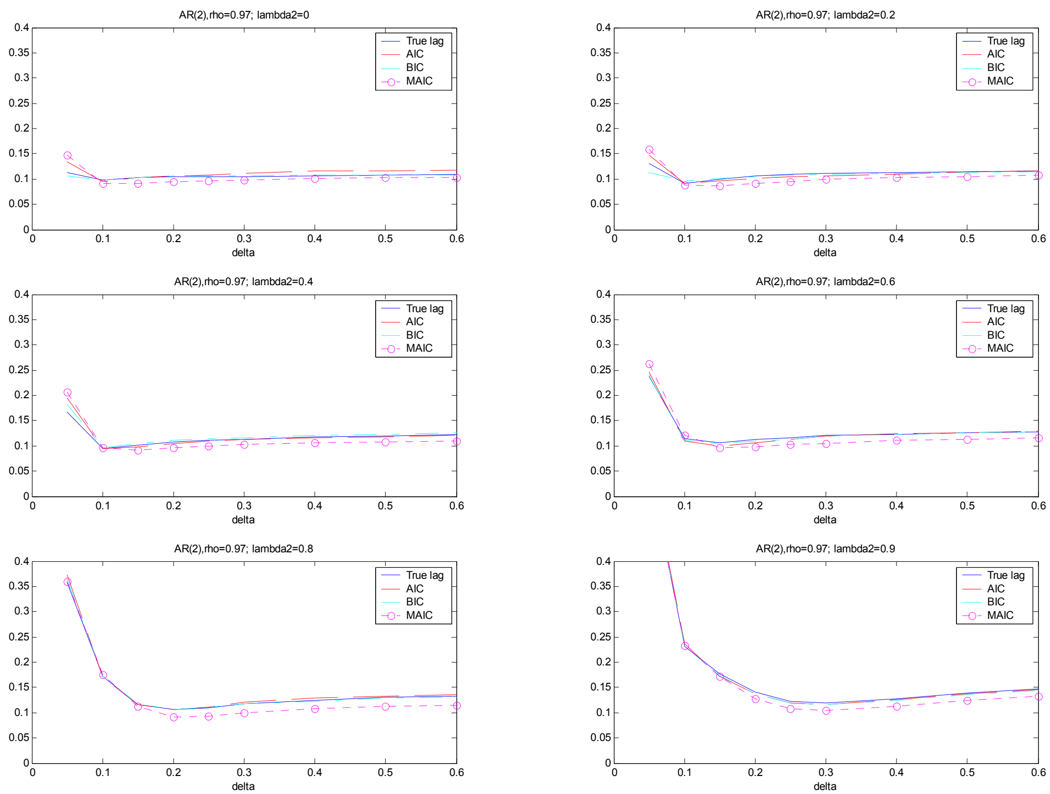
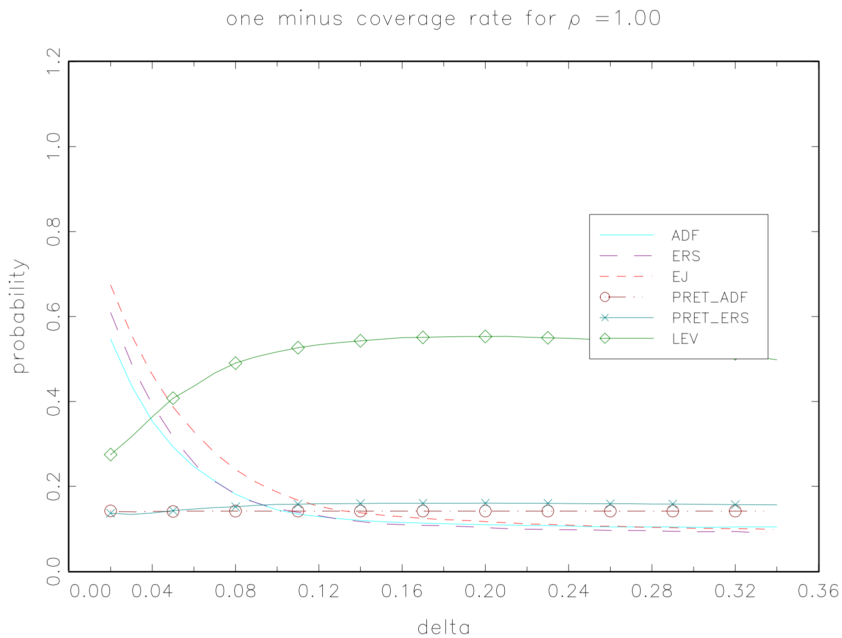


