

WHICH MODEL TO MATCH?

2012

Matteo Barigozzi, Roxana Halbleib
and David Veredas

Documentos de Trabajo
N.º 1229

BANCO DE **ESPAÑA**
Eurosistema



WHICH MODEL TO MATCH?

WHICH MODEL TO MATCH?

Matteo Barigozzi ^(*) ^(***)

LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE AND ECARES

Roxana Halbleib ^(**)

UNIVERSITY OF KONSTANZ

David Veredas ^(***)

UNIVERSITÉ LIBRE DE BRUXELLES

(*) London School of Economics and Political Science; email: m.barigozzi@lse.ac.uk.

(**) University of Konstanz; email: roxana.halbleib@uni-konstanz.de.

(***) ECARES, Solvay Brussels School of Economics and Management, Université libre de Bruxelles; email: david.veredas@ulb.ac.be.

This work was written while David Veredas was visiting the Department of Monetary and Financial Studies at the Banco de España (Madrid, Spain). Roxana Halbleib acknowledges financial support from the European Social Fund and from the Ministry of Science, Research and the Arts Baden-Württemberg. David Veredas acknowledges financial support from the Belgian National Bank and the IAP P6/07 contract, from the IAP programme (Belgian Scientific Policy), 'Economic policy and finance in the global economy'. David Veredas is member of ECORE, the association between CORE and ECARES.

We are grateful to Francisco Blasques, Christian Brownlees, Richard Davis, Miguel Ángel Delgado, Eric Ghysels, Alastair Hall, Marc Hallin, Atsushi Inoue, Paolo Santucci de Magistris, Olivier Scaillet and Enrique Sentana for insightful remarks. We also acknowledge the useful feedback from the participants of the International Symposium in Statistics and Financial Time Series (Wakayama, Japan), the Interdisciplinary Workshop on Econometric and Statistical Modelling of Multivariate Time Series (Louvain-la-neuve, Belgium), the DFHWorkshop in Econometrics (Königsfeld, Germany), the conference on Statistical Models for Financial Data (Graz, Austria), and the Workshop on Financial Time Series Analysis: High-dimensionality, Non-stationarity and the Financial Crisis (Singapore), as well as the seminar participants at the Universities of Carlos III de Madrid, Manchester, and Pompeu Fabra. Any errors and inaccuracies are ours.

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the INTERNET at the following website: <http://www.bde.es>.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

© BANCO DE ESPAÑA, Madrid, 2012

ISSN: 1579-8666 (on line)

Abstract

The asymptotic efficiency of indirect estimation methods, such as the efficient method of moments and indirect inference, depends on the choice of the auxiliary model. To date, this choice has been somewhat ad hoc and based on an educated guess. In this article we introduce a class of information criteria that helps the user to optimize the choice between nested and non-nested auxiliary models. They are the indirect analogues of the widely used Akaike-type criteria. A thorough Monte Carlo study based on two simple and illustrative models shows the usefulness of the criteria.

Keywords: Indirect inference, efficient method of moments, auxiliary model, information criteria, asymptotic efficiency.

JEL classification: C13, C52.

Resumen

La eficiencia asintótica de los estimadores de inferencia indirecta, tales como el método eficiente de los momentos, depende de la elección del modelo auxiliar. Hasta ahora, esta elección era *ad hoc* y basada en criterios subjetivos. En este artículo introducimos una clase de criterios de información que ayuda al usuario a escoger entre modelos anidados y no anidados. Esta clase es la análoga a los ampliamente utilizados criterios del tipo Akaike. Un detallado estudio de Monte Carlo basado en dos modelos simples e ilustrativos muestran la utilidad de los criterios.

Palabras clave: inferencia indirecta, método eficiente de los momentos, modelo auxiliar, criterios de información, eficiencia asintótica.

Códigos JEL: C13, C52.

1 Introduction

The title of this article intentionally resembles to Gallant and Tauchen (1996)'s *Which moments to match?* They, Gouriéroux et al. (1993), and Smith (1993) introduce an inference method based on matching sample and theoretical features of an auxiliary model in order to estimate the parameters of the model of interest. Gouriéroux et al. (1993) call it Indirect Inference, Smith (1993) Simulated Quasi Maximum Likelihood, and Gallant and Tauchen (1996) the Efficient Method of Moments (EMM). We will denote these methods generically as *indirect methods*.

Because of their flexibility, applications of indirect methods in all areas of economics abound.¹ Provided that the auxiliary estimators converge to some point in their space and that the so-called binding function is injective, then the general theory shows that any auxiliary model is valid. Validity in this context means that the estimated parameters are, under fairly weak regularity conditions, consistent, asymptotically Gaussian, and with a variance-covariance matrix that is a function of the binding function and the Fisher information matrix of the auxiliary model.

In practice, a good auxiliary model is a model that, in some sense, is close to the one of interest. Gallant and Tauchen (1996) introduce the SNP (Semi Non-Parametric) score generator as universal auxiliary model. It is based on Hermite expansions that can, in principle, span the score vector of any model of interest. The SNP score generator is a class of auxiliary models since there are as many models as terms in the expansion; so a choice has to be made. An alternative to SNP is to use an auxiliary model that is much bigger (and hence with many more parameters) than the model of interest, on the grounds that the bigger the auxiliary model, the better the approximation to the model of interest. In the limit, this leads to the sieve-based approach of Nickl and Potscher (2010) and Blasques (2011). Though theoretically it is a reasonable argument, in practice we face limited information given by the sample size, meaning that the finite sample properties of the estimated parameters may deteriorate sensibly with the dimension of the auxiliary model. This is the classical trade-off between model fit and the number of parameters.

To the best of our knowledge no systematic analysis has been done on how to choose the auxiliary model within the general theory of indirect methods. In this article we propose information criteria that can be used for choosing between a set of nested and non-nested auxiliary models. They are the indirect analogues of the widely used Akaike-type model information criteria (among those, the most widely used are AIC of Akaike (1974), BIC of Schwarz (1978), and HQIC of Hannan and Quinn (1979)). The indirect auxiliary model selection criteria (that we denote with subscript *IM* for Indirect Methods) are based on the likelihood of the estimated parameters of the model of interest, which is centered

¹A non-exhaustive list is: discrete stochastic volatility models (Monfardini (1998), Gallant et al. (1997), Lombardi and Calzolari (2009), and Corsi and Reno (2012)), continuous stochastic volatility models (Chernov et al. (2003) and Grassi and de Magistris (2012)), diffusion models (Engle and Lee (1996), Broze et al. (1998)), threshold models (Guay and Scaillet (2003)), heavy tailed distributions (Lombardi and Veredas (2009), Garcia et al. (2011)), DSGE models (Dridi et al. (2007)), survival data models (Jiang and Turnbull (2003)), stochastic agent-based models (Jang and Lux (2012)), labor income risk (Guvenen and Smith (2010)), discrete choice models (Keane and Smith (2003)), heteroskedastic factor models (Calzolari et al. (2008)), learning models (Czellar and Calvet (2011)), and price contingent claims in derivative markets (Phillips and Yu (2009)).

Gaussian regardless of the choice of the auxiliary model. The difference therefore comes in terms of precision, i.e. the variance–covariance matrix. But to account for the trade–off between the fit of the auxiliary model to the model of interest and its dimension, we need a penalty that depends on the number of parameters of the auxiliary model.

Our work relates to the existing literature on Generalized Method of Moments (GMM) for choosing the moment conditions (MC). Two issues arise in the GMM context: the choice of MC that ensures consistency of the estimates, and the choice of no redundant moment conditions. The former is related to weak identification, or to the case where there is insufficient information to yield a consistent estimator. Andrews (1999) proposes GMM analogues of the AIC, BIC and HQIC (which we denote with subscript *GMM*) for consistently selecting the correct MC.² As for the choice of no redundant MC, it relates with moment conditions that provide no incremental information about the parameters of the model of interest. Hall et al. (2007) introduce the Relevant Moment Selection Criteria (RMSC) where the word relevant denotes the subset of MC that are asymptotically efficient but that contain no redundant moment conditions. Though redundant moment conditions have no asymptotic impact, Hall and Peixe (2003) report simulation evidence that their inclusion can lead to a poor approximation of the limiting distribution to the finite sample behavior.

In the framework of indirect methods, the choice of the auxiliary model presents no concerns in terms of consistency, provided the conditions stated above. The closest to our work is therefore Hall et al. (2007). They propose RMSC based on the observation that the entropy of the limiting distribution of the GMM estimator provides a metric for information of the moment conditions. These ideas were also used by Hall et al. (2011) in the context of DSGE models, where the auxiliary model are impulse response functions (IRF hereafter) stemming from a VAR. This is alike to Dridi et al. (2007) that estimate a DSGE model with sequential partial indirect inference.

The indirect inference criteria that we present can also be used for choosing between and within other matching inference methods, such as GMM, the Method of Simulated Moments (MSM; McFadden (1989), Pakes and Pollard (1989), and Duffie and Singleton (1993)) or the Method of Simulated Quantiles (MSQ; Dominicy and Veredas (2012)). As long as the functions to match lead to consistent estimators, the only difference between them is in terms of the variance–covariance matrix, as in the indirect methods. In other words, the information criteria only depend on the variance–covariance matrix of the estimators and the number of functions to match.

We proceed with a thorough Monte Carlo study on two simple, yet important and illustrative, models: a moving average (MA) of order one and a stochastic volatility (SV) model. In this context of indirect methods, the former has been studied by Gouriéroux et al. (1993), Ghysels et al. (2003), and Genton and Ronchetti (2003) while the latter by Monfardini (1998) and Calzolari et al. (2004). For the MA, we choose as auxiliary models an AR(*r*) and the true model. The purpose of this study is to investigate the performance of the criteria when choosing among the true model and nested models. The

²See also Andrews and Lu (2001), Kitamura (2002), Hong et al. (2003) and Hong and Preston (2005) for related work on moment selection criteria, and Konishi and Kitagawa (1996) for a generalized information criteria.

criteria choose the true model as best auxiliary model. The second best is the AR model with lag order that increases with the magnitude of the moving average parameter. As for the SV model, we consider two auxiliary models: ARCH(r) and GARCH(1,1) on the observations, and AR(r) and ARMA(1,1) on the log of the square of the observations. The purpose of this study is to investigate the performance of the criteria when choosing among nested and non-nested auxiliary models. The criteria choose r that increases with the SV autoregressive parameter, and always prefer the ARCH to the AR model.

The remaining sections are laid out as follows. Section 2 introduces the notation and the basic background of indirect methods, while Section 3 describes the criteria. Section 4 covers the Monte Carlo study. Conclusions are touched upon in Section 5. Assumptions, technical lemmas, proofs, and tables are relegated to the appendices.

2 Indirect methods

The model of interest We consider a $n \times 1$ random vector \mathbf{Y} that follows a model $\mathcal{M}(\mathbf{y}, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denotes the p -dimensional vector of unknown parameters that are in an interior point of the parameter set $\Theta \subset \mathbb{R}^p$, and $\mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^n$ is a realization of \mathbf{Y} . The model is defined by a parametric probability law

$$\mathcal{M}(\mathbf{y}, \boldsymbol{\theta}) = \{P_{\boldsymbol{\theta}}(\mathbf{y}), \mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^n, \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p\}. \quad (1)$$

The true value of the parameters is denoted by $\boldsymbol{\theta}_0$ and it is an interior point of Θ . We define the Fisher information matrix for this model as $\mathcal{I}(\boldsymbol{\theta}_0)$. We assume to have T realizations of \mathbf{Y} and we denote as $\mathbf{y}_t = (y_{1t} \dots y_{nt})'$ the $n \times 1$ vector of a generic realization with $t = 1, \dots, T$.

The class of models we consider is broad, covering those quoted in the introduction, and including static and dynamic models, with unobserved components, trends, and various innovation terms. Each process and model has its own assumptions. Establishing general hypothesis on \mathbf{Y} and $\mathcal{M}(\mathbf{y}, \boldsymbol{\theta})$ is therefore idle, beyond A.1 in Appendix A on $\boldsymbol{\theta}$ and its space. It is however important that the model is such that, for a given value of the parameter vector, simulated draws of $\mathbf{y}_1(\boldsymbol{\theta}), \dots, \mathbf{y}_T(\boldsymbol{\theta})$ can be obtained. This typically implies that $\mathcal{M}(\mathbf{y}, \boldsymbol{\theta})$ contains at least one *i.i.d.* innovation with known distribution.

The auxiliary model Consider the auxiliary model for \mathbf{Y}

$$\mathcal{M}^a(\mathbf{y}, \boldsymbol{\beta}) = \{P_{\boldsymbol{\beta}}(\mathbf{y}), \mathbf{y} \in \mathcal{Y} \subset \mathbb{R}^n, \boldsymbol{\beta} \in \mathcal{B} \subset \mathbb{R}^q\}$$

that depends on a q -dimensional vector of parameters $\boldsymbol{\beta}$. We assume that for model $\mathcal{M}^a(\mathbf{y}, \boldsymbol{\beta})$ the log-likelihood $\ell^a(\boldsymbol{\beta}; \mathbf{y}_t)$ is known and available analytically. We define $\widehat{\boldsymbol{\beta}}_T$ as the Quasi Maximum Likelihood (QML) estimator of the true value of the parameters $\boldsymbol{\beta}_0$. Under the standard regularity conditions of QML estimation (see also assumptions A.1–A.4 in Appendix A), $\widehat{\boldsymbol{\beta}}_T$ has the following limiting distribution

$$\sqrt{T}(\widehat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0) \sim \mathcal{N}(\mathbf{0}_q, (\mathbf{J}(\boldsymbol{\beta}_0)\mathbf{I}^{-1}(\boldsymbol{\beta}_0)\mathbf{J}(\boldsymbol{\beta}_0))^{-1}), \text{ as } T \rightarrow \infty,$$

where $\mathbf{0}_q$ is a $q \times 1$ vector of zeros, $\mathbf{I}(\boldsymbol{\beta}_0)$ is the $q \times q$ Fisher information matrix of the auxiliary model, and $\mathbf{J}(\boldsymbol{\beta}_0)$ is minus the expectation of the Hessian of the log-likelihood.

For a given value of θ , we simulate H paths of length T : $\mathbf{y}_1^h(\theta), \dots, \mathbf{y}_T^h(\theta)$, $h = 1, \dots, H$, from the model of interest $\mathcal{M}(\mathbf{y}, \theta)$ and we estimate

$$\widehat{\beta}_{HT}(\theta) = \arg \max_{\beta \in \mathcal{B}} \frac{1}{HT} \sum_{h=1}^H \sum_{t=1}^T \ell^a(\beta, \mathbf{y}_t^h(\theta)). \quad (2)$$

Under assumptions A.1–A.4 in Appendix A, this is a consistent and asymptotically Gaussian estimator (as $T \rightarrow \infty$) of the so-called binding function $\mathbf{b}(\theta)$ linking θ to β and such that $\mathbf{b}(\theta_0) = \beta_0$. Finally, notice that we could simulate paths of length $T^* \neq T$ (typically $T^* > T$) and therefore both the estimated auxiliary parameters and their estimated asymptotic variance–covariance matrix will be functions of T^* instead of T . For simplicity of notation we omit this distinction in what follows.

Indirect estimation The indirect estimator is defined as

$$\widehat{\theta}_{HT,IM,a,q} = \arg \min_{\theta \in \Theta} \left(\widehat{\beta}_{HT}(\theta) - \widehat{\beta}_T \right)' \Omega \left(\widehat{\beta}_{HT}(\theta) - \widehat{\beta}_T \right), \quad (3)$$

where Ω is a $q \times q$ weighting matrix defining the metric. We make explicit the dependence of the estimated parameter $\widehat{\theta}_{HT,IM,a,q}$ on the auxiliary model a and the dimension of its parameter vector q . The β -based optimization (3) was introduced by Gouriéroux et al. (1993). Alternatively we may consider a $h(\beta)$ -based optimization, for h an injective and continuously differentiable function with respect to β , and after a suitable re-parametrization of the weighting matrix by means of the delta method. The vector $h(\beta)$ is of dimension q' that can be larger or equal to q . In the score-based optimization of Gallant and Tauchen (1996) $q' = q$, while in the IRF-based optimization of Hall et al. (2011) $q' > q$. A special case is when the auxiliary model is estimated under constraints on the parameter space, i.e. constrained QML that imply the estimation of multipliers (Calzolari et al. (2004)). The vector β is expanded to dimension $q' > q$ as to include the multipliers. But it is q that enters in the penalty. If the constraints are not binding, the multipliers are zero, and inference is as unconstrained. If, by contrast, the constraints bind, the bound parameters contain no information about θ , as it goes through the associated multipliers. In the sequel, we consider the case of the β -based optimization, and hence the penalty of the criteria is a function of q . If, instead, the $h(\beta)$ -based optimization is used, the criteria are easily adapted by replacing q by q' .

Under assumptions A.1–A.7 in Appendix A and the appropriate model assumptions, the estimators are consistent with limiting distribution

$$\sqrt{T}(\widehat{\theta}_{HT,IM,a,q} - \theta_0) \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_{H,\Omega,a,q}(\theta_0)), \text{ as } T \rightarrow \infty.$$

The asymptotic variance–covariance matrix depends on the true value of the parameters, the number of simulated paths, and the weighting matrix. The optimal choice for the latter is $\Omega^* = \mathbf{J}(\beta)\mathbf{I}^{-1}(\beta)\mathbf{J}(\beta)$ so that

$$\mathbf{W}_{H,\Omega^*,a,q}(\theta) = \left(1 + \frac{1}{H} \right) \left[\frac{\partial \mathbf{b}(\theta)}{\partial \theta} \mathbf{J}(\beta) \mathbf{I}^{-1}(\beta) \mathbf{J}(\beta) \frac{\partial \mathbf{b}(\theta)}{\partial \theta'} \right]^{-1}.$$

Since the value of H is arbitrarily chosen, we consider $H \rightarrow \infty$ and, as in the next section we need further notation, hereafter we skip the subindexes H and Ω^* , so $\widehat{\theta}_{HT,IM,a,q}$

becomes $\widehat{\boldsymbol{\theta}}_{T,IM,a,q}$ and $\mathbf{W}_{H,\Omega^*,a,q}(\boldsymbol{\theta})$ becomes

$$\mathbf{W}_{a,q}(\boldsymbol{\theta}) = \left[\frac{\partial \mathbf{b}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{J}(\boldsymbol{\beta}) \mathbf{I}^{-1}(\boldsymbol{\beta}) \mathbf{J}(\boldsymbol{\beta}) \frac{\partial \mathbf{b}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right]^{-1}. \quad (4)$$

A plethora of theoretical developments and improvements of the original methods have been proposed. What follows is a selection of an extensive and growing literature. Dhaene et al. (1998) introduce a notion of indirect encompassing and non-nested hypotheses tests using indirect methods. That is, the case where not only the relevant binding function does not have closed form, but also at least one of the competing models. Billio and Monfort (2003) propose kernel-based conditional expectations for the binding functions, which have the advantage that no optimization step is involved in the computation of the binding function. They also propose two criteria specific for conditional expectations as binding functions. Another area of research within indirect methods concerns robustification. Genton and Ronchetti (2003) develop robust indirect inference by deriving the influence function of the indirect estimator, as well as the level and power influence function of indirect tests. In a similar spirit, Trojani and Ortelli (2005) focus on robust EMM estimation of a general parametric time series stationary processes. Czellar and Ronchetti (2010) propose robust over-identification tests that exhibit better finite sample accuracy than classical tests. Partial and semi-nonparametric indirect methods have also been introduced. Dridi et al. (2007) introduce the sequential partial indirect inference approach, in the sense that the model of interest is partially well specified. They show the usefulness of the sequential method on the calibration of dynamic stochastic general equilibrium (DGSE) models. Also motivated by DSGE models, Blasques (2011) presents semi-nonparametric indirect methods, i.e. a sieve class of estimator, on the grounds that, at population level, an infinite dimensional auxiliary model may be needed for the indirect estimation of many economic models. Indirect methods have also been used for bias correction (see Gouriéroux et al. (2000) for general time time series models, Ghysels et al. (2003) for a MA(1), and Gouriéroux et al. (2010) for dynamic panel data models). The advantage with other bias correction techniques is that indirect methods do not require a given explicit form for the bias function of its expansion, as it is calibrated via simulations. Last, Phillips (2012) provides new limit theory for the delta method and the continuous mapping theorem for cases when the problem involves sample functions that depend on the sample size, or when the quantity of interest appears in an implicit functional form, as it happens in indirect methods.

