Escola de Pós-Graduação em Economia - EPGE
Fundação Getulio Vargas

Ensaios em Macroeconometria e Finanças

Tese submetida à Escola de Pós-Graduação em Economia
da Fundação Getulio Vargas como requisito de obtenção do
título de Doutor em Economia

Aluno: Wagner Piazza Gaglianone

Professor Orientador: João Victor Issler
Professor Co-orientador: Luiz Renato Lima

Rio de Janeiro
2007

Escola de Pós-Graduação em Economia - EPGE
Fundação Getulio Vargas

Ensaios em Macroeconometria e Finanças

Tese submetida à Escola de Pós-Graduação em Economia
da Fundação Getulio Vargas como requisito de obtenção do
título de Doutor em Economia

Aluno: Wagner Piazza Gaglianone

Banca Examinadora:
João Victor Issler (Orientador, EPGE-FGV)
Luiz Renato Lima (Co-orientador, EPGE-FGV)
Maria Cristina Terra (EPGE-FGV)
André Minella (Banco Central do Brasil)
Tomás Málaga (Banco Itaú S.A.)

Rio de Janeiro
2007

Contents

Acknowledgements

Agradeço,
Ao meu orientador, João Victor Issler, pela oportunidade de pesquisa conjunta e fundamental contribuição para a minha formação e o desenvolvimento desta tese.
Sem sua intensa parceria, dedicação e paciência inesgotável esta tese não seria possível.
Ao meu co-orientador, Luiz Renato Lima, um incansável parceiro, cuja contribuição foi também crucial para a minha formação e para o desenvolvimento desta tese.
Sem sua imensa dedicação e paciência esta tese também não seria possível.
Ao meu orientador no Banco Central do Brasil, Octávio M. Bessada Lion, que sempre me incentivou desde o início de toda a jornada.
Sem o seu fundamental apoio e incentivo, esta tese também não seria possível.
Aos membros da Banca Examinadora, pela paciência na leitura dos artigos e pelas valiosas contribuições a este trabalho.
Aos professores Oliver Linton e Myung Seo, pelo constante apoio durante meu doutorado sanduíche na LSE - The London School of Economics and Political Science, Inglaterra.
A todo corpo docente da EPGE, pelo compromisso com a excelência acadêmica, e pela vital contribuição para a minha formação, em especial,
João Victor Issler, Luiz Renato Lima, Marcelo Fernandes, Marco Bonomo e Maria Cristina Terra.
Aos funcionários da EPGE e da Biblioteca Mario Henrique Simonsen, pelo apoio técnico, indispensável.
Aos colegas da EPGE, pela amizade e companheirismo durante estes anos de convivência.
Aos meus pais Paulo e Sandra, e minha irmã Karina, que sempre me apoiaram, com muito amor e carinho, em todos os meus projetos e desafios, ao longo de toda a minha vida.
A minha esposa, Carla, pelo seu amor e compreensão incondicional.
A todos meus amigos, dos quais sempre recebi grande incentivo.
Ao Banco Central do Brasil, pelo patrocínio e pela oportunidade de uma formação continuada.
À CAPES, que através do programa de doutorado sanduíche, proporcionou-me a oportunidade de aperfeiçoar os meus estudos e artigos de tese na LSE.
Ao Programme Alban - European Union, pela possibilidade de estudos numa universidade estrangeira.

Por último e, mais importante,
A Deus, sempre, pela iluminação e proteção.

List of Tables

List of Figures

Introduction

This thesis is composed of three essays referent to the subjects of macroeconometrics and finance. In each essay, which corresponds to one chapter, the objective is to investigate and analyze advanced econometric techniques, applied to relevant macroeconomic questions, such as the capital mobility hypothesis and the sustainability of public debt. A finance topic regarding portfolio risk management is also investigated, through an econometric technique used to evaluate Value-at-Risk models.

The first chapter investigates an intertemporal optimization model to analyze the current account. Based on Campbell & Shiller's (1987) approach, a Wald test is conducted to analyze a set of restrictions imposed to a VAR used to forecast the current account. The estimation is based on three different procedures: OLS, SUR and the two-way error decomposition of Fuller & Battese (1974), due to the presence of global shocks. A note on Granger causality is also provided, which is shown to be a necessary condition to perform the Wald test with serious implications to the validation of the model. An empirical exercise for the G-7 countries is presented, and the results substantially change with the different estimation techniques. A small Monte Carlo simulation is also presented to investigate the size and power of the Wald test based on the considered estimators.

The second chapter presents a study about fiscal sustainability based on a quantile autoregression (QAR) model. A novel methodology to separate periods of nonstationarity from stationary ones is proposed, which allows one to identify trajectories of public debt that are not compatible with fiscal sustainability. Moreover, such trajectories are used to construct a debt ceiling, that is, the largest value of public debt that does not jeopardize long-run fiscal sustainability. An out-of-sample forecast of such a ceiling is also constructed, and can be used by policy makers interested in keeping the public debt on a sustainable path. An empirical exercise by using Brazilian data is conducted to show the applicability of the methodology.

In the third chapter, an alternative backtest to evaluate the performance of Value-at-Risk (VaR) models is proposed. The econometric methodology allows one to directly test the overall performance of a VaR model, as well as identify periods of an increased risk exposure, which seems to be a novelty in the literature. Quantile regressions provide an appropriate environment to investigate VaR models, since they can naturally be viewed as a conditional quantile function of a given return series. An empirical exercise is conducted for daily S&P500 series, and a Monte Carlo simulation is also presented, revealing that the proposed test might exhibit more power in comparison to other backtests.

Chapter 1

An econometric contribution to the intertemporal approach of the current account[1]

Abstract

This paper investigates an intertemporal optimization model to analyze the current account through Campbell & Shiller's (1987) approach. In this setup, a Wald test is conducted to analyze a set of restrictions imposed to a VAR, used to forecast the current account for a set of countries. We focused here on three estimation procedures: OLS, SUR and the two-way error decomposition of Fuller & Battese (1974). We also propose an original note on Granger causality, which is a necessary condition to perform the Wald test. Theoretical results show that, in the presence of global shocks, OLS and SUR estimators might lead to a biased covariance matrix, with serious implications to the validation of the model. A small Monte Carlo simulation confirms these findings and indicates the Fuller & Battese procedure in the presence of global shocks. An empirical exercise for the G-7 countries is also provided, and the results of the Wald test substantially change with different estimation techniques. In addition, global shocks can account up to 40% of the total residuals of the G-7. The model is not rejected for Canada, in sharp contrast to the literature, since the previous results might be seriously biased, due to the existence of global shocks.

JEL Classification: C31, E21, F32, F47.

Keywords: current account, capital mobility, error decomposition, common shocks.

[1] This article was jointly made with João Victor Issler. We would like to thank John C. Driscoll, Luiz Renato Lima, Fabiana Fontes Rocha, Maria Cristina Terra, and Carlos C. Gutierrez for insightful advices and helpful comments. We are also grateful to the participants of XXVII Brazilian Econometric Society Meeting (SBE, 2005), specially Sylvio Heck for helpful suggestions.

1 Introduction

The current account can be used by domestic residents to smooth consumption by borrowing from or lending to the rest of the world. Several authors have analyzed the open economy model, initially proposed by Sachs (1982) and later detailed by Obstfeld & Rogoff (1994), with a theoretical framework that defines the optimal current account from the agents' intertemporal optimization problem, supposing that agents can freely smooth consumption in the presence of shocks. The comparison of this optimal value with the observed current account allows us to test for consumption optimality.

This approach is encompassed by several classes of small open economy models,[2] and the most basic version is the present value model (PVM) of the current account. Although the literature of PVMs is relatively extensive, the following papers should be mentioned (suggesting an overall rejection of the model for developed countries): Sheffrin & Woo (1990) perform a study of the current account of Belgium, Canada, Denmark and UK. The results indicate a rejection of the model for Denmark, Canada and UK, whereas the PVM could not be rejected for Belgium. Otto (1992) tests the PVM for the USA and Canada, and rejects the model in both countries. Ghosh (1995) investigates the current account of 5 major industrialized countries: USA, Canada, Japan, Germany and UK, and the results suggest rejection of the model in all countries, except for the USA.

On the other hand, some papers document results supporting the PVM, in contrast to the previous findings, such as Ghosh & Ostry (1995) that test it for 45 developing countries and do not reject it for about 2/3 of the countries. Hussein & Mello (1999) also test the PVM for some developing countries (Chile, Greece, Ireland, Israel, Malaysia, Mexico, South Africa, South Korea and Venezuela), and find evidences to support the PVM. In the same line, Agénor et al. (1999) focus on the current account of France, concluding that the PVM holds and the analyzed country was perfectly able to smooth consumption.[3]

Notwithstanding the lack of consensus on the macroeconomic front, what happens on the "econometric side"? Is it possible that an inappropriate econometric technique leads to wrong conclusions regarding the rejection of the PVM? Unlike the mentioned literature, the objective of this paper is to provide an econometric approach to the current account debate. The methodology generally adopted in the literature to analyze the PVM was initially proposed by Campbell & Shiller (1987), and consists of estimating an unrestricted VAR,

[2] For recent developments regarding small open economy models see Grohé & Uribe (2003). In addition, see Chinn & Prasad (2003), which provide an empirical characterization of the determinants of current account for a large sample of industrial and developing countries. See also Aguiar & Gopinath (2006), which develop a quantitative model of debt and default in a small open economy. Finally, see Obstfeld & Rogoff (1996) and Bergin (2003) for a good discussion about new open economy literature and its empirical dimension.

[3] In order to deepen the debate, several authors also proposed extensions to the standard PVM model. A short list includes Ghosh & Ostry (1997), which consider precautionary saving, Gruber (2000) includes habit formation, Bergin & Sheffrin (2000) allow for a time-varying world interest rate and consider tradable and non-tradable goods, İşcan (2002) modifies the basic model introducing durables and also nontraded goods. More recently, Nason & Rogers (2006) propose a real business cycle (RBC) model, which nests the basic PVM, including non-separable preferences, shocks to fiscal policy and world interest rate, and imperfect capital mobility, explanations broadly presented in the literature for the rejection of the PVM. According to Nason & Rogers (2006), although each suspect matters in some way, none is capable to completely improve the fit of the model to the data.

whose parameters are used in the construction of the optimal current account, and perform a Wald test to investigate a set of restrictions imposed to the VAR, testing whether the optimal current account equals the observed series.

However, the presence of common shocks in the econometric model can play a crucial role, and is widely recommended in the literature to explain business cycles fluctuations. For instance, Centoni et al. (2003) investigate whether co-movements observed in the international business cycles are the consequences of common shocks or common transmission mechanisms. Similarly to most studies (such as King et al. 1991), Centoni et al. (2003) confirm that permanent shocks are the main source of the business cycles, accounting for a 50% effect in a panel of European countries. The authors also show that the domestic component is responsible for most of the business cycle effects of transitory shocks for all the G-7 countries, whereas the foreign component dominates the cyclical variability that is due to permanent shocks in France, Germany and Italy.[4]

This way, seems to exist a consensus in the literature regarding a common world component that might partially explain current account fluctuations. This common (or global) shock is ignored in the OLS estimation (widely used in the literature), but could be considered in a SUR approach. In fact, along this paper we stress the fact that in the estimation process an econometrician might consider a set of countries separately (OLS) as well as jointly (e.g., SUR), in order to capture contemporaneous correlations of the residuals of the VAR. However, due to the possible finite sample bias of the OLS and SUR covariance matrices (see Driscoll & Kraay, 1998), we also investigate the two-way error decomposition of Fuller & Battese (1974), hereafter FB, which can properly treat the existence of common shocks in the estimation process.

Therefore, we aim to contribute to the current account debate by investigating the estimation of a PVM through three different techniques (OLS, SUR and FB). In addition, we propose a quite original note on Granger causality, which is showed to be a necessary condition to perform the Wald test of Campbell & Shiller (1987). In addition, we present some theoretical results to show that (in the presence of common shocks) OLS and SUR estimators might produce a biased covariance matrix, with serious implications to the validation of the model.

A small Monte Carlo simulation confirms these findings and indicates the FB procedure in the presence of global shocks. We also provide an empirical exercise for the G-7 countries, and (indeed) the results substantially change with different estimation techniques. In addition, global shocks can account up to 40% of the total residuals of the G-7, confirming the importance of such shocks in the estimation process. The model is not rejected for Canada, in sharp contrast to the literature, since the previous results might be seriously biased, due to the existence of global shocks.

[4] In the same sense, Canova & Dellas (1993) document that after 1973 the presence of common disturbances, such as the first oil shock, plays a role in accounting for international output co-movements. Glick & Rogoff (1995) study the current account response to different productivity shocks in the G-7 countries, based on a structural model including global and country-specific shocks. Furthermore, Canova & Marrinan (1998), which investigate the generation and transmission of international cycles in a multicountry model with production and consumption interdependencies, argue that a common component to the shocks and of production interdependencies appear to be crucial in matching the data.

This paper is structured in the following way: Section 2 provides an overview of the macroeconomic model of the current account and discusses some econometric techniques that might be used in the estimation process. Section 3 presents the results of an empirical exercise for the G-7 countries, and Section 4 presents our main conclusions.

2 Methodology

2.1 Present Value Model

The Present Value Model (PVM) adopted to analyze the intertemporal optimization problem of a representative agent is based on Sachs (1982), considering the perfect capital mobility hypothesis across countries. In this context, countries save through flows of capital in their current accounts, according to their expectations of future changes in net output. Thus, the current account is used as an instrument of consumption smoothing against possible shocks to the economy, and can be expressed by

$$CA_t = B_{t+1} - B_t = Y_t + rB_t - I_t - G_t - C_t \tag{1}$$

where B_t represents foreign assets, Y_t gross domestic product (GDP), r the world interest rate, I_t total investment, G_t the government's expenses and C_t aggregated consumption.

The consumption path, related to the dynamics of the current account, can be divided into two components: the trend term, generated by the difference between the world interest rate and the rate of time preference, and the smoothing component, related to the expectations of changes in permanent income. This paper only studies the second component effect, by isolating from the current account, the trend component in consumption. Thus, the optimal current account (only associated with the consumption smoothing term) is given by

$$CA_t^* = Y_t + rB_t - I_t - G_t - \theta C_t \tag{2}$$

where θ is a parameter that removes the trend component in consumption.[5] The net output Z_t , also known in the literature as national cash flow, is defined by

$$Z_t \equiv Y_t - I_t - G_t \tag{3}$$

Substituting the optimal consumption expression in equation (2), it can be shown that the present value relationship between the current account and the future changes in net output is given by (see Ghosh & Ostry (1995) for further details):

$$CA_t^* = -\sum_{j=1}^{\infty}(\frac{1}{1+r})^j E_t(\Delta Z_{t+j} \mid R_t) \tag{4}$$

where R_t is the agent's information set. It should be mentioned that the main assumptions of the model are time-separable preferences, zero depreciation of capital, and complete asset markets. A quadratic form is also adopted for the utility function, without precautionary saving effects (see Ghosh & Ostry, 1997).

