Finite-sample simulation-based inference in VAR models with applications to order selection and causality testing*

Jean-Marie Dufour † and Tarek Jouini ‡
Université de Montréal

First version: January 1999
Revised: May 2000, July 2001, December 2003
This version: August 13, 2005
Compiled: August 13, 2005, 12:44pm

A shorter version of this paper is forthcoming in the *Journal of Econometrics*.

* The authors thank an anonymous referee and the Editors, Graham Elliott and Norman Swanson, for several useful comments. This work was supported by the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Alexander-von-Humboldt Foundation (Germany), the Institut de finance mathématique de Montréal (IFM2), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Canada Council for the Arts (Killam Fellowship), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, the Fonds de recherche sur la société et la culture (Québec), and the Fonds de recherche sur la nature et les technologies (Québec).

† Canada Research Chair Holder (Econometrics). Centre interuniversitaire de recherche en analyse des organisations (CIRANO), Centre interuniversitaire de recherche en économie quantitative (CIREQ), and Département de sciences économiques, Université de Montréal. Mailing address: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre-ville, Montréal, Québec, Canada H3C 3J7. TEL: 1 (514) 343 2400; FAX: 1 (514) 343 5831; e-mail: jean.marie.dufour@umontreal.ca . Web page: http://www.fas.umontreal.ca/SCECO/Dufour

‡ CIRANO, CIREQ, and Département de sciences économiques, Université de Montréal. Mailing address: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre-ville, Montréal, Québec, Canada H3C 3J7. TEL.: 1 (514) 343-6111, ext. 1814; FAX: 1 (514) 343 5831. E-mail: jouinit@magellan.umontreal.ca

ABSTRACT

Statistical tests in vector autoregressive (VAR) models are typically based on large-sample approximations, involving the use of asymptotic distributions or bootstrap techniques. After documenting that such methods can be very misleading even with fairly large samples, especially when the number of lags or the number of equations is not small, we propose a general simulation-based technique that allows one to control completely the level of tests in parametric VAR models. In particular, we show that maximized Monte Carlo tests [Dufour (2005, *Journal of Econometrics*)] can provide provably exact tests for such models, whether they are stationary or integrated. Applications to order selection and causality testing are considered as special cases. The technique developed is applied to quarterly and monthly VAR models of the U.S. economy, comprising income, money, interest rates and prices, over the period 1965-1996.

Key-words : Vector autoregression; VAR; exact test; Monte Carlo test; maximized Monte Carlo test; bootstrap; Granger causality; order selection; nonstationary model; macroeconomics; money and income; interest rate; inflation.

Journal of Economic Literature classification: C32, C12, C15, E4, E5.

RÉSUMÉ

Les tests statistiques sur des modèles autorégressifs multivariés (VAR) sont habituellement basés sur des approximations de grands échantillons, qui utilisent une loi asymptotique ou une technique de bootstrap. Après avoir montré que ces méthodes peuvent être très peu fiables, même avec des échantillons de taille assez grande, particulièrement lorsque le nombre des retards ou le nombre d'équations augmente, nous proposons une technique générale basée sur la simulation qui permet de contrôler parfaitement le niveau des tests dans les modèles VAR paramétriques. En particulier, nous montrons que la technique des tests de Monte Carlo maximisés [Dufour (2005, *Journal of Econometrics*)] fournit des tests exacts pour de telles modèles, que ceux-ci soient stationnaires ou intégrés. Sélectionner l'ordre du modèle ainsi que tester la causalité au sens de Granger sont étudiés comme problèmes particuliers dans ce cadre. La technique proposée est appliquée à des modèles VAR, trimestriels et mensuels, de l'économie américaine, comprenant le revenu, la monnaie, un taux d'intérêt et le niveau des prix, sur la période 1965-1996.

Mots-clés: autorégression vectorielle; VAR; test exact; test de Monte Carlo; test de Monte Carlo maximisé; bootstrap; causalité au sens de Granger; sélection de l'ordre; modèle non-stationnaire; macroéconomie; monnaie et revenu; taux d'intérêt; inflation.

Classification du Journal of Economic Literature: C32, C12, C15, E4, E5.

Contents

List of Tables

1. Introduction

Vector autoregressive (VAR) models are widely used for multivariate time series analysis, especially in econometrics; see Sims (1980), Lütkepohl (1993, 2001), Reinsel (1993), Hamilton (1994), Hendry (1995), Gouriéroux and Monfort (1997), Dhrymes (1998) and Clements and Hendry (2002). One reason for this popularity is that VAR models are easy to estimate and can account for relatively complex dynamic phenomena. Important features and applications based on such models include forecasting, causality analysis [in the sense of Wiener (1956) and Granger (1969)], impulse responses, cointegration, etc.

VAR models, however, typically involve large numbers of parameters, so that the usual statistical difficulties associated with dynamic models are compounded by high dimensionality. Not surprisingly, statistical inference (tests and confidence sets) in such models is almost universally based on large-sample approximations. Even in static multivariate linear regression models, it is well-known that asymptotic approximations can be very unreliable; see Dufour and Khalaf (2002). These problems get worse in VAR models (which can be interpreted as dynamic multivariate linear regressions), even if strong regularity assumptions (e.g., stationarity) are made: finite-sample distributions of usual test statistics are complicated and depend on large numbers of unknown nuisance parameters. Further, the presence of non-stationary variables – such as integrated processes – can affect the asymptotic distributions and lead to further reliability problems; see, for example, Sims, Stock, and Watson (1990), Johansen (1995), Hatanaka (1996), Tanaka (1996), Dhrymes (1998), Hansen and Johansen (1998), Maddala and Kim (1998), and McAleer and Oxley (1999). In particular, the appropriate asymptotic distribution may depend on unknown features of the process (e.g., the integration order or the number of cointegrating relationships). This is the case, for example, in causality testing; see Sims, Stock, and Watson (1990) and Toda and Phillips (1993, 1994).

In view of alleviating the unreliability of asymptotic distributions for inference in VAR models, bootstrap techniques [see Efron and Tibshirani (1993) and Hall (1992)] have also been proposed; see, for example, Jeong and Maddala (1993), Li and Maddala (1996), Paparoditis (1996), Berkowitz and Kilian (2000), Kilian and Demiroglu (1997), Kilian (1998b, 1998a), Caner and Kilian (1999), Inoue and Kilian (2002a, 2002b). Bootstrap methods can lead to spectacular improvements over standard asymptotic approximations, but their justification remains intrinsically asymptotic. Further, it is well known that bootstrapping can fail to provide asymptotically valid tests when the simulated test statistic has an asymptotic distribution involving nuisance parameters, especially if the asymptotic distribution has discontinuities with respect to the nuisance parameters; see Athreya (1987), Basawa, Mallik, McCormick, Reeves, and Taylor (1991), Sriram (1994), Andrews (2000), Benkwitz, Lütkepohl, and Neumann (2000), and Inoue and Kilian (2002a, 2003). This type of situation can easily occur in VAR models.

In this paper, we propose a finite-sample simulated-based inference technique applicable to parametric finite-order VAR models that allows one to control completely the level of the test, despite the presence of large numbers of nuisance parameters and without further assumptions on the structure of the process [such as stationarity or the order of integration]. The disturbances in the VAR model may follow any distribution that is specified up to a linear transformation [or covariance matrix], which allows for both Gaussian and non-Gaussian distributions. The central assumption is that the model can be simulated once a finite number of parameters have been specified. The technique proposed is based on an extension of the old technique of Monte Carlo (MC) tests [Dwass (1957), Barnard (1963), Birnbaum (1974)], which we call *maximized Monte Carlo* (MMC) tests [Dufour (2005)]. This method involves maximizing a simulated p-value function over the nuisance parameter space. Two main variants of this method are considered: the first one maximizes the simulated p-value function over the full nuisance parameter space and yields provably exact tests of general restrictions on model parameters, while the second variant considers a maximization restricted to a *consistent set estimator* of the nuisance parameters. The latter can be viewed as an approximate simplified version of the fully exact procedure (because the set over which the p-value function is maximized can be much smaller); it provides asymptotically valid tests without the need to establish the asymptotic distribution of the test statistic or to make further assumptions on the structure of the process. We also consider *local Monte Carlo* (LMC) tests which can be viewed as a degenerate version of the simplified maximized procedure, obtained by replacing the consistent set estimator by a consistent point estimate, and may be interpreted as parametric bootstrap test. Of course, the latter procedure is not exact in finite samples and requires stronger assumptions (to yield asymptotically valid tests) than the consistent set MC procedure.

The method proposed is obviously computer intensive, and an important contribution of this paper consists in showing that the proposed theoretical approach can indeed be implemented in a high-dimensional setup, such as a VAR model. For that purpose, we focus on likelihood ratio (LR) tests for two categories of hypotheses: (1) the order of a VAR model; (2) Granger non-causality. We present simulation evidence on tests for Granger non-causality which document three things. First, standard tests based on asymptotic critical values can have catastrophic size properties. Second, the LMC approach (or bootstrapping) does provide improvements from that viewpoint, but can still allow for sizeable overrejection rates. Third, under the same circumstances, the MMC approach does control level perfectly, even we only use a consistent set estimator for the nuisance parameters, and provides good power. In other words, the maximization operated by the MMC approach yields effective corrections for possible failures of the bootstrap (both in finite samples and asymptotically). We also apply the proposed method to causality testing in a quarterly model of the U.S. economy, based on data previously studied in Bernanke and Mihov (1998) and Dufour, Pelletier, and Renault (2005), involving nonborrowed reserves, the federal funds rate, real gross domestic product, and the GDP deflator.

The paper is organized as follows. Section 2 describes the model and the main test problems that will be studied. Section 3 presents the principles of MC tests and MMC tests. In section 4, we discuss how such procedures can be applied to LR-type tests in VAR models. The results of our simulation study are presented in section 5 and the empirical macroeconomic application in section 6. We conclude in section 7.

2. Framework

In this paper we consider a general k-dimensional finite-order vector autoregressive (VAR) process $\{Y_t : t \in \mathbb{Z}\}$ of the form:

$$
\begin{aligned}
Y_t &= \mu + \sum_{i=1}^{p} \varPhi_i Y_{t-i} + u_t\,, \quad t = 1,\, \ldots\,,\, T\,, & (2.1)\\
u_t &= R\varepsilon_t\,, \quad t = 1,\, \ldots\,,\, T\,, & (2.2)
\end{aligned}
$$

where the vectors $Y_t = (Y_{1t},\, \ldots\,,\, Y_{kt})'$, $t = -p+1,\, \ldots\,,\, T$, are observable, p is a specified nonnegative integer $(p \geq 1)$, $\mu = (\mu_1,\, \ldots\,,\, \mu_k)'$ is an unknown $k \times 1$ vector of intercept terms, $\varPhi_i = [\varPhi_{ijl}]_{j,l=1,\ldots,k}$ is an unknown $k \times k$ matrix of fixed coefficient matrices $(1 \leq i \leq p)$, R is an *unknown* fixed non-singular matrix, and the conditional distribution of the vector $\varepsilon(T) = \mathrm{vec}(\varepsilon_1,\, \ldots\,,\, \varepsilon_T)$, given the initial values $Y_0,\, \ldots\,,\, Y_{-p+1}$, is completely specified. A common special case here would consist in assuming that

$$
\varepsilon_t \overset{i.i.d.}{\sim} N[0\,,\, I_k]\,,\ \ t = 1,\, \ldots\,,\, T\,, \tag{2.3}
$$

given the initial values, so that the errors are independent and identically distributed (*i.i.d.*) according to a multinormal distribution $N[0\,,\, \varSigma]$ with $\varSigma = RR'$. But, from the viewpoint of implementing the procedures proposed in this paper, the assumptions (2.1) - (2.2) will be sufficient. For example, there is no need to assume normality or even the existence of moments.