3 The indirect information criteria

The principle Consider two auxiliary models $\mathcal{M}^{a_1}(\mathbf{y}, \boldsymbol{\beta}_1)$ and $\mathcal{M}^{a_2}(\mathbf{y}, \boldsymbol{\beta}_2)$ with $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ of dimension q_1 and q_2 respectively. Under the optimal choice of the weighting matrices and as $T, H \rightarrow \infty$ we have

$$\begin{aligned} \sqrt{T}(\widehat{\boldsymbol{\theta}}_{T,IM,a_1,q_1} - \boldsymbol{\theta}_0) &\sim \mathcal{N}(\mathbf{0}, \mathbf{W}_{a_1,q_1}(\boldsymbol{\theta}_0)) \quad \text{and} \\ \sqrt{T}(\widehat{\boldsymbol{\theta}}_{T,IM,a_2,q_2} - \boldsymbol{\theta}_0) &\sim \mathcal{N}(\mathbf{0}, \mathbf{W}_{a_2,q_2}(\boldsymbol{\theta}_0)). \end{aligned}$$

Both estimators are consistent and the limiting distribution of $\sqrt{T}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q} - \boldsymbol{\theta}_0)$ is centered Gaussian. The difference between them comes in terms of precision, i.e. the variance–covariance matrix. Choices of \mathcal{M}^a and q cover both nested and non–nested auxiliary models. For simplicity of notation we use the index a to indicate a generic auxiliary model \mathcal{M}^a . In particular, if the models are nested $\mathcal{M}^{a_1} \subset \mathcal{M}^{a_2}$ (in the sense that they belong to the same auxiliary class but $q_1 < q_2$), we use the same index $a = a_1 = a_2$ for both. If the models are non–nested, there is no well defined inclusion relation between \mathcal{M}^{a_1} and \mathcal{M}^{a_2} , q_1 and q_2 may or may not be equal, and we say $a_1 \neq a_2$. Finally, we define $\mathcal{A}_{q_{\max}}$ as the set of all possible auxiliary models satisfying assumptions A.1–A.7 in Appendix A, where q_{\max} is the largest dimension of the auxiliary parameter we allow for.³ Each element of $\mathcal{A}_{q_{\max}}$ is made of a couple (a, q) indicating the chosen auxiliary model and the dimension of the corresponding auxiliary parameter.

We say that a matrix \mathbf{A} is larger than another \mathbf{B} , and we denote it as $\mathbf{A} \succ \mathbf{B}$, when $\|\mathbf{A}\| > \|\mathbf{B}\|$, where $\|\cdot\|$ is a norm (such as the Frobenius norm $\|\mathbf{A}\|^2 = \text{Tr}(\mathbf{A}\mathbf{A}')$).⁴ Consider two nested (thus belonging to the same class a) auxiliary models with $q_1 < q_2$. Parzen (1959) and Carrasco and Florens (2000) show that

$$\mathbf{W}_{a,q_1}(\boldsymbol{\theta}_0) \succ \mathbf{W}_{a,q_2}(\boldsymbol{\theta}_0) \succ \mathbf{W}_{a,q_{\max}}(\boldsymbol{\theta}_0).$$

That is, the more auxiliary parameters, the more information and the smaller is the asymptotic variance–covariance of the estimated parameters. Whatever the value of q_{\max} , the lower bound for the auxiliary models is always larger than the Cramér–Rao lower bound, i.e. $\mathbf{W}_{a,q_{\max}}(\boldsymbol{\theta}_0) \succ \mathcal{I}^{-1}(\boldsymbol{\theta}_0)$ for any a and q_{\max} (Gouriéroux and Monfort (1996)). Hence, the optimal choice is the largest possible auxiliary model. But this is not the case when we replace $\boldsymbol{\theta}_0$ by $\widehat{\boldsymbol{\theta}}_{T,IM,a_1,q_1}$ as we have to take into account the increasing estimation error due to the increasing number of estimated parameters. Some penalization that is a function of q is thus required.

Simulating the estimated parameters For computing the indirect criteria that we propose below, we need realizations of the random vector $\sqrt{T}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q} - \boldsymbol{\theta}_0)$ that we do not have since $\boldsymbol{\theta}_0$ is not observed. But they are straightforward to simulate. Let $\mathbf{X}_{a,q} \equiv \sqrt{T}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q} - \boldsymbol{\theta}_0)$, so that $\mathbf{X}_{a,q} \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_{a,q}(\boldsymbol{\theta}_0))$, as $T \rightarrow \infty$. Let $\widehat{\mathbf{W}}_{T,a,q}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q})$ be the estimator of $\mathbf{W}_{a,q}(\boldsymbol{\theta}_0)$ (Gouriéroux et al. (1993)). We know that

$$p\lim_{T \rightarrow \infty} \widehat{\mathbf{W}}_{T,a,q}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q}) = \mathbf{W}_{a,q}(\boldsymbol{\theta}_0).$$

We can therefore simulate a sample of length N from a p –dimensional centered Gaussian distribution with variance–covariance matrix $\widehat{\mathbf{W}}_{T,a,q}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q})$. Let us denote the pdf of $\mathbf{X}_{a,q}$ by $\phi(\mathbf{x}_{a,q,i}; \widehat{\mathbf{W}}_{T,a,q}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q}))$, and the i –th simulated observation of $\mathbf{X}_{a,q}$ as the p –dimensional vector $\mathbf{x}_{a,q,i} = (x_{1,a,q,i} \dots x_{p,a,q,i})'$, with $i = 1, \dots, N$.

Notice that n , T and N are three independent quantities with different meanings. The first one, n , is the dimension of the random vector \mathbf{Y} , while T is the sample length of the

³For the sake of simplicity, in this section to compare two classes of auxiliary models and for each class we consider the same maximum number of parameters $q_{1,\max} = q_{2,\max} = q_{\max}$. It is however possible that $q_{1,\max} \neq q_{2,\max}$, as in the SV model in Section 4.

⁴Alternatively, \mathbf{A} is larger than a matrix \mathbf{B} if $\mathbf{A} - \mathbf{B}$ is a positive definite matrix

observed data. Once the data is given, n and T are fixed. By contrast, N is the sample length of the simulated vector $\mathbf{X}_{a,q}$, and can be chosen large enough for the Gaussian limiting distribution to hold. If we define the sample variance–covariance matrix of $\mathbf{X}_{a,q}$ as

$$\widetilde{\mathbf{W}}_{N,a,q} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{a,q,i} \mathbf{x}'_{a,q,i},$$

then the following holds

$$p \lim_{T \rightarrow \infty} \left(p \lim_{N \rightarrow \infty} \widetilde{\mathbf{W}}_{N,a,q} \right) = p \lim_{T \rightarrow \infty} \left(\widehat{\mathbf{W}}_{T,a,q}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q}) \right) = \mathbf{W}_{a,q}(\boldsymbol{\theta}_0).$$

The criterion The indirect Akaike information criterion, $AIC_{IM}(a, q)$, is given in the following proposition.

Proposition 1 *Given the models of Section 2 and the assumptions A.1–A.8 and A.10 in Appendix A, the optimal auxiliary model is the one within the model class \tilde{a} and auxiliary parameter dimension \tilde{q} such that*

$$(\tilde{a}, \tilde{q}) = \underset{(q, a) \in \mathcal{A}_{q_{\max}}}{\operatorname{argmin}} AIC_{IM}(a, q),$$

where

$$AIC_{IM}(a, q) = -2 \sum_{i=1}^N \log \phi(\mathbf{x}_{a,q,i}; \widetilde{\mathbf{W}}_{N,a,q}) + \frac{q(q+1)}{2}. \quad (5)$$

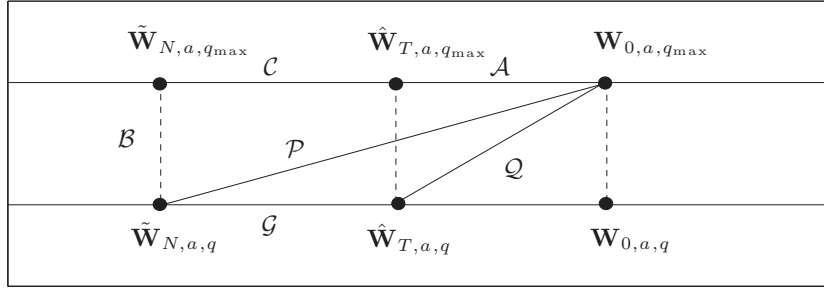
Proof: See Appendix B.

Three remarks are in order.

First, we give the intuition for the nested case, i.e. when the class of auxiliary models, a , is fixed. Figure 1 gives a geometric interpretation of this problem. The diagram is divided in two rows and three columns. The rows correspond to the largest dimension of the auxiliary model (top row) and a given dimension (bottom). The columns correspond to the variance–covariance matrices in population (right), and the sample and simulated counterparts (middle and left respectively). The calligraphic letters denote vectors.⁵ Asymptotically the auxiliary model that is the closest to the model of interest has the maximum possible dimension with variance–covariance matrix of the estimated parameters equal to $\mathbf{W}_{a,q_{\max}}(\boldsymbol{\theta}_0)$, with q_{\max} large enough. In practice however, $\mathbf{W}_{a,q_{\max}}(\boldsymbol{\theta}_0)$ is unknown and the variance–covariance matrix is estimated for a given $q < q_{\max}$: $\widehat{\mathbf{W}}_{T,a,q}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q})$. The aim is to find the value of q that minimizes the distance between $\mathbf{W}_{a,q_{\max}}(\boldsymbol{\theta}_0)$ and $\widehat{\mathbf{W}}_{T,a,q}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q})$. This value is not necessarily q_{\max} , for fixed T , as when q increases the number of parameters to estimate increases and the estimation error increases accordingly. However, since we use simulated samples $\mathbf{X}_{a,q}$, we need to use asymptotic results for large values of the sample size N . For this reason we have to consider the sample covariance $\widetilde{\mathbf{W}}_{N,a,q}$, and thus the left column in the diagram.

⁵The right and middle columns of this diagram are also the base of AIC_{ML} and AIC_{GMM} .

Figure 1: Diagrammatic representation of the nested case



The diagram is divided in two rows and three columns. The rows correspond to the largest dimension of the auxiliary model (top row) and a given dimension (bottom). The columns correspond to the variance-covariance matrices in population (right), and the sample and simulated counterparts (middle and left respectively). The calligraphic letters denote vectors.

What follows is a sketch of the proof for the nested case. Let $\mathbf{w} = \text{vech}(\mathbf{W})$ and $\mathcal{Q} = (\mathbf{w}_{0,a,q_{\max}} - \hat{\mathbf{w}}_{T,a,q})$. Then the distance between $\mathbf{W}_{0,a,q_{\max}}$ and $\hat{\mathbf{W}}_{T,a,q}$ can be written as the norm squared

$$\mathcal{Q}^2 = \|\mathbf{w}_{0,a,q_{\max}} - \hat{\mathbf{w}}_{T,a,q}\|_{\mathcal{J}}^2, \quad (6)$$

where the norm is based on Hessians \mathcal{J} as defined in Appendix B. We want the auxiliary model (a, q) that minimizes the distance \mathcal{Q}^2 . From the geometry of Figure 1, this distance equals

$$\mathcal{Q}^2 = (\mathcal{P} - \mathcal{G})^2 = \mathcal{P}^2 + \mathcal{G}^2 - 2\mathcal{G}\mathcal{P}. \quad (7)$$

From Lemma 1 in Appendix B, the term \mathcal{G}^2 is distributed as $\chi_{\frac{q(q+1)}{2}}^2/N$ so it has expectation $\frac{q(q+1)}{2N}$, while the last term has expectation zero for large N . Finally, by Pythagora's theorem we have

$$\mathcal{P}^2 = (\mathcal{A} + \mathcal{C})^2 + \mathcal{B}^2.$$

Now $\mathcal{A} + \mathcal{C}$ does not depend on q so it does not play any role in the criterion and we can treat it as a constant, say k , and \mathcal{B}^2 gives the log-likelihood term computed in $\tilde{\mathbf{W}}_{N,a,q}$. Therefore, by letting N diverging and taking expectations of (7), we can write

$$\begin{aligned} \mathbb{E}_w [\mathcal{Q}^2] &\simeq \mathbb{E}_w [(\mathcal{A} + \mathcal{C})^2 + \mathcal{B}^2 + \mathcal{G}^2 - 2\mathcal{G}\mathcal{P}] = \mathbb{E}_w [(\mathcal{A} + \mathcal{C})^2 + \mathcal{B}^2 + \mathcal{G}^2] \\ &\simeq k + \frac{2}{N} \sum_{i=1}^N \log \frac{\phi(\mathbf{x}_{i,a,q_{\max}}, \tilde{\mathbf{W}}_{N,a,q_{\max}})}{\phi(\mathbf{x}_{i,a,q}, \tilde{\mathbf{W}}_{N,a,q})} + \frac{q(q+1)}{2N}. \end{aligned} \quad (8)$$

By multiplying the last expression by N and keeping only the terms that depend on (a, q) we have the criterion. Summing up, \mathcal{Q}^2 depends on i) the distance between $\mathbf{W}_{a,q_{\max}}(\boldsymbol{\theta}_0)$ and its projection in a q dimensional subspace $\mathbf{W}_{a,q}(\boldsymbol{\theta}_0)$, i.e. the term \mathcal{B}^2 , and ii) the distance between $\tilde{\mathbf{W}}_{N,a,q}$ and the estimated variance-covariance matrix of the auxiliary model $\hat{\mathbf{W}}_{T,a,q}(\hat{\boldsymbol{\theta}}_{T,IM,a,q})$, i.e. the term \mathcal{G}^2 . While the first distance decreases as q increases, the second distance is an estimation error –or a penalty– that increases with q .

While this is the sketched proof for the nested case, AIC_{IM} compares nested and non-nested models, which implies that the best model is the true model, i.e. the one with

asymptotic variance covariance matrix $\mathcal{I}^{-1}(\boldsymbol{\theta}_0)$. But since this is never achievable, the geometric argument cannot be made. For this reason a more involving proof than original Akaike's is needed –as it is shown in detail in Appendix B– and that is valid for the nested and non-nested cases.

Second, there are important differences between AIC_{IM} and the ML and GMM counterparts. First, the likelihood term in AIC_{IM} is of the estimated parameters of the model of interest, which is centered Gaussian regardless of the choice of the auxiliary model and the model of interest (under assumptions A.1–A.8 in Appendix A and the appropriate model assumptions). By contrast, in AIC_{ML} the likelihood term depends on the model of interest and in AIC_{GMM} depends on the moment conditions. Second, in AIC_{IM} the penalty term depends on the parameters of the auxiliary model, while in AIC_{ML} it depends on the parameters of the model of interest. Third, the dimension of the parameters of the model of interest is constant over choices of the auxiliary model, whilst in AIC_{ML} the dimension varies with the model.

Finally, as it happens with AIC_{ML} and AIC_{GMM} , AIC_{IM} is not consistent when $N \rightarrow \infty$, in the sense that it tends to select large models because the penalty does not depend on the number of observations. Indeed when $N \rightarrow \infty$ the first term on the right hand side of (5) diverges to $-\infty$, thus making the penalty term not effective. As a consequence, AIC_{IM} will not necessarily select the true model if included in the set of all possible auxiliary models, and in this sense we say that it is not consistent. By modifying the penalty, we can generalize AIC_{IM} to obtain consistent criteria.

We define as (a^*, p) the couple corresponding to the true model with p parameters, as defined in (1), and we consider the set containing all possible auxiliary models and the true model, denoted as $\mathcal{A}^* = \mathcal{A}_{q_{\max}} \cup \{(a^*, p)\}$. We then have the following proposition.

Proposition 2 *Given the models of Section 2 and assumptions A.1–A.10 in Appendix A, define the model (\tilde{a}, \tilde{q}) such that*

$$(\tilde{a}, \tilde{q}) = \arg \min_{(q, a) \in \mathcal{A}^*} IC_{IM}(a, q),$$

where

$$IC_{IM}(a, q) = -2 \sum_{i=1}^N \log \phi(\mathbf{x}_{a,q,i}; \tilde{\mathbf{W}}_{N,a,q}) + p(N, q), \quad (9)$$

where the penalty $p(N, q)$ satisfies assumption A.9. Then,

$$(\tilde{a}, \tilde{q}) \xrightarrow{P} (a^*, p), \text{ as } N, T \rightarrow \infty,$$

*i.e. with probability one the criterion is asymptotically selecting the true model.*⁶

Proof: See Appendix B.

⁶The previous probabilistic limit is a symbolic way of defining consistency. Indeed, a is an index indicating a particular class of auxiliary models and not a true parameter. We could formalize it by considering a selecting vector of the same dimension as the number of models considered and taking values 1 or 0 if a given model is selected by the criterion or not, see Andrews (1999) and Hall et al. (2011). However, for simplicity of notation we prefer to avoid this choice as, once understood the meaning of a , nothing changes in the proofs.