[5] The tilt parameter (θ) is not equal to one whenever the rate of time preference differs from the world interest rate.

According to equation (4), the optimal current account is equal to minus the present value of the expected changes in net output. For instance, the representative agent will increase its current account, accumulating foreign assets, if a future decrease in income is expected, and vice-versa.

2.2 Econometric Model

The econometric model is based on the methodology developed by Campbell & Shiller (1987), which suggest an alternative way to verify a PVM when the involved variables are stationary. The idea is to test a set of restrictions imposed to a Vector Auto Regression (VAR), used to forecast the current account through equation (4). The advantage of this approach is that, although the econometrician does not observe the agent's information set, this framework allows us to summarize all the relevant information through the variables used in the construction of the VAR.

However, to apply this methodology, the VAR must be stationary. Hence, the first empirical implication is to verify whether ΔZ_t is a weakly stationary variable. The current account (in level) must also be a stationary variable, since it can be written as a lineal combination of stationary variables (via equation (4)). The stationarity of these variables can be checked later by unit root tests. Campbell & Shiller (1987) argue that series represented by a Vector Error Correction Model (VECM) can be rewritten as an unrestricted VAR. Thus, consider the following VAR representation:[6]

$$\begin{bmatrix} \Delta Z_t^i \\ CA_t^i \end{bmatrix} = \begin{bmatrix} a^i(L) & b^i(L) \\ c^i(L) & d^i(L) \end{bmatrix} \times \begin{bmatrix} \Delta Z_{t-1}^i \\ CA_{t-1}^i \end{bmatrix} + \begin{bmatrix} \mu_1^{i*} \\ \mu_2^{i*} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t}^i \\ \varepsilon_{2t}^i \end{bmatrix} \tag{5}$$

where the index i represents the analyzed country and $a^i(L)$, $b^i(L)$, $c^i(L)$ and $d^i(L)$ are polynomials of order p. Hence, the estimation of the VAR must be preceded by the estimation of θ, which occurs in the cointegration analysis between C_t and $(Y_t + rB_t - I_t - G_t)$. The model VAR(p) can be described as a VAR(1), in the following way:

$$\begin{bmatrix} \Delta Z_t^i \\ \vdots \\ \vdots \\ \Delta Z_{t-p+1}^i \\ CA_t^i \\ \vdots \\ \vdots \\ CA_{t-p+1}^i \end{bmatrix} = \begin{bmatrix} a_1^i & \cdots & a_p^i & b_1^i & \cdots & b_p^i \\ 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ c_1^i & \cdots & c_p^i & d_1^i & \cdots & d_p^i \\ & & & 1 & & \\ & 0 & & & \ddots & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} \Delta Z_{t-1}^i \\ \vdots \\ \vdots \\ \Delta Z_{t-p}^i \\ CA_{t-1}^i \\ \vdots \\ \vdots \\ CA_{t-p}^i \end{bmatrix} + \begin{bmatrix} \mu_1^{i*} \\ 0 \\ \vdots \\ 0 \\ \mu_2^{i*} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t}^i \\ 0 \\ \vdots \\ 0 \\ \varepsilon_{2t}^i \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{6}$$

or, in a compact form:

$$X_t = AX_{t-1} + \mu^* + \varepsilon_t \tag{7}$$

[6] It should be mentioned that, hereafter, CA_t will be constructed considering the parameter θ, to remove the trend component in consumption, as it follows: $CA_t = Y_t + rB_t - I_t - G_t - \theta C_t$.

where $X_t \equiv \begin{bmatrix} \Delta Z_t^i & \cdots & \Delta Z_{t-p+1}^i & CA_t^i & \cdots & CA_{t-p+1}^i \end{bmatrix}'$, A is the companion matrix, μ^* represents a vector of intercepts, and ε_t is a vector that contains the residuals. The VAR(1) is stationary by assumption, and the equation (7) can be rewritten removing the vector of means μ:

$$(X_t - \mu) = A(X_{t-1} - \mu) + \varepsilon_t \tag{8}$$

where $\mu^* = (I - A)\mu$. The forecast of the model j periods ahead is given by

$$E[(X_{t+j} - \mu \mid H_t) = A^j (X_t - \mu) \tag{9}$$

where H_t is the econometrician's information set (composed of current and past values of CA and ΔZ), contained in the agent's information set R_t. Define h' as a vector with $2p$ null elements, except the first: $h' = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix}$. Then, one can select ΔZ_t in the vector X_t, in the following way:

$$\Delta Z_t = h' X_t \therefore \Delta Z_{t+j} = h' X_{t+j} \therefore (\Delta Z_{t+j} - \mu_{\Delta Z}) = h'(X_{t+j} - \mu) \tag{10}$$

where the vector μ contains the means $\mu_{\Delta Z}$ and μ_{CA^*}. Thus, applying the conditional expectation in the previous expression, it follows that:

$$E[(\Delta Z_{t+j} - \mu_{\Delta Z}) \mid H_t] = E[h'(X_{t+j} - \mu) \mid H_t] = h' E[(X_{t+j} - \mu) \mid H_t] = h' A^j (X_t - \mu) \tag{11}$$

where the last equality comes from equation (9). In order to calculate the optimal current account CA_t^*, one can take expectations of equation (4):

$$E(CA_t^* \mid H_t) = CA_t^* = -\sum_{j=1}^{\infty} (\frac{1}{1+r})^j E(\Delta Z_{t+j} \mid H_t) \tag{12}$$

The first equality comes from the fact that CA_t^* is contained in H_t, and the second is given by the law of iterated expectations ($H_t \subseteq R_t$). Applying the unconditional expectation in the previous expression:

$$E(CA_t^*) = -\sum_{j=1}^{\infty} (\frac{1}{1+r})^j E(\Delta Z_{t+j}) \therefore \mu_{CA^*} = -\sum_{j=1}^{\infty} (\frac{1}{1+r})^j \mu_{\Delta Z} \tag{13}$$

Combining equation (12) with equation (13), it follows that:

$$(CA_t^* - \mu_{CA^*}) = -\sum_{j=1}^{\infty} (\frac{1}{1+r})^j E(\Delta Z_{t+j} - \mu_{\Delta Z} \mid H_t) \tag{14}$$

Applying the expression (11) in the equation above:

$$(CA_t^* - \mu_{CA^*}) = -\sum_{j=1}^{\infty} (\frac{1}{1+r})^j h' A^j (X_t - \mu) = -h'(\frac{A}{1+r})(I - \frac{A}{1+r})^{-1} (X_t - \mu) \tag{15}$$

where the last equality is due to the convergence of an infinite sum, since the variables ΔZ_t and CA_t are stationary. Rewriting the previous equation in a simplified form:

$$(CA_t^* - \mu_{CA^*}) = K(X_t - \mu) \tag{16}$$

$$K = -h'(\frac{A}{1+r})(I - \frac{A}{1+r})^{-1} \tag{17}$$

where the vector K is derived from the world interest rate r and the matrix A. To formally test the model, one can analyze the null hypothesis $(CA_t^* - \mu_{CA^*}) = (CA_t - \mu_{CA})$. Define g' as a vector with $2p$ null elements, except the $(p+1)th$ element, that assumes a unit value. Thus, under the null hypothesis, it follows that:

$$(CA_t^* - \mu_{CA^*}) = (CA_t - \mu_{CA}) = g'(X_t - \mu) \tag{18}$$

Combining equations (16) and (18), the model can be formally tested through a set of restrictions imposed to the coefficients of the VAR:

$$g'(X_t - \mu) = -h'(\frac{A}{1+r})(I - \frac{A}{1+r})^{-1}(X_t - \mu) \therefore g'(I - \frac{A}{1+r}) = -h'(\frac{A}{1+r}) \tag{19}$$

Applying the structure of matrix A into equation (19), the following restrictions[7] can be derived:

$$\begin{aligned} a_i &= c_i \;\; ; i = 1...p \\ b_i &= d_i \;\; ; i = 2...p \\ b_1 &= d_1 - (1+r) \end{aligned} \tag{20}$$

Another important implication of the model is that the current account Granger-cause changes in net output, or in other words, CA_t helps to forecast ΔZ_t. This causality can be tested by means of the statistical significance of the $b(L)$ coefficients. Therefore, the implications of the intertemporal optimization model, according to Otto (1992), can be summarized by:[8]

1. Verifying the stationarity of CA_t and ΔZ_t, through unit root tests;
2. Checking if CA_t Granger-cause ΔZ_t;
3. Analyzing the cointegration between C_t and $(Y_t + rB_t - I_t - G_t)$, and calculating the parameter θ;
4. Formally investigating, by means of a Wald test, the equality of the optimal and observed current accounts, given by restrictions (20).

2.3 A note on Granger Causality and Wald Tests

The optimal current account is generated from the vector K (see expressions (16) and (17)), which depends on matrix A and the world interest rate r. However, it should be noted that an estimated coefficient for matrix A could not be statistically significant. These results could seriously compromise the subsequent optimal current account analysis, as it follows.

The Granger causality between the current account and net output (CA_t Granger-cause ΔZ_t) is a primordial implication of the theoretical model, and as argued before, can be alternatively tested through the significance of the $b(L)$ coefficients.[9] Moreover, if this implication is not empirically observed, the model

[7] These restrictions can be verified by a Wald test.

[8] It is in fact a set of testable implications of the PVM. Therefore, the statistical acceptance of the model occurs only if all of these implications could be verified.

[9] Presented in equation (5).

should be rejected irrespective of any other results, since equation (4) is the theoretical foundation of the whole study. In this case, the current account could not help to predict variations in net output, suggesting that the agents are badly described by the model. Thus, one should not construct the optimal current account and perform a comparison with the observed series. Unfortunately, this is done in several papers presented in the literature.

To study this topic more carefully, a simple VAR(1) is initially presented. The Granger causality between CA_t and ΔZ_t, in this case, will be determined by the statistical significance of the b_1 coefficient.

$$\begin{bmatrix} \Delta Z_t \\ CA_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \times \begin{bmatrix} \Delta Z_{t-1} \\ CA_{t-1} \end{bmatrix} + \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}. \tag{21}$$

In this case, the VAR is represented in a compact form by $X_t = AX_{t-1} + \mu + \varepsilon_t$ and, after some algebraic manipulations, the vector K takes the form:

$$K = -h'(\frac{A}{1+r})(I_2 - \frac{A}{1+r})^{-1} = \begin{bmatrix} \alpha & \beta \end{bmatrix}, \tag{22}$$

where

$$\alpha = \frac{-a_1(1+r-d_1) - b_1c_1}{(1+2r-d_1+r^2-rd_1-a_1-a_1r+a_1d_1-b_1c_1)}, \tag{23}$$

$$\beta = \frac{-a_1b_1 - b_1(1+r-a_1)}{(1+2r-d_1+r^2-rd_1-a_1-a_1r+a_1d_1-b_1c_1)}. \tag{24}$$

If the Granger causality is rejected by the data (e.g., b_1 is not significant), then equation (25) indicates that $\beta = 0$, or in other words, CA_t^* is not a function of CA_t. In this case, the optimal current account would be given by

$$(CA_t^* - \mu_{CA^*}) = K(X_t - \mu) = \begin{bmatrix} \alpha & 0 \end{bmatrix} \begin{bmatrix} \Delta Z_t - \mu_{\Delta Z} \\ CA_t - \mu_{CA} \end{bmatrix} = \alpha\,(\Delta Z_t - \mu_{\Delta Z})\,. \tag{25}$$

Hence, if $\beta = 0$ the null hypothesis $(CA_t^* - \mu_{CA}^*) = (CA_t - \mu_{CA})$ is always rejected, since under Ho β should be equal to one (and α should be zero). A further analysis of the vector K for a VAR(2) is presented in appendix, in a similar way. The generalization of this cautionary note for a VAR(p) is straightforward, and can be summarized by Proposition 1. According to Hamilton (1994), in the context of a bivariate VAR(p), if one of the two variables does not Granger-cause the other, then the companion matrix is lower triangular (e.g., $b(L) = 0$). Thus, the β_i coefficients of the vector K $(i = 1, \ldots, p)$ are always zero, because of the algebraic structure of the vector, as also detailed in appendix.

Proposition 1 *Consider the VAR representation (5) of the intertemporal model of current account. The Granger causality from the current account (CA_t) to the first difference of the net output (ΔZ_t) is a necessary condition to perform the Wald test and verify the validation of the model, i.e., if the $b(L)$ coefficients of the VAR(p) model are not statistically significant, then, the Wald test is not applicable and the model should be rejected.*

Proof. See Appendix. ∎

Therefore, if the Granger causality could not be confirmed by the data set, neither a Wald test should be performed nor the optimal current account should be generated, since the basic assumption of the model is not verified,[10] as summarized in table 1.

Table 1 - A note on Granger causality and Wald tests

Result of Granger causality	Wald test	Model	Conclusion
CA_t not Granger-cause ΔZ_t	not applicable (β=0)	rejected	model cannot generate CA_t^*(*)
CA_t Granger-cause ΔZ_t	rejects Ho ($\beta\neq$1)	rejected	$CA_t^* \neq CA_t$ (**)
	does not reject Ho (β=1)	not rejected	$CA_t^* = CA_t$ (***)

Notes: (*) indicates that CA_t^* only depends on ΔZ, instead of CA_t

(**) suggests that agents do not smooth consumption;

(***) means that agents perfectly smooth consumption.