Setting

$$
\varPhi\,(z) = I_k - \sum_{i=1}^{p} \varPhi_i z^i\,, \quad z \in \mathbb{C}\,, \tag{2.4}
$$

the model is said to be *stationary* if

$$
\det\{\varPhi\,(z)\} \neq 0 \text{ for all } |z| \leq 1\,, \tag{2.5}
$$

and it is *stable* (non-explosive) if

$$
\det\{\varPhi\,(z)\} \neq 0 \text{ for all } |z| < 1\,; \tag{2.6}
$$

stable models allow for the presence of roots on the unit circle for the equation $\det\{\varPhi\,(z)\} = 0$. Note, however, that stationarity (or stability) assumptions will not be needed for the validity of the procedures proposed in this paper, so cointegrating relations may be present. The central feature we shall exploit is the fact that the model can be easily simulated, once a *finite* number of parameters is specified.

In this paper, we consider the problem of testing general hypotheses of the form

$$
H_0 : \mathrm{vec}(\varPhi) \in \varGamma_0 \tag{2.7}
$$

where $\varPhi = [\varPhi_1,\, \ldots\,,\, \varPhi_p]$ and $\varGamma_0 \subseteq \mathbb{R}^{k^2 p}$. This covers both linear and nonlinear hypotheses on model coefficients. In our simulations and applications, however, we shall focus on linear hypotheses, more precisely:

1. hypotheses on individual coefficients:

$$H_0(\Phi^0_{ijl}) : \Phi_{ijl} = \Phi^0_{ijl} \; ; \tag{2.8}$$

2. hypotheses on the order of the process:

$$H_0(i) : \Phi_i = 0 \tag{2.9}$$

   and

$$H_0[i\,,\,p] : \Phi_i = \cdots = \Phi_p = 0\,; \tag{2.10}$$

3. non-causality in the sense of Granger (1969):

$$H_0(Y_l \nrightarrow Y_j) : \Phi_{ijl} = 0\,, \quad i = 1,\,\ldots\,,\,p\,, \tag{2.11}$$

The distribution of most standard test statistics [such as Wald-type, LM-type or LR-type statistics] for hypotheses on the coefficients of VAR models typically depends (under the null hypothesis) on both the hypothesis tested and unknown nuisance parameters. To be more precise, if we denote by $\mathcal{H}_0$ the set of data distributions F – or data generating processes (DGP's) – compatible with H_0, the null hypothesis can be written in the form

$$H_0 : F \in \mathcal{H}_0\,. \tag{2.12}$$

Then a test of H_0 has *level* α iff

$$\mathsf{P}_F[\text{Rejecting } H_0] \leq \alpha \text{ for all } F \in \mathcal{H}_0 \tag{2.13}$$

or, equivalently,

$$\sup_{F\in\mathcal{H}_0} \mathsf{P}_F[\text{Rejecting } H_0] \leq \alpha\,, \tag{2.14}$$

and the test has *size* α iff

$$\sup_{F\in\mathcal{H}_0} \mathsf{P}_F[\text{Rejecting } H_0] = \alpha\,; \tag{2.15}$$

see Lehmann (1986, Chapter 3). If we also had

$$\mathsf{P}_F[\text{Rejecting } H_0] = \alpha \text{ for all } F \in \mathcal{H}_0\;, \tag{2.16}$$

the test would be *similar*. But, in complex models, this appears extremely difficult to achieve with any reasonable procedure that depends on the data. So we will focus on designing tests that satisfy as closely as possible the level restriction (2.13) in finite samples. So one needs a method that can adapt readily to both these features. We will now describe such a method.

3. Monte Carlo tests

In this section, we describe in a succinct way the general approach that will allow us to construct finite-sample tests for any VAR model, such as the one described in section 2. To ensure clarity, we describe first the basic principle underlying Monte Carlo tests by considering two basic cases: (1) the distribution of the test statistic under the null hypothesis does not depend on nuisance parameters; (2) the distribution of the test statistic depends on nuisance parameters. Of course, the second case is the relevant one for inference in VAR models. To deal with it, we consider three alternative approaches: (a) *maximized Monte Carlo* (MMC) tests over the full nuisance parameter case; (b) MMC tests over a consistent set estimator of the nuisance parameters; (c) *local Monte Carlo* tests, *i.e.* Monte Carlo tests obtained after replacing the unknown nuisance parameters by a point estimate. To simplify exposition, we limit ourselves to the case where the test statistic has a continuous distribution, although it is relatively easy to extend Monte Carlo test methods to situations where the statistic follows a discrete distribution under the null hypothesis. Further details and proofs are provided in Dufour (2005) and Dufour and Khalaf (2001).

3.1. Monte Carlo tests without nuisance parameters

Let $S \equiv S(Y_1, \ldots, Y_T)$ be a test statistic for testing an hypothesis H_0, with critical region of the form:

$$S \geq c\,. \tag{3.1}$$

We will denote by S_0 the value of the test statistic based on the observed data. Suppose now that the distribution of S under H_0 does not depend on unknown nuisance parameters (and is continuous). The test has level α if

$$\mathsf{P}[S_0 \geq c] \leq \alpha \tag{3.2}$$

and size α if

$$\mathsf{P}[S_0 \geq c] = \alpha\,. \tag{3.3}$$

Suppose now we can generate by simulation N i.i.d. replications of S under H_0,

$$S_1, \ldots, S_N \tag{3.4}$$

independently of S_0. We can then estimate the survival function

$$G(x) = \mathsf{P}[S \geq x] \tag{3.5}$$

from the simulated samples:

$$\hat{G}_N\left[x;\, S\left(N\right)\right] = \frac{1}{N}\sum_{i=1}^{N} s\left(S_i - x\right) \tag{3.6}$$

where

$$S\left(N\right) = \left(S_1, \ldots, S_N\right)'\,, \tag{3.7}$$

$$
\begin{aligned}
s(x) &= 1\,, \text{ if } x \geq 0\,, \\
&= 0\,, \text{ if } x < 0\,.
\end{aligned}
\tag{3.8}
$$

Let us also consider

$$
\hat{p}_N(x) = \frac{N\hat{G}_N[x;\, S(N)] + 1}{N+1}
\tag{3.9}
$$

the simulated p-value function associated with $S(N)$. Then, if N is chosen so that $\alpha(N+1)$ is an integer, it can be shown that, under H_0,

$$
\mathsf{P}\big[\hat{p}_N(S_0) \leq \alpha\big] = \alpha\,.
\tag{3.10}
$$

In other words, the test which rejects H_0 when $\hat{p}(S_0) \leq \alpha$ has level α exactly.

3.2. Monte Carlo tests with nuisance parameters

We will now study the case where the distribution of the test statistic depends on nuisance parameters. In other words, we consider a model $\{(\mathcal{Z}\,,\, \mathcal{A}_{\mathcal{Z}}\,,\, \mathsf{P}_\theta) : \theta \in \Omega\}$ where we assume that the distribution of S is determined by $\mathsf{P}_{\bar{\theta}}$ [i.e., $\bar{\theta}$ is the "true" parameter vector]. We wish to test the hypothesis

$$
H_0 : \bar{\theta} \in \Omega_0\,,
\tag{3.11}
$$

where $\Omega_0 \subset \Omega$. The critical region $\{S \geq c\}$, where c is a constant which does not depend of θ, has level α if and only if

$$
\mathsf{P}_\theta[S \geq c] \leq \alpha\,,\ \forall \theta \in \Omega_0\,,
\tag{3.12}
$$

or equivalently

$$
\sup_{\theta \in \Omega_0} G[S \mid \theta] \leq \alpha
\tag{3.13}
$$

where

$$
G[x \mid \theta] = \mathsf{P}_\theta[S \geq x]\ .
\tag{3.14}
$$

Suppose now that, for each $\theta \in \Omega_0$, we generate N i.i.d. replications of S,

$$
S(N,\, \theta) = [S_1(\theta)\,,\, \ldots\,,\, S_N(\theta)]'
$$

and compute a simulated p-value function:

$$
\hat{p}_N[x \mid \theta] = \frac{N\hat{G}_N[x;\, S(N,\, \theta)] + 1}{N+1}\,.
\tag{3.15}
$$

If $\alpha(N+1)$ is an integer, then, under H_0,

$$
\mathsf{P}\big[\sup\{\hat{p}_N[S_0 \mid \theta] : \theta \in \Omega_0\} \leq \alpha\big] \leq \alpha\,,
\tag{3.16}
$$

which means that the critical region $\sup\{\hat{p}_N[S_0 \mid \theta] : \theta \in \Omega_0\} \leq \alpha$ has level α. This procedure will be called a *maximized Monte Carlo test*. It allows one to obtain provably exact tests based on any statistic that can be simulated once a finite number of nuisance parameters have been specified.

The simulated p-value function $\hat{p}_N [S_0 \mid \theta]$ is not continuous, so standard gradient based methods cannot be used to maximize it. But search methods applicable to non-differentiable functions are applicable, e.g. *simulated annealing* [see Goffe, Ferrier, and Rogers (1994)].

3.3. MMC tests based on consistent set estimators

Suppose now that the test statistic depends on a sample of size T,

$$S = S_T \,, \tag{3.17}$$

and we have a consistent set estimator C_T of θ (under H_0) :

$$\lim_{T \to \infty} \mathsf{P}\left[\bar{\theta} \in C_T\right] = 1\,. \tag{3.18}$$

For example, if $\hat{\theta}_T$ is a consistent point estimate of $\bar{\theta}$ and c is any positive constant, the set

$$C_T = \{\theta \in \Omega : \| \,\hat{\theta}_T - \theta\, \| < c\} \tag{3.19}$$

is a consistent set estimator of $\bar{\theta}$:

$$\lim_{T \to \infty} \mathsf{P}\big[\bar{\theta} \in C_T\big] = \lim_{T \to \infty} \mathsf{P}\big[\|\hat{\theta}_T - \bar{\theta}\| < \varepsilon\big] = 1\,,\ \forall \varepsilon > 0\,. \tag{3.20}$$

Assuming that, for each $\theta \in \Omega_0$, we can generate N i.i.d. replications of S_T, say $S_{T1}(\theta)\,, \ldots ,$ $S_{TN}(\theta)$, we have, under H_0 :

$$\lim_{T \to \infty} \mathsf{P}\big[\sup \{\hat{p}_{TN}[S_{T0} \mid \theta] : \theta \in C_T\} \leq \alpha\big] \leq \alpha \tag{3.21}$$

where S_{T0} is the value of S_T based on the observed data, and

$$\hat{p}_{TN}[x \mid \theta] = \frac{N\hat{G}_N[x;\, S_T(N,\, \theta)] + 1}{N+1}\,, \quad S_T(N,\, \theta) = [S_{T1}(\theta)\,, \ldots,\, S_{TN}(\theta)]'\,. \tag{3.22}$$

In other words, the critical region $\sup \{\hat{p}_{TN}[S_{T0} \mid \theta] : \theta \in C_T\} \leq \alpha$ has level α. No assumption on the asymptotic distribution of S_T is needed.