The assumed form of the penalty is $p(N, q) = h(q)K_N$, as postulated by Andrews (1999). In $AIC_{IM}(a, q)$ we have $h(q) = q(q+1)/2$ and $K_N = 1$, hence $p(N, q) = p(q)$ and it does not diverge with N . Penalties similar to BIC and $HQIC$ are potential candidates. Andrews (2000) and Hall et al. (2007) report simulation evidence that the BIC -type penalty –i.e. $K_N = \log N$ – works best in their context, since $K_N = \log \log N$ appears to be a too slow rate. This is also our experience. Bozdogan (1987) suggests $K_N = (\log N)^b$ for $b \geq 1$ as a generalization of BIC . The Monte Carlo study in the next section is performed with $K_N = \log N$.

The criteria and other matching-based methods Other matching-based estimation methods, such as GMM, MSM and MSQ, also require the choice of functions. Provided that these functions ensure consistency of the estimates, all methods deliver asymptotically Gaussian estimates; their precision being the only difference. In GMM and MSM q is the number of moment conditions that provide consistent estimators. Then our criteria become similar to the RMSC of Hall et al. (2007). In MSQ the issue of weak identification is not relevant since, in theory, any q -dimensional vector of functions of quantiles provide consistent estimators. We can therefore, in principle, estimate the model of interest with different matching methods (nested and non-nested models, different moment conditions, and different functions of quantiles), and select the one that provides the smallest indirect criteria. A deeper understanding of the selection process between and within inference methods remains open and deserves further investigation.

4 Monte Carlo Study

To investigate the performance of the indirect criteria, we proceed with a thorough Monte Carlo study on two simple and widely used models: a moving average of order one and a stochastic volatility model. Appendix C shows the step-by-step description of the implementation.

Moving Average Consider the following MA(1) model

$$y_t = \mu + \theta u_{t-1} + u_t, \quad t = 1 \dots, T,$$

where u_t is a standardized Gaussian white noise. The number of parameters is $p = 2$. We set $\mu = 0.1$ and we simulate $R = 500$ series for $T = \{1000, 10000\}$ and $\theta = \{0.2, 0.5, 0.8\}$. The purpose of this study is to investigate the performance of the criteria when choosing among the true and nested auxiliary models, for which we choose a set of nested autoregressive models with intercept:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \dots + \phi_r y_{t-r} + \varepsilon_t,$$

where ε_t is a standardized Gaussian white noise and $r = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. The sample size of the simulated series $\mathbf{x}_{a,q,i}$ is set to $N = \{100, 500, 1000\}$, and $H = \{1, 10\}$. The aim of varying N is to study the differences between the penalties. The number of parameters of the auxiliary models is given by $q = 2$ for the MA(1) and $q = r + 1$ for the AR(r).

Tables 1–3 and 4–6 in Appendix D report the criteria for $T = 1000$ and 10000 respectively. The column Rate shows the percentage of times (out of $R = 500$) that an auxiliary model is chosen, while the columns Mean and Var show the sample mean and variance of the $R = 500$ replications.

We extract five conclusions. First, regardless of T , N and H , on average (column Mean) IC_{IM} chooses the true model (numbers in bold) followed by an AR (numbers in bold and italics) with lags that increase with θ . An increase in θ increases the discrepancy in the criteria between the MA(1) and the nested AR models. This may be due to the fact that a small persistence in the MA(1) process is easily captured by few AR lags, and that's also why the criteria chooses the AR when θ is small. However, large values of θ require a large number of lags in the AR model. This effect becomes more apparent once N increases. Second, and related with the previous conclusion, the criteria have a very small dispersion (column Var), and in the majority of cases the IC_{IM} criterion chooses the true model (column Rate) for any value of N . This is not the case for AIC_{IM} . As N increases the frequency at which AIC_{IM} selects models others than the true one increases. Given these two findings, we conclude that the penalty with $K_N = \log N$ works better than with $K_N = 1$. Third, on average IC_{IM} always chooses an auxiliary AR model with equal or smaller lags than the AIC_{IM} does. This is in line with the simulation evidence reported in Andrews (1999) and Hall et al. (2007) for choosing consistent and non-redundant moment conditions respectively. Fourth, an increase in H stabilizes the results and increases the quality of the criteria. This is particularly clear for AIC_{IM} and is a natural effect of the increase in the precision of the estimates leading automatically to an increase in the values of the likelihoods, and to a decrease of the criteria. Also an increase in T stabilizes the quality of the criteria, as expected from Proposition 2.

Stochastic Volatility Consider the following model:

$$\begin{aligned} y_t &= e^{h_t/2} u_t, & u_t &\sim \mathcal{N}(0, 1), \\ h_t &= \mu + \rho h_{t-1} + v_t, & v_t &\sim \mathcal{N}(0, \sigma^2), \quad t = 1 \dots, T. \end{aligned}$$

We set $\mu = 0$, $\sigma^2 = 0.01$ and we simulate three series, each of size $T = 10000$, for three choices of $\rho = \{0.2, 0.4, 0.9\}$. Other choices of T (smaller) are possible but the results for the MA(1) have shown that T does not play a significant role. The first two values correspond to processes with mild volatility clustering, while the third is a common choice in the existing literature (e.g. Shephard (1996) and Monfardini (1998)) and close to values often obtained for daily financial returns. Following Broto and Ruiz (2004), we refrain from estimating μ . Therefore, the number of parameters is $p = 2$. Similar to the MA(1), we choose $N = \{100, 500, 1000\}$ and $H = \{1, 10\}$. The main purpose of this study is to investigate the performance of the criteria when choosing among nested and non-nested auxiliary models. The first class of auxiliary models consists of ARCH(r) models:

$$\begin{aligned} y_t &= d_t^{1/2} z_t, & z_t &\sim \mathcal{N}(0, 1), \\ d_t &= \alpha_0 + \alpha_1 y_{t-1}^2 + \dots + \alpha_r y_{t-r}^2, \end{aligned}$$

for $r = \{1, \dots, 9\}$. The number of auxiliary parameters is $q = r + 1$. This class of auxiliary models was proposed by Engle and Lee (1996) and Calzolari et al. (2008). We

also consider a GARCH(1,1) $d_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2 d_{t-1}$ and hence $q = 3$. The second class consists of AR(r) models on the log of the squared observations:

$$\ln(y_t^2) = \beta_0 + \beta_1 \ln(y_{t-1}^2) + \dots + \beta_r \ln(y_{t-r}^2) + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \varphi)$$

with $r = \{1, \dots, 9\}$. The number of auxiliary parameters is $q = r + 2$. This class of auxiliary models was proposed by Monfardini (1998). We also consider an ARMA(1,1) $\ln(y_t^2) = \beta_0 + \beta_1 \ln(y_{t-1}^2) + \beta_2 \eta_{t-1} + \eta_t$ and hence $q = 3$.

Tables 7–9 in Appendix D report the criteria. We extract three conclusions, all in line with the findings of the MA(1) model. First, similarly to the results reported by Monfardini (1998), the lag selected by the criteria, for both ARCH(r) and AR(r), increases with the persistence of the volatility. More importantly, regardless of the value of H , N , and the penalty, the (G)ARCH model is always preferred (to a large extent) to the AR(MA), suggesting that the former is closer to the SV than the latter. Moreover, GARCH(1,1) is almost always preferred to an ARCH(r), specially for large values of ρ . Second, increasing H stabilizes the results and decreases the values of the criteria. Third, between the ARCH(r) and AR(r) models, IC_{IM} always chooses a model equal or smaller than the AIC_{IM} does. This confirms once more that $K_N = \log N$ works better than $K_N = 1$.

5 Conclusions

Indirect estimation methods depend on the choice of an auxiliary model. The difference between them comes in terms of precision of the estimators and, theoretically, the larger the number of parameters of the auxiliary model, the more precise are the estimates of the model of interest. But to account for the trade-off between the fit of the auxiliary model to the model of interest and its dimension, we need a penalty that depends on the number of parameters in the auxiliary model. In this article we develop information criteria for selecting among nested and non-nested auxiliary models. The criteria can also be used for choosing functions to match between and within other inference methods, such as GMM, the Method of Simulated Moments and the Method of Simulated Quantiles. We proceed with a thorough Monte Carlo study on a MA(1) and a stochastic volatility model. For the former we choose as auxiliary models a set of AR(r) and the true model. The criteria choose the true model as the best auxiliary model. As for the SV model, we consider two auxiliary classes of models: (G)ARCH and AR(MA) on the log of the square of the observations. Among the ARCH and AR type of models, the lag chosen by the criteria increases with the SV autoregressive parameter. Moreover, the criteria always prefer the (G)ARCH to the AR(MA) models.

References

- Akaike, H. (1974), 'A new look at statistical model identification', *IEEE Transactions on Automatic Control* **19**(6), 716–723.
- Andrews, D. and Lu, B. (2001), 'Consistent model and moment selection procedures for GMM estimation with applications to dynamic panel data models', *Journal of Econometrics* **101**, 123–164.
- Andrews, D. W. (2000), Consistent moment selection procedures for GMM estimation: strong consistency and simulation results. Discussion paper, Cowles Foundation for Economics.
- Andrews, D. W. K. (1999), 'Consistent moment selection procedures for generalized method of moments estimation', *Econometrica* **67**, 543–564.
- Billio, M. and Monfort, A. (2003), 'Kernel-based indirect inference', *Journal of Financial Econometrics* **1**, 297–326.
- Blasques, F. (2011), *Semi-Nonparametric Indirect Inference*, PhD thesis. University of Maastricht.
- Bozdogan, H. (1987), 'Model selections and akaike's information criteria (AIC): the general theory and its analytical extensions', *Psychometrika* **52**, 345–370.
- Broto, C. and Ruiz, E. (2004), 'Estimation methods for stochastic volatility models: a survey', *Journal of Economic Surveys* **18**, 613–649.
- Broze, L., Scaillet, O. and Zakoian, J. M. (1998), 'Quasi indirect inference for diffusion processes', *Econometric Theory* **14**, 161–186.
- Calzolari, G., Fiorentini, G. and Sentana, E. (2004), 'Constrained indirect estimation', *Review of Economic Studies* **71**, 945–973.
- Calzolari, G., Fiorentini, G. and Sentana, E. (2008), 'Indirect estimation of large conditionally heteroskedastic factor models, with an application to the dow 30 stocks', *Journal of Econometrics* **146**, 10–25.
- Carrasco, M. and Florens, J.-P. (2000), 'Generalization of GMM to a continuum of moment conditions', *Econometric Theory* **16**, 797–834.
- Chernov, M., Galland, R., Ghysels, E. and Tauchen, G. (2003), 'Alternative models for stock price dynamics', *Journal of Econometrics* **116**, 225–257.
- Corsi, F. and Reno, R. (2012), 'Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling', *Journal of Business and Economic Statistics* **Forthcoming**.
- Czellar, V. and Calvet, L. (2011), Indirect inference for learning models. HEC Mimeo.
- Czellar, V. and Ronchetti, E. (2010), 'Accurate and robust tests for indirect inference', *Biometrika* **97**, 621–630.
- Dhaene, G., Gouriéroux, C. and Scaillet, O. (1998), 'Instrumental models and indirect encompassing', *Econometrica* **66**, 673–688.

- Dominicy, Y. and Veredas, D. (2012), ‘The method of simulated quantiles’, *Journal of Econometrics* **Forthcoming**.
- Dridi, R., Guay, A. and Renault, E. (2007), ‘Indirect inference and calibration of dynamic stochastic general equilibrium models’, *Journal of Econometrics* **136**, 397–430.
- Duffie, D. and Singleton, K. (1993), ‘Simulated moments estimation of Markov models of asset prices’, *Econometrica* **61**, 929–952.
- Engle, R. F. and Lee, G. (1996), Estimating diffusions models of stochastic volatility, in P. E. Rossi, ed., ‘Modelling Stock market Volatility: Bridging the Gap to Continuous Time’, San Diego: Academic Press, pp. 333–335.
- Gallant, A. R. and Tauchen, G. (1996), ‘Which moments to match?’, *Econometric Theory* **12**, 657–681.
- Gallant, R., Hsieh, D. and Tauchen, G. (1997), ‘Estimation of stochastic volatility models with diagnostics’, *Journal of Econometrics* **81**, 159–192.
- Garcia, R., Renault, E. and Veredas, D. (2011), ‘Estimation of stable distributions with indirect inference’, *Journal of Econometrics* **161**, 325–337.
- Genton, M. G. and Ronchetti, E. (2003), ‘Robust indirect inference’, *Journal of the American Statistical Association* **98**, 67–76.
- Ghysels, E., Khalaf, L. and Vodounou, C. (2003), ‘Simulation based inference in moving average models’, *Annales d’Economie et Statistique* **69**, 85–99.
- Gouriéroux, C. and Monfort, A. (1996), *Simulation-based Estimation Methods*, CORE Lectures, Oxford University Press.
- Gouriéroux, C., Monfort, A. and Renault, E. (1993), ‘Indirect inference’, *Journal of Applied Econometrics* **8**, S85–S118.
- Gouriéroux, C., Phillips, P. and Yu, J. (2010), ‘Indirect inference for dynamic panel models’, *Journal of Econometrics* **157**, 68–77.
- Gourieroux, C., Renault, E. and Touzi, N. (2000), Calibration by simulation for small sample bias correction, in R. Mariano, T. Schuerman and M. Weeks, eds, ‘Simulation-based Inference in Econometrics: Methods and Applications’, Cambridge University Press, chapter 13, pp. 328–358.
- Grassi, S. and de Magistris, P. S. (2012), A dynamic multi factor model for high and low frequency volatility activity. Mimeo.
- Guay, A. and Scaillet, O. (2003), ‘Indirect inference, nuisance parameter and threshold moving average models’, *Journal of Business and Economic Statistics* **21**, 122–132.
- Guvenen, F. and Smith, A. (2010), ‘Inferring labor income risk from economic choices: an indirect inference approach’, *Federal research bank of Minneapolis. Research department staff report 450*.
- Hall, A., Inoue, A., Jana, K. and Shin, C. (2007), ‘Information in generalized method of moments estimation and entropy-based moment selection’, *Journal of Econometrics* **138**, 488–512.

- Hall, A., Inoue, A., Nason, J. M. and Rossi, B. (2011), 'Information criteria for impulse response function matching estimation of DSGE models', *Journal of Econometrics* **Forthcoming**.
- Hall, A. and Peixe, F. (2003), 'A consistent method for the selection of relevant instruments', *Econometric Reviews* **22**, 269–288.
- Hannan, E. and Quinn, B. (1979), 'The determination of order of an autoregression', *Journal of the Royal Statistical Society Series B* **41** (2), 190–195.
- Hong, H. and Preston, B. (2005), Non-nested model selection criteria. Mimeo.
- Hong, H., Preston, B. and Shum, B. (2003), 'Generalized empirical likelihood-based model selection criteria for moment condition moments', *Econometric Theory* **19**, 923–943.
- Jang, T.-S. and Lux, T. (2012), Identification of social interaction effects in financial data: inference of herd behaviour via Monte Carlo simulations. Mimeo.
- Jiang, W. and Turnbull, B. (2003), 'Indirect inference for survival data', *Statistics and Operations Research Transactions* **27**, 79–94.
- Keane, M. and Smith, A. (2003), Generalized indirect inference for discrete choice models. Yale University, department of economics, mimeo.
- Kitamura, Y. (2002), A likelihood-based approach to the analysis of a class of nested and non-nested models. Mimeo department of economics, Yale University.
- Konishi, S. and Kitagawa, G. (1996), 'Generalized information criteria in model selection', *Biometrika* **83**, 875–890.
- Lombardi, M. and Calzolari, G. (2009), 'Indirect estimation of α -stable stochastic volatility models', *Computational Statistics and Data Analysis* **53**(6), 2298–2308.
- Lombardi, M. and Veredas, D. (2009), 'Indirect estimation of elliptical stable distributions', *Computational Statistics and Data Analysis* **53**, 2309–2324.
- McFadden, D. (1989), 'A method of simulated moments for estimation of discrete response models without numerical integration', *Econometrica* **57**, 995–1026.
- Monfardini, C. (1998), 'Estimating stochastic volatility models through indirect inference', *Econometrics Journal* **1**, 113–128.
- Nickl, R. and Potscher, B. (2010), 'Efficient simulation-based minimum distance estimation', *Mathematical methods of statistics* **19**, 327–364.
- Pakes, A. and Pollard, D. (1989), 'Simulation and the asymptotics of optimization estimators', *Econometrica* **57**, 1027–1057.
- Parzen, E. (1959), Statistical inference on time series by Hilbert space methods, i. Technical report 23, Applied Mathematics and Statistics Laboratory, Stanford.
- Phillips, P. C. (2012), 'Folklore theorems, implicit maps, and indirect inference', *Econometrica* **80**, 425–454.
- Phillips, P. C. and Yu, J. (2009), 'Simulation-based estimation of contingent-claims prices', *Review of Financial Studies* **22**, 3669–3705.

- Schwarz, G. (1978), 'Estimating the dimension of a model', *Annals of Statistics* **6**, 461–464.
- Shephard, N. (1996), Statistical aspects of ARCH and stochastic volatility, in D. Cox, O. Barndorff-Nielsen and D. V. Hinkley, eds, 'Time Series Models in Econometrics, Finance and Other Fields', Chapman & Hall, London, pp. 1–67.
- Smith, A. (1993), 'Estimating non-linear time series models using simulated vector autoregressions: two approaches', *Journal of Applied Econometrics* **8**, 63 – 84.
- Trojani, F. and Ortelli, C. (2005), 'Robust efficient method of moments estimation', *Journal of Econometrics* **128**, 69–97.

A Assumptions

A.1 The set of parameters Θ is non-empty, compact, and $\theta_0 \in \text{int}(\Theta)$. The same holds for the auxiliary model: \mathcal{B} is non-empty, compact, and $\beta_0 \in \text{int}(\mathcal{B})$.

A.2 The log-likelihoods ℓ and ℓ^a are twice-differentiable, i.e. $\ell \in \mathcal{C}^2(\Theta)$ and $\ell^a \in \mathcal{C}^2(\mathcal{B})$.

A.3 The true parameter vector $\theta_0 \in \Theta$ is the unique maximum for $E[\ell(\theta; \mathbf{Y})]$ and β_0 is the unique maximum of $E[\ell^a(\beta; \mathbf{Y})]$. Moreover, $\mathbf{b}(\theta_0)$ is the unique maximum of $E[\ell^a(\mathbf{b}(\theta_0); \mathbf{Y})]$, which implies $\beta_0 = \mathbf{b}(\theta_0)$.