2.4 Estimation Method

2.4.1 SUR estimation

The VAR model (5) is usually estimated in the literature, equation-by-equation, using OLS. However, the Seemingly Unrelated Regressions (SUR) technique, originally developed by Zellner (1962), can also be adopted, since it is based on a Generalized Least Squares (GLS) estimation applied to a system of equations as a whole, in which the data for several countries is examined simultaneously. The joint estimation is given by stacking the system of equations that compose the VAR (for each country $i = 1, \ldots, N$) in the following way:

$$\begin{bmatrix} \Delta Z_t^i \\ CA_t^i \end{bmatrix} = \begin{bmatrix} a^i(L) & b^i(L) \\ c^i(L) & d^i(L) \end{bmatrix} \times \begin{bmatrix} \Delta Z_{t-1}^i \\ CA_{t-1}^i \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t}^i \\ \varepsilon_{2t}^i \end{bmatrix} \tag{26}$$

Then, define $Y_t^i = \begin{pmatrix} \Delta Z_t^i \\ CA_t^i \end{pmatrix}_{(2\times1)}$, $X_t^i = \begin{pmatrix} \Delta Z_{t-1}^i \\ \vdots \\ \Delta Z_{t-p_i}^i \\ CA_{t-1}^i \\ \vdots \\ CA_{t-p_i}^i \end{pmatrix}_{(2p_i\times1)}$, $\varepsilon_t^i = \begin{pmatrix} \varepsilon_{1t}^i \\ \varepsilon_{2t}^i \end{pmatrix}_{(2\times1)}$

and $\beta_i' = \begin{pmatrix} a_1^i & \ldots & a_{p_i}^i & b_1^i & \ldots & b_{p_i}^i & c_1^i & \ldots & c_{p_i}^i & d_1^i & \ldots & d_{p_i}^i \end{pmatrix}_{(1\times4p_i)}$

[10] Recall Ghosh & Ostry (1995) results, in which the authors test the PVM for 45 developing countries and do not reject it for 29 countries. However, a careful analysis of the tests reveals that only 25 countries (from the entire set of countries) in fact support the Granger causality implication (at 5% level). This way, the paper should conclude that (at most) in only 18 countries (instead of 29) the model could not be rejected, since only 18 countries indeed exhibit good results for both the Wald and Granger causality tests.

This way, the VAR for a country i can be represented by

$$\left(Y_t^i\right)_{(2 \text{ x } 1)} = \begin{pmatrix} X_t^{1\prime} & 0 \\ 0 & X_t^{1\prime} \end{pmatrix}_{(2 \text{ x } 4\text{p}_i)} \begin{pmatrix} \beta_i \end{pmatrix}_{(4\text{p}_i \text{ x } 1)} + \begin{pmatrix} \varepsilon_t^i \end{pmatrix}_{(2 \text{ x } 1)} \tag{27}$$

Furthermore, the system of equations for a given set of N countries can be expressed by

$$\begin{pmatrix} Y_t^1 \\ \vdots \\ Y_t^N \end{pmatrix}_{(2\text{N x } 1)} = \begin{pmatrix} X_t^{1\prime} & 0 & \dots & & 0 \\ 0 & X_t^{1\prime} & & & \\ \vdots & & \ddots & 0 & \vdots \\ & & 0 & X_t^{N\prime} & 0 \\ 0 & & \dots & 0 & X_t^{N\prime} \end{pmatrix}_{(2\text{N x } 4\text{P})} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_N \end{pmatrix}_{(4\text{P x } 1)} + \begin{pmatrix} \varepsilon_t^1 \\ \vdots \\ \varepsilon_t^N \end{pmatrix}_{(2\text{N x } 1)} \tag{28}$$

or in a compact form $Y = X\beta + \varepsilon$. The name SUR comes from the fact that each equation in the previous system has its own vector of coefficients, which might suggest that the equations are unrelated. Nevertheless, correlation across the errors in different equations can provide links that can be exploited in estimation. It should be noted that $\sum_{i=1}^{N} p_i = P$, where p_i is the number of lags of the VAR, for a country i. The residuals ε have mean zero and are serially uncorrelated, with covariance matrix given by $E(\varepsilon\varepsilon') = \sigma^2\Omega$. Hence, the GLS estimator of β and its variance-covariance matrix are given by

$$\widetilde{\beta} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y \tag{29}$$

$$E(\widetilde{\beta} - \beta)(\widetilde{\beta} - \beta)' = \sigma^2 \left(X'\Omega^{-1}X\right)^{-1} \tag{30}$$

In general, the $(N \times N)$ matrix Ω is unknown and the last expression cannot be directly applied. However, $\widetilde{\beta}$ can be calculated by an estimate of the ijth element of Ω, given by

$$\widehat{w}_{ij} = \frac{e_i' e_j}{T}, \quad \text{where } i, j = 1, \dots, N \tag{31}$$

where e_i is a $(T \times 1)$ vector containing the residuals of the ith equation estimated by OLS. In this case, a feasible SUR estimator of β is obtained as

$$\widetilde{\beta}^* = \left(X'\widehat{\Omega}^{-1}X\right)^{-1} X'\widehat{\Omega}^{-1}Y \tag{32}$$

$$E(\widetilde{\beta}^* - \beta)(\widetilde{\beta}^* - \beta)' = \sigma^2 \left(X'\widehat{\Omega}^{-1}X\right)^{-1} \tag{33}$$

The OLS estimator, on the other hand, is given by

$$\widehat{\beta} = \left(X'X\right)^{-1} X'Y \tag{34}$$

$$E(\widehat{\beta} - \beta)(\widehat{\beta} - \beta)' = \sigma^2 \left(X'X\right)^{-1} X'\Omega X \left(X'X\right)^{-1} \tag{35}$$

The difference between their variance-covariance matrices is a positive semidefinite matrix, and can be expressed by

$$E(\widehat{\beta} - \beta)(\widehat{\beta} - \beta)' - E(\widetilde{\beta} - \beta)(\widetilde{\beta} - \beta)' = \sigma^2 \eta \Omega \eta' \tag{36}$$

where $\eta = \left(X'X\right)^{-1} X' - \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}$, indicating the gain in efficiency of SUR estimators in comparison to the OLS counterpart.

2.4.2 Caveats of SUR estimation

The SUR estimator generally exhibits a good performance when N is small relative to T, but in fact becomes not feasible when $T < (N+1)/2$. Even when the SUR model is correctly specified, its performance might be poor due to a very large number of free parameters to be estimated, in comparison to the time dimension. In other words, as N becomes large for a fixed value of T, the estimated covariance matrix becomes "nearly" singular, introducing a bias into the standard error estimates. This way, the finite sample performance of the OLS and SUR estimators deteriorates rapidly as the size of the cross-sectional dimension increases.

Driscoll & Kraay (1998) investigate finite-sample properties of variance estimators, concluding that both OLS and SUR estimators indeed exhibit substantial downward finite sample bias, even for moderate values of cross-sectional dependence, and are outperformed by a spatial correlation consistent estimator proposed in their article, based on the nonparametric technique of Newey & West (1987) and Andrews (1991).

The main idea is to obtain consistent estimates of the $N \times N$ matrix of cross-sectional correlations by averaging over the time dimension. This way, the estimated cross-sectional covariance matrix can be used to construct standard errors, which are robust to the presence of spatial correlation. Driscoll & Kraay (1998)'s approach, in contrast to SUR, might be applicable in situations such as cross-country panel data models with a relatively large number of countries.

In this paper, however, we focus on a panel model with small N and large T, but in order to deal with possible finite sample bias of the covariance matrix, we also investigate a two-way error decomposition (next described) that can properly deal with cross-country correlations.

2.4.3 Fuller & Battese (1974) and the two-way error decomposition

The performance of any estimation procedure depends on the statistical characteristics of the error components in the model. In this section, we adopt the Fuller & Battese (1974) method to consider individual and time-specific random effects into the error disturbances, in which parameters can efficiently be estimated by using a feasible GLS framework.

In dynamic panel models, the presence of lagged dependent variables might lead to a non-zero correlation between regressors and error term. This could render OLS estimator for a dynamic error-component model to be biased and inconsistent (see Baltagi, 2001, p. 130), due to the correlation between the lagged dependent variable and the individual specific effect. In addition, a feasible GLS estimator for the random-effects model under the assumption of independence between the effects and explanatory variables would also be biased. In these cases (with large N and short T), Andersen & Hsiao (1981) suggests first differencing the model to get rid of the individual effect. On a different approach, but still in a framework of dynamic models with large N and short T, Holtz-Eakin et alli (1988) investigate panel VAR (PVAR) models, in order to provide more flexibility to the VAR modeling for panel data. See also Hsiao (2003, p. 70,107) for further details.

In this paper, due to the specific structure of our VAR, we take a different route. Since we are interested here in weakly stationary variables, in a random-effects model with short N and large T, we apply the Fuller

& Battese (1974) approach to our system of equations (28). These authors establish sufficient conditions for a feasible GLS estimator to be unbiased and exhibit the same asymptotic properties of the GLS estimator in a crossed-error model, i.e., in which an error decomposition is considered to allow for individual effects that are constant over cross sections or time periods. To do so, initially consider the stacked model $Y = X\beta + \varepsilon$, from the system of equations (28), where $Y = (y_{1,1}; y_{1,2}; ...; y_{1,T}; ...; y_{2N,T})$; $X = (x_{1,1}; x_{1,2}; ...; x_{1,T}; ...; x_{2N,T})$; β and $x_{i,t}$ are $p \times 1$ vectors. The Fuller & Battese (1974) two-way random error decomposition is given by $\varepsilon_{i,t} = v_i + e_t + \epsilon_{i,t}$, in which $E(\varepsilon\varepsilon' \mid X) \equiv \Omega$.

Thus, the model is a variance components model, with the variance components $\sigma^2_\epsilon; \sigma^2_v; \sigma^2_e$ to be estimated. A crucial implication of such a specification is that the effects are not correlated with the regressors. For random effects models, the estimation method is a feasible generalized least squares (FGLS) procedure that involves estimating the variance components in the first stage and using the estimated variance covariance matrix thus obtained to apply generalized least squares (GLS) to the data. It is also assumed that $E(v_i) = 0; E(v_i^2) = \sigma^2_v$; $E(v_i v_j) = 0, \forall i \neq j; v_i$ is uncorrelated with $\epsilon_{i,t}, \forall i, t$; and also $E(e_t) = 0; E(e_t^2) = \sigma^2_e$; $E(e_t e_s) = 0, \forall t \neq s; e_t$ is uncorrelated with v_i and $\epsilon_{i,t}, \forall i, t$.[11]

Contrary to Wallace & Hussain (1969) or Swamy & Arora (1972), Fuller & Battese (1974) also consider the case in which σ^2_v and/or σ^2_e are equal to zero.[12] The estimators for the variance components are obtained by the fitting-of-constants method, with the provision that any negative variance components is set to zero for parameter estimation purposes. First, the least square residuals are defined by: $\widehat{\varepsilon} = M_{12}(I - X[X'M_{12}X]^{-1}X'M_{12}]Y; \widehat{v} = (M_{12} + M_{1.})(I - X[X'(M_{12} + M_{1.})X]^{-1}X'(M_{12} + M_{1.})]Y; \widehat{e} = (M_{12} + M_{.2})(I - X[X'(M_{12} + M_{.2})X]^{-1}X'(M_{12} + M_{.2})]Y$. Next, Fuller & Battese compute the unbiased estimators for the variance components: $\widehat{\sigma}^2_\epsilon = \frac{\widehat{\varepsilon}\widehat{\varepsilon}'}{(N-1)(T-1)-\delta_1}$; $\widehat{\sigma}^2_v = \frac{\widehat{v}\widehat{v}' - [T(N-1)-\delta_2]\widehat{\sigma}^2_\epsilon}{T(N-1)-T\phi_1}$; $\widehat{\sigma}^2_e = \frac{\widehat{e}\widehat{e}' - [N(T-1)-\delta_3]\widehat{\sigma}^2_\epsilon}{N(T-1)-N\phi_2}$ where $\delta_1 \equiv rank(X'M_{12}X)$; $\delta_2 \equiv rank(X'M_{.2}X)$; $\delta_3 \equiv rank(X'M_{1.}X)$; $\phi_1 \equiv tr\{[X'(M_{12} + M_{1.})X]^{-1}X'M_{1.}X\}$; $\phi_2 \equiv tr\{[X'(M_{12} + M_{.2})X]^{-1}X'M_{.2}X\}$.

Once the component variances have been estimated, we form an estimator of the composite residual covariance, and then GLS transform the dependent and regressor data. The respective GLS estimator is given by $\widehat{\beta}_{FB} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$ and, thus, the FB estimator is the related feasible GLS estimator $\widehat{\widehat{\beta}}_{FB}$, in which Ω is estimated through $\widehat{\sigma}^2_\epsilon, \widehat{\sigma}^2_v$, and $\widehat{\sigma}^2_e$. Fuller & Battese (1974) show that their estimator is consistent,

[11] The authors also define the following mutually orthogonal, symmetric and idempotent matrices $M_{..} = \frac{J_{2NT}}{2NT}$; $M_{1.} = \frac{I_{2N} \otimes J_T}{T} - M_{..}$; $M_{.2} = \frac{J_{2N} \otimes I_T}{2N} - M_{..}$; $M_{12} = I_{2NT} - \frac{I_{2N} \otimes J_T}{T} - \frac{J_{2N} \otimes I_T}{2N} + M_{..}$; where I_{2N} and I_T are identity matrices of order $2NT$ and T, respectively; and J_{2N} and J_T are $(2N \times 2N)$ and $(T \times T)$ matrices having all elements equal to one. The covariance matrix $\Omega = \sigma^2_\epsilon I_{2NT} + \sigma^2_v(I_{2N} \otimes J_T) + \sigma^2_e(J_{2N} \otimes I_T)$ can be expressed by $\Omega \equiv \sigma^2_\epsilon M_{12} + (\sigma^2_\epsilon + T\sigma^2_v)M_{1.} + (\sigma^2_\epsilon + 2N\sigma^2_e)M_{.2} + (\sigma^2_\epsilon + T\sigma^2_v + 2N\sigma^2_e)M_{..}$, or even, $\Omega = \gamma_1 M_{12} + \gamma_2 M_{1.} + \gamma_3 M_{.2} + \gamma_4 M_{..}$, where $\gamma_1 \equiv \sigma^2_\epsilon$; $\gamma_2 \equiv (\sigma^2_\epsilon + T\sigma^2_v)$; $\gamma_3 \equiv (\sigma^2_\epsilon + 2N\sigma^2_e)$; $\gamma_4 \equiv (\sigma^2_\epsilon + T\sigma^2_v + 2N\sigma^2_e)$.

[12] Baltagi (1981) performed a Monte Carlo study on a single regression equation with two-way error component disturbances and studied the properties of several estimators, including OLS and six feasible GLS estimators: Fuller & Battese (1974), Swamy-Arora (1972), Wallace & Hussain (1969), among others. The results suggest that OLS standard errors are biased and all FGLS are asymptotically efficient and performed relatively well in finite samples, making it difficult to choose among them. The methods differ only in the specifications estimated in evaluating the residuals: The Swamy-Arora estimator of the component variances uses residuals from the within (fixed effect) and between (means) regressions, while the Wallace-Hussain estimator uses only OLS residuals. In general, they provide similar answers, especially in large samples. Additional details on random effects models are provided in Baltagi (2001).

unbiased and asymptotically equivalent to the GLS estimator. In addition, the estimated covariance matrix of coefficients is unbiased, since it is based on unbiased and consistent estimators $\widehat{\sigma}_{\epsilon}^{2}, \widehat{\sigma}_{v}^{2}, \widehat{\sigma}_{e}^{2}$.