An obvious alternative would consist in taking

$$C_T = \{\hat{\theta}_T\}\,, \tag{3.23}$$

which suggests one to use a critical region of the form

$$\hat{p}_{TN}[S_{T0} \mid \hat{\theta}_T] \leq \alpha\,. \tag{3.24}$$

We shall call this procedure a *local Monte Carlo* test. Under additional regularity conditions, it is possible to show that

$$\lim_{T\to\infty} \mathsf{P}\big[\hat{p}_{TN}[S_{T0} \mid \hat{\theta}_T] \le \alpha\big] \le \alpha \tag{3.25}$$

but the conditions under which this holds are notably more restrictive than those under which (3.21) obtains. This procedure may also be interpreted as a *parametric bootstrap*, except for the fact that the number of replications N is explicitly taken into account.

A good consistent restricted estimate $\hat{\theta}_T$ is typically a reasonable starting point for computing the MMC p-value. Since

$$\hat{p}_{TN}\left[S_{T0} \mid \hat{\theta}_T\right] \le \sup\left\{\hat{p}_{TN}\left[S_{T0} \mid \theta\right] : \theta \in \Omega_0\right\}\,, \tag{3.26}$$

it is clear that

$$\hat{p}_{TN}\left[S_{T0} \mid \hat{\theta}_T\right] > \alpha \tag{3.27}$$

implies

$$\sup\left\{\hat{p}_{TN}\left[S_{T0} \mid \theta\right] : \theta \in \Omega_0\right\} > \alpha\,. \tag{3.28}$$

A non-significant bootstrap p-value entails a non-significant MMC p-value. The MMC procedure offers protection against failures of the bootstrap.

4. Tests in VAR models

We will now consider the problem of testing restrictions on the coefficients Φ of model (2.1) - (2.2). Even though various procedures, such as Wald-type, LM-type or LR-type tests, may be used, we will concentrate here on LR tests based on statistics of the form:

$$LR = 2\left[\ln L(\hat{\delta}) - \ln L(\hat{\delta}^0)\right] \tag{4.1}$$

where $L(\cdot)$ is the likelihood function, $\hat{\delta}$ is the unconstrained maximum likelihood (ML) estimator of parameter vector $\delta \equiv \mathrm{vec}[\mu\,,\, \Phi\,,\, R]$ obtained by maximizing the likelihood function over the full feasible parameter space, and $\hat{\delta}^0$ is the constrained ML estimator. Since a specific likelihood function must be specified, we shall focus on Gaussian LR statistics.

Under the assumption that the errors $u_t\,,\ t = 1, \dots, T$, conditional on the initial values $Y_{-p} = vec[Y_0, \dots, Y_{-p+1}]$, are i.i.d. $N[0\,,\,\Sigma]$, the likelihood function is

$$L(\delta) = \kappa - \frac{T}{2}\ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{T} u_t'\Sigma^{-1}u_t \tag{4.2}$$

where κ is a constant and

$$u_t = Y_t - \mu - \sum_{i=1}^{p} \Phi_i Y_{t-i}\,, \quad t = 1, \dots, T\,. \tag{4.3}$$

Then the (conditional) LR statistic for testing any hypothesis of the form $H_0 : \mathrm{vec}(\Phi) \in \Gamma_0$, is

$$LR_G = T \ln \left(\Lambda_T^0 \right) \tag{4.4}$$

with

$$\Lambda_T^0 = \left| \hat{\Sigma}_T^0 \right| / \left| \hat{\Sigma}_T \right| \tag{4.5}$$

where $\hat{\Sigma}_T^0$ and $\hat{\Sigma}_T$ are respectively the restricted and unrestricted ML estimators of the error covariance matrix Σ. For stationary processes, under standard regularity conditions, the asymptotic distribution of the LR statistic under the null hypothesis is chi-square with number of degrees equal to the number of (linearly independent) constraints. This will be the case in particular for zero coefficient restrictions (2.8), restrictions on the order of the process (2.10) and Granger non-causality restrictions (2.11).

For example, consider the problem of testing a Granger non-causality hypothesis, such as $H_0(Y_l \nrightarrow Y_j)$ in (2.11). Here, all the coefficients of the VAR which are not fixed by this null hypothesis – *i.e.* the unknown coefficients of μ, Φ or R – may appear as nuisance parameters in the distribution of LR_G. Further, once the nuisance parameters are set, the model (2.1) - (2.2) and the corresponding test statistic can be simulated. So we propose using Monte Carlo test procedures adapted to the presence of nuisance parameters, in particular maximized Monte Carlo tests. This means applying the MMC procedures described in section 3 with $S(\theta)$ replaced by LR_G, where θ may stand for the elements of δ which are fixed by the null hypothesis.

Such procedures are obviously computer intensive and require dynamic simulations of the process under various parameter configurations (compatible with the null hypothesis). Explosive parameter are not necessarily excluded by the estimation procedure or the model considered. But parameter values can lead to numerical problems (overflows), so one may wish to exclude explosive processes. In VAR models, such restrictions may not be easy to impose. For that purpose, it is useful to note that the roots of the determinantal equation $\det[\Phi(z)] = 0$ are identical with the inverses of the eigenvalues of the matrix

$$\bar{\Phi} = \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_k & 0 & \cdots & 0 & 0 \\ 0 & I_k & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & I_k & 0 \end{bmatrix} . \tag{4.6}$$

The corresponding VAR(p) process is stationary when these eigenvalues are all *inside* the unit circle [see Lütkepohl (1993, Chapter 2) and, for a proof, Appendix A below]. Given the availability of efficient algorithms for computing eigenvalues, this can provide a useful way of excluding explosive processes or limiting the degree of explosion.

The algorithm for computing the MC p-values can be described as follows:

1. choose the restricted subset of the parameter space Ω_0 over which the maximization required by the MMC procedure will be performed; Ω_0 may be the whole parameter space restricted

by the null hypothesis (and, eventually, by stationarity or stability restrictions) or a consistent restricted set estimator;

2. compute the test statistic $LR^{(0)}$ based on the observed data;

3. generate i.i.d. error vectors $\varepsilon^{(l)} = \big[\varepsilon_1^{(l)}, \ldots, \varepsilon_T^{(l)}\big]$, $l = 1, \ldots, N$, according to the selected distribution – for example, $\varepsilon_t^{(l)} \overset{i.i.d.}{\sim} N[0, I_k]$, $t = 1, \ldots, T$ – and construct pseudo-samples as functions of the model parameter vector $\delta = \mathrm{vec}[\mu, \varPhi, R]$:

$$Y_t^{(l)} = \mu + \sum_{i=1}^{p} \varPhi_i Y_{t-i}^{(l)} + R\varepsilon_t^{(l)}, \quad t = 1, \ldots, T,\ l = 1, \ldots, N; \tag{4.7}$$

4. compute the corresponding test statistics as $LR^{(l)}(\delta),\ l = 1, \ldots, N$, which in turn can be viewed as functions of δ and $\varepsilon^{(l)}$;

5. compute the simulated p-value function

$$\hat{p}_N[x \mid \delta] = \frac{N\hat{G}_N[x \mid \delta] + 1}{N+1}, \quad \hat{G}_N[x \mid \delta] = \frac{1}{N}\sum_{l=1}^{N} s\big(LR^{(l)}(\delta) - x\big); \tag{4.8}$$

6. compute the LMC p-value $\hat{p}_N\big[LR^{(0)} \mid \hat{\delta}_T^0\big]$ where $\hat{\delta}_T^0$ is the constrained estimator based on the observed data; if $\hat{p}_N\big[LR^{(0)} \mid \hat{\delta}_T^0\big] > \alpha$, the LMC test is not significant at level α and it is clear the MMC test does not reject either at level α (so the process can be stopped).

The procedure just described can be interpreted as the generation of a parametric bootstrap p-value. Of course, to the extent that the point estimate is typically different from the "true" parameter, the test so obtained is not exact. The MMC procedure involves maximizing the p-value function over the nuisance parameter space, as follows:

1. compute the maximized p-value

$$\hat{p}_{MMC} = \sup\big\{\hat{p}_N\big[LR^{(0)} \mid \delta\big] : \delta \in \varOmega_0\big\}; \tag{4.9}$$

2. reject the null hypothesis at level α if $\hat{p}_{MMC} \leq \alpha$.

When evaluating $\hat{p}_{MMC}$, it is important to note that δ is the only free variable; the observed test statistic and the simulated errors $\varepsilon^{(l)}$, $l = 1, \ldots, N$, are treated as fixed. Even if the LMC test procedure is not significant, it may still be of interest to compute $\hat{p}_{MMC}$ to get a better idea how strongly the null is accepted. As indicated above, the maximization yields a procedure such that the probability of rejection under the null hypothesis is not larger than the level, irrespective of the unknown value of the nuisance parameters. In practice, a reasonable strategy would consist in maximizing the p-value function by taking $\delta = \hat{\delta}_T^0$ as the starting value: even if the maximization is not complete, this provides an immediate safeguard against bootstrap p-values that would be highly sensitive to nuisance parameters. As described in section 3.3, if the region over which we maximize is properly designed (as a consistent set estimator), this yields an asymptotically valid test even if the parametric bootstrap test does not.

5. Simulation experiment

In this section, we present simulation evidence on the performance of three basic types of procedures for hypothesis testing in VAR models: (1) standard tests based on asymptotic chi-square critical values; (2) local Monte Carlo tests (or parametric bootstrap tests), based on a single consistent restricted estimator of model nuisance parameters; (3) maximized Monte Carlo tests. In view of allowing for VAR processes which are non-stationary (integrated) or with roots close to unit circle, we also consider lag-augmented Wald tests proposed by Dolado and Lütkepohl (1996), Toda and Yamamoto (1995), Yamada and Toda (1998) and Dufour, Pelletier, and Renault (2005).[1] The latter procedure have the feature of leading to usual normal (or chi-square) asymptotic distributions, even when the process is integrated, so that we can expect smaller size distortions. Below, we shall consider parametric and nonparametric bootstrap versions of this procedure. A detailed description of the way lag-augmented Wald tests were implemented in this study is presented in Appendix B.[2]

For the purpose of this experiment, we considered standard VAR(p) models with Gaussian disturbances:

$$Y_t = \sum_{i=1}^{p} \Phi_i Y_{t-i} + u_t\,, \quad t = 1,\, \dots\,,\, T\,, \tag{5.1}$$

$$u_t = R\varepsilon_t\,, \quad t = 1,\, \dots\,,\, T\,, \tag{5.2}$$

$$\varepsilon_t \overset{i.i.d.}{\sim} N[0\,,\, I_k]\,, \quad t = 1,\, \dots\,,\, T\,. \tag{5.3}$$

The null hypothesis tested is Granger non-causality

$$H_0 : (Y_2,\, \dots\,,\, Y_k) \nrightarrow Y_1 \tag{5.4}$$

which is equivalent to

$$H_0 : \Phi_{i1l} = 0\,,\; i = 1,\, \dots\,, p\,,\; l = 2\,,\, \dots\,,\, k\,. \tag{5.5}$$

Various dimensions $(k = 2,\, \dots\,,\, 6)$, autoregressive orders $(p = 1,\, \dots\,,\, 5)$, sample sizes $(T = 30,\, 50,\, 100,\, 200,\, 300)$ and parameter structures (Φ) were considered. Under the null hypothesis, the data generating processes have the following relatively simple structure:

$$\Phi\,(L) = (1 - \varphi L)^p \otimes I_k \tag{5.6}$$

[1]For related results, see also Sims, Stock, and Watson (1990), Park and Phillips (1989), Choi (1993), Yamamoto (1996), and Kurozumi and Yamamoto (2000).