Figure 3



AR(2),rho=0.97; lambda2=0
AR(2),rho=0.97; lambda2=0.2
AR(2),rho=0.97; lambda2=0.4
AR(2),rho=0.97; lambda2=0.6
AR(2),rho=0.97; lambda2=0.8
AR(2),rho=0.97; lambda2=0.9
True lag
AIC
BIC
MAIC
delta

Figure 4(*a*) : one minus coverage rate for various values of ρ, $R^2 = 0.5$



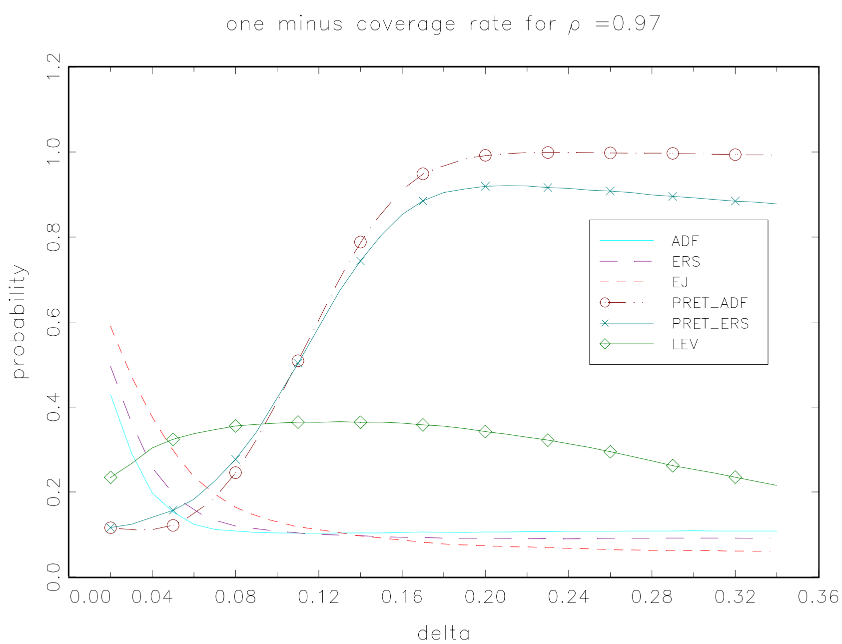
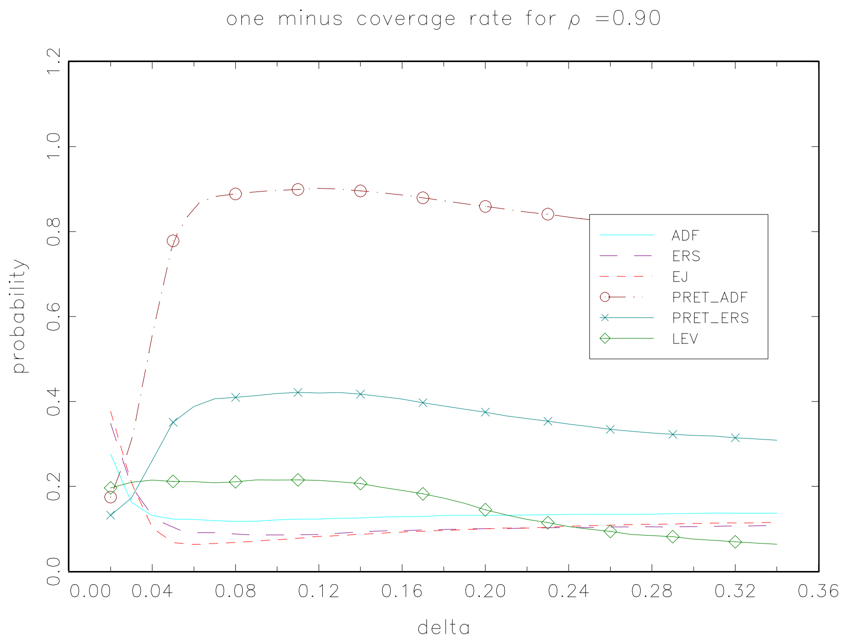