Table 2 - Comparison of OLS, SUR and FB covariance matrix of residuals

$E(\varepsilon_{i,t}\varepsilon_{j,s} \mid X)$	OLS	SUR	FB
$i = j; t = s$	σ_i^2	σ_i^2	$(\sigma_v^2 + \sigma_e^2 + \sigma_\epsilon^2)$
$i \neq j; t = s$	0	$\sigma_{i,j}^2$	σ_e^2
$i = j; t \neq s$	0	0	σ_v^2
$i \neq j; t \neq s$	0	0	0

In the SUR approach, the covariance structure allows for conditional correlation between the contemporaneous residuals for cross-section, but restricts residuals in different periods to be uncorrelated. On the other hand, following the argument of Wooldridge (2002, p. 259), rather than depending on $N(N+1)/2$ variances and covariances, as would be the case in a SUR analysis, Ω of the Fuller & Battese (1974) approach only depends on three parameters, $\sigma_{\epsilon}^2, \sigma_v^2, \sigma_e^2$, regardless of the size of N. This parsimonious feature might be useful for a large panel model, with $N, T \to \infty$. See Baltagi (1980), which investigates a SUR model with error components, and also Pesaran and Smith (1995) and Phillips and Moon (1999) for panel data with large T and N. [13] We next show some important results of the OLS and SUR estimators for the system of equations (28), under the Fuller & Battese error decomposition.

Definition 3: Define the bias on the estimated covariance matrix by: $B \equiv Var(\beta) - E(\widehat{Var}(\beta) \mid X)$;

Definition 4: Define $\Psi \equiv 2N\sigma_e^2/(\sigma_\epsilon^2 + T\sigma_v^2 + 2N\sigma_e^2)$;

Assumption A1: $(X'X)$ and $(X'JX)$ are positive definite matrices, where $J \equiv J_{2N} \otimes I_T$;

Assumption A2: (i) $tr[(X'X)^{-1}X'JX - I] > 0$; and (ii) $T > 2p$, where p is the number of lags of the VAR[14];

Proposition 2 *(OLS) Assume the Fuller & Battese (1974) two-way random error decomposition. (i) If* $(\sigma_\epsilon^2; \sigma_e^2; \sigma_v^2) > 0$ *, then,* $\widehat{\beta}_{OLS}$ *is inconsistent; (ii) if* $\sigma_v^2 = 0$ *and* $(\sigma_\epsilon^2; \sigma_e^2) > 0$*, then,* $\widehat{Var}(\beta_{OLS})$ *is biased; (iii) if* $\sigma_v^2 = 0$*,* $(\sigma_\epsilon^2; \sigma_e^2) > 0$*, and A1-A2 hold, then,* $diag(\frac{\partial B_{OLS}}{\partial \sigma_e^2}) > 0$*, i.e., an increase of the common shock* σ_e^2 *(and, thus,* Ψ*) induces an upward bias on all estimated (OLS) variances; and (iv) if* $(\sigma_v^2; \sigma_e^2) = 0$ *and* $\sigma_\epsilon^2 > 0$ *, then,* $\widehat{Var}(\beta_{OLS})$ *is unbiased.*

Proof. See Appendix. ■

As already expected, substituting OLS residuals instead of the true disturbances introduces bias in the corresponding estimates of the variance components and, thus, on the covariance matrix of coefficients. See

[13] For panels with large N and T, several approaches might be considered: (i) sequential limits, in which a sequential limit theory is considered; diagonal-path limits, which allows the two indexes to pass to infinity along a specific diagonal path in the two dimensional array; or joint limits, in which both indexes pass to infinity simultaneously. See Hsiao (2003, p.295) for further details, and also Phillips & Moon (1999), which provide sufficient conditions that ensures the sequential limits to be equivalent to joint limits.

[14] Recall from (28) that $\sum_{i=1}^{N} p_i = P$, where p_i is the number of lags of the VAR for a given country i. By assuming that $p_i = p$, $\forall i$, then it follows that $k \equiv 4P = 4Np$, and thus $T > 2p$ means $2NT > 4Np = k$.

Maddala (1971) and Baltagi (2001, p.35) for further details. On the other hand, if assumptions A1-A2 do not hold, then, we cannot guarantee that all estimated variances are upwarded biased. It could be the case that some variances do exhibit a positive bias, whereas others are not affected at all, or even show a negative bias.[15]

Assumption A3: $\sigma_\epsilon^2 I_{2N} + \sigma_e^2 J_{2N} - \widehat{\Sigma} > 0$;

Assumption A4: $\sigma_\epsilon^2 I_{2N} + \sigma_e^2 J_{2N} = \widehat{\Sigma}$;

Assumption A5: $\widehat{\Sigma}$ is a positive semidefinite matrix; and $\Psi \geqslant \frac{2N(2NT-k)}{(tr((X'X)^{-1}X'JX)-2NT)+2N(2NT-k)}$; where $k = 4Np$ and $J \equiv J_{2N} \otimes I_T$;

Proposition 3 *(SUR) Assume the Fuller & Battese (1974) two-way random error decomposition. (i) If* $(\sigma_\epsilon^2;\sigma_e^2;\sigma_v^2) > 0$ *, then,* $\widehat{\beta}_{SUR}$ *is inconsistent; (ii) if* $\sigma_v^2 = 0$, $(\sigma_\epsilon^2;\sigma_e^2) > 0$ *and A3 hold, then,* $\widehat{Var}(\beta_{SUR})$ *is biased. In addition, if A2(ii) and A5 also hold, then,* $(B_{OLS} - B_{SUR}) \geqslant 0$*, i.e., the bias of SUR estimated variances of* β *is not greater than the respective OLS bias; (iii) if* $\sigma_v^2 = 0$, $(\sigma_\epsilon^2;\sigma_e^2) > 0$ *and A4 hold, then,* $\widehat{Var}(\beta_{SUR})$ *is unbiased; and (iv) if* $(\sigma_v^2;\sigma_e^2) = 0$ *and A4 hold, then,* $\widehat{Var}(\beta_{SUR}) = \widehat{Var}(\beta_{OLS})$ *is unbiased.*

Proof. See Appendix. ■

Note that as long as $tr(X(X'X)^{-1}X'J)$ increases, the exigency for Ψ decreases, in order to guarantee that $(B_{OLS} - B_{SUR}) \geqslant 0$. In other words, if the common shock is relatively significant in the disturbance term, then, the SUR technique might produce a less biased covariance matrix in respect to the OLS approach. Now, we present sufficient conditions for the Fuller & Battese (1974) estimator to be unbiased and consistent when applied to our setup. Initially, lets define $\widehat{\beta}_{FB}$ as the unfeasible Fuller & Battese (1974) estimator, and $\widehat{\widehat{\beta}}_{FB}$ as the feasible FB estimator, based on the estimated effects $\widehat{\sigma}_v^2;\widehat{\sigma}_e^2$ and $\widehat{\sigma}_\epsilon^2$.

Assumption A6: $X'\Omega^{-1}X$ is nonsingular, and $plim(\frac{X'\Omega^{-1}X}{T}) = Q_*$, when $T \to \infty$, with fixed N; where Q_* is a finite positive definite matrix;

Assumption A7: e_t and $\epsilon_{i,t}$ are independent and normally distributed;

Assumption A8: $plim[(\frac{X'\widehat{\Omega}^{-1}X}{T}) - (\frac{X'\Omega^{-1}X}{T})] = 0$ and $plim[(\frac{X'\widehat{\Omega}^{-1}\varepsilon}{\sqrt{T}}) - (\frac{X'\Omega^{-1}\varepsilon}{\sqrt{T}})] = 0$; where $\widehat{\Omega} = \widehat{\sigma}_\epsilon^2 I_{2NT} + \widehat{\sigma}_v^2(I_{2N} \otimes J_T) + \widehat{\sigma}_e^2(J_{2N} \otimes I_T)$ is the estimated Fuller & Battese (1974) covariance matrix.

Proposition 4 *Assume the Fuller & Battese (1974) two-way random error decomposition. Thus, it follows that: (i) if* $\sigma_v^2 = 0$, $(\sigma_\epsilon^2;\sigma_e^2) > 0$, *and A6 holds, then, (i) the FB estimator* $\widehat{\beta}_{FB}$ *is unbiased and consistent; (ii) if* $\sigma_v^2 = 0$, $(\sigma_\epsilon^2;\sigma_e^2) > 0$, *and A6-A7 hold, then, the FB estimator* $\widehat{\beta}_{FB}$ *is asymptotically normally distributed, i.e.,* $\widehat{\beta}_{FB} \sim N(\beta;(X'\Omega^{-1}X)^{-1})$*; and (iii) if A6, A7 and A8 hold, then, the feasible FB estimator* $\widehat{\widehat{\beta}}_{FB}$ *is asymptotically equivalent to* $\widehat{\beta}_{FB}$.

[15] Note that if $X = [1,...,1]'$, then, it follows that $(X'X)^{-1} = 1/2NT$ and $tr[XX'J] = tr[(J_{2NT})(J_{2N} \otimes I_T)] = tr[(J_{2N} \otimes J_T)(J_{2N} \otimes I_T)] = tr[(J_{2N}J_{2N}) \otimes (J_T I_T)] = tr(J_{2N}J_{2N})tr(J_T I_T) = Ttr(J_{2N}J_{2N}) = T(2N)(2N)$. Thus, $tr[(X'X)^{-1}X'JX - I] = tr[(X'X)^{-1}X'JX] - tr(I) = tr[X'JX]/2NT - 2NT = tr[XX'J]/2NT - 2NT = 2N(2NT)/2NT - 2NT = 2N(1-T) < 0$. In other words, if A1-A2 do not hold we cannot guarantee that $diag(\frac{\partial B}{\partial \sigma_e^2}) > 0$.

Proof. See Appendix. ■

Note that if one assumes that the residuals of each country are serially uncorrelated, then, $E(\varepsilon_{i,t}\varepsilon_{i,s} \mid X) = 0$; $\forall t \neq s$, which means that $\sigma_v^2 = 0$.[16] In addition, a fixed effect for v_i would be more appropriate in this framework, since we are focused on a specific set of N countries (see Baltagi, 2001 p.12). On the other hand, for the time component e_t we assume a random effect in order to avoid a significant loss of degrees of freedom due to the large T setup.

In order to verify the finite sample performance of the competing OLS, SUR and FB estimators, we next conduct a small Monte Carlo simulation.

2.5 Monte Carlo simulation

The econometric methodology described in section 2.2 suggests a Wald test to investigate a set of restrictions imposed to the coefficients of a VAR, used to forecast the current account. The Wald test verifies whether or not the optimal current account is statistically equal to the observed current account. Moreover, the Wald test could be conducted based on OLS, SUR and FB estimations for the coefficients of the VAR. Therefore, the goal of our experiment is to investigate the power and the size of the Wald test, comparing different techniques used for the estimation of the VAR (reproduced below for a generic country i):

$$\begin{bmatrix} \Delta Z_t^i \\ CA_t^i \end{bmatrix} = \begin{bmatrix} a^i(L) & b^i(L) \\ c^i(L) & d^i(L) \end{bmatrix} \times \begin{bmatrix} \Delta Z_{t-1}^i \\ CA_{t-1}^i \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t}^i \\ \varepsilon_{2t}^i \end{bmatrix}. \tag{37}$$

One of the critical issues regarding Monte Carlo experiments is that of Data-Generating Processes (DGPs). In our experiment, we construct 100 DGPs, and for each DGP we generate 1,000 samples of the series $\begin{bmatrix} \Delta Z_t & CA_t \end{bmatrix}'$, by sampling random series of ε_t's. Moreover, each sample contains 1,000 observations, but, in order to reduce the impact of initial values we consider only the last $T = 100$ or 200 observations. Thus, the Monte Carlo simulation performs 100,000 replications of the experiment.[17]

Two important issues regarding the companion matrix of the generated series must be addressed at this point. The first one is related to the null hypothesis to be checked by the Wald test: In our simulation, we impose Ho to be true or false by just controlling the impact of the theoretical restrictions into the companion matrix. The magnitude of the theoretical restrictions is given by the gamma parameter, in which $\gamma = 1$ imposes Ho to be true, whereas $\gamma \neq 1$ leads to a false Ho (see appendix for details). The second issue is related to the stationarity of the VAR: In order to apply the econometric methodology, each sample of the experiment must be constructed to generate a covariance-stationary VAR. This way, we show (see also

[16] In addition, the "individual effect" translates in practice into an individual intercept; which is also expected to be zero in our setup, since all series are supposed to be weakly stationary and previously demeaned.

[17] A hybrid solution using E-Views, R and MatLab environments is adopted, since the proposed simulation is extremely computational intensive. We proceed as follows: an E-Views code initially generates the time series ΔZ_t and CA_t for each DGP. Then, it estimates the VAR coefficients based on OLS and SUR techniques and save all the replications and the Wald test results in the hard disk. Next, an R code computes the FB estimator and conducts the respective Wald test, also saving the results in a text file. Finally, a MatLab code reads all the results from the hard disk, computes the size of the Wald test based on the three estimators and also constructs the power results.

appendix for details) how to guarantee the stationarity of the VAR by properly "choosing" the eigenvalues of the companion matrix inside the unit circle, and then calculating the coefficients of the companion matrix that generate those eigenvalues.

Results

Concerning the size of the Wald test, we calculate the estimated significance level by simply observing the frequency of rejection of the null hypothesis in the 100,000 replications of the experiment under conditions where the null hypothesis is imposed to be true. Regarding the power of the test, we also compute the rejection frequencies, but under conditions where the null hypothesis is now imposed to be false.

Table 3 - Size of the Wald test

Model	OLS	SUR	FB
(a) N=2 ; T=100	0.0171	0.0653	0.0534
(b) N=2 ; T=200	0.0154	0.0559	0.0520
(c) N=5 ; T=200	0.0175	0.0620	0.0503

Note: The nominal size of the test is α =5%, in which

empirical size = (frequency of p-values below the α (nominal size) / (MC*DGP);

where (MC*DGP) = 100,000 = total number of replications.