A.4 The following matrices exist finite and are positive definite:

$$\begin{aligned} \mathcal{I}(\theta_0) &= E_{\mathbf{Y}} \left[\begin{array}{cc} \frac{\partial \ell(\theta; \mathbf{Y})}{\partial \theta} \Big|_{\theta=\theta_0} & \frac{\partial \ell(\theta; \mathbf{Y})}{\partial \theta'} \Big|_{\theta=\theta_0} \\ \hline \frac{\partial \ell^a(\beta; \mathbf{Y})}{\partial \beta} \Big|_{\beta=\beta_0} & \frac{\partial \ell^a(\beta; \mathbf{Y})}{\partial \beta'} \Big|_{\beta=\beta_0} \end{array} \right], \\ \mathbf{I}(\beta_0) &= E_{\mathbf{Y}} \left[\begin{array}{cc} \frac{\partial \ell^a(\beta; \mathbf{Y})}{\partial \beta} \Big|_{\beta=\beta_0} & \frac{\partial \ell^a(\beta; \mathbf{Y})}{\partial \beta'} \Big|_{\beta=\beta_0} \\ \hline \frac{\partial^2 \ell^a(\beta; \mathbf{Y})}{\partial \beta \partial \beta'} \Big|_{\beta=\beta_0} \end{array} \right] \text{ and} \\ \mathbf{J}(\beta_0) &= -E_{\mathbf{Y}} \left[\frac{\partial^2 \ell^a(\beta; \mathbf{Y})}{\partial \beta \partial \beta'} \Big|_{\beta=\beta_0} \right]. \end{aligned}$$

And the corresponding sample matrices converge in probability to them.

A.5 The binding function $\mathbf{b}(\theta)$ is continuously differentiable, i.e. $\mathbf{b} \in \mathcal{C}^1(\Theta)$.

A.6 The binding function $\mathbf{b}(\theta)$ is injective.

A.7 The $q \times p$ matrix $\frac{\partial \mathbf{b}(\theta_0)}{\partial \theta'}$ has full-column rank p .

A.8 $AIC_{IM}(a, q)$ has a unique minimum over $\mathcal{A}_{q_{\max}}$.

A.9 The penalty function is such that $p(N, q) = h(q)K_N$, with $h(q)$ strictly increasing and K_N satisfying $K_N \rightarrow \infty$ and $K_N/N \rightarrow 0$ as $N \rightarrow \infty$.

A.10 The following matrices exist, are finite and positive definite:

$$\begin{aligned} \mathcal{J}_{a,q}(\widehat{\mathbf{w}}_{T,a,q}) &= -E_{\mathbf{x}_{a,q}} \left[\frac{\partial^2 \log \phi(\mathbf{X}_{a,q}; \mathbf{w}_{a,q})}{\partial \mathbf{w}_{a,q} \partial \mathbf{w}'_{a,q}} \Big|_{\mathbf{w}_{a,q}=\widehat{\mathbf{w}}_{T,a,q}} \right] \text{ and} \\ \mathcal{J}^*(\widehat{\mathbf{w}}_T^*) &= -E_{\mathbf{x}^*} \left[\frac{\partial^2 \log \phi(\mathbf{X}^*; \mathbf{w}^*)}{\partial \mathbf{w}^* \partial \mathbf{w}'^*} \Big|_{\mathbf{w}^*=\widehat{\mathbf{w}}_T^*} \right]. \end{aligned}$$

And the corresponding sample matrices converge in probability to them.

Assumptions A.1–A.4 are usual for ML estimation, and in a conditional case all log-likelihoods are to be intended as conditional on other variables and possibly on past information. Assumptions A.5–A.7 characterize the binding function and, in particular, A.6 and A.7 are the global and local identifiability conditions. A.8 and A.9 characterize the objective function of the information criterion. In particular they guarantee the existence of a global minimum and consistency respectively. Assumption A.10 is the analogous of A.4 but for the simulated data $\mathbf{X}_{a,q}$ and \mathbf{X}^* used in the Propositions 1 and 2.

B Technical appendix

We first define some quantities needed for the proofs, followed by two preliminary lemmas, and the proofs of Propositions 1 and 2.

Preliminary results

Definitions For simplicity of notation we omit, unless otherwise stated, all indexes referring to the class of auxiliary models a , the size of the auxiliary parameter q , and the sample size T . We also omit any reference to the value of the parameters in which the matrices defined below are computed and we implicitly assume that everything is computed in the true value of the parameters θ_0 and β_0 . We start with some definitions and then we prove two lemmas necessary to prove Proposition 1.

We re-write the asymptotic variance-covariance matrix given in (4) as

$$\mathbf{W} = (\mathbf{B}'\mathbf{V}\mathbf{B})^{-1}, \quad (10)$$

where $\mathbf{V} = \mathbf{J}(\beta_0)\mathbf{I}^{-1}(\beta_0)\mathbf{J}(\beta_0)$ and $\mathbf{B} = \left. \frac{\partial \mathbf{b}(\theta)}{\partial \theta'} \right|_{\theta=\theta_0}$. Recall that \mathbf{V} is $q \times q$, \mathbf{B} is $q \times p$ with full column rank p (see assumption A7). We define the $q \times p$ right-inverse matrix of \mathbf{B}'

$$(\mathbf{B}')^+ = \mathbf{B}(\mathbf{B}'\mathbf{B})^{-1}, \quad (11)$$

such that $\mathbf{B}'(\mathbf{B}')^+ = \mathbf{I}_p$. Notice that this matrix always exists as long as \mathbf{B}' has full column rank p . Then, consider the $q \times q$ matrix $\mathbf{B}\mathbf{B}'$ and its eigenvalue decomposition $\mathbf{B}\mathbf{B}' = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$, where the columns of \mathbf{P} are the eigenvectors ($\mathbf{P}\mathbf{P}' = \mathbf{I}_q$) and $\mathbf{\Lambda}$ is a $q \times q$ diagonal matrix of eigenvalues. Since from assumption A.6 we have $q \geq p$, $\mathbf{\Lambda}$ has entries $\lambda_i > 0$ for $i = 1, \dots, p$ and $\lambda_i = 0$ for $i = p + 1, \dots, q$. Then, we define the generalized inverse matrix

$$(\mathbf{B}\mathbf{B}')^+ = \mathbf{P}\mathbf{\Lambda}^+\mathbf{P}', \quad (12)$$

where

$$\mathbf{\Lambda}^+ = \left(\begin{array}{c|c} \mathbf{L}^{-1} & \mathbf{0}_{p \times (q-p)} \\ \hline \mathbf{0}_{(q-p) \times p} & \mathbf{0}_{(q-p) \times (q-p)} \end{array} \right) \text{ and } \mathbf{L}^{-1} = \text{diag}(\lambda_1^{-1} \dots \lambda_p^{-1}). \quad (13)$$

Using these definitions we define a $q \times p$ matrix

$$\mathbf{B}^\dagger = \mathbf{B}'(\mathbf{B}\mathbf{B}')^+. \quad (14)$$

Finally, for the $q \times q$ matrix \mathbf{V}^{-1} , we define the duplication matrix, \mathbf{D}_q of dimension $q^2 \times \frac{q(q+1)}{2}$, while for the $p \times p$ matrix \mathbf{W} , we define the duplication matrix \mathbf{D}_p of dimension $p^2 \times \frac{p(p+1)}{2}$ which has a left-inverse $\mathbf{D}_p^+ = (\mathbf{D}_p'\mathbf{D}_p)^{-1}\mathbf{D}_p'$, such that $\mathbf{D}_p^+\mathbf{D}_p = \mathbf{I}_{\frac{p(p+1)}{2}}$. These matrices are such that

$$\text{vec}(\mathbf{V}^{-1}) = \mathbf{D}_q \text{vech}(\mathbf{V}^{-1}), \quad \text{vech}(\mathbf{W}) = \mathbf{D}_p^+ \text{vec}(\mathbf{W}). \quad (15)$$

The following Lemma decomposes the asymptotic variance-covariance matrix \mathbf{W} into the binding function and the information matrix of the auxiliary model.

Lemma 1 *Let $\mathbf{w} = \text{vech}(\mathbf{W})$ be a $\frac{p(p+1)}{2}$ -dimensional vector and \mathbf{W} the asymptotic variance-covariance matrix (4). Then $\mathbf{w} = \mathbf{A}\mathbf{v}$, where \mathbf{A} is a $\frac{p(p+1)}{2} \times \frac{q(q+1)}{2}$ matrix defined as $\mathbf{A} =$*

$\mathbf{D}_p^+(\mathbf{B}^\dagger \otimes \mathbf{B}^+) \mathbf{D}_q$, and \mathbf{v} is a $\frac{q(q+1)}{2} \times 1$ vector defined as $\mathbf{v} = \text{vech}(\mathbf{V}^{-1})$.

Proof. From (10), we have

$$\mathbf{W} = (\mathbf{B}'\mathbf{V}\mathbf{B})^{-1} = \mathbf{C}_1\mathbf{V}^{-1}\mathbf{C}_2,$$

where \mathbf{C}_1 and \mathbf{C}_2 must be such that

$$\mathbf{B}'\mathbf{V}\mathbf{B}\mathbf{C}_1\mathbf{V}^{-1}\mathbf{C}_2 = \mathbf{I}_p, \quad (16)$$

and

$$\mathbf{C}_1\mathbf{V}^{-1}\mathbf{C}_2\mathbf{B}'\mathbf{V}\mathbf{B} = \mathbf{I}_p. \quad (17)$$

We first proof the equality in (16). By substituting in (16) $\mathbf{C}_1 = \mathbf{B}^\dagger$ (see (14)) and $\mathbf{C}_2 = (\mathbf{B}')^+$ (see (11)), and using (12) and (13), we have

$$\begin{aligned} \mathbf{B}'\mathbf{V}\mathbf{B}\mathbf{B}^\dagger\mathbf{V}^{-1}\mathbf{B}^+ &= \mathbf{B}'\mathbf{V}\mathbf{B}\mathbf{B}'(\mathbf{B}\mathbf{B}')^+\mathbf{V}^{-1}\mathbf{B}^+ \\ &= \mathbf{B}'\mathbf{V}\mathbf{P}\mathbf{A}\mathbf{P}'\mathbf{P}\mathbf{A}^+\mathbf{P}'\mathbf{V}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1} \\ &= \mathbf{B}'\mathbf{V}\mathbf{P} \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_{p \times (q-p)} \\ \mathbf{0}_{(q-p) \times p} & \mathbf{0}_{(q-p) \times (q-p)} \end{pmatrix} \mathbf{P}'\mathbf{V}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1} \\ &= \mathbf{B}'\mathbf{V}\mathbf{P}^{(p)}\mathbf{P}^{(p)T}\mathbf{V}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1} \\ &= \mathbf{B}'\mathbf{V}\mathbf{V}^{-1}\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1} \\ &= \mathbf{B}'\mathbf{B}(\mathbf{B}'\mathbf{B})^{-1} \\ &= \mathbf{I}_p, \end{aligned}$$

where $\mathbf{P}^{(p)}$ is a $q \times p$ matrix with just the first p columns of \mathbf{P} . The equality in (17) is proven analogously.

Then

$$\mathbf{W} = \mathbf{B}^\dagger\mathbf{V}^{-1}(\mathbf{B}')^+ = \mathbf{B}^\dagger\mathbf{V}^{-1}(\mathbf{B}^+)',$$

and, from the properties of the $\text{vec}(\cdot)$ operator,

$$\text{vec}(\mathbf{W}) = \text{vec}(\mathbf{B}^\dagger\mathbf{V}^{-1}(\mathbf{B}^+)') = (\mathbf{B}^\dagger \otimes \mathbf{B}^+) \text{vec}(\mathbf{V}^{-1}). \quad (18)$$

Finally, from (18), and using the duplication matrices defined in (15), we have

$$\begin{aligned} \mathbf{w} &= \text{vech}(\mathbf{W}) = \mathbf{D}_p^+ \text{vec}(\mathbf{W}) \\ &= \mathbf{D}_p^+ (\mathbf{B}^\dagger \otimes \mathbf{B}^+) \text{vec}(\mathbf{V}^{-1}) \\ &= \mathbf{D}_p^+ (\mathbf{B}^\dagger \otimes \mathbf{B}^+) \mathbf{D}_q \text{vech}(\mathbf{V}^{-1}) \\ &= \mathbf{A}\mathbf{v}. \end{aligned}$$

□

In order to estimate the binding function, we have to simulate H paths $(\mathbf{y}_1^h(\boldsymbol{\theta}), \dots, \mathbf{y}_{T^*}^h(\boldsymbol{\theta}))$ of length T^* not necessarily equal to T . Thus, $\mathbf{b}(\boldsymbol{\theta})$ is estimated using $\widehat{\boldsymbol{\beta}}_{HT^*}(\boldsymbol{\theta})$ as defined in (2). The error of this estimator can be made as small as we wish by taking $H \rightarrow \infty$ and $T^* \rightarrow \infty$. Moreover, we estimate derivatives numerically as

$$\frac{\partial \widehat{\boldsymbol{\beta}}_{HT^*}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \simeq \frac{-\widehat{\boldsymbol{\beta}}_{HT^*}(\boldsymbol{\theta} + 2\Delta) + 8\widehat{\boldsymbol{\beta}}_{HT^*}(\boldsymbol{\theta} + \Delta) - 8\widehat{\boldsymbol{\beta}}_{HT^*}(\boldsymbol{\theta} - \Delta) + \widehat{\boldsymbol{\beta}}_{HT^*}(\boldsymbol{\theta} - 2\Delta)}{12\Delta}$$

for some arbitrarily small Δ . Thus,

$$p \lim_{H, T^* \rightarrow \infty} \left(\lim_{\Delta \rightarrow 0} \frac{\partial \widehat{\beta}_{HT^*}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) = \frac{\partial \mathbf{b}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{B}.$$

Since Δ , H , and T^* can be arbitrarily fixed, in simulations we can neglect the error made in estimating \mathbf{B} and thus consider \mathbf{A} , which involves \mathbf{B} , as non random. Intuitively, \mathbf{A} does not involve any additional information with respect to the information contained in the elements of the asymptotic variance–covariance matrix of the auxiliary model \mathbf{V}^{-1} .

We now re–establish the complete notation, and the estimator of the asymptotic variance–covariance matrix is such that

$$\widehat{\mathbf{w}}_{T,a,q} = \mathbf{A} \widehat{\mathbf{v}}_{T,a,q}.$$

Recall that for the simulated vector $\mathbf{X}_{a,q} \equiv \sqrt{T}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q} - \boldsymbol{\theta}_0)$ we have a sample covariance matrix $\widetilde{\mathbf{w}}_{N,a,q}$. If we define a new vector $\mathbf{Y}_{a,q} \equiv \sqrt{T}(\widehat{\boldsymbol{\beta}}_{T,IM,a,q} - \boldsymbol{\beta}_0)$, then the sample covariance matrix of $\mathbf{Y}_{a,q}$ is $\widetilde{\mathbf{v}}_{N,a,q}$ such that

$$\widetilde{\mathbf{w}}_{N,a,q} = \mathbf{A} \widetilde{\mathbf{v}}_{N,a,q}. \quad (19)$$

The following Lemma shows the asymptotic distribution of $\widetilde{\mathbf{v}}_{N,a,q}$.

Lemma 2 *For any given auxiliary model a with parameters of dimension q , we have, for $N \rightarrow \infty$,*

$$\sqrt{N}(\widetilde{\mathbf{v}}_{N,a,q} - \widehat{\mathbf{v}}_{T,a,q}) \sim \mathcal{N} \left(\mathbf{0}, \mathbf{A}^\dagger \mathcal{J}_{a,q}^{-1}(\widehat{\mathbf{w}}_{T,a,q}) \mathbf{A}^{\dagger'} \right),$$

where N is the size of the simulated process $(\mathbf{x}_{a,q,1}, \dots, \mathbf{x}_{a,q,N})$,

$$\mathcal{J}_{a,q}(\widehat{\mathbf{w}}_{T,a,q}) = -\mathbb{E}_{\mathbf{x}_{a,q}} \left[\frac{\partial^2 \log \phi(\mathbf{X}_{a,q}; \mathbf{w}_{a,q})}{\partial \mathbf{w}_{a,q} \partial \mathbf{w}'_{a,q}} \Big|_{\mathbf{w}_{a,q} = \widehat{\mathbf{w}}_{T,a,q}} \right],$$

and $\mathbf{A}^\dagger = (\mathbf{A}'\mathbf{A})^+ \mathbf{A}$ with $(\mathbf{A}'\mathbf{A})^+$ defined analogously as in (12), but this time with $\frac{p(p+1)}{2}$ non–zero eigenvalues, thus $\mathbf{A}^\dagger \mathbf{A} = \mathbf{I}_{\frac{q(q+1)}{2}}$.

Proof. We defined $\widetilde{\mathbf{w}}_{N,a,q}$ to be the vector containing the entries of the sample covariance matrix of $\mathbf{X}_{a,q}$ which in turn is distributed as a Gaussian with zero mean and covariance matrix $\widehat{\mathbf{W}}_{T,a,q}$. Then by defining $\widehat{\mathbf{w}}_{T,a,q} = \text{vech}(\widehat{\mathbf{W}}_{T,a,q})$, we have, as $N \rightarrow \infty$,

$$\sqrt{N}(\widetilde{\mathbf{w}}_{N,a,q} - \widehat{\mathbf{w}}_{T,a,q}) \sim \mathcal{N} \left(\mathbf{0}, \mathcal{J}_{a,q}^{-1}(\widehat{\mathbf{w}}_{T,a,q}) \right).$$

Using Lemma 1, the estimator in (19), and the delta method, we obtain the result. \square

Proof of Proposition 1

In this and the following proofs, we reintroduce those indexes referring to the class of auxiliary models a , the size of the auxiliary parameter q , and the observed and simulated samples sizes T and N . We also indicate the value of the parameters in which the matrices are computed, but we simplify the notation using the following definitions of the indirect methods and ML asymptotic

variance–covariance matrices

$$\mathbf{W}_{0,a,q} \equiv \mathbf{W}_{a,q}(\boldsymbol{\theta}_0), \quad \mathbf{W}_{a,q} \equiv \mathbf{W}_{a,q}(\boldsymbol{\theta}), \quad \mathbf{W}_0^* \equiv \mathcal{I}^{-1}(\boldsymbol{\theta}_0), \quad \text{and} \quad \mathbf{W}^* \equiv \mathcal{I}^{-1}(\boldsymbol{\theta}). \quad (20)$$

Moreover, analogously to the q dimensional simulated vector $\mathbf{X}_{a,q} \equiv \sqrt{T}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q} - \boldsymbol{\theta}_0)$, we define the p dimensional random vector $\mathbf{X}^* \equiv \sqrt{T}(\widehat{\boldsymbol{\theta}}_{T,ML} - \boldsymbol{\theta}_0)$. For $\mathbf{X}_{a,q}$ and \mathbf{X}^* we can compute their sample variance–covariance matrices $\widehat{\mathbf{W}}_{N,a,q}$ and $\widehat{\mathbf{W}}_N^*$ respectively. These matrices are the consistent ML estimators of $\widehat{\mathbf{W}}_{T,a,q} \equiv \widehat{\mathbf{W}}_{T,a,q}(\widehat{\boldsymbol{\theta}}_{T,IM,a,q})$ and $\widehat{\mathbf{W}}_T^* \equiv \widehat{\mathbf{W}}_T^*(\widehat{\boldsymbol{\theta}}_{T,ML})$, which in turn are estimators of $\mathbf{W}_{0,a,q}$ and \mathbf{W}_0^* . For any of the matrices \mathbf{W} defined above, we have a vectorized form, generically denoted as $\mathbf{w} \equiv \text{vech}(\mathbf{W})$.