[2]We are grateful to a referee for suggesting that we study such a method in the context of our simulation.

where φ is scalar which determines the degree of persistence in the series. Clearly, the process is stationary when $|\varphi| < 1$. R is a nonsingular lower triangular matrix (the values of R used in this experiment are given in Appendix C). The nominal level of the tests is 0.05. The test statistic considered is the LR-type statistic described in section 4. Monte Carlo tests (local and maximized) are based on $N = 99$ replications for tables 1, 5 and 6, $N = 999$ for tables 2 - 4, while the number of trials used for evaluating rejection frequencies is 1000.[3] Local Monte Carlo tests are based on the restricted ML estimator, while the MMC tests are based on maximizing the p-value in a box obtained by taking 5 units on each side of the restricted ML estimator. Some of the results of our experiment are presented in tables 1 to 5 (rejection proportions are expressed in percentages). In Table 1, models M0, M1, M2, M3, M4 are based respectively on the following values of the persistence parameter: $\varphi = 0.9,\ 0.95,\ 0.99,\ (0.95)^{kp}$, $(0.99)^{kp}$. In Table 5, models M2 (panel A) and M0 (panels B, C) are the basic models used, except for modifications to Φ_{i1l} in order to evaluate power. Namely, power is assessed by changing the values of the coefficients Φ_{i1l} , $i = 1, \ldots , p$, as follows: $\Phi_{i1l} = \bar{\Phi}_{i1l}(p) \neq 0$, $l = 2, \ldots , k$, where $\bar{\Phi}_{i1l}(p)$ depends on the order p of the process: $\bar{\Phi}_{i1l}(1) = 0.1$, $\bar{\Phi}_{i1l}(2) = 0.02$, and $\bar{\Phi}_{i1l}(p) = 0.01$ for $p = 3,\ 4,\ 5$. Initial values were set equal to zero. The simulations were all run with GAUSS.

From the results in tables 1 and 5, we see clearly that asymptotic tests based on standard chi-square critical values can have catastrophic size properties, with rejection frequencies as high as 0.97 (instead of 0.05). Using local Monte Carlo (or bootstrap) test does provide important improvements, but overrejections can still be much higher than the level (for example, 0.59 rather than 0.05). Interestingly, the lag-augmented Wald tests can also severely over-reject, even when they are implemented using bootstrap methods (tables 2 - 4). By contrast, in Table 5, we can see that the MMC procedure controls very well the level of the test allows and provides good power under the alternative. Indeed it is the only method that allows one to do that.[4] Finally, Table 6 shows that the MMC has power which behaves in a normal way, increasing as the model moves away from the null hypothesis.

6. Application to a VAR model of the U.S. economy

In this section, we present an application of the techniques proposed above to test Granger causality in the context of a VAR model involving four U.S. macroeconomic variables. The data used come from a study of U.S. monetary policy due to Bernanke and Mihov (1998); see also Dufour, Pelletier, and Renault (2005). This data set consists of monthly observations on nonborrowed reserves (NBR, also denoted M), the federal funds rate (FFR, r), real gross domestic product (GDP, y), the GDP deflator (GDPD, P). The monthly data on GDP and GDP deflator were constructed by state space methods from quarterly observations [see Bernanke and Mihov (1998) for more details]. The sample goes from January 1965 to December 1996. In what follows, all the variables were transformed by a logarithmic transformation. For the purpose of the study, the data were also aggregated to get

[3]This relatively small number was used because the restricted model requires a nonlinear estimation and the simulation-based tests are themselves computer intensive.

[4]We do not report power evaluations for the asymptotic and bootstrap tests, because the level of these procedures cannot be controlled in practice.

Table 1. Empirical levels of Granger causality tests with nominal level $\alpha = 0.05$

(A) VAR(1) models with different dimensions $k = 2,\ 3,\ 4,\ 5, 6\,;\ T = 30$

	Model M0		Model M1		Model M2	
k	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$
2	9.9	6.1	12.7	7.1	15.3	7.7
3	13.4	6.5	16.3	7.6	19.5	8.6
4	17.7	7.1	21.3	8.7	26.2	9.3
5	21.8	7.9	25.2	9.4	29.6	10.7
6	26.3	8.8	32.4	10.2	35.1	12.1

(B) Bivariate VAR(p) models, $p = 1,\ 2,\ 3,\ 4,\ 5;\ T = 30$

	Model M0		Model M3		Model M4	
p	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$
1	10.0	5.3	14.1	7.6	16.3	7.8
2	25.9	10.6	28.0	10.4	32.5	10.2
3	44.9	17.9	39.8	13.3	50.1	18.0
4	64.8	26.5	47.8	14.7	64.5	25.9
5	76.7	36.4	60.0	16.8	75.3	30.8

(C) VAR models [(5.6) with $\varphi = 0.90$]

	k	2		3		4		5	
T	p	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$
50	1	7.9	4.8	10.6	5.9	13.1	7.8	16.5	7.2
	2	16.8	7.8	30.6	9.8	45.8	13.6	60.5	18.6
	3	35.3	11.8	62.2	19.8	81.7	26.2	91.7	37.9
	4	57.0	18.9	83.2	34.1	93.5	49.5	97.0	59.0
	5	69.4	26.5	91.0	49.8	96.9	58.2	97.7	59.3
100	1	7.2	5.2	9.5	5.9	10.3	6.7	10.5	4.9
	2	11.0	5.9	16.7	7.2	24.6	8.0	36.7	10.9
	3	20.2	8.8	34.9	10.0	60.6	14.1	78.0	20.3
	4	34.4	10.2	68.7	16.6	88.8	33.5	96.8	48.6
	5	53.4	15.0	87.2	33.5	98.8	45.4	98.1	50.2
200	1	5.7	5.5	6.0	4.3	7.2	5.1	7.5	5.1
	2	8.9	5.7	10.5	6.3	14.4	7.4	17.8	6.0
	3	11.7	5.9	18.8	7.2	26.5	7.5	42.6	10.7
	4	16.2	6.8	30.6	8.8	54.4	11.5	73.9	18.2
	5	24.5	9.0	51.4	12.0	82.6	19.1	96.0	25.0
300	1	5.5	4.4	6.5	5.4	6.6	4.6	8.3	6.1
	2	7.4	5.5	9.0	6.3	9.3	4.3	14.0	6.7
	3	8.3	4.8	13.1	5.7	18.0	7.3	23.9	7.5
	4	10.4	4.8	17.0	5.2	28.3	5.8	49.7	9.8
	5	13.1	5.8	27.0	7.1	55.6	11.8	78.9	16.1

Note – ASY_{lr} stays for the asymptotic test based on the likelihood ratio statistic while $\text{LMC}_{\text{lr}}^{\text{pa}}$ is the corresponding local MC (parametric bootstrap) p-value. Models M0, M1, M2, M3 and M4 correspond to (5.6) with $\varphi = 0.90,\ 0.95,\ 0.99$, $(0.95)^{kp}$, $(0.99)^{kp}$ respectively. The proportions in this table as well subsequent tables are written in percentages.

Table 2. Causality tests based on lag-augmented Wald statistics.
Empirical levels of asymptotic, parametric bootstrap and nonparametric bootstrap procedures.($\alpha = 0.05$).
Autoregressive matrix polynomials of the form (5.6) with $\varphi = 0.90$

T	50			100			200			300		
$(k,\ p)$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{pa}}$	$\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{np}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{pa}}$	$\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{np}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{pa}}$	$\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{np}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{pa}}$	$\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{np}}$
(2, 1)	8.1	5.4	5.3	6.9	5.9	5.4	4.8	4.5	4.5	4.7	4.3	4.5
(2, 2)	11.6	5.0	5.2	9.0	5.7	5.5	6.6	5.0	4.8	4.7	4.1	4.1
(2, 3)	30.7	9.3	9.6	15.8	6.4	6.5	11.0	7.0	7.1	7.1	4.8	4.7
(2, 4)	58.3	16.1	15.9	30.0	9.3	9.1	11.5	4.7	4.5	9.1	5.5	5.2
(2, 5)	79.5	22.9	23.2	50.9	13.2	13.0	19.4	4.7	4.8	11.6	4.3	4.5
(3, 1)	9.9	5.7	5.5	7.5	5.3	5.9	6.5	5.4	4.8	6.2	5.6	5.9
(3, 2)	24.9	8.4	8.4	9.9	5.1	4.8	6.7	3.8	4.0	7.0	5.0	4.6
(3, 3)	63.1	13.1	13.6	32.8	8.7	8.7	14.2	5.5	5.9	10.1	5.0	4.8
(3, 4)	90.8	29.6	28.9	63.1	15.3	15.4	31.1	9.0	8.9	17.3	5.9	6.0
(3, 5)	98.4	41.3	40.4	89.6	31.4	32.0	53.4	10.4	10.4	25.6	6.1	6.2
(4, 1)	13.2	5.4	4.8	8.3	5.1	5.6	7.2	5.5	5.7	5.3	4.4	4.4
(4, 2)	40.6	9.3	9.6	19.9	6.6	6.4	9.9	4.6	5.1	7.3	4.9	4.8
(4, 3)	84.4	21.1	21.0	53.9	12.9	12.5	23.9	6.6	7.1	15.2	6.2	6.2
(4, 4)	99.0	38.5	37.4	88.6	26.7	26.4	49.5	10.4	10.4	28.3	7.1	7.1
(4, 5)	100	49.8	49.6	98.7	50.0	49.7	83.6	19.4	19.5	52.9	8.2	8.7
(5, 1)	16.2	6.1	6.4	6.8	3.7	3.5	7.6	5.9	5.8	6.6	4.8	5.0
(5, 2)	54.5	9.8	9.8	29.0	8.4	8.3	12.6	5.4	5.7	9.9	5.7	5.7
(5, 3)	95.2	25.1	24.8	73.0	16.2	17.3	37.9	7.9	7.8	22.0	5.7	5.9
(5, 4)	100	46.6	47.3	97.4	39.0	38.9	75.5	16.4	16.2	41.7	8.6	8.7
(5, 5)	100	61.4	60.9	100	65.9	65.5	96.5	33.5	33.6	76.7	14.4	14.8

Note – $\mathrm{ASY}_{\mathrm{aw}}$ represents asymptotic p-values for the lag-augmented Wald statistic, while $\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{pa}}$ and $\mathrm{LMC}_{\mathrm{aw}}^{\mathrm{np}}$ are the related parametric and nonparametric bootstrap p-values, respectively.