Overall, the results suggest that FB-based test has an adequate size in all cases, whereas the OLS-test exhibits a serious finite sample bias, as already predicted by Proposition 2. In the same line, the SUR-test also seems to show a (small) non-zero bias due to the presence of global shocks, as previously discussed in Proposition 3.

The power investigation can be conducted by controlling the experiment under conditions where the null is imposed to be false, i.e., $\gamma \neq 1$. If the null hypothesis were only "slightly false" (e.g., $\gamma = 0.9$), one would expect power to be lower than if it were "grossly false" (e.g., $\gamma = 0.5$). The results of the power investigation corroborate this expectation and are presented in next tables. To adjust for size distortion, we also report a "size-corrected" power.[18] The results with size correction are quite similar, suggesting that the Wald test might exhibit similar power across the considered estimation procedures.

[18] "Size-corrected power" is just power using the critical values that would have yielded correct size under the null hypothesis.

Table 4 - Power of the Wald test

(a) N=2;T=100	without size-correction			size-corrected power		
γ	OLS	SUR	FB	OLS	SUR	FB
0.9	0.051	0.150	0.126	0.082	0.087	0.104
0.7	0.439	0.541	0.545	0.520	0.533	0.527
0.5	0.725	0.795	0.782	0.729	0.727	0.742
(b) N=2;T=200	without size-correction			size-corrected power		
γ	OLS	SUR	FB	OLS	SUR	FB
0.9	0.165	0.264	0.252	0.241	0.246	0.242
0.7	0.568	0.645	0.635	0.628	0.643	0.627
0.5	0.847	0.887	0.878	0.872	0.881	0.874
(c) N=5;T=200	without size-correction			size-corrected power		
γ	OLS	SUR	FB	OLS	SUR	FB
0.9	0.113	0.226	0.200	0.184	0.205	0.194
0.7	0.544	0.674	0.638	0.642	0.660	0.633
0.5	0.795	0.851	0.834	0.827	0.823	0.826

Notes: a) Power = (frequency of p-values below the α% nominal size) / (MC*DGP),

b) Gamma < 1 indicates a false Ho.

3 Empirical Results

3.1 Data

All data are from the national accounts of IFS – International Financial Statistics (IMF). The CA_t and ΔZ_t series for the G-7 countries are constructed from seasonally adjusted quarterly data (at annual rates), and are expressed in 2000 local currency.[19] In addition, all data are converted in per capita real terms, by dividing it by the implicit GDP deflator and the population. It is worth mentioning that the current account data are not directly obtained from the balance of payments data sets, since these series are not available for all of the countries for an extensive period of time, and it would lead to an arbitrary allocation of "net errors and omissions" in the current account.

Sample 1 (G-7): USA, Canada, Japan, United Kingdom, Germany, Italy, and France.
Period: 1980:q1–2007:q1 (set of 7 countries, 109 time periods, with a total amount of $2NT = 1,526$ observations).

Sample 2: USA, Canada, Japan, United Kingdom, and Germany.
Period: 1960:q1–2007:q1 (set of 5 countries, 189 periods, with a total amount of $2NT = 1,890$ observations).

[19] Based on IMF's World Economic Outlook (April, 2005 - statistical appendix), we adopt the following fixed conversion rates (after 31/12/1998) between the Euro and the currencies of Germany, France and Italy: 1 Euro = 1.95583 Deutsche mark = 6.55957 French francs = 1,936.27 Italian lire.

3.2 Granger Causality

One of the four implications of the theoretical model,[20] listed by Otto (1992), is that CA_t helps to forecast ΔZ_t. According to our results, one can verify that the null hypothesis (CA_t does not Granger-cause ΔZ_t) is rejected at 5% level for Canada and Japan.

Table 5 - Comparison of Results (Ho: CA_t does not Granger-cause ΔZ_t)

	Granger Causality (p-value)			
Country \ Author	**IG**	**O**	**G**	**A**
USA	0.16944	0.0001	0.0004	-
CAN	0.04617	0.21	0.40	-
JPN	0.02426	-	0.62	-
UK	0.47515	-	0.68	-
GER	0.83352	-	0.76	-
ITA	0.91321	-	-	-
FRA	0.59218	-	-	0.10

Notes: IG means Issler & Gaglianone (our results), O refers to Otto (92), G indicates Ghosh (95), A refers to Agénor et al. (1999).

Our results are quite in contrast to the literature, probably due to the different sample periods. For instance, Otto (1992) rejects Ho for the USA (at 1% level), but does not reject it for Canada. Ghosh (1995) also rejects Ho for the USA (at 1% level) and does not reject it for Canada, Japan, UK and Germany, and Agénor et al. (1999) present a p-value of 0.10 for France. The different results could possibly be explained by the broader range of our sample period, in comparison to the previous studies, which do not account for all global and idiosyncratic shocks occurred in the last decades: Our sample period covers quarterly data from 1960 until 2007, whereas Otto (1992) considers the period 1950-88, Ghosh(1995) studies the period 1960-88, and Agénor et al. (1999) covers 1970-96.

Thus, our results indicate that, with the exception of Canada and Japan, the current account does not help to forecast the net output of the G-7 countries, indicating that the agents possibly do not have any additional information to predict ΔZ_t, other than those contained in the past of their own series. Recall Proposition 1, which states that if the Granger causality implication is not verified, then, the optimal current account should not be generated, since it would lead to spurious results. In the present work, the Granger causality is only verified for Canada and Japan. Thus, for the other countries, the model should be rejected and the optimal current account should not be generated. For instance, the case of UK (sample 1) can be analyzed as an example of spurious result. The VAR(1) estimated for this country is given by

VAR coefficients for **UK**

$$\begin{bmatrix} \Delta Z_t \\ CA_t \end{bmatrix} = \begin{bmatrix} a_1 = -0.211322 \ \text{(-2.40)} & b_1 = -0.066236 \ \text{(-1.76)} \\ c_1 = -0.133631 \ \text{(-1.12)} & d_1 = 0.854978 \ \text{(16.99)} \end{bmatrix} \times \begin{bmatrix} \Delta Z_{t-1} \\ CA_{t-1} \end{bmatrix} + \begin{bmatrix} 5.477849 \ \text{(1.67)} \\ 2.626709 \ \text{(0.59)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

Note: t-statistics in parentheses

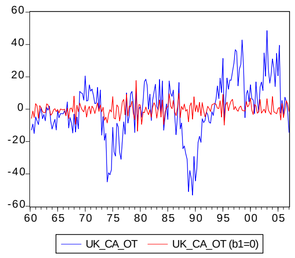
[20] The results of ADF unit root test and the cointegration analysis are presented in appendix.

where a_1 and d_1 are statistically significant, but this is not the case for b_1 and c_1. As described in table 6, the vector K (recall equation (22)) is extremely sensitive to variations in b_1 and could generate completely different CA_t^* series.[21] Assuming $b_1 = -0.066$ (instead of zero, since b_1 is not statistically significant), the model indicates that $\beta = 0.348$ (unlike the correct value of $\beta = 0$).

Table 6 - Vector $K = [\alpha; \beta]$ for UK

	Assume $b_1 = -0.066$	Assume $b_1 = 0$
α	0.134	0.172
β	0.348	0.000

Figure 1 - CA_t^* for UK



Note: The picture above exhibits two optimal current accounts for UK (sample 1), generated from different K vectors of table 6.

3.3 Correlation matrix

The residual correlation matrix obtained in the joint estimation of the VAR could be a starting point to justify the SUR technique, since the contemporaneous correlation across the G-7 countries should not be ignored.

Table 7 - Residual Correlation Matrix (sample 1)

	USA_DZ	USA_CA	CAN_DZ	CAN_CA	JPN_DZ	JPN_CA	FRA_DZ	FRA_CA	UK_DZ	UK_CA	GER_DZ	GER_CA	ITA_DZ	ITA_CA
USA_DZ	1.00	0.26												
USA_CA	0.26	1.00												
CAN_DZ	0.15	0.04	1.00	0.53										
CAN_CA	0.02	0.18	0.53	1.00										
JPN_DZ	0.13	0.15	-0.09	0.01	1.00	0.06								
JPN_CA	0.11	0.00	0.05	0.00	0.06	1.00								
FRA_DZ	0.07	0.04	-0.03	-0.11	0.25	-0.06	1.00	0.32						
FRA_CA	0.10	0.08	-0.13	-0.12	0.17	-0.02	0.32	1.00						
UK_DZ	0.02	-0.05	0.13	0.15	-0.12	0.00	-0.10	0.01	1.00	0.44				
UK_CA	-0.13	-0.08	0.06	0.08	-0.01	-0.06	0.04	0.05	0.44	1.00				
GER_DZ	0.10	0.16	-0.05	0.07	0.05	0.10	0.15	-0.04	-0.08	-0.06	1.00	0.20		
GER_CA	0.21	0.01	0.19	0.05	-0.04	0.04	0.11	0.09	0.17	0.12	0.20	1.00		
ITA_DZ	0.09	0.21	0.04	-0.06	-0.02	0.06	0.16	0.17	0.07	0.09	0.08	0.14	1.00	0.49
ITA_CA	0.02	0.28	0.07	0.03	0.13	-0.03	0.08	0.30	0.08	-0.08	0.11	0.07	0.49	1.00

Furthermore, a Likelihood Ratio (LR) test is able to provide a formal argument to adopt the SUR approach, instead of the OLS technique. Under the null hypothesis, the residuals covariance matrix (Ω) is a diagonal (band)[22] matrix, suggesting the OLS method. On the other hand, the alternative specification (H_1) supposes that Ω is a non-diagonal (band) matrix, recommending the SUR approach. This way, Ho imposes a set of restrictions on the residuals covariance matrix, since all elements out of the diagonal (band) are set to zero. In this case, according to Hamilton (1994), twice the log likelihood ratio for a Gaussian VAR is given by

$$2\left(L_1^* - L_0^*\right) = T\left(\ln\left|\widehat{\Omega}_0\right| - \ln\left|\widehat{\Omega}_1\right|\right), \tag{38}$$

[21] Optimal current account is generated by equations (16) and (17), and depends on the estimated coefficients of the VAR and the world interest rate (supposed 2% per year).

[22] The residuals covariance matrix is diagonal-band, in OLS estimation, because of the structure of the VAR, since each country has two equations (CA_t and ΔZ_t).

where L_0^* is the maximum value for the log likelihood under Ho (and L_1^* under the alternative hypothesis), T is the number of effective observations, $\left|\widehat{\Omega}_0\right|$ is the determinant of the residuals covariance matrix estimated by OLS, and $\left|\widehat{\Omega}_1\right|$ is the determinant of the same matrix estimated by SUR. Under the null hypothesis, the difference between L_1^* and L_0^* is statistically zero, and the LR statistic asymptotically follows a χ^2 distribution, with degrees of freedom equal to the number of restrictions imposed under Ho.

Table 8 - Results of the LR test

	Sample 1	Sample 2
T	1,520	1,870
$\left\|\widehat{\Omega}_0\right\|_{OLS}$	1.11734E+51	1.53125E+38
$\left\|\widehat{\Omega}_1\right\|_{SUR}$	4.47295E+50	1.23388E+38
ζ_{LR}	1,391.53	403.77
ζ_{LR}^*	1,385.13	402.91
5% critical value	$\chi^2_{(80)}$=101.88 ; $\chi^2_{(90)}$=113.15	$\chi^2_{(40)}$=55.76
1% critical value	$\chi^2_{(80)}$=112.33 ; $\chi^2_{(90)}$=124.12	$\chi^2_{(40)}$=63.69

Notes: a) T is the number of effective observations, and $\zeta_{LR} = T\left(\ln\left|\widehat{\Omega}_0\right| - \ln\left|\widehat{\Omega}_1\right|\right)$ is the LR statistic.

b) ζ_{LR}^* is a modification to the LR test to take into account small-sample bias, replacing T by $(T-k)$, where k is the number of parameters estimated per equation.

c) In sample 1, the degrees of freedom (dof)= 84, and in sample 2, dof=40.

Hence, the null hypothesis could be rejected in both samples, since the LR statistics are larger than the critical values. Therefore, the residuals covariance matrices are non-diagonal (band), and the SUR approach is better recommended than the OLS method.

3.4 Wald test

A formal comparison between $(CA_t^* - \mu_{CA}^*)$ and $(CA_t - \mu_{CA})$, to measure the fit of the model with the data, is provided by the restrictions (20) imposed to the coefficients of the VAR, through a Wald test, which asymptotically follows a χ^2 distribution (with the degrees of freedom equal to the number of restrictions). The acceptance of those restrictions in the Wald test means that both series of current account (optimal and observed) are statistically the same.[23]

We perform the Wald test for the G-7 countries based on two different types of time series. The first one is the usual time series suggested by the literature (e.g., Ghosh, 1995), in which the Wald test is conducted from a seasonally adjusted quarterly data (at annual rates), expressed in 2000 local currency, converted in per capita real terms, by dividing it by the implicit GDP deflator and the population. In order to compute common shocks among the considered countries, we also convert all series to 2000 U.S. dollars. In the second approach, however, data is not expressed in per capita terms and is converted to U.S. dollars by using a

[23] The Wald test can be implemented for several values of the world interest rate (r). However, the results are almost the same for values of r ranging from 1% to 6%. This way, we have adopted $r = 2\%$ following the literature.

proper exchange rate time series, due to the fact that the global shocks must be computed from a proper set of scaled and comparable time series. Since both methodologies lead to very similar results, we next present the results only for the later approach, based on different estimation procedures: OLS, SUR, FB.1 and FB.2 (see tables 13 and 14 in appendix for further details).[24]

The first estimator is merely OLS equation-by-equation, ignoring possible crossed effects among countries. The second approach, SUR, is just a feasible GLS estimator applied to the considered system of equations and, as already mentioned, might improve the efficiency when compared to the OLS method and, thus, the covariance matrix of a SUR estimation could result in a rejection of the model previously accepted in the OLS framework. The last estimators, FB.1 and FB.2, are based on a two-way random error decomposition procedure of Fuller & Battese (1974), in which the residual term is decomposed into individual effects, common shocks, and idiosyncratic terms, i.e., $\varepsilon_{i,t} = v_i + e_t + \epsilon_{i,t}$. The only difference between these two estimators is that FB.1 assumes a unique global shock e_t for the whole system of equations, whereas, FB.2 considers a common shock e_t^{CA} for the CA_t system of equations, and a different shock e_t^{DZ} for the ΔZ_t equations.

A comparison of these results with the empirical evidence found in the literature is presented in Table 9: Otto (1992) rejects the model for the USA and Canada, Ghosh (1995) rejects the model for Canada, Japan, the UK and Germany, but does not reject it for the USA, and Agénor et al. (1999) do not reject it for France.