Finally, notice that, since ML estimation of the model of interest is unfeasible, we cannot compute $\widehat{\boldsymbol{\theta}}_{T,ML}$ and therefore \mathbf{X}^* cannot be simulated. This, however, is an object needed only for the proof and plays no role in the final formulation of the criteria.

Given consistency of the sample variance–covariance matrices we have, as $N \rightarrow \infty$,

$$\sqrt{N}(\widetilde{\mathbf{w}}_{N,a,q} - \widehat{\mathbf{w}}_{T,a,q}) \sim \mathcal{N}(\mathbf{0}, \mathcal{J}_{a,q}^{-1}(\widehat{\mathbf{w}}_{T,a,q})) \quad (21)$$

and

$$\sqrt{N}(\widetilde{\mathbf{w}}_N^* - \widehat{\mathbf{w}}_T^*) \sim \mathcal{N}(\mathbf{0}, \mathcal{J}^{*-1}(\widehat{\mathbf{w}}_T^*)), \quad (22)$$

where the asymptotic variance–covariance matrices have now size $\frac{p(p+1)}{2} \times \frac{p(p+1)}{2}$ and are defined as

$$\begin{aligned} \mathcal{J}_{a,q}(\widehat{\mathbf{w}}_{T,a,q}) &= -\mathbb{E}_{\mathbf{x}_{a,q}} \left[\frac{\partial^2 \log \phi(\mathbf{X}_{a,q}; \mathbf{w}_{a,q})}{\partial \mathbf{w}_{a,q} \partial \mathbf{w}'_{a,q}} \Big|_{\mathbf{w}_{a,q} = \widehat{\mathbf{w}}_{T,a,q}} \right] \quad \text{and} \\ \mathcal{J}^*(\widehat{\mathbf{w}}_T^*) &= -\mathbb{E}_{\mathbf{x}^*} \left[\frac{\partial^2 \log \phi(\mathbf{X}^*; \mathbf{w}^*)}{\partial \mathbf{w}^* \partial \mathbf{w}^{*'}} \Big|_{\mathbf{w}^* = \widehat{\mathbf{w}}_T^*} \right]. \end{aligned}$$

The sample analogues of the two previous expressions are:

$$\begin{aligned} \widetilde{\mathcal{J}}_{N,a,q}(\widehat{\mathbf{w}}_{T,a,q}) &= -\frac{1}{N} \sum_{i=1}^N \frac{\partial^2 \log \phi(\mathbf{x}_{a,q,i}; \mathbf{w}_{a,q})}{\partial \mathbf{w}_{a,q} \partial \mathbf{w}'_{a,q}} \Big|_{\mathbf{w}_{a,q} = \widehat{\mathbf{w}}_{T,a,q}} \quad \text{and} \\ \widetilde{\mathcal{J}}_N^*(\widehat{\mathbf{w}}_T^*) &= -\frac{1}{N} \sum_{i=1}^N \frac{\partial^2 \log \phi(\mathbf{x}_i^*; \mathbf{w}^*)}{\partial \mathbf{w}^* \partial \mathbf{w}^{*'}} \Big|_{\mathbf{w}^* = \widehat{\mathbf{w}}_T^*}. \end{aligned}$$

Once $\mathbf{X}_{a,q}$ is given, we have to find an auxiliary model with asymptotic variance–covariance matrix $\mathbf{W}_{a,q}$ that better approximates (in terms of information) the best achievable value \mathbf{W}_0^* . In order to do this, we minimize the Kullback–Leibler information (or negentropy)

$$\mathcal{D}(\mathbf{w}_0^*, \mathbf{w}_{a,q}) = 2 \mathbb{E}_{\mathbf{x}^*} \left[\log \frac{\phi(\mathbf{X}^*; \mathbf{w}_0^*)}{\phi(\mathbf{X}_{a,q}; \mathbf{w}_{a,q})} \right], \quad (23)$$

If we replace the expectation by its sample counterpart, we have to minimize

$$\widehat{\mathcal{D}}_N(\mathbf{w}_0^*, \mathbf{w}_{a,q}) = \frac{2}{N} \sum_{i=1}^N \log \frac{\phi(\mathbf{x}_i^*; \mathbf{w}_0^*)}{\phi(\mathbf{x}_{a,q,i}; \mathbf{w}_{a,q})}. \quad (24)$$

Since we can only estimate $\mathbf{w}_{a,q}$, the best we can do is to use its indirect estimator $\widehat{\mathbf{w}}_{T,a,q}$. Therefore, instead of minimizing (24) we look for the auxiliary model defined by the couple (a, q) that minimizes

$$\mathbb{E}_{\mathbf{w}} \left[\widehat{\mathcal{D}}_N(\mathbf{w}_0^*, \widehat{\mathbf{w}}_{T,a,q}) \right] = \mathbb{E}_{\mathbf{w}} \left[\frac{2}{N} \sum_{i=1}^N \log \frac{\phi(\mathbf{x}_i^*; \mathbf{w}_0^*)}{\phi(\mathbf{x}_{a,q,i}; \widehat{\mathbf{w}}_{T,a,q})} \right]. \quad (25)$$

We decompose the objective function as the sum of two terms

$$\mathcal{Q} \equiv \widehat{\mathcal{D}}_N(\mathbf{w}_0^*, \widehat{\mathbf{w}}_{T,a,q}) = \underbrace{\widehat{\mathcal{D}}_N(\mathbf{w}_0^*, \widetilde{\mathbf{w}}_{N,a,q})}_{\mathcal{P}} + \underbrace{\widehat{\mathcal{D}}_N(\widetilde{\mathbf{w}}_{N,a,q}, \widehat{\mathbf{w}}_{T,a,q})}_{\mathcal{G}}. \quad (26)$$

Term \mathcal{P} of (26) can be written as

$$\mathcal{P} = \widehat{\mathcal{D}}_N(\mathbf{w}_0^*, \widetilde{\mathbf{w}}_{N,a,q}) = \underbrace{\widehat{\mathcal{D}}_N(\mathbf{w}_0^*, \widehat{\mathbf{w}}_T^*)}_{\mathcal{A}} + \underbrace{\widehat{\mathcal{D}}_N(\widehat{\mathbf{w}}_T^*, \widetilde{\mathbf{w}}_N^*)}_{\mathcal{C}} + \underbrace{\widehat{\mathcal{D}}_N(\widetilde{\mathbf{w}}_N^*, \widetilde{\mathbf{w}}_{N,a,q})}_{\mathcal{B}}.$$

So (26) can be written as

$$\mathcal{Q} = \mathcal{A} + \mathcal{C} + \mathcal{B} + \mathcal{G}. \quad (27)$$

Both term \mathcal{A} and \mathcal{C} do not depend on the chosen auxiliary model thus it can be treated as constants in the criterion. Let us consider the other two terms \mathcal{B} and \mathcal{G} separately.

\mathcal{B} : This term is

$$\mathcal{B} = \frac{2}{N} \sum_{i=1}^N \log \frac{\phi(\mathbf{x}_i^*; \widetilde{\mathbf{w}}_N^*)}{\phi(\mathbf{x}_{a,q,i}; \widetilde{\mathbf{w}}_{N,a,q})} = \frac{2}{N} \sum_{i=1}^N \{ \log \phi(\mathbf{x}_i^*; \widetilde{\mathbf{w}}_N^*) - \log \phi(\mathbf{x}_{a,q,i}; \widetilde{\mathbf{w}}_{N,a,q}) \}.$$

Only the second term on the right hand side depends on the chosen auxiliary model, (a, q) , while the first term can be treated as a constant.

\mathcal{G} : This term can be approximated using a Taylor expansion in a neighborhood of $\widetilde{\mathbf{w}}_{N,a,q}$:

$$\begin{aligned} \mathcal{G} &\simeq -(\widehat{\mathbf{w}}_{T,a,q} - \widetilde{\mathbf{w}}_{N,a,q})' \frac{1}{N} \sum_{i=1}^N \frac{\partial^2 \log \phi(\mathbf{x}_{i,a,q}; \mathbf{w}_{a,q})}{\partial \mathbf{w}_{a,q} \partial \mathbf{w}'_{a,q}} \Big|_{\mathbf{w}_{a,q} = \widetilde{\mathbf{w}}_{N,a,q}} (\widehat{\mathbf{w}}_{T,a,q} - \widetilde{\mathbf{w}}_{N,a,q}) \\ &= (\widehat{\mathbf{w}}_{T,a,q} - \widetilde{\mathbf{w}}_{N,a,q})' \widetilde{\mathcal{J}}_{N,a,q}(\widetilde{\mathbf{w}}_{N,a,q}) (\widehat{\mathbf{w}}_{T,a,q} - \widetilde{\mathbf{w}}_{N,a,q}). \end{aligned}$$

Since

$$p \lim_{N \rightarrow \infty} \widetilde{\mathcal{J}}_{N,a,q}(\widetilde{\mathbf{w}}_{N,a,q}) = \mathcal{J}_{a,q}(\widehat{\mathbf{w}}_{T,a,q}),$$

and given (21) and lemma 2, we have for $N \rightarrow \infty$

$$\begin{aligned} N\mathcal{G} &\simeq N(\widehat{\mathbf{w}}_{T,a,q} - \widetilde{\mathbf{w}}_{N,a,q})' \mathcal{J}_{a,q}(\widehat{\mathbf{w}}_{T,a,q}) (\widehat{\mathbf{w}}_{T,a,q} - \widetilde{\mathbf{w}}_{N,a,q}) \\ &= N(\widehat{\mathbf{v}}_{T,a,q} - \widetilde{\mathbf{v}}_{N,a,q})' \mathbf{A}^\dagger \mathcal{J}_{a,q}(\widehat{\mathbf{w}}_{T,a,q}) \mathbf{A}^\dagger (\widehat{\mathbf{v}}_{T,a,q} - \widetilde{\mathbf{v}}_{N,a,q}) \sim \chi_{\frac{q(q+1)}{2}}^2. \end{aligned}$$

By combining the previous results, we have that minimizing (25) is equivalent to minimize

$$\begin{aligned} \mathbb{E}_{\mathbf{w}} [N\mathcal{Q}] &= \mathbb{E}_{\mathbf{w}} [N\mathcal{A} + N\mathcal{C} + N\mathcal{B} + N\mathcal{G}] \\ &= k - 2 \sum_{i=1}^N \log \phi(\mathbf{x}_{a,q,i}; \widetilde{\mathbf{w}}_{N,a,q}) + \frac{q(q+1)}{2}. \end{aligned}$$

where the first term k on the right hand side contains \mathcal{A} , \mathcal{C} , and the first term of \mathcal{B} , therefore it does not depend on the chosen auxiliary model, (a, q) . Thus, minimizing (25) is equivalent to minimize with respect to all possible auxiliary models (a, q) the following criterion

$$AIC_{IM}(a, q) = -2 \sum_{i=1}^N \log \phi(\mathbf{x}_{a,q,i}; \tilde{\mathbf{w}}_{N,a,q}) + \frac{q(q+1)}{2},$$

where $\tilde{\mathbf{w}}_{N,a,q} = \text{vech}(\tilde{\mathbf{W}}_{N,a,q})$ is the vector containing the entries of the sample variance–covariance matrix of the simulated vector $\mathbf{X}_{a,q}$. \square

Proof of Proposition 2

For each model $(a, q) \in \mathcal{A}^*$, we have the criterion

$$IC_{IM}(a, q) = -2 \sum_{i=1}^N \log \phi(\mathbf{x}_{a,q,i}; \tilde{\mathbf{w}}_{N,a,q}) + h(q)K_N,$$

If we could compute the criterion in correspondence of the true model, we would have

$$IC_{IM}(a^*, p) = -2 \sum_{i=1}^N \log \phi(\mathbf{x}_i^*; \tilde{\mathbf{w}}_N^*) + h(p)K_N.$$

Therefore,

$$\begin{aligned} \frac{1}{N}[IC_{IM}(a, q) - IC_{IM}(a^*, p)] &= \frac{2}{N} \sum_{i=1}^N \log \frac{\phi(\mathbf{x}_i^*; \tilde{\mathbf{w}}_N^*)}{\phi(\mathbf{x}_{a,q,i}; \tilde{\mathbf{w}}_{N,a,q})} + (h(q) - h(p)) \frac{K_N}{N} \\ &= \hat{\mathcal{D}}_N(\tilde{\mathbf{w}}_N^*, \tilde{\mathbf{w}}_{N,a,q}) + (h(q) - h(p)) \frac{K_N}{N}. \end{aligned} \quad (28)$$

The first term on the right–hand–side of (28) can be decomposed as

$$\begin{aligned} \hat{\mathcal{D}}_N(\tilde{\mathbf{w}}_N^*, \tilde{\mathbf{w}}_{N,a,q}) &= \hat{\mathcal{D}}_N(\tilde{\mathbf{w}}_N^*, \hat{\mathbf{w}}_T^*) + \hat{\mathcal{D}}_N(\hat{\mathbf{w}}_T^*, \mathbf{w}_0^*) + \hat{\mathcal{D}}_N(\mathbf{w}_0^*, \mathbf{w}_{0,a,q}) + \\ &\quad + \hat{\mathcal{D}}_N(\mathbf{w}_{0,a,q}, \hat{\mathbf{w}}_{T,a,q}) + \hat{\mathcal{D}}_N(\hat{\mathbf{w}}_{T,a,q}, \tilde{\mathbf{w}}_{N,a,q}). \end{aligned} \quad (29)$$

From (21) and (22), and using Taylor expansions, we have that as $N \rightarrow \infty$ the first and last term of (29) are $O_p(N^{-1})$. In a similar way, using Taylor approximations we write

$$\hat{\mathcal{D}}_N(\mathbf{w}_{0,a,q}, \hat{\mathbf{w}}_{T,a,q}) \simeq (\mathbf{w}_{0,a,q} - \hat{\mathbf{w}}_{T,a,q})' \hat{\mathcal{J}}_{T,a,q}(\hat{\mathbf{w}}_{T,a,q})(\mathbf{w}_{0,a,q} - \hat{\mathbf{w}}_{T,a,q}).$$

From Gouriéroux et al. (1993) we have, as $T \rightarrow \infty$,

$$p \lim_{T \rightarrow \infty} \hat{\mathbf{w}}_{T,a,q} = \mathbf{w}_{0,a,q},$$

which, provided that $\hat{\mathcal{J}}_{T,a,q}$ is bounded by A.10, implies $\|\mathbf{w}_{0,a,q} - \hat{\mathbf{w}}_{T,a,q}\| = O_p(T^{-1/2})$ and therefore the fourth term of (29) is $O_p(T^{-1})$. Analogously from Maximum Likelihood theory we have that the second term in (29) can be approximated as

$$\hat{\mathcal{D}}_N(\hat{\mathbf{w}}_T^*, \mathbf{w}_0^*) \simeq (\mathbf{w}_0^* - \hat{\mathbf{w}}_T^*)' \hat{\mathcal{J}}_T^*(\hat{\mathbf{w}}_T^*)(\mathbf{w}_0^* - \hat{\mathbf{w}}_T^*),$$

which, given consistency of $\widehat{\boldsymbol{w}}_T^*$ as $T \rightarrow \infty$, and boundedness of $\widehat{\mathcal{J}}_T^*$ (see A.10), is also $O_p(T^{-1})$.

Therefore, (28) becomes

$$\frac{1}{N}[IC_{IM}(a, q) - IC_{IM}(a^*, p)] = \widehat{\mathcal{D}}_N(\boldsymbol{w}_0^*, \boldsymbol{w}_{0,a,q}) + O_p(N^{-1}) + O_p(T^{-1}) + (h(q) - h(p))\frac{K_N}{N}.$$

And, since $K_N/N = o(1)$,

$$\begin{aligned} p \lim_{N \rightarrow \infty} \left\{ \frac{1}{N}[IC_{IM}(a, q) - IC_{IM}(a^*, p)] \right\} &= p \lim_{N \rightarrow \infty} \left\{ \widehat{\mathcal{D}}_N(\boldsymbol{w}_0^*, \boldsymbol{w}_{0,a,q}) \right\} + O_p(T^{-1}) \\ &= \mathcal{D}(\boldsymbol{w}_0^*, \boldsymbol{w}_{0,a,q}) + O_p(T^{-1}). \end{aligned}$$

If we also let $T \rightarrow \infty$, we have

$$p \lim_{N, T \rightarrow \infty} \left\{ \frac{1}{N}[IC_{IM}(a, q) - IC_{IM}(a^*, p)] \right\} = \mathcal{D}(\boldsymbol{w}_0^*, \boldsymbol{w}_{0,a,q}) \geq 0, \quad (30)$$

where the last inequality follows from the definition of Kullback–Leibler information which would be zero when $\boldsymbol{w}_{0,a,q} = \boldsymbol{w}_0^*$.

From (30), we have that, as $N, T \rightarrow \infty$,

$$IC_{IM}(a^*, p) \leq IC_{IM}(a, q), \text{ for any } (a, q) \in \mathcal{A}^*,$$

with probability one. Moreover, given the definition of (\tilde{a}, \tilde{q}) and A.8, we also have

$$IC_{IM}(\tilde{a}, \tilde{q}) < IC_{IM}(a, q), \text{ for any } (a, q) \in \mathcal{A}_{q_{\max}}.$$

Therefore, as $N, T \rightarrow \infty$, we have that with probability one and when considering all models in \mathcal{A}^* , the models (\tilde{a}, \tilde{q}) and (a^*, p) will coincide. In other words we would select the true model if it is included in the set of models we are considering. \square

C Step-by-step implementation

Step 1: Choose an auxiliary model $\mathcal{M}^a(\mathbf{y}, \beta)$ and estimate β from the observations to obtain $\hat{\beta}_T$.