Table 3. Causality tests based on lag-augmented Wald statistics.
Empirical levels of asymptotic, parametric bootstrap and nonparametric bootstrap procedures.($\alpha = 0.05$).
Autoregressive matrix polynomials of the form (5.6) with $\varphi = 0.95$

T	50			100			200			300		
$(k,\ p)$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{np}}_{\mathrm{aw}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{np}}_{\mathrm{aw}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{np}}_{\mathrm{aw}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{np}}_{\mathrm{aw}}$
(2, 1)	6.5	3.8	3.9	5.2	4.2	3.6	5.6	5.3	5.4	5.6	5.5	5.2
(2, 2)	18.3	7.2	7.0	8.9	4.8	4.8	6.4	4.2	4.3	5.9	5.1	5.2
(2, 3)	38.0	11.4	11.6	22.3	8.3	8.2	14.2	6.1	6.3	10.7	6.3	6.2
(2, 4)	70.4	24.1	24.1	51.1	18.0	17.4	25.0	8.0	8.2	15.4	7.6	6.6
(2, 5)	85.5	31.4	31.3	68.7	22.8	23.1	41.3	11.4	11.3	28.4	8.5	8.1
(3, 1)	11.9	6.6	6.1	7.0	4.8	4.8	4.8	4.1	3.8	6.5	5.6	5.6
(3, 2)	28.9	8.2	8.2	17.9	8.0	7.0	10.4	6.5	6.5	7.6	5.2	5.4
(3, 3)	72.5	19.8	19.7	48.3	15.3	15.3	24.0	6.9	7.0	16.5	7.9	7.9
(3, 4)	92.9	37.1	37.3	79.5	31.0	31.0	54.8	17.7	17.9	35.1	10.5	10.7
(3, 5)	99.0	43.7	43.5	94.9	46.4	46.2	81.9	24.2	24.0	62.5	12.1	12.5
(4, 1)	13.8	5.5	6.2	9.5	6.3	6.1	5.0	3.4	3.7	6.3	5.4	5.5
(4, 2)	47.6	10.9	10.9	24.6	6.8	7.2	14.1	6.6	6.5	10.4	5.7	5.7
(4, 3)	89.1	27.7	27.6	68.3	18.1	18.7	45.3	13.5	13.6	28.2	9.5	9.5
(4, 4)	99.1	44.0	44.6	94.7	42.6	43.3	83.5	25.6	24.5	64.2	15.1	15.3
(5, 1)	17.4	6.5	6.5	10.4	5.2	5.1	5.6	3.6	3.4	6.3	5.3	5.1
(5, 2)	60.2	12.7	12.1	37.3	10.0	9.7	19.1	6.5	6.4	15.3	6.4	6.2
(5, 3)	97.7	33.9	34.3	81.8	27.3	27.2	63.3	17.2	17.2	43.1	11.6	11.7
(5, 4)	100	56.0	55.9	98.7	56.0	55.7	94.7	38.1	37.2	82.3	22.0	21.8

Table 4. Causality tests based on lag-augmented Wald statistics.
Empirical levels of asymptotic, parametric bootstrap and nonparametric bootstrap procedures.($\alpha = 0.05$).
Autoregressive matrix polynomials of the form (5.6) with $(\varphi,\ T) = (0.99,\ 50)$, $(0.98,\ 100)$, $(0.97,\ 200)$ and $(0.96,\ 300)$

T	50			100			200			300		
$(k,\ p)$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{np}}_{\mathrm{aw}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{np}}_{\mathrm{aw}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{np}}_{\mathrm{aw}}$	$\mathrm{ASY}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{aw}}$	$\mathrm{LMC}^{\mathrm{np}}_{\mathrm{aw}}$
(2, 1)	7.9	4.4	4.5	6.8	5.5	5.2	5.5	5.2	5.3	4.9	4.5	4.8
(2, 2)	21.2	8.7	8.3	12.4	5.7	5.7	9.1	5.4	6.0	6.6	4.9	4.8
(2, 3)	48.5	16.4	16.6	36.6	14.6	15.1	18.2	6.8	7.2	12.3	6.7	6.9
(2, 4)	70.2	29.1	29.0	61.2	24.1	24.4	38.2	14.3	13.8	22.0	8.9	9.0
(2, 5)	84.4	27.4	27.8	73.8	23.6	23.8	57.0	14.2	14.4	31.7	10.4	9.7
(3, 1)	10.8	5.4	5.3	6.8	4.6	4.6	6.9	5.9	5.8	6.3	5.6	6.1
(3, 2)	34.8	11.3	11.5	22.9	9.5	9.8	11.7	4.9	5.0	9.4	6.1	6.6
(3, 3)	75.9	27.0	26.8	60.2	22.2	22.3	40.4	14.6	14.6	21.3	6.2	6.3
(3, 4)	95.2	45.5	45.2	86.5	38.1	38.5	70.8	24.0	23.2	46.7	13.4	14.0
(3, 5)	98.4	42.1	41.7	94.2	36.9	37.4	87.2	30.2	30.2	72.1	16.4	16.0
(4, 1)	15.3	6.5	6.5	8.7	5.8	5.8	7.5	6.0	6.2	6.5	5.4	5.4
(4, 2)	50.3	13.3	13.1	32.5	9.9	10.5	19.9	9.1	9.2	13.1	7.1	6.7
(4, 3)	92.0	34.9	34.6	79.9	29.5	29.3	57.0	17.5	18.3	35.8	9.4	10.2
(4, 4)	99.1	49.9	51.1	95.4	52.1	51.7	88.9	38.0	37.8	73.6	19.7	19.5
(5, 1)	17.2	5.2	5.2	9.2	4.3	4.4	8.1	5.4	5.6	6.5	4.8	5.1
(5, 2)	66.3	16.1	15.1	42.0	13.6	13.5	24.2	8.6	8.8	17.9	7.6	7.8
(5, 3)	98.3	38.4	38.7	90.2	35.7	35.6	71.9	22.6	22.1	51.7	14.4	14.4
(5, 4)	100	56.5	56.1	99.6	59.6	59.7	97.5	47.1	47.6	91.1	29.9	30.4

Table 5. Empirical levels of asymptotic, LMC and MMC Granger causality tests with nominal level $\alpha = 0.05$

(A) VAR(1) models with different dimensions $k = 2,\ 3,\ 4,\ 5, 6;\ T = 30$

k	Level			Power
	ASY_{lr}	$\text{LMC}^{\text{pa}}_{\text{lr}}$	$\text{MMC}^{\text{pa}}_{\text{lr}}$	$\text{MMC}^{\text{pa}}_{\text{lr}}$
2	15.5	7.7	3.2	84.7
3	22.0	9.0	3.3	95.1
4	24.7	9.7	2.8	99.8
5	32.2	11.9	3.1	99.4
6	35.1	12.1	2.7	92.1

(B) Bivariate VAR(p) models of different orders $p = 1,\ 2,\ 3,\ 4,\ 5;\ T = 30$

p	Level			Power
	ASY_{lr}	$\text{LMC}^{\text{pa}}_{\text{lr}}$	$\text{MMC}^{\text{pa}}_{\text{lr}}$	$\text{MMC}^{\text{pa}}_{\text{lr}}$
1	10.5	5.1	0.4	70.0
2	25.7	8.9	0.9	56.4
3	45.2	15.8	2.3	74.2
4	64.3	25.3	4.7	85.1
5	78.0	39.2	4.2	96.2

(C) k-dimensional VAR(p) models with different sample sizes

T	50			100			200			300		
$(k,\ p)$	ASY_{lr}	$\text{LMC}^{\text{pa}}_{\text{lr}}$	$\text{MMC}^{\text{pa}}_{\text{lr}}$	ASY_{lr}	$\text{LMC}^{\text{pa}}_{\text{lr}}$	$\text{MMC}^{\text{pa}}_{\text{lr}}$	ASY_{lr}	$\text{LMC}^{\text{pa}}_{\text{lr}}$	$\text{MMC}^{\text{pa}}_{\text{lr}}$	ASY_{lr}	$\text{LMC}^{\text{pa}}_{\text{lr}}$	$\text{MMC}^{\text{pa}}_{\text{lr}}$
(2, 1)	10.2	6.5	1.6	6.4	5.3	1.1	5.9	4.9	0.8	6.4	5.2	1.0
(2, 2)	18.9	8.9	1.2	11.1	5.2	0.6	8.4	6.2	0.7	7.7	4.8	0.4
(2, 3)	36.8	11.9	2.2	19.2	6.1	0.3	10.5	5.1	1.2	7.5	4.9	0.6
(2, 4)	60.1	18.6	4.4	33.2	10.7	1.5	16.2	5.5	0.3	11.1	6.9	0.3
(2, 5)	69.2	25.9	3.8	51.8	15.3	2.7	24.8	8.2	1.2	12.4	5.6	0.8
(3, 1)	12.2	7.2	0.8	9.0	5.4	1.4	6.1	5.7	1.1	6.5	5.0	1.8
(3, 2)	29.2	9.9	1.5	18.0	7.8	2.3	10.3	5.2	1.1	8.1	5.0	0.9
(3, 3)	63.8	18.5	2.0	37.6	10.5	1.7	19.5	6.5	0.3	11.7	5.7	0.2
(3, 4)	85.2	36.5	2.2	69.3	18.9	4.1	31.2	8.6	1.2	18.3	5.9	0.4
(3, 5)	92.1	48.3	3.6	88.1	32.7	4.3	50.8	11.4	1.7	26.3	7.4	1.3
(4, 1)	15.1	6.7	1.2	9.3	6.1	1.0	7.0	4.4	1.0	6.2	5.1	1.7
(4, 2)	45.4	13.4	3.1	26.6	8.6	1.4	12.7	5.4	0.1	10.6	5.5	0.1
(4, 3)	82.0	28.6	3.3	59.4	15.5	0.2	28.7	6.7	0.1	17.2	6.4	0.1
(4, 4)	93.8	48.9	3.2	86.9	34.1	2.9	52.7	10.9	0.6	27.3	6.2	0.9
(4, 5)	97.8	58.5	4.1	98.9	46.3	3.1	84.1	18.7	2.1	57.9	11.2	1.9

Note – ASY_{lr} stays for the asymptotic test based on the likelihood ratio statistic while $\text{LMC}^{\text{pa}}_{\text{lr}}$ and $\text{MMC}^{\text{pa}}_{\text{lr}}$ are the corresponding local MC (parametric bootstrap) and maximized MC p-values, respectively. Panel A is based on model M2 ($\varphi = 0.99$), while panels B and C are based on model M0 ($\varphi = 0.90$). Power is obtained under alternatives where $\Phi_{i1l} = \bar{\Phi}_{i1l}(p) \neq 0$, $l = 2, \ldots, k$, $i = 1, \ldots, p$, where $\bar{\Phi}_{i1l}(p)$ depends on the order p of the process: $\bar{\Phi}_{i1l}(1) = 0.1$, $\bar{\Phi}_{i1l}(2) = 0.02$, and $\bar{\Phi}_{i1l}(p) = 0.01$ for $p = 3,\ 4,\ 5$.

quarterly observations (using arithmetic averages) and put in first differences so that we roughly consider growth rates. Results based on both quarterly and monthly data are presented below. Of course, the monthly models involves considerably more nuisance parameters. Monte Carlo tests in this example are based on $N = 999$ replications, while the MMC tests are based on maximizing the p-value in a box obtained by taking 5 units on each side of the restricted ML estimator.

The first problem we face consists in specifying the order of the VAR. Using quarterly data, we found that the MC tests reject much less often than the asymptotic procedure: LMC tests are significant procedure at level 5% for the orders 0, 1, 2, 3, 8, 9 [plus 7 at level 10%], while the MMC tests are significant at level 5% only for the orders 0, 1, 2, 3 [plus the orders 7, 8 and 9 at level 10%]. In view of these results and the quarterly frequency of the data, we present here results based on a VAR(4) for Granger causality testing.[5] The results on testing Granger causality are presented in table 8.

Based on the VAR(4) model, we can identify the following significant relationships (according to MMC tests): at level 5%,

$$M \quad \longrightarrow \quad r \quad \longleftrightarrow \quad y \tag{6.1}$$

and at level 10%

$$\begin{array}{ccccc} M & \longrightarrow & r & \longleftrightarrow & y \\ & & \downarrow & & \\ & & P & & \end{array} . \tag{6.2}$$

Interestingly, these results appear to be consistent with a monetarist interpretation of the relationship between money and income, where money Granger causes interest rates which in turn causes (and is caused by) income.