Table 9 - Comparison with the literature

	Wald Test (p-value)						
Country	**OLS**	**SUR**	**FB.1**	**FB.2**	**Otto (92)**	**Ghosh (95)**	**Agénor et al.(99)**
USA (1)	0.1400	0.0207 (*)	0.034 (*)	0.0394 (*)	0.0041 (**)	1.19	-
CAN (1)	0.2402	0.0312 (*)	0.8692	0.8766	0.0020 (**)	95 (**)	-
JPN (1)	0.0291 (*)	0.0205 (*)	0.0032 (**)	0.0051 (**)	-	75 (*)	-
UK (1)	0.1796	0.1550	0.9970	0.9979	-	464 (**)	-
GER (1)	0.048 (*)	0.0097 (**)	0.6667	0.6898	-	90 (**)	-
ITA (1)	0.1425	0.0789	0.9771	0.9594	-	-	-
FRA (1)	0.0295 (*)	0.0000 (**)	0.7957	0.7525	-	-	0.314
USA (2)	0.0291 (*)	0.0069 (**)	0.0000 (**)	0.0000 (**)			
CAN (2)	0.1499	0.0123 (*)	0.8288	0.8361			
JPN (2)	0.0201 (*)	0.0160 (*)	0.0363 (*)	0.3019			
UK (2)	0.0139 (*)	0.0001 (**)	0.8727	0.6861			
GER (2)	0.0143 (*)	0.0078 (**)	0.7161	0.3260			

Notes: Ghosh (1995) presents chi-squared values; $\text{USA}_{(1)}$ indicates sample 1 and $\text{USA}_{(2)}$ means sample 2; (**) means rejection at 1% level, and (*) at 5% level.

First of all, note that the SUR estimator generally over rejects the model in comparison to OLS.[25] However, recall from Proposition 1 that a fair analysis of Table 9 should only consider countries in which

[24] EViews 5.1 was used to obtain the OLS and SUR estimators, whereas a code in R was developed for the FB estimators, which could alternatively be computed in SAS.

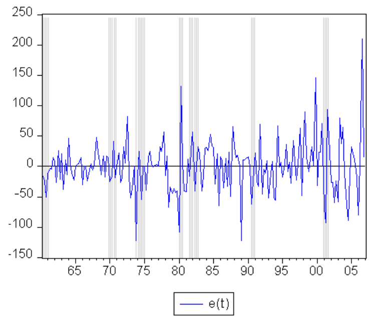
[25] An important remark is provided by Ghosh and Ostry (1995), which argue that the non-rejection of the model for a given country can occur because of the magnitude of the standard deviations in the coefficients of the VAR. High values for the standard errors could lead to a statistical equality between the optimal and observed current accounts, even if these series are graphically different.

the Granger causality can (indeed) be verified. Since from Table 5 it only occurs for Canada and Japan, and the later country failed at the Wald test, our empirical results suggest, contrary to the previous results found in the literature, that the PVM model of the current account cannot be rejected for Canada.

Secondly, note that due to the possible finite sample bias of OLS and SUR estimators, the results are quite different from the FB results, as already expected from Propositions 2-4. In fact, regarding the FB results, the Hausman (1978) m-statistic, that provides information about the appropriateness of the random effects specification, do not indicate a rejection[26] of the null hypothesis of zero correlation between regressors and effects.

More importantly, in both samples the FB approach suggest that $\sigma_v^2 = 0$, and that the global shock component (e_t) should not be ignored in the estimation process, since its relative importance in respect to the total residual is estimated as $\Psi = 0.42$ in sample 1, and $\Psi = 0.36$ in sample 2. Recall that $\Psi \equiv 2N\sigma_e^2/(\sigma_\epsilon^2 + T\sigma_v^2 + 2N\sigma_e^2)$ should be zero, in the case of no global shocks, and also recall Proposition 2(ii), which states that the finite sample bias of the OLS estimated covariance matrix increases as long as the common shocks become more present in the data. A global shock (e_t) time series is depicted in next figure and compared (for illustrative purposes) to U.S. recessions:

Figure 2 - Global shocks (sample 2) and U.S. recessions



Note: Gray bars represent the U.S. (NBER) recessions.

Note that some of the U.S. recessions indeed coincide with the negative peaks of e_t, including the most recent period in 2001, which is a natural result since the global shocks in the current account of the G-7 are expected to be (at least) partially driven by the world´s biggest economy movements.

4 Conclusions

The standard intertemporal optimization model of the current account is adopted to analyze the G-7 countries. In this framework, the perfect capital mobility allows the agents to smooth consumption via current account. The econometric approach of the model, developed by Campbell & Shiller (1987), consists of estimating an unrestricted VAR to verify the adherence of the theoretical framework onto the data. Furthermore,

[26] For instance, in sample 1, the FB.1 estimator, with 28 degrees of freedom, exhibit the m statistic equal to 11.19779 (p-value: 0.9980223); and for FB.2, with 14 degrees of freedom in each system, m=4.139241 (p-value: 0.9945687).

a Wald test is used to investigate a set of restrictions imposed to the VAR, used to forecast the current account, testing whether or not the optimal current account is equal to the observed series (null hypothesis).

In spite of all theoretical advances regarding the current account PVM models in recent years, we have opted to investigate some econometric techniques that could be used in the estimation process, in order to answer an important related question: Could an inappropriate econometric technique lead to wrong conclusions regarding the rejection of the current account PVM model? We focused on three estimation techniques (OLS, SUR and FB).

Firstly, a SUR technique could be recommended (instead of OLS) in order to properly consider contemporaneous correlations across the considered countries, that might be caused by global shocks such as the oil shocks of 70s or the financial crises of 90s. However, due to potential pitfalls associated with the SUR estimation, we also provide an application of the Fuller & Battese (1974) two-way random error decomposition, due to the existence of common shocks and the possible bias in the covariance matrix estimated with OLS and SUR techniques.

We investigate these estimators in a Monte Carlo experiment, and evaluate the power and size of the Wald test in the presence of global shocks, concluding that the FB-based test exhibits a good performance in the size investigation. On the other hand, the OLS-based test performs unsatisfactorily, and a small finite sample bias can also be detected for the SUR approach. The numerical simulations also indicate that there is no clear difference in the size-corrected power of the considered Wald test based on different estimators.

To summarize, the theoretical methodology, as well as the Monte Carlo simulations, suggests the adoption of the FB estimator (instead of OLS) in the presence of global shocks, small N and relatively large T. The SUR approach could be used in this case, but only when the common shock exhibits a low magnitude, since it directly leads to a biased estimated covariance matrix.

This paper also proposes a note on Granger causality and Wald tests: the Granger causality is addressed as a "sine qua non" condition for the entire validation of the model, since the construction of the optimal current account leads to spurious results when this condition is not verified. We further provide an empirical exercise by estimating the model for the G-7 countries. Indeed, the results substantially change with the application of the different estimation techniques.

More importantly, the FB framework indicates that the common shock in the G-7 countries can account for almost 40% of the total residuals, suggesting that the previous (OLS) estimations might be seriously biased, and the familiar inference procedures would no longer be appropriate. The error-decomposition procedure only indicates a rejection of the model for the USA and Japan, which is in sharp contrast to the previous literature. Putting all together, these findings suggest that the PVM cannot be rejected for Canada, and cast serious doubts to some results presented in the literature, based on OLS estimation, that ignore the presence of common shocks among countries.

Appendix A. Note on Granger Causality and Wald tests

A VAR(2) model can be rewritten as a VAR(1) in the following way:

$$\begin{bmatrix} \Delta Z_t \\ \Delta Z_{t-1} \\ CA_t \\ CA_{t-1} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & b_1 & b_2 \\ 1 & 0 & 0 & 0 \\ c_1 & c_2 & d_1 & d_2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \Delta Z_{t-1} \\ \Delta Z_{t-2} \\ CA_{t-1} \\ CA_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ 0 \\ \varepsilon_{2t} \\ 0 \end{bmatrix}$$

or in a compact form $X_t = AX_{t-1} + \varepsilon_t$. Hence, the vector K is given by

$$K = \begin{bmatrix} \alpha_1 & \alpha_2 & \beta_1 & \beta_2 \end{bmatrix} = -h'(\frac{A}{1+r})(I_4 - \frac{A}{1+r})^{-1}$$

$$\therefore K = -\begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} (\frac{1}{1+r} \begin{bmatrix} a_1 & a_2 & b_1 & b_2 \\ 1 & 0 & 0 & 0 \\ c_1 & c_2 & d_1 & d_2 \\ 0 & 0 & 1 & 0 \end{bmatrix})(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{1+r} \begin{bmatrix} a_1 & a_2 & b_1 & b_2 \\ 1 & 0 & 0 & 0 \\ c_1 & c_2 & d_1 & d_2 \\ 0 & 0 & 1 & 0 \end{bmatrix})^{-1}$$

Thus, after some algebraic manipulations, the β_1 coefficient of the vector K is given by:

$$\begin{aligned} \beta_1 &= -\frac{a_1(1+r)(b1 + b1r + b2)}{\phi} - \frac{a_2(b1 + b1r + b2)}{\phi} \\ &\quad -\frac{b1(1+r)(1 + 2r + r^2 - a1 - a1r - a2)}{\phi} - \frac{b2(1 + 2r + r^2 - a1 - a1r - a2)}{\phi} \end{aligned}$$

$$\begin{aligned} \text{where } \phi &= 1 + 2a_1d_1r + 4r + a_1d_1r^2 - c_1b_2r - 2c_1b_1r + a_2d_1r - d_1 - d_2 - a_2 - a_1 - c_2b_1r + 6r^2 \\ &\quad -3d_1r - 2a_2r - a_2r^2 + a_2d_1 + a_2d_2 - c_2b_1 - c_2b_2 + a_1d_2r - c_1b_1r^2 + 4r^3 + r^4 \\ &\quad -3d_1r^2 - 2rd_2 - 3a_1r - 3a_1r^2 + a_1d_1 + a_1d_2 - d_1r^3 - r^2d_2 - a_1r^3 - c_1b_1 - c_1b_2 \end{aligned}$$

In this case, the optimal current account is given by

$$(CA_t^* - \mu_{CA^*}) = K(X_t - \mu) = \begin{bmatrix} \alpha_1 & \alpha_2 & \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} \Delta Z_t - \mu_{\Delta Z} \\ \Delta Z_{t-1} - \mu_{\Delta Z} \\ CA_t - \mu_{CA} \\ CA_{t-1} - \mu_{CA} \end{bmatrix}$$

and the Wald test analyzes the joint restrictions: $\alpha_1 = \alpha_2 = \beta_2 = 0$ and $\beta_1 = 1$. Again, one should note that if the Granger causality is not verified (e.g., $b_1 = b_2 = 0$) then $\beta_1 = 0$. In this manner, the implication of Granger causality becomes a necessary condition for the validation of the VAR(2) model. The generalization of this result to a VAR(p) framework is straightforward, as presented in the proof of Proposition 1.

Appendix B. Proof of Propositions

Proof of Proposition 1. The VAR(p) can be rewritten as a VAR(1), in the following way

$$\begin{bmatrix} \Delta Z_t \\ \vdots \\ \vdots \\ \Delta Z_{t-p+1} \\ CA_t \\ \vdots \\ \vdots \\ CA_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & \cdots & a_p & b_1 & \cdots & \cdots & b_p \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ c_1 & \cdots & \cdots & c_p & d_1 & \cdots & \cdots & d_p \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \Delta Z_{t-1} \\ \vdots \\ \vdots \\ \Delta Z_{t-p} \\ CA_{t-1} \\ \vdots \\ \vdots \\ CA_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ 0 \\ \vdots \\ 0 \\ \varepsilon_{2t} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or in the same compact form $X_t = AX_{t-1} + \varepsilon_t$. Supposing that the Granger causality is not provided by the data set, it follows that $b(L) = 0$, and the companion matrix A becomes:

$$A = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$$

where $A_{11} = \begin{bmatrix} a_1 & \cdots & \cdots & a_p \\ 1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$; $A_{21} = \begin{bmatrix} c_1 & \cdots & \cdots & c_p \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$; $A_{22} = \begin{bmatrix} d_1 & \cdots & \cdots & d_p \\ 1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

In this case, vector K can be written as $K = -h'BC^{-1}$, where the matrices B and C are defined as $B = (\frac{A}{1+r})$ and $C = (I_{2p} - \frac{A}{1+r})$. It should be noted that matrices B and C are also partitioned matrices with a null upper-right block. According to Simon & Blume (1994, p.182), in this case, the inverse of the partitioned matrix C will also result in a null upper-right block matrix (see Theorem reproduced below).