Step 2: Find the Indirect Inference estimator of θ :

$$\hat{\theta}_{HT,IM,a,q} = \arg \min_{\theta \in \Theta} \left(\hat{\beta}_{HT}(\theta) - \hat{\beta}_T \right)' \Omega^* \left(\hat{\beta}_{HT}(\theta) - \hat{\beta}_T \right),$$

where Ω^* is the optimal weighting matrix. The algorithm for the minimization starts at some initial value θ^0 . The corresponding simulated H paths of length T are $\mathbf{y}_1^h(\theta^0), \dots, \mathbf{y}_T^h(\theta^0)$, $h = 1, \dots, H$, and the estimates $\hat{\beta}_{HT}(\theta^0)$. Note that the simulated paths can be of size $T^* > T$.

Step 3: Let $\mathbf{X}_{a,q} \equiv \sqrt{T}(\hat{\theta}_{T,IM,a,q} - \theta_0)$. Simulate a sample of length N from the $\mathcal{N}(\mathbf{0}, \widehat{\mathbf{W}}_{T,a,q}(\hat{\theta}_{T,IM,a,q}))$. Denote the i -th simulated observation of $\mathbf{X}_{a,q}$ as the p -dimensional vector $\mathbf{x}_{a,q,i} = (x_{1,a,q,i} \dots x_{p,a,q,i})'$.

Step 5: From the N simulated series $\mathbf{x}_{a,q,i}$ compute the sample variance-covariance matrix $\widetilde{\mathbf{W}}_{N,a,q}$ and the preferred information criterion given in (5) or (9).

Step 6: Select another (nested or non-nested) auxiliary model and repeat steps 1–5.

Step 7: Choose the auxiliary model that provides the smallest value for the preferred information criterion.

D Tables

Table 1: $MA(1)$ with $\theta = 0.2$ and $T = 1000$

N	Aux. Model	$H = 1$						$H = 10$					
		AIC_{IM}			IC_{IM}			AIC_{IM}			IC_{IM}		
		Rate	Mean	Var	Rate	Mean	Var	Rate	Mean	Var	Rate	Mean	Var
100	$MA(1)$	96	21.1940	0.0031	100	21.3022	0.0031	81.4	20.3790	0.0031	99.4	20.4871	0.0031
	$AR(1)$	0	21.6609	0.0007	0	21.7691	0.0007	0	20.7759	0.0007	0	20.8841	0.0007
	$AR(2)$	0	21.3547	0.0009	0	21.5710	0.0009	18.6	20.4183	0.0010	0.6	20.6346	0.0010
	$AR(3)$	4	21.2641	0.0012	0	21.6247	0.0012	0	20.4727	0.0014	0	20.8332	0.0014
	$AR(4)$	0	21.2768	0.0014	0	21.8176	0.0014	0	20.5421	0.0015	0	21.0828	0.0015
	$AR(5)$	0	21.3901	0.0014	0	22.1471	0.0014	0	20.6048	0.0015	0	21.3619	0.0015
	$AR(6)$	0	21.4455	0.0015	0	22.4550	0.0015	0	20.6785	0.0015	0	21.6880	0.0015
	$AR(7)$	0	21.6441	0.0015	0	22.9420	0.0015	0	20.7660	0.0015	0	22.0638	0.0015
	$AR(8)$	0	21.6871	0.0015	0	23.3094	0.0015	0	20.8721	0.0015	0	22.4944	0.0015
	$AR(9)$	0	21.8460	0.0015	0	23.8288	0.0015	0	20.9760	0.0015	0	22.9589	0.0015
500	$MA(1)$	38.2	21.2867	0.0031	98.2	21.3180	0.0031	63.6	20.1565	0.0031	85	20.1878	0.0031
	$AR(1)$	0	21.7536	0.0007	0	21.7849	0.0007	0	20.5535	0.0007	0	20.5848	0.0007
	$AR(2)$	0	21.4234	0.0009	0	21.4859	0.0009	35.6	20.1718	0.0010	15	20.2344	0.0010
	$AR(3)$	0	21.3008	0.0012	1.8	21.4051	0.0012	0.8	20.1942	0.0014	0	20.2985	0.0014
	$AR(4)$	61.8	21.2735	0.0014	0	21.4300	0.0014	0	20.2236	0.0015	0	20.3801	0.0015
	$AR(5)$	0	21.3387	0.0014	0	21.5578	0.0014	0	20.2384	0.0015	0	20.4574	0.0015
	$AR(6)$	0	21.3382	0.0015	0	21.6302	0.0015	0	20.2561	0.0015	0	20.5481	0.0015
	$AR(7)$	0	21.4728	0.0015	0	21.8482	0.0015	0	20.2795	0.0015	0	20.6550	0.0015
	$AR(8)$	0	21.4438	0.0015	0	21.9131	0.0015	0	20.3137	0.0015	0	20.7830	0.0015
	$AR(9)$	0	21.5226	0.0015	0	22.0962	0.0015	0	20.3376	0.0015	0	20.9112	0.0015
1000	$MA(1)$	26.2	21.2258	0.0031	87.2	21.2436	0.0031	62	20.1152	0.0031	75.6	20.1329	0.0031
	$AR(1)$	0	21.6928	0.0007	0	21.7105	0.0007	0	20.5122	0.0007	0	20.5299	0.0007
	$AR(2)$	0	21.3595	0.0009	0	21.3950	0.0009	36.6	20.1275	0.0010	24.4	20.1630	0.0010
	$AR(3)$	0	21.2330	0.0012	8	21.2921	0.0012	1.4	20.1459	0.0014	0	20.2050	0.0014
	$AR(4)$	73.8	21.2007	0.0014	4.8	21.2893	0.0014	0	20.1703	0.0015	0	20.2589	0.0015
	$AR(5)$	0	21.2599	0.0014	0	21.3840	0.0014	0	20.1790	0.0015	0	20.3031	0.0015
	$AR(6)$	0	21.2524	0.0015	0	21.4178	0.0015	0	20.1898	0.0015	0	20.3552	0.0015
	$AR(7)$	0	21.3789	0.0015	0	21.5916	0.0015	0	20.2052	0.0015	0	20.4179	0.0015
	$AR(8)$	0	21.3409	0.0015	0	21.6068	0.0015	0	20.2303	0.0015	0	20.4962	0.0015
	$AR(9)$	0	21.4098	0.0015	0	21.7347	0.0015	0	20.2443	0.0015	0	20.5692	0.0015

The model of interest is a $MA(1)$ with intercept 0.1, moving average parameter 0.2 and variance of the errors equal to 1. The total number of replications is $R = 500$ and the simulated time series contains $T = 1000$ observations. The number of simulated paths H is 1 and 10, and N (first column) equals 100, 500, and 1000. The auxiliary model is either the model of interest itself or an $AR(r)$ (column Aux. Model). The table reports the percentage of choosing each model as the best (column Rate), the average value and the variance of the criteria (columns Mean and Var respectively) over the 500 replications. Numbers in bold denote the model chosen by the criteria. Numbers in bold and italics denote the second best model chosen by the criteria.

Table 2: $MA(1)$ with $\theta = 0.5$ and $T = 1000$

N	Aux. Model	$H = 1$						$H = 10$					
		Rate	AIC_{IM}		Rate	IC_{IM}		Rate	AIC_{IM}		Rate	IC_{IM}	
			Mean	Var		Mean	Var		Mean	Var		Mean	Var
100	$MA(1)$	100	20.0470	0.0047	100	20.1552	0.0047	100	19.2059	0.0049	100	19.3141	0.0049
	$AR(1)$	0	22.6439	0.0104	0	22.7520	0.0104	0	21.7143	0.0093	0	21.8225	0.0093
	$AR(2)$	0	21.2741	0.0014	0	21.4904	0.0014	0	20.2017	0.0009	0	20.4180	0.0009
	$AR(3)$	0	20.6332	0.0007	0	20.9937	0.0007	0	19.6834	0.0008	0	20.0439	0.0008
	$AR(4)$	0	20.3518	0.0013	0	20.8926	0.0013	0	19.5052	0.0015	0	20.0460	0.0015
	$AR(5)$	0	20.3282	0.0019	0	21.0853	0.0019	0	19.4662	0.0021	0	20.2233	0.0021
	$AR(6)$	0	20.3027	0.0024	0	21.3122	0.0024	0	19.4939	0.0026	0	20.5033	0.0026
	$AR(7)$	0	20.5041	0.0028	0	21.8020	0.0028	0	19.5567	0.0029	0	20.8546	0.0029
	$AR(8)$	0	20.5578	0.0029	0	22.1801	0.0029	0	19.6529	0.0030	0	21.2752	0.0030
	$AR(9)$	0	20.7524	0.0028	0	22.7353	0.0028	0	19.7547	0.0031	0	21.7376	0.0031
500	$MA(1)$	88.8	20.1397	0.0047	100	20.1710	0.0047	94.6	18.9835	0.0049	100	19.0148	0.0049
	$AR(1)$	0	22.7366	0.0104	0	22.7678	0.0104	0	21.4919	0.0093	0	21.5232	0.0093
	$AR(2)$	0	21.3427	0.0014	0	21.4053	0.0014	0	19.9553	0.0009	0	20.0178	0.0009
	$AR(3)$	0	20.6699	0.0007	0	20.7742	0.0007	0	19.4049	0.0008	0	19.5092	0.0008
	$AR(4)$	0	20.3485	0.0013	0	20.5049	0.0013	0	19.1867	0.0015	0	19.3432	0.0015
	$AR(5)$	0	20.2768	0.0019	0	20.4959	0.0019	0	19.0997	0.0021	0	19.3188	0.0021
	$AR(6)$	11.2	20.1954	0.0024	0	20.4874	0.0024	2.4	19.0714	0.0026	0	19.3635	0.0026
	$AR(7)$	0	20.3328	0.0028	0	20.7083	0.0028	3	19.0703	0.0029	0	19.4458	0.0029
	$AR(8)$	0	20.3145	0.0029	0	20.7838	0.0029	0	19.0945	0.0030	0	19.5638	0.0030
	$AR(9)$	0	20.4291	0.0028	0	21.0027	0.0028	0	19.1163	0.0031	0	19.6899	0.0031
1000	$MA(1)$	76	20.0789	0.0047	100	20.0966	0.0047	87.2	18.9422	0.0049	100	18.9599	0.0049
	$AR(1)$	0	22.6757	0.0104	0	22.6934	0.0104	0	21.4506	0.0093	0	21.4683	0.0093
	$AR(2)$	0	21.2789	0.0014	0	21.3143	0.0014	0	19.9109	0.0009	0	19.9464	0.0009
	$AR(3)$	0	20.6020	0.0007	0	20.6611	0.0007	0	19.3566	0.0008	0	19.4157	0.0008
	$AR(4)$	0	20.2756	0.0013	0	20.3643	0.0013	0	19.1334	0.0015	0	19.2220	0.0015
	$AR(5)$	0	20.1980	0.0019	0	20.3221	0.0019	0	19.0404	0.0021	0	19.1645	0.0021
	$AR(6)$	24	20.1096	0.0024	0	20.2750	0.0024	1	19.0051	0.0026	0	19.1705	0.0026
	$AR(7)$	0	20.2390	0.0028	0	20.4516	0.0028	11.4	18.9960	0.0029	0	19.2086	0.0029
	$AR(8)$	0	20.2116	0.0029	0	20.4775	0.0029	0.4	19.0111	0.0030	0	19.2770	0.0030
	$AR(9)$	0	20.3163	0.0028	0	20.6412	0.0028	0	19.0230	0.0031	0	19.3479	0.0031

The model of interest is a $MA(1)$ with intercept 0.1, moving average parameter 0.5 and variance of the errors equal to 1. The total number of replications is $R = 500$ and the simulated time series contains $T = 1000$ observations. The number of simulated paths H is 1 and 10, and N (first column) equals 100, 500, and 1000. The auxiliary model is either the model of interest itself or an $AR(r)$ (column Aux. Model). The table reports the percentage of choosing each model as the best (column Rate), the average value and the variance of the criteria (columns Mean and Var respectively) over the 500 replications. Numbers in bold denote the model chosen by the criteria. Numbers in bold and italics denote the second best model chosen by the criteria.

Table 3: $MA(1)$ with $\theta = 0.8$ and $T = 1000$

N	Aux. Model	$H = 1$						$H = 10$					
		AIC_{IM}			IC_{IM}			AIC_{IM}			IC_{IM}		
		Rate	Mean	Var	Rate	Mean	Var	Rate	Mean	Var	Rate	Mean	Var
100	$MA(1)$	100	17.5395	0.0136	100	17.6477	0.0136	100	16.5575	0.0148	100	16.6657	0.0148
	$AR(1)$	0	29.9585	8.7360	0	30.0667	8.7360	0	28.3374	11.1184	0	28.4456	11.1184
	$AR(2)$	0	25.5198	2.6937	0	25.7361	2.6937	0	21.7528	0.0609	0	21.9691	0.0609
	$AR(3)$	0	22.0062	0.0791	0	22.3667	0.0791	0	20.5391	0.0285	0	20.8996	0.0285
	$AR(4)$	0	20.6593	0.0117	0	21.2001	0.0117	0	19.7656	0.0145	0	20.3064	0.0145
	$AR(5)$	0	20.0218	0.0048	0	20.7789	0.0048	0	19.1448	0.0059	0	19.9019	0.0059
	$AR(6)$	0	19.4169	0.0011	0	20.4264	0.0011	0	18.6881	0.0024	0	19.6976	0.0024
	$AR(7)$	0	19.2517	0.0008	0	20.5496	0.0008	0	18.3511	0.0010	0	19.6490	0.0010
	$AR(8)$	0	19.0154	0.0008	0	20.6377	0.0008	0	18.1354	0.0007	0	19.7577	0.0007
	$AR(9)$	0	19.0711	0.0008	0	21.0540	0.0008	0	17.9984	0.0008	0	19.9813	0.0008
500	$MA(1)$	100	17.6322	0.0136	100	17.6635	0.0136	100	16.3351	0.0148	100	16.3664	0.0148
	$AR(1)$	0	30.0512	8.7360	0	30.0825	8.7360	0	28.1150	11.1184	0	28.1463	11.1184
	$AR(2)$	0	25.5885	2.6937	0	25.6510	2.6937	0	21.5064	0.0609	0	21.5690	0.0609
	$AR(3)$	0	22.0428	0.0791	0	22.1471	0.0791	0	20.2607	0.0285	0	20.3650	0.0285
	$AR(4)$	0	20.6560	0.0117	0	20.8124	0.0117	0	19.4472	0.0145	0	19.6036	0.0145
	$AR(5)$	0	19.9705	0.0048	0	20.1895	0.0048	0	18.7784	0.0059	0	18.9974	0.0059
	$AR(6)$	0	19.3096	0.0011	0	19.6016	0.0011	0	18.2657	0.0024	0	18.5577	0.0024
	$AR(7)$	0	19.0804	0.0008	0	19.4558	0.0008	0	17.8647	0.0010	0	18.2402	0.0010
	$AR(8)$	0	18.7720	0.0008	0	19.2413	0.0008	0	17.5769	0.0007	0	18.0462	0.0007
	$AR(9)$	0	18.7478	0.0008	0	19.3214	0.0008	0	17.3600	0.0008	0	17.9336	0.0008
1000	$MA(1)$	100	17.5714	0.0136	100	17.5891	0.0136	100	16.2937	0.0148	100	16.3115	0.0148
	$AR(1)$	0	29.9903	8.7360	0	30.0081	8.7360	0	28.0736	11.1184	0	28.0913	11.1184
	$AR(2)$	0	25.5246	2.6937	0	25.5601	2.6937	0	21.4620	0.0609	0	21.4975	0.0609
	$AR(3)$	0	21.9750	0.0791	0	22.0341	0.0791	0	20.2123	0.0285	0	20.2714	0.0285
	$AR(4)$	0	20.5831	0.0117	0	20.6717	0.0117	0	19.3939	0.0145	0	19.4825	0.0145
	$AR(5)$	0	19.8917	0.0048	0	20.0157	0.0048	0	18.7191	0.0059	0	18.8431	0.0059
	$AR(6)$	0	19.2238	0.0011	0	19.3892	0.0011	0	18.1994	0.0024	0	18.3648	0.0024
	$AR(7)$	0	18.9865	0.0008	0	19.1992	0.0008	0	17.7904	0.0010	0	18.0030	0.0010
	$AR(8)$	0	18.6692	0.0008	0	18.9350	0.0008	0	17.4936	0.0007	0	17.7594	0.0007
	$AR(9)$	0	18.6350	0.0008	0	18.9599	0.0008	0	17.2667	0.0008	0	17.5916	0.0008

The model of interest is a $MA(1)$ with intercept 0.1, moving average parameter 0.8 and variance of the errors equal to 1. The total number of replications is $R = 500$ and the simulated time series contains $T = 1000$ observations. The number of simulated paths H is 1 and 10, and N (first column) equals 100, 500, and 1000. The auxiliary model is either the model of interest itself or an $AR(r)$ (column Aux. Model). The table reports the percentage of choosing each model as the best (column Rate), the average value and the variance of the criteria (columns Mean and Var respectively) over the 500 replications. Numbers in bold denote the model chosen by the criteria. Numbers in bold and italics denote the second best model chosen by the criteria.