For monthly data, we first studied the appropriate order and found that LMC and MMC tests are both significant at level 5% for the orders 0, 1, 2, 4, 5, 8, 14, 15 [plus the orders 7 and 11at level 10%].[6] This suggests choosing VAR(16) for analyzing the Granger causality structure between these four macroeconomic variables. The results on pairwise Granger causality tests are presented in Table 10. From these results, we identify the following significant relationships (according to MMC tests): at level 5%,

$$\begin{array}{ccccc} M & \longrightarrow & r & \longleftarrow & y \\ & \searrow & & & \\ & & P & & \end{array} \tag{6.3}$$

and at level 10%

$$\begin{array}{ccccc} M & \longrightarrow & r & \longleftarrow & y \\ & \searrow & \downarrow & & \\ & & P & & \end{array} \tag{6.4}$$

Note that the results provided by these different models are quite similar to those of the VAR(4) model, except for the fact that we now have unidirectional (rather than bidirectional) causality from income to interest rates ($y \longrightarrow r$) and unidirectional causality from money to prices ($M \longrightarrow P$). It is interesting to note that money (when measured by bank reserves) appears not to be Granger caused

[5]Detailed results on order selection tests are available in a discussion paper [Dufour and Jouini (2004)].

[6]Detailed results are available from the authors.

Table 6. Power of the MMC causality tests.
VAR(1) models, $T = 30$

Φ_{i1l}	$\mathrm{MMC}^{\mathrm{pa}}_{\mathrm{lr}}$				
	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
0.01	3.4	1.9	3.8	2.9	1.9
0.02	3.4	4.2	7.1	6.4	4.2
0.03	7.3	7.7	23.3	15.7	6.7
0.04	10.2	15.4	41.3	30.1	14.0
0.05	18.6	22.6	65.4	49.2	25.1
0.06	26.6	37.3	78.5	65.9	41.1
0.07	40.3	54.8	88.8	77.4	53.4
0.08	50.1	63.6	95.2	87.7	63.9
0.09	62.0	77.1	97.9	94.1	75.3
0.10	70.0	84.5	98.8	96.8	82.0
0.15	91.8	98.3	100	99.9	98.8
0.20	98.9	99.8	100	100	99.7

Note – These results are based on model M0 ($\varphi = 0.90$). Power is obtained under alternatives where $\Phi_{i1l} = \bar{\Phi}_{i1l}(p) \neq 0$, $l = 2, \ldots, k$, $i = 1, \ldots, p$, where $\bar{\Phi}_{i1l}(p)$ depends on the order p of the process: $\bar{\Phi}_{i1l}(1) = 0.1$, $\bar{\Phi}_{i1l}(2) = 0.02$, and $\bar{\Phi}_{i1l}(p) = 0.01$ for $p = 3, 4, 5$.

Table 7. Tests for VAR order (quarterly data)
AR(p) $vs.$ AR$(p+1)$

p	$\mathrm{ASY}_{\mathrm{lr}}$	$\mathrm{LMC}^{\mathrm{pa}}_{\mathrm{lr}}$	$\mathrm{MMC}^{\mathrm{pa}}_{\mathrm{lr}}$
17	0.016***	24.100	24.100
16	0.155***	33.400	33.400
15	0.312***	35.700	35.700
14	7.244*	65.100	65.100
13	12.483	70.100	70.100
12	22.961	79.000	79.000
11	2.205**	29.100	29.100
10	0.356***	11.700	11.700
9	0.058***	4.800**	6.300*
8	2.195**	4.600**	6.400*
7	4.526**	6.200*	7.900*
6	2.055**	11.300	11.300
5	8.226*	24.500	24.500
4	8.302*	20.100	20.100
3	0.916***	2.800**	4.600**
2	0.995***	2.500**	4.200**
1	0.251***	0.500***	1.100**
0	0.000***	0.100***	0.100***

Note – The numbers in the table are p-values in percentage. *** and ** highlight p-values not larger than 1.00% and 5.00%, respectively, while * highlights a p-values not larger than 10%.
– The quarterly data is computed using arithmetic average over successive blocks of three consecutive observations.

Table 8. Pairwise Granger non-causality tests based on a quarterly VAR(4) model

H_0			ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	$\text{MMC}_{\text{lr}}^{\text{pa}}$
M	$\nrightarrow$	r	0.495***	1.300**	2.100**
	$\nrightarrow$	P	42.195	49.600	49.600
	$\nrightarrow$	y	61.352	69.500	69.500
r	$\nrightarrow$	M	88.927	92.200	92.200
	$\nrightarrow$	P	2.108**	3.900**	5.300*
	$\nrightarrow$	y	1.671**	2.400**	3.800**
P	$\nrightarrow$	M	55.120	61.000	61.000
	$\nrightarrow$	r	22.472	29.400	29.400
	$\nrightarrow$	y	65.790	72.700	72.700
y	$\nrightarrow$	M	33.619	41.600	41.600
	$\nrightarrow$	r	0.021***	0.100***	0.200***
	$\nrightarrow$	P	25.144	33.100	33.100

Note – The numbers in the table are p-values in percentage. *** and ** highlight p-values not larger than 1.00% and 5.00%, respectively, while * highlights p-values not larger than 10%.

Table 9. Tests for VAR order (monthly data)
AR(p) $vs.$ AR$(p+1)$

p	ASY_{lr}	$\text{LMC}_{\text{lr}}^{\text{pa}}$	$\text{MMC}_{\text{lr}}^{\text{pa}}$
18	3.353**	16.600	16.600
17	96.756	99.200	99.200
16	7.833*	25.800	25.800
15	0.039***	0.500***	0.800***
14	0.348***	2.000**	3.300**
13	30.411	52.900	52.900
12	8.084*	20.900	20.900
11	2.279**	7.400*	8.400*
10	5.901*	13.800	13.800
9	5.549*	10.600	10.600
8	0.077***	0.300***	0.400***
7	1.984**	6.400*	7.500*
6	44.546	54.300	54.300
5	0.000***	0.100***	0.100***
4	1.938**	3.700**	4.400**
3	10.975	15.200	15.200
2	0.000***	0.100***	0.100***
1	0.000***	0.100***	0.100***
0	0.000***	0.100***	0.100***

Note – The numbers in the table are p-values in percentage. *** and ** highlight p-values not larger than 1.00% and 5.00%, respectively, while * highlights a p-values not larger than 10%.

Table 10. Pairwise Granger non-causality tests based on a monthly VAR(16) model

H_0			$\mathrm{ASY_{lr}}$	$\mathrm{LMC_{lr}^{pa}}$	$\mathrm{MMC_{lr}^{pa}}$
M	$\nrightarrow$	r	0.004***	0.200***	0.200***
	$\nrightarrow$	P	0.000***	0.200***	0.200***
	$\nrightarrow$	y	2.710**	12.400	12.400
r	$\nrightarrow$	M	14.961	33.500	33.500
	$\nrightarrow$	P	1.997**	7.900*	8.800*
	$\nrightarrow$	y	4.915**	16.500	16.500
P	$\nrightarrow$	M	49.853	71.100	71.100
	$\nrightarrow$	r	21.040	40.200	40.200
	$\nrightarrow$	y	19.954	43.100	43.100
y	$\nrightarrow$	M	21.543	43.700	43.700
	$\nrightarrow$	r	0.000***	0.100***	0.100***
	$\nrightarrow$	P	14.593	36.500	36.500

Note – The numbers in the table are p-values in percentage. *** and ** highlight p-values not larger than 1.00% and 5.00%, respectively, while * highlights a p-values not larger than 10%.

by the other variables of the system. Further, money does not Granger cause income (y) directly but has an effect on the interest rate (r) which in turn Granger causes the prices. So reserves may have an indirect effect on income [see Dufour and Renault (1998) and Dufour, Pelletier, and Renault (2005)].

7. Conclusion

In this paper, we have proposed a general simulation-based method to produce finite-sample tests in parametric VAR models with known lag order (or a known upper bound on the order of the process). The method has the important feature that no other assumption is needed on the structure of the underlying process: all that is required is the possibility of simulating the model once a finite number of parameters has been specified. For example, the VAR process may be integrated of any order. We also showed that the proposed method can be implemented in practice, despite the presence of a large number of nuisance parameters. In a simulation experiment, we saw clearly that both standard asymptotic as well as bootstrap procedures can suffer from severe size distortions, while, under the same conditions, the MMC method controls the level of the test perfectly (as expected), although its size could be lower than the test. To best of our knowledge, no other available procedure has these features. We also provided an application to Granger causality testing in a four-variable macroeconomic model of the U.S. economy.

Even though we have focused here on tests on the order of a VAR and Granger causality, the approach proposed here can be applied in principle to any set of restrictions on the model, such as unit root or cointegration hypotheses. In such cases, even though the unit root hypothesis (for example) could be taken into account by an asymptotic distributional theory or a bootstrap procedure, large roots in the stationary region but close to the unit-circle could still lead to large size distortions. By construction, the MMC procedure remains valid irrespective of the structure of the VAR. It is also important to note that the error distribution need not be normal: any assumption that specifies completely the distribution of $\varepsilon(T) = \mathrm{vec}(\varepsilon_1, \ldots, \varepsilon_T)$, *i.e.* the disturbance distribution up to an unknown linear transformation (or covariance matrix) can be used. No assumption on the existence of moments is needed, so one could consider distributions with heavy tails. One could also introduce further free parameters in the error distribution: such parameters can be treated as extra nuisance parameters.

The main limitations of the approach proposed here lie in the parametric setup required to perform the MC tests and the computational cost. On the first issue, it is important to note that parametric assumptions involve putting a bound on the maximal order of the process (which is equivalent to assuming that the order of VAR process is "known"). In the case of testing Granger non-causality (as well as for many hypotheses of interest), this means that the lag order is an integral part of the null hypothesis: there is no way to "separate" Granger non-causality from an assumption on the order of the process. Allowing for a data-based order selection would require simulating as well the model selection procedure. Note, however, that producing finite-sample inference without putting an explicit upper bound on the order of the process is fundamentally an impossible task [see the discussions in Sims (1971a, 1971b), Cochrane (1991), Blough (1992), Faust (1996, 1999), Pötscher (2002) and Dufour (2003)]. So, from the viewpoint of developing valid tests in finite samples, the assumption of a "known order" is unavoidable.

If one is prepared to accept a procedure which has only an "asymptotic justification", it is also important to note that the proposed "exact procedures" remain asymptotically valid (in the usual sense of pointwise asymptotic validity) under much weaker assumptions, including an "unknown" order which may be "consistently estimated". As long as the MC tests are performed using a distribution which is covered by the assumptions of the limiting distributional theory, the probability of type I error will satisfy the level condition asymptotically. Of course, under usual assumptions, such a convergence will not typically be uniform – which opens the possibility of arbitrary deviations from the nominal level of the test – but this simply reflects the fact that typical regularity assumptions are simply too weak to even allow the existence of provably valid finite-sample procedures [see Dufour (2003)]. It is worthwhile to note also that the MMC procedure automatically adapts to possible dependence of the distribution of the test statistic upon the autoregressive coefficients.

On the second issue, it is clear that MMC tests are computer intensive. The code that we used to perform the simulations and applications presented is certainly not optimal [given that these were performed with GAUSS] and we are working on improving it. Given the regular improvements in computer speeds, the importance of this type of consideration should decline in the future.