Figure 4(*b*) : one minus coverage rate for various values of ρ, $R^2 = 0.5$



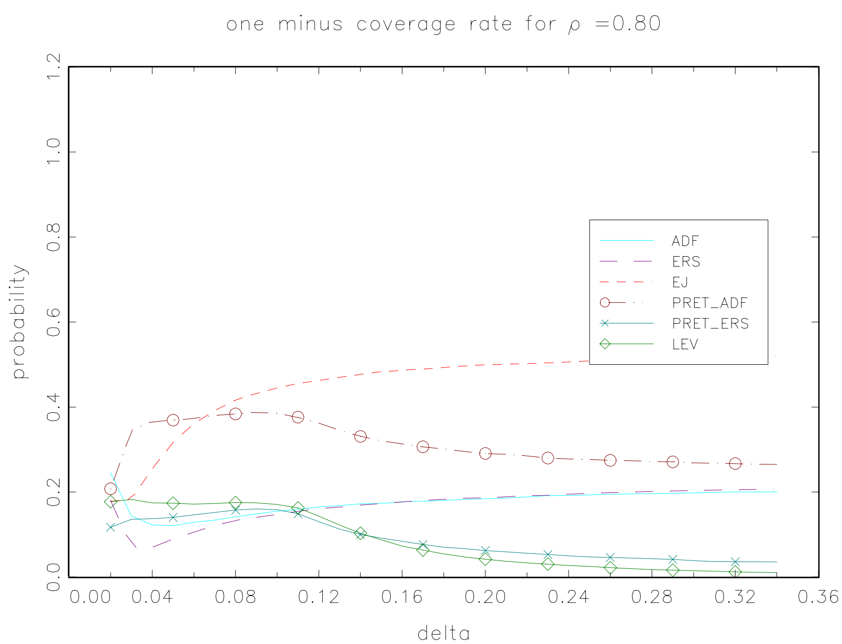
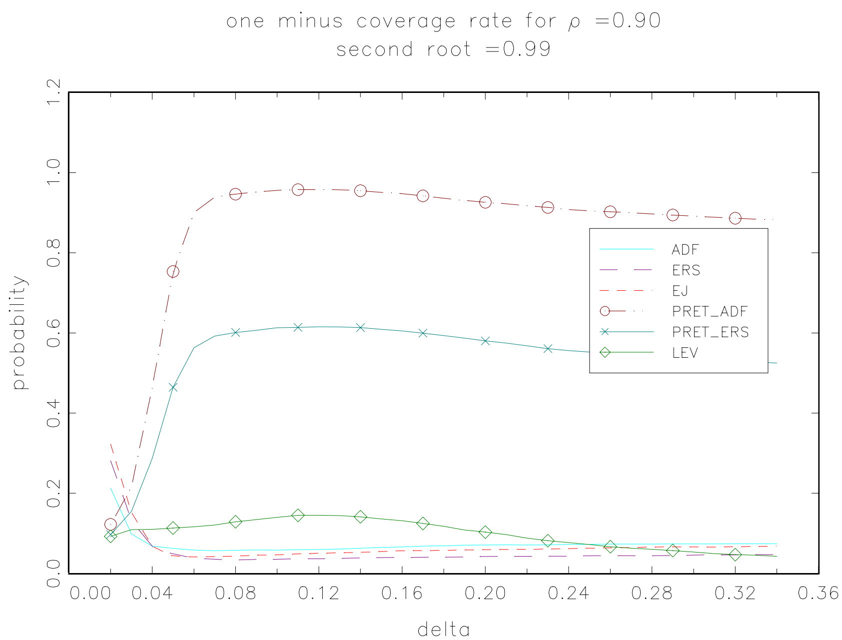