Table 4: $MA(1)$ with $\theta = 0.2$ and $T = 10000$

N	Aux. Model	$H = 1$						$H = 10$					
		AIC_{IM}			IC_{IM}			AIC_{IM}			IC_{IM}		
		Rate	Mean	Var	Rate	Mean	Var	Rate	Mean	Var	Rate	Mean	Var
100	$MA(1)$	90.2	21.5141	0.0031	99.8	21.6223	0.0031	93.6	20.1979	0.0032	100	20.3061	0.0032
	$AR(1)$	0	21.8561	0.0006	0	21.9643	0.0006	0	20.5594	0.0007	0	20.6675	0.0007
	$AR(2)$	9.6	21.5727	0.0011	0.2	21.7890	0.0011	5.2	20.2633	0.0010	0	20.4796	0.0010
	$AR(3)$	0.2	21.5946	0.0014	0	21.9551	0.0014	1.2	20.2706	0.0014	0	20.6311	0.0014
	$AR(4)$	0	21.6421	0.0014	0	22.1829	0.0014	0	20.3355	0.0014	0	20.8763	0.0014
	$AR(5)$	0	21.6947	0.0014	0	22.4518	0.0014	0	20.4003	0.0014	0	21.1574	0.0014
	$AR(6)$	0	21.7487	0.0014	0	22.7581	0.0014	0	20.4621	0.0014	0	21.4716	0.0014
	$AR(7)$	0	21.8635	0.0015	0	23.1613	0.0015	0	20.5420	0.0014	0	21.8398	0.0014
	$AR(8)$	0	21.9378	0.0014	0	23.5601	0.0014	0	20.6291	0.0014	0	22.2514	0.0014
	$AR(9)$	0	22.0211	0.0015	0	24.0040	0.0015	0	20.7277	0.0014	0	22.7106	0.0014
500	$MA(1)$	71.2	21.4412	0.0031	92.8	21.4725	0.0031	64.6	20.1232	0.0032	95.2	20.1545	0.0032
	$AR(1)$	0	21.7832	0.0006	0	21.8144	0.0006	0	20.4846	0.0007	0	20.5159	0.0007
	$AR(2)$	4.4	21.4757	0.0011	7.2	21.5383	0.0011	0	20.1646	0.0010	4.6	20.2272	0.0010
	$AR(3)$	22.6	21.4657	0.0014	0	21.5700	0.0014	35.4	20.1399	0.0014	0.2	20.2442	0.0014
	$AR(4)$	1.2	21.4731	0.0014	0	21.6296	0.0014	0	20.1648	0.0014	0	20.3212	0.0014
	$AR(5)$	0	21.4777	0.0014	0	21.6967	0.0014	0	20.1816	0.0014	0	20.4007	0.0014
	$AR(6)$	0.6	21.4757	0.0014	0	21.7677	0.0014	0	20.1874	0.0014	0	20.4795	0.0014
	$AR(7)$	0	21.5265	0.0015	0	21.9020	0.0015	0	20.2032	0.0014	0	20.5787	0.0014
	$AR(8)$	0	21.5288	0.0014	0	21.9981	0.0014	0	20.2184	0.0014	0	20.6877	0.0014
	$AR(9)$	0	21.5322	0.0015	0	22.1058	0.0015	0	20.2370	0.0014	0	20.8106	0.0014
1000	$MA(1)$	59	21.4201	0.0031	86	21.4378	0.0031	60.6	20.0884	0.0032	87.2	20.1061	0.0032
	$AR(1)$	0	21.7621	0.0006	0	21.7798	0.0006	0	20.4498	0.0007	0	20.4675	0.0007
	$AR(2)$	0.6	21.4516	0.0011	12.2	21.4871	0.0011	0	20.1268	0.0010	5.6	20.1622	0.0010
	$AR(3)$	2	21.4376	0.0014	1.8	21.4967	0.0014	39.4	20.0981	0.0014	7.2	20.1572	0.0014
	$AR(4)$	0.6	21.4401	0.0014	0	21.5287	0.0014	0	20.1180	0.0014	0	20.2066	0.0014
	$AR(5)$	1.2	21.4386	0.0014	0	21.5627	0.0014	0	20.1288	0.0014	0	20.2529	0.0014
	$AR(6)$	36.6	21.4296	0.0014	0	21.5951	0.0014	0	20.1276	0.0014	0	20.2930	0.0014
	$AR(7)$	0	21.4724	0.0015	0	21.6851	0.0015	0	20.1354	0.0014	0	20.3481	0.0014
	$AR(8)$	0	21.4657	0.0014	0	21.7316	0.0014	0	20.1416	0.0014	0	20.4074	0.0014
	$AR(9)$	0	21.4591	0.0015	0	21.7840	0.0015	0	20.1502	0.0014	0	20.4751	0.0014

The model of interest is a $MA(1)$ with intercept 0.1, moving average parameter 0.2 and variance of the errors equal to 1. The total number of replications is $R = 500$ and the simulated time series contains $T = 10000$ observations. The number of simulated paths H is 1 and 10, and N (first column) equals 100, 500, and 1000. The auxiliary model is either the model of interest itself or an $AR(r)$ (column Aux. Model). The table reports the percentage of choosing each model as the best (column Rate), the average value and the variance of the criteria (columns Mean and Var respectively) over the 500 replications. Numbers in bold denote the model chosen by the criteria. Numbers in bold and italics denote the second best model chosen by the criteria.

Table 5: $MA(1)$ with $\theta = 0.5$ and $T = 10000$

N	Aux. Model	$H = 1$						$H = 10$					
		Rate	AIC_{IM}		Rate	IC_{IM}		Rate	AIC_{IM}		Rate	IC_{IM}	
			Mean	Var		Mean	Var		Mean	Var		Mean	Var
100	$MA(1)$	100	20.3296	0.0050	100	20.4377	0.0050	100	19.0198	0.0050	100	19.1279	0.0050
	$AR(1)$	0	22.6906	0.0071	0	22.7988	0.0071	0	21.4221	0.0077	0	21.5303	0.0077
	$AR(2)$	0	21.3424	0.0010	0	21.5587	0.0010	0	20.0489	0.0010	0	20.2653	0.0010
	$AR(3)$	0	20.8298	0.0008	0	21.1903	0.0008	0	19.5051	0.0008	0	19.8657	0.0008
	$AR(4)$	0	20.6412	0.0015	0	21.1820	0.0015	0	19.3201	0.0015	0	19.8609	0.0015
	$AR(5)$	0	20.5965	0.0020	0	21.3536	0.0020	0	19.2888	0.0021	0	20.0459	0.0021
	$AR(6)$	0	20.5932	0.0025	0	21.6026	0.0025	0	19.3106	0.0024	0	20.3200	0.0024
	$AR(7)$	0	20.6886	0.0027	0	21.9865	0.0027	0	19.3722	0.0027	0	20.6701	0.0027
	$AR(8)$	0	20.7678	0.0028	0	22.3901	0.0028	0	19.4513	0.0028	0	21.0736	0.0028
	$AR(9)$	0	20.8405	0.0029	0	22.8233	0.0029	0	19.5456	0.0029	0	21.5285	0.0029
500	$MA(1)$	91.4	20.2566	0.0050	100	20.2879	0.0050	95.2	18.9450	0.0050	100	18.9763	0.0050
	$AR(1)$	0	22.6176	0.0071	0	22.6489	0.0071	0	21.3474	0.0077	0	21.3787	0.0077
	$AR(2)$	0	21.2455	0.0010	0	21.3080	0.0010	0	19.9502	0.0010	0	20.0128	0.0010
	$AR(3)$	0	20.7008	0.0008	0	20.8051	0.0008	0	19.3744	0.0008	0	19.4787	0.0008
	$AR(4)$	0	20.4723	0.0015	0	20.6287	0.0015	0	19.1494	0.0015	0	19.3059	0.0015
	$AR(5)$	0	20.3796	0.0020	0	20.5986	0.0020	0	19.0701	0.0021	0	19.2891	0.0021
	$AR(6)$	8.6	20.3202	0.0025	0	20.6122	0.0025	1.6	19.0358	0.0024	0	19.3279	0.0024
	$AR(7)$	0	20.3516	0.0027	0	20.7271	0.0027	2.6	19.0335	0.0027	0	19.4090	0.0027
	$AR(8)$	0	20.3589	0.0028	0	20.8282	0.0028	0.6	19.0406	0.0028	0	19.5099	0.0028
	$AR(9)$	0	0.0000	0.0000	0	0.0000	0.0000	0	0.0000	0.0000	0	0.0000	0.0000
1000	$MA(1)$	79.2	20.2355	0.0050	100	20.2533	0.0050	87.2	18.9102	0.0050	100	18.9279	0.0050
	$AR(1)$	0	22.5966	0.0071	0	22.6143	0.0071	0	21.3126	0.0077	0	21.3303	0.0077
	$AR(2)$	0	21.2214	0.0010	0	21.2568	0.0010	0	19.9124	0.0010	0	19.9479	0.0010
	$AR(3)$	0	20.6727	0.0008	0	20.7318	0.0008	0	19.3326	0.0008	0	19.3917	0.0008
	$AR(4)$	0	20.4392	0.0015	0	20.5278	0.0015	0	19.1026	0.0015	0	19.1912	0.0015
	$AR(5)$	0	20.3405	0.0020	0	20.4646	0.0020	0	19.0173	0.0021	0	19.1414	0.0021
	$AR(6)$	15.2	20.2741	0.0025	0	20.4395	0.0025	0.6	18.9760	0.0024	0	19.1414	0.0024
	$AR(7)$	0	20.2976	0.0027	0	20.5102	0.0027	3.2	18.9657	0.0027	0	19.1784	0.0027
	$AR(8)$	0	20.2958	0.0028	0	20.5616	0.0028	8	18.9637	0.0028	0	19.2296	0.0028
	$AR(9)$	0	20.2784	0.0029	0	20.6033	0.0029	0	18.9681	0.0029	0	19.2930	0.0029

The model of interest is a $MA(1)$ with intercept 0.1, moving average parameter 0.5 and variance of the errors equal to 1. The total number of replications is $R = 500$ and the simulated time series contains $T = 10000$ observations. The number of simulated paths H is 1 and 10, and N (first column) equals 100, 500, and 1000. The auxiliary model is either the model of interest itself or an $AR(r)$ (column Aux. Model). The table reports the percentage of choosing each model as the best (column Rate), the average value and the variance of the criteria (columns Mean and Var respectively) over the 500 replications. Numbers in bold denote the model chosen by the criteria. Numbers in bold and italics denote the second best model chosen by the criteria.

Table 6: $MA(1)$ with $\theta = 0.8$ and $T = 10000$

N	Aux. Model	H = 1						H = 10					
		Rate	AIC_{IM}		Rate	IC_{IM}		Rate	AIC_{IM}		Rate	IC_{IM}	
			Mean	Var		Mean	Var		Mean	Var		Mean	Var
100	$MA(1)$	100	17.7548	0.0135	100	17.8629	0.0135	100	16.4297	0.0136	100	16.5378	0.0136
	$AR(1)$	0	25.9499	4.6340	0	26.0580	4.6340	0	25.4503	7.6719	0	25.5585	7.6719
	$AR(2)$	0	22.9971	0.0906	0	23.2134	0.0906	0	21.7752	0.1058	0	21.9915	0.1058
	$AR(3)$	0	21.7800	0.0385	0	22.1405	0.0385	0	20.3860	0.0304	0	20.7466	0.0304
	$AR(4)$	0	20.9349	0.0159	0	21.4756	0.0159	0	19.5385	0.0125	0	20.0792	0.0125
	$AR(5)$	0	20.3177	0.0067	0	21.0748	0.0067	0	18.9529	0.0055	0	19.7100	0.0055
	$AR(6)$	0	19.7852	0.0021	0	20.7946	0.0021	0	18.4994	0.0022	0	19.5089	0.0022
	$AR(7)$	0	19.5012	0.0010	0	20.7990	0.0010	0	18.1794	0.0009	0	19.4772	0.0009
	$AR(8)$	0	19.2882	0.0008	0	20.9106	0.0008	0	17.9553	0.0007	0	19.5777	0.0007
	$AR(9)$	0	19.1130	0.0008	0	21.0958	0.0008	0	17.8115	0.0007	0	19.7944	0.0007
500	$MA(1)$	100	17.6818	0.0135	100	17.7131	0.0135	100	16.3550	0.0136	100	16.3862	0.0136
	$AR(1)$	0	25.8769	4.6340	0	25.9082	4.6340	0	25.3756	7.6719	0	25.4069	7.6719
	$AR(2)$	0	22.9001	0.0906	0	22.9627	0.0906	0	21.6765	0.1058	0	21.7390	0.1058
	$AR(3)$	0	21.6510	0.0385	0	21.7553	0.0385	0	20.2553	0.0304	0	20.3596	0.0304
	$AR(4)$	0	20.7659	0.0159	0	20.9223	0.0159	0	19.3678	0.0125	0	19.5242	0.0125
	$AR(5)$	0	20.1007	0.0067	0	20.3197	0.0067	0	18.7342	0.0055	0	18.9532	0.0055
	$AR(6)$	0	19.5122	0.0021	0	19.8042	0.0021	0	18.2247	0.0022	0	18.5167	0.0022
	$AR(7)$	0	19.1642	0.0010	0	19.5397	0.0010	0	17.8407	0.0009	0	18.2161	0.0009
	$AR(8)$	0	18.8793	0.0008	0	19.3486	0.0008	0	17.5446	0.0007	0	18.0140	0.0007
	$AR(9)$	0	18.6240	0.0008	0	19.1976	0.0008	0	17.3208	0.0007	0	17.8944	0.0007
1000	$MA(1)$	100	17.6607	0.0135	100	17.6784	0.0135	100	16.3201	0.0136	100	16.3379	0.0136
	$AR(1)$	0	25.8558	4.6340	0	25.8736	4.6340	0	25.3408	7.6719	0	25.3585	7.6719
	$AR(2)$	0	22.8760	0.0906	0	22.9115	0.0906	0	21.6386	0.1058	0	21.6741	0.1058
	$AR(3)$	0	21.6229	0.0385	0	21.6820	0.0385	0	20.2135	0.0304	0	20.2726	0.0304
	$AR(4)$	0	20.7328	0.0159	0	20.8214	0.0159	0	19.3209	0.0125	0	19.4096	0.0125
	$AR(5)$	0	20.0616	0.0067	0	20.1857	0.0067	0	18.6813	0.0055	0	18.8054	0.0055
	$AR(6)$	0	19.4661	0.0021	0	19.6315	0.0021	0	18.1649	0.0022	0	18.3303	0.0022
	$AR(7)$	0	19.1101	0.0010	0	19.3228	0.0010	0	17.7729	0.0009	0	17.9855	0.0009
	$AR(8)$	0	18.8162	0.0008	0	19.0820	0.0008	0	17.4678	0.0007	0	17.7337	0.0007
	$AR(9)$	0	18.5509	0.0008	0	18.8759	0.0008	0	17.2340	0.0007	0	17.5589	0.0007

The model of interest is a $MA(1)$ with intercept 0.1, moving average parameter 0.8 and variance of the errors equal to 1. The total number of replications is $R = 500$ and the simulated time series contains $T = 10000$ observations. The number of simulated paths H is 1 and 10, and N (first column) equals 100, 500, and 1000. The auxiliary model is either the model of interest itself or an $AR(r)$ (column Aux. Model). The table reports the percentage of choosing each model as the best (column Rate), the average value and the variance of the criteria (columns Mean and Var respectively) over the 500 replications. Numbers in bold denote the model chosen by the criteria. Numbers in bold and italics denote the second best model chosen by the criteria.

Table 7: SV with $\rho = 0.2$

N	Aux. Model	H = 1		H = 10	
		AIC_{IM}	IC_{IM}	AIC_{IM}	IC_{IM}
100	ARCH(1)	33.2831	33.3913	32.1805	32.2887
	ARCH(2)	33.3475	33.5638	32.2851	32.5014
	ARCH(3)	33.4058	33.7664	32.3426	32.7031
	ARCH(4)	33.7246	34.2654	32.5310	33.0717
	ARCH(5)	34.0544	34.8115	32.6334	33.3905
	ARCH(6)	33.8899	34.8994	32.6871	33.6966
	ARCH(7)	33.8705	35.1683	32.8249	34.1228
	ARCH(8)	33.9082	35.5306	33.1147	34.7370
	ARCH(9)	34.0972	36.0800	33.2071	35.1900
	GARCH(1,1)	33.0184	33.2347	32.1916	32.4079
500	ARCH(1)	33.3465	33.3778	32.2439	32.2751
	ARCH(2)	33.3868	33.4494	32.3244	32.3870
	ARCH(3)	33.4132	33.5175	32.3499	32.4542
	ARCH(4)	33.6919	33.8484	32.4983	32.6548
	ARCH(5)	33.9737	34.1928	32.5528	32.7718
	ARCH(6)	33.7533	34.0453	32.5505	32.8425
	ARCH(7)	33.6698	34.0453	32.6243	32.9997
	ARCH(8)	33.6356	34.1049	32.8420	33.3113
	ARCH(9)	33.7445	34.3182	32.8545	33.4281
	GARCH(1,1)	33.0577	33.1203	32.2309	32.2935
1000	ARCH(1)	33.3710	33.3887	32.2684	32.2861
	ARCH(2)	33.4083	33.4438	32.3459	32.3814
	ARCH(3)	33.4307	33.4898	32.3674	32.4265
	ARCH(4)	33.7045	33.7931	32.5108	32.5994
	ARCH(5)	33.9802	34.1043	32.5593	32.6834
	ARCH(6)	33.7528	33.9182	32.5500	32.7154
	ARCH(7)	33.6613	33.8740	32.6158	32.8285
	ARCH(8)	33.6181	33.8839	32.8245	33.0904
	ARCH(9)	33.7171	34.0420	32.8270	33.1519
	GARCH(1,1)	33.0792	33.1147	32.2524	32.2879
100	AR(1)	35.3864	35.6028	35.4460	35.6623
	AR(2)	35.2903	35.6508	35.2713	35.6318
	AR(3)	35.3676	35.9083	35.3018	35.8425
	AR(4)	35.4884	36.2455	35.3790	36.1361
	AR(5)	35.5886	36.5981	35.4819	36.4913
	AR(6)	35.8711	37.1690	35.5469	36.8448
	AR(7)	36.1076	37.7300	35.8613	37.4836
	AR(8)	36.6213	38.6042	35.6990	37.6819
	AR(9)	36.6022	38.9816	35.8123	38.1917
	ARMA(1,1)	34.1004	34.4609	34.9697	35.3302
500	AR(1)	35.4258	35.4884	35.4853	35.5479
	AR(2)	35.2977	35.4020	35.2787	35.3830
	AR(3)	35.3349	35.4914	35.2691	35.4256
	AR(4)	35.4078	35.6268	35.2983	35.5174
	AR(5)	35.4520	35.7440	35.3452	35.6373
	AR(6)	35.6705	36.0459	35.3463	35.7217
	AR(7)	35.8350	36.3043	35.5887	36.0580
	AR(8)	36.2687	36.8423	35.3464	35.9200
	AR(9)	36.1616	36.8499	35.3717	36.0600
	ARMA(1,1)	34.1078	34.2120	34.9771	35.0814
1000	AR(1)	35.4473	35.4828	35.5068	35.5423
	AR(2)	35.3152	35.3743	35.2962	35.3553
	AR(3)	35.3474	35.4360	35.2816	35.3702
	AR(4)	35.4143	35.5384	35.3049	35.4289
	AR(5)	35.4515	35.6169	35.3447	35.5102
	AR(6)	35.6620	35.8747	35.3378	35.5504
	AR(7)	35.8175	36.0834	35.5712	35.8370
	AR(8)	36.2412	36.5661	35.3189	35.6438
	AR(9)	36.1231	36.5130	35.3332	35.7231
	ARMA(1,1)	34.1253	34.1843	34.9946	35.0536

The model of interest is a SV with intercept 0, autoregressive parameter 0.2, and variances of the errors 1 and 0.01 for the observation and volatility equations respectively. The number of simulated paths H is 1 and 10, N (first column) equals 100, 500, and 1000, and $T = 10000$. The auxiliary models (column Aux. Model) are ARCH(r) and GARCH(1,1) on the observations (top three panels), and AR(r) and ARMA(1,1) in the log of the square observations (bottom three panels). Numbers in bold denote the model chosen by the criteria.