A. Appendix: Equivalence between roots and eigenvalues of a VAR process

By the definition,

$$\bar{\Phi} = \begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \cdots & \Phi_{p-1} & \Phi_p \\ I_k & 0 & 0 & \cdots & 0 & 0 \\ 0 & I_k & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & I_k & 0 \end{bmatrix}, \tag{A.1}$$

we see that, for any $z \in \mathbb{C}$,

$$\det\left[I_{kp} - \bar{\Phi} z\right] = \det \begin{bmatrix} I_k - \Phi_1 z & -\Phi_2 z & -\Phi_3 z & \cdots & -\Phi_{p-1} z & -\Phi_p z \\ -I_k z & I_k & 0 & \cdots & 0 & 0 \\ 0 & -I_k z & I_k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & I_k & 0 \\ 0 & 0 & 0 & \cdots & -I_k z & I_k \end{bmatrix}. \tag{A.2}$$

Now, on multiplying by z the second block of k columns and adding it to the first block, we get

$$\det\left[I_{kp} - \bar{\Phi} z\right] = \det \begin{bmatrix} I_k - \sum_{i=1}^{2} \Phi_i z^i & -\Phi_2 z & -\Phi_3 z & \cdots & -\Phi_{p-1} z & -\Phi_p z \\ 0 & I_k & 0 & \cdots & 0 & 0 \\ -I_k z^2 & -I_k z & I_k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_k & 0 \\ 0 & 0 & 0 & \cdots & -I_k z & I_k \end{bmatrix}. \tag{A.3}$$

and, repeating this process up to the p-th block (*i.e.*, multiplying the j-th block by z^{j-1} and adding the result to the first block, for $j = 2, \ldots, p$), we obtain:

$$\begin{aligned} \det\left[I_{kp} - \bar{\Phi} z\right] &= \det \begin{bmatrix} I_k - \sum_{i=1}^{p} \Phi_i z^i & -\Phi_2 z & -\Phi_3 z & \cdots & -\Phi_{p-1} z & -\Phi_p z \\ 0 & I_k & 0 & \cdots & 0 & 0 \\ 0 & -I_k z & I_k & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_k & 0 \\ 0 & 0 & 0 & \cdots & -I_k z & I_k \end{bmatrix} \\ &= \det\left[I_k - \sum_{i=1}^{p} \Phi_i z^i\right] = \det\left\{\Phi\left(z\right)\right\}. \end{aligned} \tag{A.4}$$

Since, for $z \neq 0$,

$$\det\left[I_{kp} - \bar{\Phi} z\right] = |z|^{kp} \det\left[z^{-1} I_{kp} - \bar{\Phi}\right] = (-1)^{kp} |z|^{kp} \det\left[\bar{\Phi} - z^{-1} I_{kp}\right] , \tag{A.5}$$

the stationarity condition

$$\det\left[\Phi\left(z\right)\right] = 0 \Leftrightarrow |z| > 1 \tag{A.6}$$

is equivalent to

$$\det\left[\bar{\Phi} - z^{-1} I_{kp}\right] = 0 \Leftrightarrow |z| > 1 \tag{A.7}$$

and, setting $\lambda = 1/z$,

$$\det\left[\bar{\Phi} - \lambda I_{kp}\right] = 0 \Leftrightarrow |\lambda| < 1\,. \tag{A.8}$$

This means that, to have stationarity, the eigenvalues of the matrix $\bar{\Phi}$ should be *inside* the unit circle.

B. Appendix: Lag-augmented Wald tests for causality

We give here a brief description of the lag-augmented Wald tests considered in the simulation. In order to test Granger non-causality, the first step consists in estimating by ordinary least squares an unrestricted VAR$(p+1)$ model, rather than a VAR(p) model:

$$Y_t = \sum_{i=1}^{p+1} \Phi_i Y_{t-i} + u_t\,, \quad t = 1,\, \dots\,,\, T\,. \tag{B.1}$$

Even though we know that $\Phi_{p+1} = 0$, this restriction is not used in order to compute the test statistic. Second, we consider the hypothesis (5.5) which is equivalent to $H_0 : (Y_2,\, \dots\,,\, Y_k) \nrightarrow Y_1$ under the VAR(p) model (leaving Φ_{p+1} as a free coefficient), and compute the corresponding Wald-type test statistic $\big[$say $W_G^{(0)}\big]$. In accordance with (5.5), H_0 may be expressed as a set of zero restrictions on the $(p+1)\,k^2 \times 1$ coefficient vector $\phi_{p+1} = \mathrm{vec}\big(\big[\Phi_1,\, \Phi_2, \dots\,,\, \Phi_p,\, \Phi_{p+1}\big]'\big)$, *i.e.* $H_0^{'} : C_{p+1}\phi_{p+1} = 0$, where C_{p+1} is a full-rank $p\,(k-1) \times (p+1)\,k^2$ matrix containing only 0 and 1. The Wald statistic then has the form:

$$W_G^{(0)} = T\big(C_{p+1}\hat{\phi}_{p+1}\big)'\big[C_{p+1}\hat{\Sigma}(\hat{\phi}_{p+1})C_{p+1}'\big]^{-1}\big(C_{p+1}\hat{\phi}_{p+1}\big) \tag{B.2}$$

where $\hat{\phi}_{p+1} = \mathrm{vec}\big(\big[\hat{\Phi}_1,\, \hat{\Phi}_2, \dots\,,\, \hat{\Phi}_p,\, \hat{\Phi}_{p+1}\big]'\big)$ and $\hat{\Phi}_i$, $i = 1, \dots\,, p+1$, are the unconstrained least squares estimates for (B.1),[7] $\hat{\Sigma}(\hat{\phi}_{p+1})$ is the usual asymptotic covariance estimator for $T^{1/2}\big(\hat{\phi}_{p+1} - \phi_{p+1}\big)$, namely $\hat{\Sigma}(\hat{\phi}_{p+1}) = \hat{\Sigma}_{p+1} \otimes \hat{\Gamma}_{p+1}^{-1}$ with

$$\hat{\Sigma}_{p+1} = \frac{1}{T}\sum_{t=p+2}^{T} \hat{u}_t\,(p+1)\,\hat{u}_t\,\,(p+1)'\,, \quad \hat{\Gamma}_{p+1} = \frac{1}{T} Y\,(p+1, T)\,Y\,(p+1, T)'\,, \tag{B.3}$$

$$\hat{u}_t\,(p+1) = Y_t - \sum_{i=1}^{p+1} \hat{\Phi}_i Y_{t-i}\,, \quad t = p+2,\, \dots\,,\, T\,, \tag{B.4}$$

$$Y\,(p+1, T) = \left[Y_{p+2}\,(p+1)\,, Y_{p+3}\,(p+1)\;, \dots\,, Y_T\,(p+1)\right]\,, \tag{B.5}$$

and $Y_t\,(p+1) = \big[Y_{t-1}', Y_{t-2}', \dots, Y_{t-p-1}'\big]'$.

Under the VAR(p) specification with H_0, this statistic follows a chi-square distribution asymptotically (with number of degrees of freedom equal to the number of restrictions) even if the process Y_t is integrated; see, for example, Toda and Yamamoto (1995) and Dolado and Lütkepohl (1996). Of course, the finite-sample distribution of the lag-augmented Wald statistic depends on nuisance parameters (the coefficients which are not fixed by the hypotheses). So the chi-square approximation may be quite unreliable in finite samples, and improvements (such as bootstrapping) may be very important in this model. We consider here two ways of bootstrapping such lag-augmented Wald tests, a "parametric" bootstrap and a "nonparametric" one.

In the parametric case, we first obtain consistent restricted estimates $\tilde{\Phi}_i^c$, $i = 1,\, \dots\,,\, p$, of

[7] Such estimates can easily be obtained by applying OLS to each equation.

the VAR(p) model [*i.e.*, (5.1) with (5.5)] along with the Cholesky factor $\tilde{R}_T^c$ associated with the estimated error covariance matrix. In the present case, the restricted estimates are obtained through maximization of the Gaussian likelihood $L(\delta)$ in (4.2). These values are then used to generate pseudo-samples $Y^{(l)} = [Y_1^{(l)}, \ldots, Y_T^{(l)}]$, according to the equation

$$Y_t^{(l)} = \sum_{i=1}^{p} \tilde{\Phi}_i^c Y_{t-i}^{(l)} + \tilde{R}^c \, \varepsilon_t^{(l)}, \quad t = 1, \ldots, T, \quad l = 1, \ldots, N, \tag{B.6}$$

where the $\varepsilon_t^{(l)}$ are simulated according to the distribution

$$\varepsilon_t^{(l)} \overset{i.i.d.}{\sim} N[0, I_k], \;\; t = 1, \ldots, T. \tag{B.7}$$

From each simulated sample $Y^{(l)}$, a VAR$(p+1)$ model is estimated and the corresponding lag-augmented Wald statistic $W_G^{(l)}$ for H_0 is computed. The initial values are kept fixed at the realized values from the observed sample. The corresponding simulated p-value $\hat{p}_N(S_0)$ then follows according to formula (3.9) with $S_l = W_G^{(l)}$, $l = 0, 1, \ldots, N$. The null hypothesis is rejected when $\hat{p}_N(S_0) \le \alpha$.

In the nonparametric case, we start from the estimated residuals

$$\tilde{u}_t^c = Y_t - \sum_{i=1}^{p} \tilde{\Phi}_i^c Y_{t-i}, \quad t = 1, \ldots, T. \tag{B.8}$$

New residuals $\tilde{u}_1^{(l)}, \ldots, \tilde{u}_T^{(l)}$ are then drawn at random (with replacement) from the set $\{\tilde{u}_1^c, \ldots, \tilde{u}_T^c\}$, a pseudo-sample is built following the equation

$$\tilde{Y}_t^{(l)} = \sum_{i=1}^{p} \tilde{\Phi}_i^c \tilde{Y}_{t-i}^{(l)} + \tilde{u}_t^{(l)}, \quad t = 1, \ldots, T, \tag{B.9}$$

and the corresponding lag-augmented Wald statistic for H_0 – say $\tilde{W}_G^{(l)}$ – is computed. On repeating this operation N times, the bootstrap p-value and test are finally obtained as for the parametric bootstrap.

C. Appendix: Covariance matrix coefficients used in the simulation

In section 5, the lower triangular matrices R which determine error covariance matrices $\Sigma = RR'$ were defined as follows:

$$R = \begin{bmatrix} 0.01 & 0.00 \\ -0.02 & 0.03 \end{bmatrix}, \quad \text{for } k = 2,$$

$$R = \begin{bmatrix} 0.01 & 0.00 & 0.00 \\ -0.02 & 0.03 & 0.00 \\ -0.01 & 0.01 & 0.02 \end{bmatrix}, \quad \text{for } k = 3,$$

$$R = \begin{bmatrix} 0.01 & 0.00 & 0.00 & 0.00 \\ -0.02 & 0.03 & 0.00 & 0.00 \\ -0.01 & 0.01 & 0.02 & 0.00 \\ -0.03 & 0.02 & 0.01 & 0.01 \end{bmatrix}, \quad \text{for } k = 4,$$

$$R = \begin{bmatrix} 0.01 & 0.00 & 0.00 & 0.00 & 0.00 \\ -0.02 & 0.03 & 0.00 & 0.00 & 0.00 \\ -0.01 & 0.01 & 0.02 & 0.00 & 0.00 \\ -0.03 & 0.02 & 0.01 & 0.01 & 0.00 \\ 0.01 & -0.02 & 0.03 & -0.01 & 0.02 \end{bmatrix}, \quad \text{for } k = 5,$$

$$R = \begin{bmatrix} 0.01 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -0.02 & 0.03 & 0.00 & 0.00 & 0.00 & 0.00 \\ -0.01 & 0.01 & 0.02 & 0.00 & 0.00 & 0.00 \\ -0.03 & 0.02 & 0.01 & 0.01 & 0.00 & 0.00 \\ 0.01 & -0.02 & 0.03 & -0.01 & 0.02 & 0.00 \\ 0.02 & -0.01 & -0.03 & 0.02 & 0.01 & 0.03 \end{bmatrix}, \quad \text{for } k = 6.$$

References

ANDREWS, D. W. K. (2000): "Inconsistency of the Bootstrap When a Parameter is on the Boundary of the Parameter Space," *Econometrica*, 68, 399–405.