Theorem 8.15 (Simon & Blume, 1994): Let C be a square matrix partitioned as $C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$ where C_{11} and C_{22} are square submatrices. If both C_{22} and $D \equiv C_{11} - C_{12}C_{22}^{-1}C_{21}$ are nonsingular, then C is nonsingular and

$$C^{-1} = \begin{bmatrix} D^{-1} & -D^{-1}C_{12}C_{22}^{-1} \\ -C_{22}^{-1}C_{21}D^{-1} & C_{22}^{-1}\left(I + C_{21}D^{-1}C_{12}C_{22}^{-1}\right) \end{bmatrix}$$

Thus, in our case, $C_{12} = 0$ and the term $-D^{-1}C_{12}C_{22}^{-1}$ becomes a null submatrix, suggesting that C^{-1} and the following product BC^{-1} are also NURB matrices. Finally, the vector $K = \begin{bmatrix} \alpha_1 & \ldots & \alpha_p & \beta_1 & \ldots & \beta_p \end{bmatrix}$ is given by selecting (through the vector $-h'$) the first line from the matrix (BC^{-1}), suggesting that all β_i $(i = 1, \ldots, p)$ coefficients are zero. ■

Proof of Proposition 2. (i) Consider the system of equations for any country i:
$\begin{cases} \Delta Z_{i,t} = \mu_i^{\Delta Z} + a_i(L)\Delta Z_{i,t} + b_i(L)CA_{i,t} + \varepsilon_{i,t}^{\Delta Z} \\ CA_{i,t} = \mu_i^{CA} + c_i(L)\Delta Z_{i,t} + d_i(L)CA_{i,t} + \varepsilon_{i,t}^{CA} \end{cases}$; where $a_i(L)$, $b_i(L)$, $c_i(L)$ and $d_i(L)$ are polynomials of order p, and $\varepsilon_{i,t}^{\Delta Z} = (v_i^{\Delta Z} + e_t^{\Delta Z} + \epsilon_{i,t}^{\Delta Z})$. The equation for $\Delta Z_{i,t-1}$ can be expressed by: $(a_{i,2}\Delta Z_{i,t-2} + a_{i,3}\Delta Z_{i,t-3} + ... + a_{i,p+1}\Delta Z_{i,t-p+1}) + b_i(L)CA_{i,t-1} + \varepsilon_{i,t-1}^{\Delta Z}$. Note that $E(\varepsilon_{i,t}^{\Delta Z}\Delta Z_{i,t-1}) = E(v_i^{\Delta Z} + e_t^{\Delta Z} + \epsilon_{i,t}^{\Delta Z})(a_{i,2}\Delta Z_{i,t-2} + a_{i,3}\Delta Z_{i,t-3} + ... + a_{i,p+1}\Delta Z_{i,t-p+1} + b_i(L)CA_{i,t-1} + v_i^{\Delta Z} + e_{t-1}^{\Delta Z} + \epsilon_{i,t-1}^{\Delta Z}) = E(v_i^{\Delta Z} v_i^{\Delta Z}) = \sigma_{v,\Delta Z}^2 > 0 \therefore E(X'\varepsilon) \neq 0$ and $\widehat{\beta}_{ols}$ becomes inconsistent;

(ii) Recall that $Var(\beta_{OLS}) = \sigma^2(X'X)^{-1}$ and $\widehat{Var}(\beta_{OLS}) = s^2(X'X)^{-1} = (X'X)^{-1}[\frac{e'e}{2NT-k}]$; where e is the estimated OLS residual and $k = 4Np$. Thus, $\widehat{Var}(\beta_{OLS}) = \frac{(X'X)^{-1}}{2NT-k}[Y'(I - X(X'X)^{-1}X')Y] = \frac{(X'X)^{-1}}{2NT-k}tr[Y'MY]$; where $M \equiv I - X(X'X)^{-1}X'$ is an idempotent and symmetric matrix, in which $MX = 0$. This way, $tr[Y'MY] = tr[Y'M'MY] = tr[\varepsilon' M' M\varepsilon] = tr[M\varepsilon\varepsilon']$; and $\widehat{Var}(\beta_{OLS}) = \frac{(X'X)^{-1}}{2NT-k}tr[M\varepsilon\varepsilon']$. Therefore, $E[\widehat{Var}(\beta_{OLS}) \mid X] = E[\frac{(X'X)^{-1}}{2NT-k}tr(M\varepsilon\varepsilon') \mid X] = \frac{(X'X)^{-1}}{2NT-k}tr[M(E(\varepsilon\varepsilon' \mid X))] = (X'X)^{-1}\frac{tr[M\Omega]}{2NT-k} = (X'X)^{-1}\varphi$, where $\varphi \equiv \frac{tr[M\Omega]}{2NT-k}$. One can also rewrite last expression as $E[\widehat{Var}(\beta_{OLS}) \mid X] = (X'X)^{-1}\{\varphi X'X\}$ $(X'X)^{-1} = (X'X)^{-1}\{\varphi X'M_{12}X + \varphi X'M_{1.}X + \varphi X'M_{.2}X + \varphi X'M_{..}X)\}(X'X)^{-1}$, where the decomposition of X, based on the idempotent and symmetric matrices $\{M_{12}; M_{1.}; M_{.2}; M_{..}\}$, was applied in last equality. Now, assume that the considered bias in the variance estimator is zero, i.e., $E[\widehat{Var}(\beta_{OLS}) \mid X] = Var(\beta_{OLS})$. Note that this is true if and only if $\varphi = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4$. But this is a contradiction since, by assumption, $(\sigma_\epsilon^2; \sigma_e^2) > 0$ and $\sigma_v^2 = 0$, and thus $\gamma_1 = \gamma_2$ and $\gamma_3 = \gamma_4$, but $\gamma_1 \neq \gamma_3$;

(iii) It follows that $\frac{\partial Var(\beta_{OLS})}{\partial\sigma_e^2} = (X'X)^{-1}\{2N(X'(M_{.2} + M_{..})X)\}(X'X)^{-1} = (X'X)^{-1}\{2N(X'(\frac{J_{2N}\otimes I_T}{2N})X)\}$ $(X'X)^{-1} = (X'X)^{-1}X'JX\}(X'X)^{-1}$; where $J \equiv J_{2N}\otimes I_T$. Recall that $E[\widehat{Var}(\beta_{OLS}) \mid X] = (X'X)^{-1}\frac{tr[M\Omega]}{2NT-k}$ and, thus, $\frac{\partial E[\widehat{Var}(\beta_{OLS})|X]}{\partial\sigma_e^2} = \frac{(X'X)^{-1}}{2NT-k}(\frac{\partial tr[M\Omega]}{\partial\sigma_e^2}) = \frac{(X'X)^{-1}}{2NT-k}(\frac{\partial vec'(\Omega)}{\partial\sigma_e^2})vec(M') = \frac{(X'X)^{-1}}{2NT-k}(2N(vec'(M_{.2} + M_{..})))vec(M') = \frac{(X'X)^{-1}}{2NT-k}vec'(J)vec(M') = \frac{(X'X)^{-1}}{2NT-k}tr(MJ) = (X'X)^{-1}\frac{tr(MJ)}{tr(M)}$. Therefore, $\frac{\partial B_{OLS}}{\partial\sigma_e^2} = \frac{\partial Var(\beta_{OLS})}{\partial\sigma_e^2}$ $-\frac{\partial E[\widehat{Var}(\beta_{OLS})|X]}{\partial\sigma_e^2} = (X'X)^{-1}X'JX\}(X'X)^{-1} - (X'X)^{-1}\frac{tr(MJ)}{tr(M)} = (X'X)^{-1}\{X'JX - \frac{tr(MJ)}{tr(M)}X'X\}(X'X)^{-1}$ $\equiv B'AB$, where $A \equiv X'JX - \frac{tr(MJ)}{tr(M)}X'X$ and $B \equiv (X'X)^{-1}$. Now, let $\alpha \equiv \frac{tr(MJ)}{tr(M)} = \frac{tr[(I - X(X'X)^{-1}X')J]}{2NT-k}$ $= \frac{tr[J - X(X'X)^{-1}X'J]}{2NT-k} = \frac{tr[J] - tr[X(X'X)^{-1}X'J]}{2NT-k} = \frac{tr[J_{2N}\otimes I_T] - tr[X(X'X)^{-1}X'J]}{2NT-k} = \frac{2NT - tr[X(X'X)^{-1}X'J]}{2NT-k}$. By assumption A2(ii), it follows that $T > 2p \therefore 2NT > 4Np = k \therefore 2NT - k > 0$. By A2(i), we have that $tr[(X'X)^{-1}X'JX] - tr(I_{2NT}) > 0 \therefore tr[X(X'X)^{-1}X'J] > 2NT$. Thus, by A2 it follows that the scalar $\alpha < 0$. Now, recall that $A \equiv X'JX - \frac{tr(MJ)}{tr(M)}X'X = X'JX - \alpha X'X = X'JX + \delta X'X$; where $\delta \equiv (-\alpha) > 0$. By A1 we have that $(X'X) > 0$ and $(X'JX) > 0$, thus, it follows that $\delta X'X > 0$ and $A > 0$. Therefore, since B is non-singular, and A is a positive definite matrix, then, $\frac{\partial B_{OLS}}{\partial\sigma_e^2} = B'AB > 0 \therefore diag(\frac{\partial B_{OLS}}{\partial\sigma_e^2}) > 0$;

(iv) If $\sigma_v^2 = \sigma_e^2 = 0$, then, $\Omega \equiv \sigma_\epsilon^2 M_{12} + (\sigma_\epsilon^2 + T\sigma_v^2)M_{1.} + (\sigma_\epsilon^2 + N\sigma_e^2)M_{.2} + (\sigma_\epsilon^2 + T\sigma_v^2 + N\sigma_e^2)M_{..} = \sigma_\epsilon^2(M_{12} + M_{1.} + M_{.2} + M_{..}) = \sigma_\epsilon^2 I_{2NT}$; and thus $tr(M\Omega) = \sigma_\epsilon^2 tr(M)$. Therefore, $E[\widehat{Var}(\beta_{OLS}) \mid X] = (X'X)^{-1}\frac{tr[M\Omega]}{2NT-k} = \sigma_\epsilon^2(X'X)^{-1}\frac{tr[M]}{2NT-k} = \sigma_\epsilon^2(X'X)^{-1}\frac{tr[I_{2NT} - X(X'X)^{-1}X']}{2NT-k} = \sigma_\epsilon^2(X'X)^{-1}\frac{(2NT-k)}{2NT-k} = \sigma_\epsilon^2(X'X)^{-1} = Var(\beta_{OLS}) \therefore$ $B_{OLS} = 0$. ■

Proof of Proposition 3. (i) Since the OLS estimator is inconsistent, the OLS residuals would lead to a feasible GLS estimator (and thus SUR) to be inconsistent as well, since the model contains variables that are correlated with the residuals;

(ii) From the Fuller & Battese (1974) two-way random error decomposition, it follows that $\Omega = \gamma_1 M_{12} + \gamma_2 M_{1.} + \gamma_3 M_{.2} + \gamma_4 M_{..}$. By applying the definition of the previous M_{ii} matrices, it follows that $\Omega = \gamma_1(I_{2NT} - \frac{I_{2N}\otimes J_T}{T} - \frac{J_{2N}\otimes I_T}{2N} + \frac{J_{2NT}}{2NT}) + \gamma_2(\frac{I_{2N}\otimes J_T}{T} - \frac{J_{2NT}}{2NT}) + \gamma_3(\frac{J_{2N}\otimes I_T}{2N} - \frac{J_{2NT}}{2NT}) + \gamma_4(\frac{J_{2NT}}{2NT}) = [\frac{\gamma_3-\gamma_1}{2N}J_{2N} + \gamma_1 I_{2N}] \otimes I_T + [\frac{\gamma_2-\gamma_1}{T}I_{2N} + \frac{\gamma_1-\gamma_2-\gamma_3+\gamma_4}{2NT}J_{2N}] \otimes J_T$. By assumption, it follows that $(\sigma_\epsilon^2; \sigma_e^2) > 0; \sigma_v^2 = 0$; and thus $\gamma_1 = \gamma_2$ and $\gamma_3 = \gamma_4$, but $\gamma_1 \neq \gamma_3$. Therefore, $\Omega = [\gamma_1 I_{2N} + \frac{\gamma_3-\gamma_1}{2N}J_{2N}] \otimes I_T \equiv \Phi_1 \otimes I_T$. On the other hand, in the SUR approach, it follows that $\widehat{\Omega} = \widehat{\Sigma} \otimes I_T$, in which $\widehat{\Sigma}$ is the estimated $(2NT \times 2NT)$ matrix, with ij-th element given by $\widehat{\sigma}_{ij} = \frac{e_i' e_j}{T}$, and e_i is a $(T \times 1)$ vector containing the residuals of the i-th equation estimated by OLS. By assumption A3, it follows that $[\sigma_\epsilon^2 I_{2N} + \sigma_e^2 J_{2N} - \widehat{\Sigma}] > 0 \therefore [\gamma_1 I_{2N} + \frac{\gamma_3-\gamma_1}{2N}J_{2N} - \widehat{\Sigma}] > 0$ $\therefore (\Phi_1 - \widehat{\Sigma}) > 0 \therefore (\widehat{\Sigma}^{-1} - \Phi_1^{-1}) > 0 \therefore X'\{(\widehat{\Sigma}^{-1} - \Phi_1^{-1}) \otimes I_T)\}X > 0 \therefore X'\{(\widehat{\Sigma}^{-1} \otimes I_T) - (\Phi_1^{-1} \otimes I_T)\}X > 0 \therefore$ $X'\{(\widehat{\Sigma} \otimes I_T)^{-1} - (\Phi_1 \otimes I_T)^{-1}\}X > 0 \therefore \Big(X'(\widehat{\Sigma} \otimes I_T)^{-1}X\Big) - \big(X'(\Phi_1 \otimes I_T)^{-1}X\big) > 0 \therefore \big(X'(\Phi_1 \otimes I_T)^{-1}X\big)^{-1} - \Big(X'(\widehat{\Sigma} \otimes I_T)^{-1}X\Big)^{-1} > 0 \therefore \big(X'\Omega^{-1}X\big)^{-1} - \widehat{Var}(\beta_{SUR}) > 0 \therefore Var(\beta_{SUR}) - E[\widehat{Var}(\beta_{SUR}) \mid X] > 0$ $\therefore B_{SUR} > 0$.