Figure 5 : one minus coverage rate for various values of $\rho = 0.90$, $R^2 = 0.5$ for different values of the second root



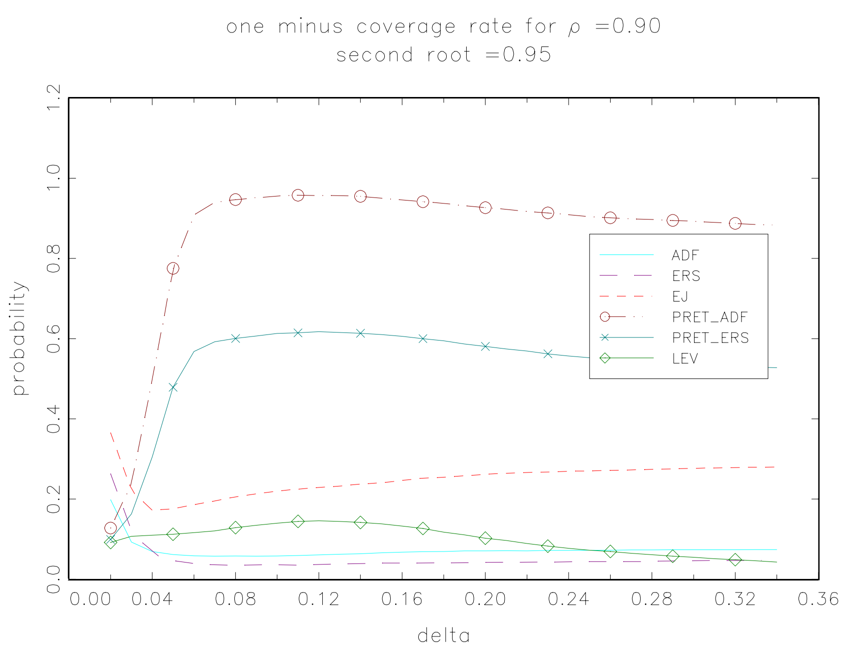
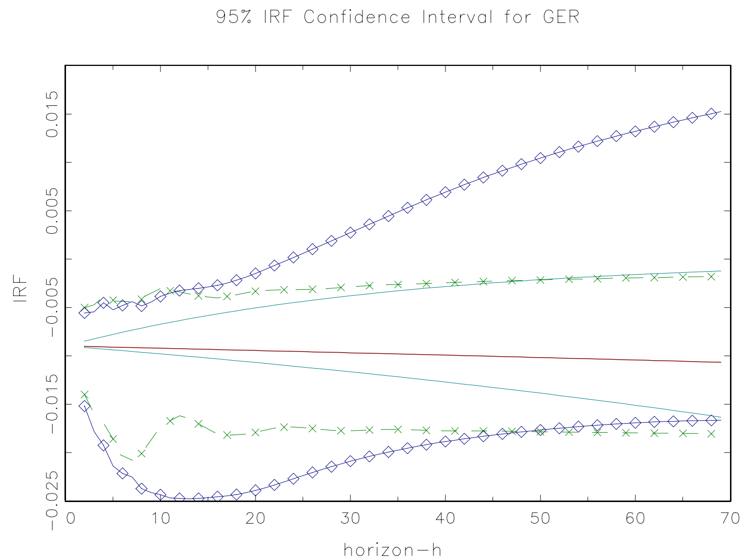
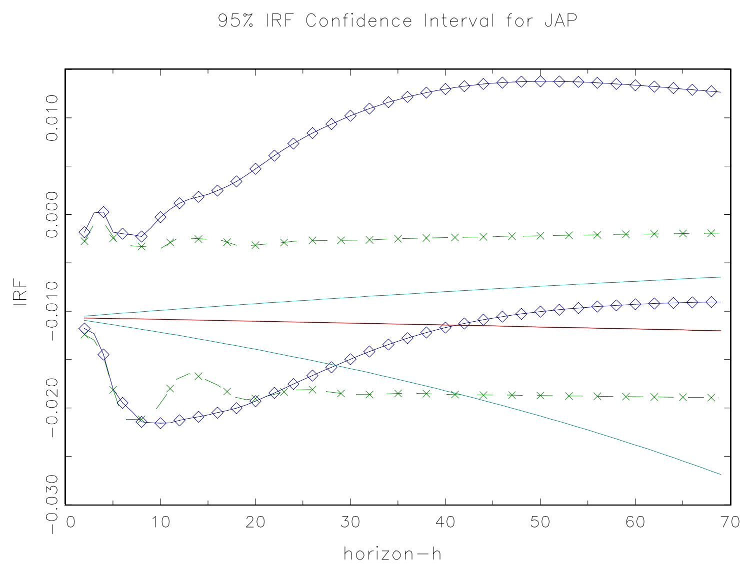