Table 8: SV with $\rho = 0.4$

N	Aux. Model	H = 1		H = 10	
		AIC_{IM}	IC_{IM}	AIC_{IM}	IC_{IM}
100	ARCH(1)	32.6909	32.7991	31.4918	31.5999
	ARCH(2)	32.7029	32.9192	31.3327	31.5490
	ARCH(3)	32.6209	32.9815	31.2061	31.5666
	ARCH(4)	32.8166	33.3574	30.9933	31.5341
	ARCH(5)	33.0946	33.8516	31.1796	31.9367
	ARCH(6)	33.0483	34.0577	31.2265	32.2359
	ARCH(7)	33.1338	34.4316	31.3767	32.6746
	ARCH(8)	33.2123	34.8346	31.4546	33.0769
	ARCH(9)	33.4839	35.4668	31.5286	33.5114
	GARCH(1, 1)	32.3074	32.5237	31.6147	31.8310
500	ARCH(1)	32.7543	32.7856	31.5551	31.5864
	ARCH(2)	32.7422	32.8048	31.3721	31.4347
	ARCH(3)	32.6283	32.7326	31.2135	31.3177
	ARCH(4)	32.7840	32.9404	30.9607	31.1171
	ARCH(5)	33.0139	33.2329	31.0990	31.3180
	ARCH(6)	32.9116	33.2036	31.0899	31.3819
	ARCH(7)	32.9331	33.3086	31.1761	31.5515
	ARCH(8)	32.9397	33.4090	31.1819	31.6512
	ARCH(9)	33.1313	33.7049	31.1759	31.7495
	GARCH(1, 1)	32.3468	32.4094	31.6541	31.7166
1000	ARCH(1)	32.7788	32.7965	31.5796	31.5974
	ARCH(2)	32.7637	32.7992	31.3936	31.4290
	ARCH(3)	32.6458	32.7049	31.2310	31.2900
	ARCH(4)	32.7965	32.8851	30.9732	31.0618
	ARCH(5)	33.0204	33.1445	31.1055	31.2296
	ARCH(6)	32.9111	33.0766	31.0894	31.2548
	ARCH(7)	32.9246	33.1373	31.1676	31.3803
	ARCH(8)	32.9222	33.1880	31.1644	31.4303
	ARCH(9)	33.1038	33.4287	31.1484	31.4734
	GARCH(1, 1)	32.3683	32.4037	31.6756	31.7110
100	AR(1)	35.0729	35.2892	35.3750	35.5914
	AR(2)	35.0370	35.3976	34.3425	34.7031
	AR(3)	35.1162	35.6570	34.2341	34.7749
	AR(4)	35.2200	35.9771	35.8620	36.6191
	AR(5)	35.3136	36.3231	39.0823	40.0917
	AR(6)	35.5694	36.8673	35.6689	36.9667
	AR(7)	35.8210	37.4434	34.4348	36.0571
	AR(8)	36.2516	38.2344	34.7760	36.7588
	AR(9)	36.2286	38.6080	34.6435	37.0229
	ARMA(1, 1)	35.0282	35.3888	33.3871	33.7476
500	AR(1)	35.1123	35.1749	35.4144	35.4770
	AR(2)	35.0444	35.1487	34.3499	34.4542
	AR(3)	35.0835	35.2400	34.2015	34.3579
	AR(4)	35.1394	35.3584	35.7814	36.0004
	AR(5)	35.1770	35.4690	38.9456	39.2377
	AR(6)	35.3688	35.7442	35.4682	35.8437
	AR(7)	35.5484	36.0177	34.1621	34.6314
	AR(8)	35.8990	36.4726	34.4233	34.9969
	AR(9)	35.7879	36.4763	34.2028	34.8912
	ARMA(1, 1)	0.0000	0.0000	0.0000	0.0000
1000	AR(1)	35.1338	35.1692	35.4359	35.4714
	AR(2)	35.0619	35.1210	34.3674	34.4265
	AR(3)	35.0960	35.1847	34.2140	34.3026
	AR(4)	35.1459	35.2700	35.7879	35.9119
	AR(5)	35.1765	35.3419	38.9451	39.1106
	AR(6)	35.3603	35.5729	35.4597	35.6724
	AR(7)	35.5309	35.7967	34.1446	34.4105
	AR(8)	35.8715	36.1964	34.3958	34.7207
	AR(9)	35.7494	36.1393	34.1643	34.5543
	ARMA(1, 1)	35.0531	35.1122	33.4120	33.4711

The model of interest is a SV with intercept 0, autoregressive parameter 0.4, and variances of the errors 1 and 0.01 for the observation and volatility equations respectively. The number of simulated paths H is 1 and 10, N (first column) equals 100, 500, and 1000, and $T = 10000$. The auxiliary models (column Aux. Model) are ARCH(r) and GARCH(1,1) on the observations (top three panels), and AR(r) and ARMA(1,1) in the log of the square observations (bottom three panels). Numbers in bold denote the model chosen by the criteria.

Table 9: SV with $\rho = 0.9$

N	Aux. Model	H = 1		H = 10	
		AIC_{IM}	IC_{IM}	AIC_{IM}	IC_{IM}
100	ARCH(1)	29.9909	30.0991	27.5533	27.6615
	ARCH(2)	26.8011	27.0174	26.0786	26.2949
	ARCH(3)	26.2736	26.6341	25.2836	25.6441
	ARCH(4)	25.4216	25.9623	24.4728	25.0135
	ARCH(5)	25.5117	26.2688	24.3088	25.0659
	ARCH(6)	25.3190	26.3284	24.1498	25.1592
	ARCH(7)	25.0921	26.3899	23.9541	25.2520
	ARCH(8)	24.5769	26.1992	23.7065	25.3288
	ARCH(9)	24.6328	26.6157	23.7469	25.7298
	GARCH(1, 1)	24.0947	24.3110	23.1882	23.4045
500	ARCH(1)	30.0543	30.0856	27.6167	27.6480
	ARCH(2)	26.8404	26.9030	26.1179	26.1805
	ARCH(3)	26.2810	26.3852	25.2909	25.3952
	ARCH(4)	25.3889	25.5454	24.4401	24.5966
	ARCH(5)	25.4311	25.6501	24.2282	24.4472
	ARCH(6)	25.1823	25.4743	24.0131	24.3052
	ARCH(7)	24.8914	25.2669	23.7535	24.1289
	ARCH(8)	24.3042	24.7736	23.4338	23.9032
	ARCH(9)	24.2802	24.8538	23.3943	23.9679
	GARCH(1, 1)	24.1341	24.1966	23.2275	23.2901
1000	ARCH(1)	30.0788	30.0965	27.6412	27.6589
	ARCH(2)	26.8619	26.8974	26.1394	26.1749
	ARCH(3)	26.2985	26.3575	25.3084	25.3675
	ARCH(4)	25.4014	25.4900	24.4526	24.5412
	ARCH(5)	25.4376	25.5617	24.2347	24.3587
	ARCH(6)	25.1818	25.3472	24.0126	24.1781
	ARCH(7)	24.8829	25.0956	23.7450	23.9576
	ARCH(8)	24.2867	24.5526	23.4163	23.6822
	ARCH(9)	24.2527	24.5776	23.3668	23.6917
	GARCH(1, 1)	24.1556	24.1910	23.2490	23.2845
100	AR(1)	30.9197	31.1361	29.6700	29.8863
	AR(2)	29.8641	30.2246	28.2187	28.5792
	AR(3)	28.9847	29.5255	27.2016	27.7424
	AR(4)	27.9531	28.7102	26.4934	27.2505
	AR(5)	27.6784	28.6878	26.0477	27.0572
	AR(6)	27.0890	28.3868	25.7066	27.0044
	AR(7)	27.1236	28.7460	25.6673	27.2896
	AR(8)	27.2477	29.2305	25.5106	27.4934
	AR(9)	27.1341	29.5135	25.5319	27.9113
	ARMA(1, 1)	26.3813	26.7418	24.4414	24.8019
500	AR(1)	30.9591	31.0217	29.7093	29.7719
	AR(2)	29.8714	29.9757	28.2261	28.3304
	AR(3)	28.9521	29.1085	27.1690	27.3254
	AR(4)	27.8725	28.0915	26.4128	26.6318
	AR(5)	27.5417	27.8338	25.9111	26.2031
	AR(6)	26.8883	27.2638	25.5059	25.8814
	AR(7)	26.8510	27.3203	25.3946	25.8639
	AR(8)	26.8950	27.4686	25.1580	25.7316
	AR(9)	26.6935	27.3818	25.0912	25.7796
	ARMA(1, 1)	26.3886	26.4929	24.4488	24.5531
1000	AR(1)	30.9806	31.0161	29.7308	29.7663
	AR(2)	29.8889	29.9480	28.2436	28.3026
	AR(3)	28.9646	29.0532	27.1815	27.2701
	AR(4)	27.8790	28.0030	26.4193	26.5433
	AR(5)	27.5412	27.7067	25.9106	26.0760
	AR(6)	26.8798	27.0925	25.4975	25.7101
	AR(7)	26.8335	27.0993	25.3771	25.6430
	AR(8)	26.8675	27.1924	25.1305	25.4554
	AR(9)	26.6550	27.0449	25.0527	25.4426
	ARMA(1, 1)	26.4061	26.4652	24.4663	24.5254

The model of interest is a SV with intercept 0, autoregressive parameter 0.9, and variances of the errors 1 and 0.01 for the observation and volatility equations respectively. The number of simulated paths H is 1 and 10, N (first column) equals 100, 500, and 1000, and $T = 10000$. The auxiliary models (column Aux. Model) are ARCH(r) and GARCH(1,1) on the observations (top three panels), and AR(r) and ARMA(1,1) in the log of the square observations (bottom three panels). Numbers in bold denote the model chosen by the criteria.

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

- 1101 GIACOMO MASIER and ERNESTO VILLANUEVA: Consumption and initial mortgage conditions: evidence from survey data.
- 1102 PABLO HERNÁNDEZ DE COS and ENRIQUE MORAL-BENITO: Endogenous fiscal consolidations
- 1103 CÉSAR CALDERÓN, ENRIQUE MORAL-BENITO and LUIS SERVÉN: Is infrastructure capital productive? A dynamic heterogeneous approach.
- 1104 MICHAEL DANQUAH, ENRIQUE MORAL-BENITO and BAZOUMANA OUATTARA: TFP growth and its determinants: nonparametrics and model averaging.
- 1105 JUAN CARLOS BERGANZA and CARMEN BROTO: Flexible inflation targets, forex interventions and exchange rate volatility in emerging countries.
- 1106 FRANCISCO DE CASTRO, JAVIER J. PÉREZ and MARTA RODRÍGUEZ VIVES: Fiscal data revisions in Europe.
- 1107 ANGEL GAVILÁN, PABLO HERNÁNDEZ DE COS, JUAN F. JIMENO and JUAN A. ROJAS: Fiscal policy, structural reforms and external imbalances: a quantitative evaluation for Spain.
- 1108 EVA ORTEGA, MARGARITA RUBIO and CARLOS THOMAS: House purchase versus rental in Spain.
- 1109 ENRIQUE MORAL-BENITO: Dynamic panels with predetermined regressors: likelihood-based estimation and Bayesian averaging with an application to cross-country growth.
- 1110 NIKOLAI STÄHLER and CARLOS THOMAS: FiMod – a DSGE model for fiscal policy simulations.
- 1111 ÁLVARO CARTEA and JOSÉ PENALVA: Where is the value in high frequency trading?
- 1112 FILIPA SÁ and FRANCESCA VIANI: Shifts in portfolio preferences of international investors: an application to sovereign wealth funds.
- 1113 REBECA ANGUREN MARTÍN: Credit cycles: Evidence based on a non-linear model for developed countries.
- 1114 LAURA HOSPIDO: Estimating non-linear models with multiple fixed effects: A computational note.
- 1115 ENRIQUE MORAL-BENITO and CRISTIAN BARTOLUCCI: Income and democracy: Revisiting the evidence.
- 1116 AGUSTÍN MARAVALL HERRERO and DOMINGO PÉREZ CAÑETE: Applying and interpreting model-based seasonal adjustment. The euro-area industrial production series.
- 1117 JULIO CÁCERES-DELPIANO: Is there a cost associated with an increase in family size beyond child investment? Evidence from developing countries.
- 1118 DANIEL PÉREZ, VICENTE SALAS-FUMÁS and JESÚS SAURINA: Do dynamic provisions reduce income smoothing using loan loss provisions?
- 1119 GALO NUÑO, PEDRO TEDDE and ALESSIO MORO: Money dynamics with multiple banks of issue: evidence from Spain 1856-1874.
- 1120 RAQUEL CARRASCO, JUAN F. JIMENO and A. CAROLINA ORTEGA: Accounting for changes in the Spanish wage distribution: the role of employment composition effects.
- 1121 FRANCISCO DE CASTRO and LAURA FERNÁNDEZ-CABALLERO: The effects of fiscal shocks on the exchange rate in Spain.
- 1122 JAMES COSTAIN and ANTON NAKOV: Precautionary price stickiness.
- 1123 ENRIQUE MORAL-BENITO: Model averaging in economics.
- 1124 GABRIEL JIMÉNEZ, ATIF MIAN, JOSÉ-LUIS PEYDRÓ AND JESÚS SAURINA: Local versus aggregate lending channels: the effects of securitization on corporate credit supply.
- 1125 ANTON NAKOV and GALO NUÑO: A general equilibrium model of the oil market.
- 1126 DANIEL C. HARDY and MARÍA J. NIETO: Cross-border coordination of prudential supervision and deposit guarantees.
- 1127 LAURA FERNÁNDEZ-CABALLERO, DIEGO J. PEDREGAL and JAVIER J. PÉREZ: Monitoring sub-central government spending in Spain.
- 1128 CARLOS PÉREZ MONTES: Optimal capital structure and regulatory control.
- 1129 JAVIER ANDRÉS, JOSÉ E. BOSCA and JAVIER FERRI: Household debt and labour market fluctuations.
- 1130 ANTON NAKOV and CARLOS THOMAS: Optimal monetary policy with state-dependent pricing.
- 1131 JUAN F. JIMENO and CARLOS THOMAS: Collective bargaining, firm heterogeneity and unemployment.
- 1132 ANTON NAKOV and GALO NUÑO: Learning from experience in the stock market.
- 1133 ALESSIO MORO and GALO NUÑO: Does TFP drive housing prices? A growth accounting exercise for four countries.
- 1201 CARLOS PÉREZ MONTES: Regulatory bias in the price structure of local telephone services.

- 1202 MAXIMO CAMACHO, GABRIEL PEREZ-QUIROS and PILAR PONCELA: Extracting non-linear signals from several economic indicators.
- 1203 MARCOS DAL BIANCO, MAXIMO CAMACHO and GABRIEL PEREZ-QUIROS: Short-run forecasting of the euro-dollar exchange rate with economic fundamentals.
- 1204 ROCIO ALVAREZ, MAXIMO CAMACHO and GABRIEL PEREZ-QUIROS: Finite sample performance of small versus large scale dynamic factor models.
- 1205 MAXIMO CAMACHO, GABRIEL PEREZ-QUIROS and PILAR PONCELA: Markov-switching dynamic factor models in real time.
- 1206 IGNACIO HERNANDO and ERNESTO VILLANUEVA: The recent slowdown of bank lending in Spain: are supply-side factors relevant?
- 1207 JAMES COSTAIN and BEATRIZ DE BLAS: Smoothing shocks and balancing budgets in a currency union.
- 1208 AITOR LACUESTA, SERGIO PUENTE and ERNESTO VILLANUEVA: The schooling response to a sustained increase in low-skill wages: evidence from Spain 1989-2009.
- 1209 GABOR PULA and DANIEL SANTABÁRBARA: Is China climbing up the quality ladder?
- 1210 ROBERTO BLANCO and RICARDO GIMENO: Determinants of default ratios in the segment of loans to households in Spain.
- 1211 ENRIQUE ALBEROLA, AITOR ERCE and JOSÉ MARÍA SERENA: International reserves and gross capital flows. Dynamics during financial stress.
- 1212 GIANCARLO CORSETTI, LUCA DEDOLA and FRANCESCA VIANI: The international risk-sharing puzzle is at business-cycle and lower frequency.
- 1213 FRANCISCO ALVAREZ-CUADRADO, JOSE MARIA CASADO, JOSE MARIA LABEAGA and DHANOOS SUTTHIPHISAL: Envy and habits: panel data estimates of interdependent preferences.
- 1214 JOSE MARIA CASADO: Consumption partial insurance of Spanish households.
- 1215 J. ANDRÉS, J. E. BOSCA and J. FERRI: Household leverage and fiscal multipliers.
- 1216 JAMES COSTAIN and BEATRIZ DE BLAS: The role of fiscal delegation in a monetary union: a survey of the political economy issues.
- 1217 ARTURO MACÍAS and MARIANO MATILLA-GARCÍA: Net energy analysis in a Ramsey-Hotelling growth model.
- 1218 ALFREDO MARTÍN-OLIVER, SONIA RUANO and VICENTE SALAS-FUMÁS: Effects of equity capital on the interest rate and the demand for credit. Empirical evidence from Spanish banks.
- 1219 PALOMA LÓPEZ-GARCÍA, JOSÉ MANUEL MONTERO and ENRIQUE MORAL-BENITO: Business cycles and investment in intangibles: evidence from Spanish firms.
- 1220 ENRIQUE ALBEROLA, LUIS MOLINA and PEDRO DEL RÍO: Boom-bust cycles, imbalances and discipline in Europe.
- 1221 CARLOS GONZÁLEZ-AGUADO and ENRIQUE MORAL-BENITO: Determinants of corporate default: a BMA approach.
- 1222 GALO NUÑO and CARLOS THOMAS: Bank leverage cycles.
- 1223 YUNUS AKSOY and HENRIQUE S. BASSO: Liquidity, term spreads and monetary policy.
- 1224 FRANCISCO DE CASTRO and DANIEL GARROTE: The effects of fiscal shocks on the exchange rate in the EMU and differences with the US.
- 1225 STÉPHANE BONHOMME and LAURA HOSPIDO: The cycle of earnings inequality: evidence from Spanish social security data.
- 1226 CARMEN BROTO: The effectiveness of forex interventions in four Latin American countries.
- 1227 LORENZO RICCI and DAVID VEREDAS: TailCoR.
- 1228 YVES DOMINICY, SIEGFRIED HÖRMANN, HIROAKI OGATA and DAVID VEREDAS: Marginal quantiles for stationary processes.
- 1229 MATTEO BARIGOZZI, ROXANA HALBLEIB and DAVID VEREDAS: Which model to match?

BANCO DE ESPAÑA
Eurosistema

Unidad de Servicios Auxiliares
Alcalá, 48 - 28014 Madrid
Telephone +34 91 338 6360
E-mail: publicaciones@bde.es
www.bde.es