ATHREYA, K. B. (1987): "Bootstrap of the Mean in the Infinite Variance Case," *The Annals of Statistics*, 15, 724–731.

BALTAGI, B. (ed.) (2001): *Companion to Theoretical Econometrics*, Blackwell Companions to Contemporary Economics. Basil Blackwell, Oxford, U.K.

BARNARD, G. A. (1963): "Comment on 'The Spectral Analysis of Point Processes' by M. S. Bartlett," *Journal of the Royal Statistical Society, Series B*, 25, 294.

BASAWA, I. V., A. K. MALLIK, W. P. MCCORMICK, J. H. REEVES, AND R. L. TAYLOR (1991): "Bootstrapping Unstable First-Order Autoregressive Processes," *The Annals of Statistics*, 19, 1098–1101.

BENKWITZ, A., H. LÜTKEPOHL, AND M. H. NEUMANN (2000): "Problems Related to Confidence Intervals for Impulse Responses of Autoregressive Processes," *Econometric Reviews*, 19, 69–103.

BERKOWITZ, J., AND L. KILIAN (2000): "Recent Developments in Bootstrapping Time Series," *Econometric Reviews*, 19, 1–48, With Comment by Russell Davidson, 49-54.

BERNANKE, B. S., AND I. MIHOV (1998): "Measuring Monetary Policy," *The Quarterly Journal of Economics*, 113(3), 869–902.

BIRNBAUM, Z. W. (1974): "Computers and Unconventional Test-Statistics," in *Reliability and Biometry*, ed. by F. Proschan, and R. J. Serfling, pp. 441–458. SIAM, Philadelphia, PA.

BLOUGH, S. R. (1992): "The Relationship between Power and Level for Generic Unit Root Tests in Finite Samples," *Journal of Applied Econometrics*, 7, 295–308.

CANER, M., AND L. KILIAN (1999): "Size Distortions of Tests of the Null Hypothesis of Stationarity: Evidence and Implications for the PPP Debate," Technical report, Department of Economics, University of Michigan, Ann Arbor, MI 48109-1220, ikilian@umich.edu.

CHOI, I. (1993): "Asymptotic Normality of the Least-Squares Estimates for Higher Order Autoregressive Integrated Processes with some Applications," *Econometric Theory*, 9, 263–282.

CLEMENTS, M. P., AND D. F. HENDRY (eds.) (2002): *A Companion to Economic Forecasting*, Blackwell Companions to Contemporary Economics. Blackwell, Oxford, U.K.

COCHRANE, J. H. (1991): "A Critique of the Application of Unit Root Tests," *Journal of Economic Dynamics and Control*, 15, 275–284.

DHRYMES, P. J. (1998): *Time Series, Unit Roots, and Cointegration*. Academic Press, San Diego, California, USA.

DOLADO, J. J., AND H. LÜTKEPOHL (1996): "Making Wald Tests Work for Cointegrated VAR Systems," *Econometric Reviews*, 15, 369–386.

DUFOUR, J.-M. (2003): "Identification, Weak Instruments and Statistical Inference in Econometrics," *Canadian Journal of Economics*, 36(4), 767–808.

——— (2005): "Monte Carlo Tests with Nuisance Parameters: A General Approach to Finite-Sample Inference and Nonstandard Asymptotics in Econometrics," *Journal of Econometrics*, forthcoming.

DUFOUR, J.-M., AND T. JOUINI (2004): "Finite-Sample Simulation-Based Tests in VAR Models with Applications to Order Selection and Causality Testing," Discussion paper, Centre interuniversitaire de recherche en analyse des organisations (CIRANO) and Centre interuniversitaire de recherche en économie quantitative (CIREQ), Université de Montréal.

DUFOUR, J.-M., AND L. KHALAF (2001): "Monte Carlo Test Methods in Econometrics," in Baltagi (2001), chap. 23, pp. 494–519.

——— (2002): "Simulation Based Finite and Large Sample Tests in Multivariate Regressions," *Journal of Econometrics*, 111(2), 303–322.

DUFOUR, J.-M., D. PELLETIER, AND É. RENAULT (2005): "Short Run and Long Run Causality in Time Series: Inference," *Journal of Econometrics*, forthcoming.

DUFOUR, J.-M., AND E. RENAULT (1998): "Short-Run and Long-Run Causality in Time Series: Theory," *Econometrica*, 66, 1099–1125.

DWASS, M. (1957): "Modified Randomization Tests for Nonparametric Hypotheses," *Annals of Mathematical Statistics*, 28, 181–187.

EFRON, B., AND R. J. TIBSHIRANI (1993): *An Introduction to the Bootstrap*, vol. 57 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, New York.

FAUST, J. (1996): "Near Observational Equivalence Problems and Theoretical Size Problems with Unit Root Tests," *Econometric Theory*, 12, 724–732.

FAUST, J. (1999): "Theoretical Confidence Level Problems with Confidence Intervals for the Spectrum of a Time Series," *Econometrica*, 67, 629–637.

GOFFE, W. L., G. D. FERRIER, AND J. ROGERS (1994): "Global Optimization of Statistical Functions with Simulated Annealing," *Journal of Econometrics*, 60, 65–99.

GOURIÉROUX, C., AND A. MONFORT (1997): *Time Series and Dynamic Models*. Cambridge University Press, Cambridge, U.K.

GRANGER, C. W. J. (1969): "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37, 424–459.

HALL, P. (1992): *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York.

HAMILTON, J. D. (1994): *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

HANSEN, P. R., AND S. JOHANSEN (1998): *Workbook on Cointegration*. Oxford University Press, Oxford, U.K.

HATANAKA, M. (1996): *Time-Series-Based Econometrics*, Advanced Texts in Econometrics. Oxford University Press, Oxford, U.K.

HENDRY, D. F. (1995): *Dynamic Econometrics*. Oxford University Press, Oxford, U.K.

INOUE, A., AND L. KILIAN (2002a): "Bootstrapping Autoregressive Processs with Possible Unit Roots," *Econometrica*, 70(1), 377–391.

INOUE, A., AND L. KILIAN (2002b): "Bootstrapping Smooth Functions of Slope Parameters and Innovation Variances in $VAR(\infty)$ Models," *International Economic Review*, 43(2), 309–332.

INOUE, A., AND L. KILIAN (2003): "The Continuity of the Limit Distribution in the Parameter of Interest is Not Essential for the Validity of the Bootstrap," *Econometric Theory*, 19(6), 944–961.

JEONG, J., AND G. S. MADDALA (1993): "A Perspective on Application of Bootstrap Methods in Econometrics," in *Handbook of Statistics 11: Econometrics*, ed. by G. S. Maddala, C. R. Rao, and H. D. Vinod, pp. 573–610. North-Holland, Amsterdam.

JOHANSEN, S. (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Advanced Texts in Econometrics. Oxford University Press, Oxford, U.K.

KILIAN, L. (1998a): "Confidence Intervals for Impulse Responses Under Departures From Normality," *Econometric Reviews*, 17(1), 1–29.

KILIAN, L. (1998b): "Small-Sample Confidence Intervals for Impulse Response Functions," *Review of Economics and Statistics*, 80, 218–230.

KILIAN, L., AND U. DEMIROGLU (1997): "Residual-Based Bootstrap Tests for Normality in Autoregresssions," Working Papers 14, Department of Economics, University of Michigan, Ann Arbor, Michigan.

KUROZUMI, E., AND T. YAMAMOTO (2000): "Modified Lag Augmented Vector Autoregressions," *Econometric Reviews*, 19, 207–231.

LEHMANN, E. L. (1986): *Testing Statistical Hypotheses, 2nd edition*. John Wiley & Sons, New York.

LI, H., AND G. S. MADDALA (1996): "Bootstrapping Time Series Models," *Econometric Reviews*, 15, 115–158.

LÜTKEPOHL, H. (1993): *Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin, second edn.

——— (2001): "Vector Autoregressions," in Baltagi (2001), chap. 32, pp. 678–699.

MADDALA, G. S., AND I.-M. KIM (1998): *Unit Roots, Cointegration and Structural Change*. Cambridge University Press, Cambridge, U.K.

MCALEER, M., AND L. OXLEY (eds.) (1999): *Practical Issues in Cointegration Analysis*. Blackwell, Oxford, U.K.

PAPARODITIS, E. (1996): "Bootstrapping Autoregressive and Moving Average Parameter Estimates of Infinite Order Vector Autoregressive Processes," *Journal of Multivariate Analysis*, 57, 277–296.

PARK, J. Y., AND P. C. B. PHILLIPS (1989): "Statistical Inference in Regressions with Integrated Processes: Part 2," *Econometric Theory*, 5, 95–131.

PÖTSCHER, B. (2002): "Lower Risk Bounds and Properties of Confidence Sets for Ill-Posed Estimation Problems with Applications to Spectral Density and Persistence Estimation, Unit Roots and Estimation of Long Memory Parameters," *Econometrica*, 70(3), 1035–1065.

REINSEL, G. C. (1993): *Elements of Multivariate Time Series Analysis*. Springer-Verlag, New York.

SIMS, C. (1971a): "Distributed Lag Estimation When the Parameter Space is Explicitly Infinite-Dimensional," *Annals of Mathematical Statistics*, 42, 1622–1636.

——— (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

SIMS, C. A. (1971b): "Discrete Approximations to Continuous Time Distributed Lags in Econometrics," *Econometrica*, 39, 545–563.

SIMS, C. A., J. H. STOCK, AND M. W. WATSON (1990): "Inference in Linear Time Series Models with some Unit Roots," *Econometrica*, 58, 113–144.

SRIRAM, T. N. (1994): "Invalidity of Bootstrap for Critical Branching Processes with Immigration," *The Annals of Statistics*, 22, 1013–1023.

TANAKA, K. (1996): *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory*. John Wiley & Sons, New York.

TODA, H. Y., AND P. C. B. PHILLIPS (1993): "Vector Autoregressions and Causality," *Econometrica*, 61, 1367–1393.

TODA, H. Y., AND P. C. B. PHILLIPS (1994): "Vector Autoregression and Causality: A Theoretical Overview and Simulation Study," *Econometric Reviews*, 13, 259–285.

TODA, H. Y., AND T. YAMAMOTO (1995): “Statistical Inference in Vector Autoregressions with Possibly Integrated Processes,” *Journal of Econometrics*, 66, 225–250.

WIENER, N. (1956): “The Theory of Prediction,” in *The Theory of Prediction*, ed. by E. F. Beckenback, chap. 8. McGraw-Hill, New York.

YAMADA, H., AND H. Y. TODA (1998): “Inference in Possibly Integrated Vector Autoregressive Models: Some Finite Sample Evidence,” *Journal of Econometrics*, 86, 55–95.

YAMAMOTO, T. (1996): “A Simple Approach to the Statistical Inference in Linear Time Series Models Which May Have Some Unit Roots,” *Hitotsubashi Journal of Economics*, 37, 87–100.