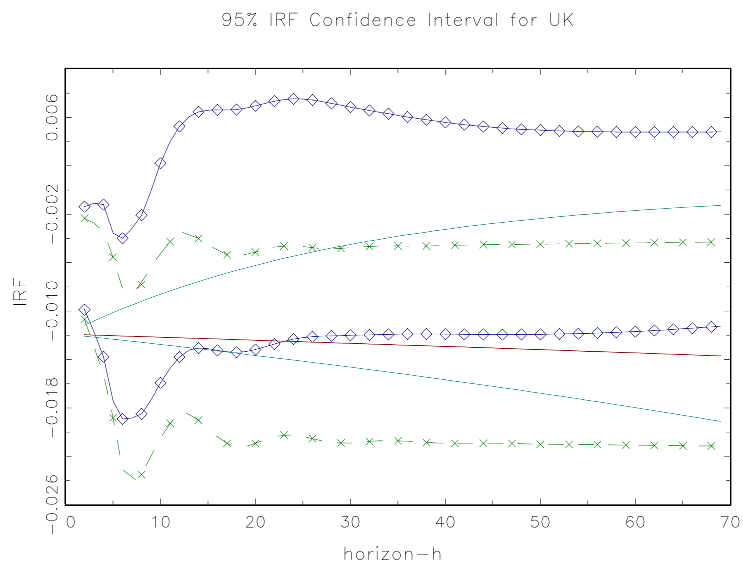
Figure 6(*a*) : Confidence Intervals for a response of q_t to a monetary shock†

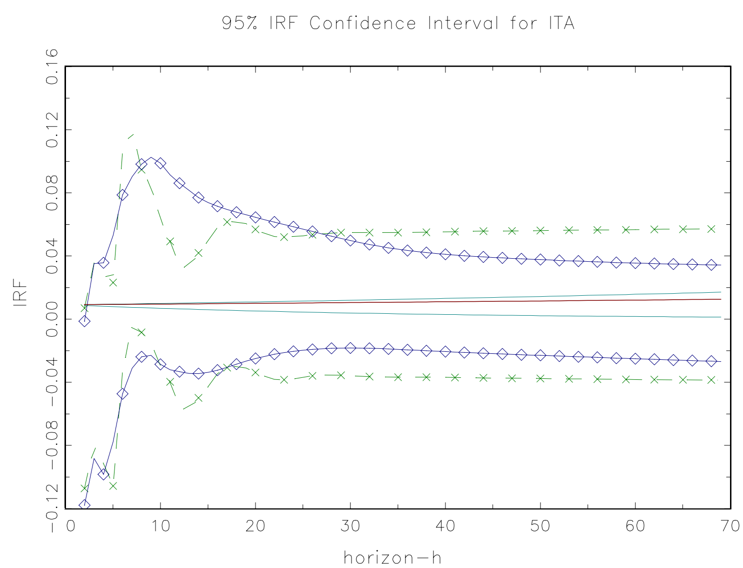




†The confidence bands are the following: VAR levels (solid line with diamonds), ERS (solid line - the central, thickest line is the median unbiased IRF estimate), VAR first differences (dotted line with stars).

Figure 6(*b*) : Confidence Intervals for a response of q_t to a monetary shock†





†The confidence bands are the following: VAR levels (solid line with diamonds), ERS (solid line - the central, thickest line is the median unbiased IRF estimate), VAR first difference (dotted line with stars).