Now, we must show that $(B_{OLS} - B_{SUR}) \geqslant 0$. First, define $C \equiv [Var(\beta_{ols}) - Var(\beta_{sur})]$ and $D \equiv [E(\widehat{Var}(\beta_{sur}) \mid X) - E(\widehat{Var}(\beta_{ols}) \mid X)$. In other words, we must prove that $(C + D) \geqslant 0$, i.e., $C \geqslant 0$ and $D \geqslant 0$. The first part is just the well-known result that the GLS estimator will in most cases be more efficient, and will never be less efficient than the OLS estimator. Regarding the second part, define $\lambda = \frac{2NT-k}{tr[M\Omega]}$. Thus, it follows that $D \equiv [E(\widehat{Var}(\beta_{sur}) \mid X) - E(\widehat{Var}(\beta_{ols}) \mid X) \geqslant 0 \Leftrightarrow \Big(X'(\widehat{\Sigma} \otimes I_T)^{-1}X\Big)^{-1} - (X'\lambda I X)^{-1} \geqslant 0$ $\Leftrightarrow (X'\lambda I X) - \Big(X'(\widehat{\Sigma} \otimes I_T)^{-1}X\Big) \geqslant 0 \Leftrightarrow X'[\lambda I - (\widehat{\Sigma} \otimes I_T)^{-1}]X \geqslant 0 \Leftrightarrow \lambda I - (\widehat{\Sigma} \otimes I_T)^{-1} \geqslant 0 \Leftrightarrow (\widehat{\Sigma} \otimes I_T) - \frac{1}{\lambda}(I_{2N} \otimes I_T) \geqslant 0 \Leftrightarrow (\widehat{\Sigma} - \frac{1}{\lambda}I_{2N}) \otimes I_T \geqslant 0 \Leftrightarrow (\widehat{\Sigma} - \frac{1}{\lambda}I_{2N}) \geqslant 0$. By assumption A5, we have that $\widehat{\Sigma} \geqslant 0$ and $\Psi \geqslant \frac{2N(2NT-k)}{(tr((X'X)^{-1}X'JX)-2NT)+2N(2NT-k)}$. Now, if we define $\pi \equiv (tr(X(X'X)^{-1}X'J) - 2NT)/(2NT-k)$, it follows that $(\pi + 2N) = \frac{(tr(X(X'X)^{-1}X'J)-2NT)+2N(2NT-k)}{(2NT-k)} \therefore \frac{2N}{(\pi+2N)} = \frac{2N(2NT-k)}{(tr(X(X'X)^{-1}X'J)-2NT)+2N(2NT-k)}$ and, thus, $\Psi \geqslant \frac{2N}{(\pi+2N)}$. This way, $2N(\sigma_\epsilon^2 + 2N\sigma_e^2) \leqslant 2N\sigma_e^2(\pi + 2N) \therefore (\sigma_\epsilon^2 + 2N\sigma_e^2) \leqslant \sigma_e^2(\pi + 2N) \therefore \sigma_\epsilon^2 + 2N\sigma_e^2 \leqslant \pi\sigma_e^2 + 2N\sigma_e^2 \therefore \sigma_\epsilon^2 \leqslant \pi\sigma_e^2 \therefore \sigma_\epsilon^2(2NT - k) \leqslant \sigma_e^2(tr(X(X'X)^{-1}X'J) - 2NT) \therefore \sigma_\epsilon^2(2NT - k) + \sigma_e^2(2NT - tr(X(X'X)^{-1}X'J)) \leqslant 0 \therefore 2NT\sigma_\epsilon^2 + \sigma_e^2 2NT - \sigma_\epsilon^2 k - \sigma_e^2 tr(X(X'X)^{-1}X'J) \leqslant 0 \therefore 2NT\gamma_1 + (\frac{\gamma_3-\gamma_1}{2N})tr(J) - \gamma_1 tr(X(X'X)^{-1}X') - (\frac{\gamma_3-\gamma_1}{2N})tr(X(X'X)^{-1}X'J) \leqslant 0 \therefore tr\{\gamma_1 I_{2NT} + \frac{\gamma_3-\gamma_1}{2N}J - X(X'X)^{-1}X'(\gamma_1 I_{2NT}) - X(X'X)^{-1}X'(\frac{\gamma_3-\gamma_1}{2N})J\} \leqslant 0 \therefore tr\{\gamma_1 I_{2N} \otimes I_T + \frac{\gamma_3-\gamma_1}{2N}J_{2N} \otimes I_T - X(X'X)^{-1}X'([\gamma_1 I_{2N} \otimes I_T] + [\frac{\gamma_3-\gamma_1}{2N}J_{2N} \otimes I_T)\} \leqslant 0 \therefore tr\{(\gamma_1 I_{2N} + \frac{\gamma_3-\gamma_1}{2N}J_{2N}) \otimes I_T - X(X'X)^{-1}X'([\gamma_1 I_{2N} + \frac{\gamma_3-\gamma_1}{2N}J_{2N}] \otimes I_T)\} \leqslant 0 \therefore tr\{(I - X(X'X)^{-1}X')([\gamma_1 I_{2N} + \frac{\gamma_3-\gamma_1}{2N}J_{2N}] \otimes I_T)\} \leqslant 0 \therefore tr[M\Omega] \leqslant 0$. Therefore, by A2(ii), $T > 2p \therefore (2NT - k) > 0$, and, thus, $\lambda = \frac{2NT-k}{tr[M\Omega]} \leqslant 0 \therefore (\widehat{\Sigma} - \frac{1}{\lambda}I_{2N}) \geqslant 0 \therefore (C + D) \geqslant 0$;

(iii) By assumption A4, it follows that $\sigma_\epsilon^2 I_{2N} + \sigma_e^2 J_{2N} = \widehat{\Sigma}$ or $\Phi_1 = \widehat{\Sigma}$; where $\Phi_1 \equiv \sigma_\epsilon^2 I_{2N} + \sigma_e^2 J_{2N}$. This way, $(\Phi_1 \otimes I_T) = (\widehat{\Sigma} \otimes I_T) \therefore X'(\Phi_1 \otimes I_T)^{-1}X = X'(\widehat{\Sigma} \otimes I_T)^{-1}X \therefore E[(X'(\Phi_1 \otimes I_T)^{-1}X)^{-1} \mid X] = E[(X'(\widehat{\Sigma} \otimes I_T)^{-1}X)^{-1} \mid X] \therefore (X'(\Phi_1 \otimes I_T)^{-1}X)^{-1} = E[(X'(\widehat{\Sigma} \otimes I_T)^{-1}X)^{-1} \mid X] \therefore Var(\beta_{SUR}) = E[\widehat{Var}(\beta_{SUR}) \mid X] \therefore B_{SUR} = 0$;

(iv) If $\sigma_v^2 = \sigma_e^2 = 0$, then, $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \sigma_\epsilon^2$; and the covariance matrix is simplified to $\Omega = \sigma_\epsilon^2(I_{2NT})$, which is a diagonal matrix. By assumption A4, it follows that $\widehat{\Sigma} = \sigma_\epsilon^2 I_{2N}$, which is a diagonal matrix. According to Wooldridge (2002, Theorem 7.5), if $\widehat{\Omega} = \widehat{\Sigma} \otimes I_T$ is a diagonal matrix, then OLS equation by equation is identical to Feasible GLS. In addition, it follows that $(\sigma_\epsilon^2 I_{2N} \otimes I_T) = (\widehat{\Sigma} \otimes I_T)$ $\therefore X'(\sigma_\epsilon^2 I_{2N} \otimes I_T)^{-1}X = X'(\widehat{\Sigma} \otimes I_T)^{-1}X \therefore E[(X'(\sigma_\epsilon^2(I_{2NT}))^{-1}X)^{-1} \mid X] = E[(X'(\widehat{\Sigma} \otimes I_T)^{-1}X)^{-1} \mid X]$

$\therefore (X'\Omega^{-1}X)^{-1} = E[(X'(\widehat{\Sigma} \otimes I_T)^{-1}X)^{-1} \mid X] \therefore Var(\beta_{SUR}) = E[\widehat{Var}(\beta_{SUR}) \mid X] \therefore B_{SUR} = 0.$ ■

Proof of Proposition 4. (ia) Following Greene (2003, p. 207), since $\Omega > 0$ and $\Omega' = \Omega$, then, it can be factored into $\Omega = C\Lambda C'$, where the columns of C are the eigenvectors of Ω, and Λ is a diagonal matrix with the eigenvalues of Ω. Let $\Lambda^{1/2}$ be a diagonal matrix with i-th diagonal element $\sqrt{\lambda_i}$, and let $T = C\Lambda^{1/2}$. Therefore, $\Omega = TT'$; and if we define $P' = C\Lambda^{-1/2}$, then, $\Omega^{-1} = P'P$. This way, pre-multiplying the considered model $Y = X\beta + \varepsilon$ by P, it follows that $PY = PX\beta + P\varepsilon$ or $Y_* = X_*\beta + \varepsilon_*$, in which $E(\varepsilon_*\varepsilon_*' \mid X) = P\Omega P'$. By considering (at this point) that Ω is known, it follows that Y_* and X_* are observed data. The FB estimator of the transformed model is given by $\widehat{\beta}_{FB} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y = (X_*'X_*)^{-1}X_*'Y_* = (X_*'X_*)^{-1}X_*'(X_*\beta + \varepsilon_*) = \beta + (X_*'X_*)^{-1}X_*'\varepsilon_*$. Thus, $E(\widehat{\beta}_{FB} \mid X_*) = \beta + E[(X_*'X_*)^{-1}X_*'\varepsilon_* \mid X_*] = \beta$ if $E(\varepsilon_* \mid X_*) = E(P\varepsilon \mid PX) = 0$. However, note that since P is a matrix of known constants, we in fact just required that $E(\varepsilon \mid X) = 0$. By applying the error decomposition $\varepsilon_{i,t} = v_i + e_t + \epsilon_{i,t}$, it follows that $E(\varepsilon_{i,t} \mid X) = E(v_i \mid X) + E(e_t \mid X) + E(\epsilon_{i,t} \mid X)$. Now, recall that vector X is composed of lagged values of $\Delta Z_{i,t}$ and $CA_{i,t}$, which could generate a correlation between $\varepsilon_{i,t}$ and these lagged variables through the v_i random effect. However, since $E(v_i) = 0$ and, by assumption $\sigma_v^2 = 0$, it follows that $E(\varepsilon \mid X) = 0$ indeed holds, since e_t and $\epsilon_{i,t}$ are assumed to be uncorrelated with the regressors;

(ib) By considering $T \to \infty$ with fixed N, note that $plim(\widehat{\beta}_{FB}) = \beta + plim((X_*'X_*)^{-1}X_*'\varepsilon_*) = \beta + plim((\frac{X'\Omega^{-1}X}{T})^{-1}\frac{X'\Omega^{-1}\varepsilon}{T}) = \beta + Q_*^{-1}plim(\frac{X'\Omega^{-1}\varepsilon}{T})$, since by assumption A6 we have that $plim[(1/T)X'\Omega^{-1}X]^{-1} = Q_*^{-1}$. Thus, in order to obtain consistency, and apply the product rule of White (1984, Lemma 4.6), we need to show that $plim(\frac{X'\Omega^{-1}\varepsilon}{T}) = 0$. Following Greene (2003, p. 67), let $\frac{X'\Omega^{-1}\varepsilon}{T} = \frac{X_*'\varepsilon_*}{T} = \frac{1}{T}\sum_{t=1}^{T} w_t \equiv \overline{w}$, which is a $k \times 1$ vector; where $w_t \equiv \{\phi(L)(\Delta Z_t^i + CA_t^i)\sum_{j=1}^{2NT} s_{i,j}\varepsilon_{j,t}\}$; $\phi(L)$ is a lag polynomial of order p; and $s_{i,j}$ is the ij-th element of Ω^{-1}, which is a $(2NT \times 2NT)$ matrix. This way, $plim(\widehat{\beta}_{FB}) = \beta + Q_*^{-1}plim(\overline{w})$. Now, note that $E[w_t] = E(E[w_t \mid X]) = E(E[\{\phi(L)(\Delta Z_t^i + CA_t^i)\sum_{j=1}^{2NT} s_{i,j}\varepsilon_{j,t}\} \mid X]) = E(\phi(L)(\Delta Z_t^i + CA_t^i)\sum_{j=1}^{2NT} s_{i,j}E[\varepsilon_{j,t} \mid X]) = 0$, since from the exogeneity assumptions, it follows that $E(\varepsilon_{j,t} \mid X) = 0; \forall j, t \therefore E[w_t] = 0$ and, thus, $E[\overline{w}] = 0$. On the other hand, note that $Var(\overline{w}) = E(Var(\overline{w} \mid X)) + Var(E(\overline{w} \mid X))$, where the second term is zero, since $E(\varepsilon_{i,t} \mid X) = 0$ and, thus, $E(\overline{w} \mid X) = \frac{1}{T}\sum_{t=1}^{T} E(w_t \mid X) = \frac{1}{T}\sum_{t=1}^{T} E(\{\phi(L)(\Delta Z_t^i + CA_t^i)\sum_{j=1}^{2NT} s_{i,j}\varepsilon_{j,t}\} \mid X) = \frac{1}{T}\sum_{t=1}^{T}\phi(L)(\Delta Z_t^i + CA_t^i)\sum_{j=1}^{2NT} s_{i,j}E(\varepsilon_{j,t} \mid X) = 0$. Now, note that $Var(\overline{w} \mid X) = E(\overline{w}\overline{w}' \mid X) = E[(\frac{1}{T}\sum_{t=1}^{T} w_t)(\frac{1}{T}\sum_{t=1}^{T} w_t') \mid X] = E[(\frac{1}{T}X_*'\varepsilon_*)(\frac{1}{T}\varepsilon_*'X_*) \mid X] = \frac{X_*'}{T}E(\varepsilon_*\varepsilon_*' \mid X_*)\frac{X_*}{T} = \frac{X_*'}{T}E(P\varepsilon\varepsilon'P' \mid X)\frac{X_*}{T} = \frac{X_*'}{T}PE(\varepsilon\varepsilon' \mid X)P'\frac{X_*}{T} = \frac{X_*'}{T}P\Omega P'\frac{X_*}{T} = \frac{X'}{T}P'P\Omega P'P\frac{X}{T} = \frac{X'}{T}\Omega^{-1}\Omega\Omega^{-1}\frac{X}{T} = \frac{X'}{T}\Omega^{-1}\frac{X}{T}$. Thus, it follows that $Var(\overline{w}) = E(Var(\overline{w} \mid X)) = E(\frac{X'}{T}\Omega^{-1}\frac{X}{T} \mid X) = \frac{1}{T}E(\frac{X'\Omega^{-1}X}{T} \mid X)$. The variance of $\overline{w}$ will collapse to zero if the conditional expectation in parentheses converges to a constant matrix, so that the leading scalar $(1/T)$ will dominate the product as T increases. Thus, by considering assumption A6 (and a WLLN), it follows that $plim[Var(\overline{w})] = plim(\frac{1}{T})Q_* = 0$. Therefore, since $E[\overline{w}] = 0$ and $plim[Var(\overline{w})] = 0 \therefore plim(\overline{w}) = 0$ and, thus, $plim(\widehat{\beta}_{FB}) = \beta$;

(ii) Given that $\sigma_v^2 = 0$ and assumption A7 holds, it follows that $\varepsilon_{i,t} = e_t + \epsilon_{i,t}$ is also normally distributed, i.e., $\varepsilon \sim N(0,\Omega)$, where $E(\varepsilon\varepsilon')=\Omega$ is finite and nonsingular. Therefore, by applying White (1984, Theorem 1.3), the result is straightforward: $\widehat{\beta}_{FB} \sim N(\beta;(X'\Omega^{-1}X)^{-1})$;

(iii) By assumptions A6-A7, and Theorem 1 of Fuller & Battese (1974), it follows that (when $T \to \infty$ and N is fixed), $\widehat{\sigma}_e^2$ and $\widehat{\sigma}_\epsilon^2$ are consistent estimators of σ_e^2 and σ_ϵ^2 respectively. Thus, $\widehat{\Omega} = \widehat{\sigma}_\epsilon^2 I_{2NT} + \widehat{\sigma}_e^2(J_{2N} \otimes I_T)$ is asymptotically equivalent to Ω. Based on assumption A8, we can apply the asymptotic equivalence Lemma 4.7 of White (1984), in which $\widehat{\widehat{\beta}}_{FB}$ is asymptotically equivalent to $\widehat{\beta}_{FB}$. See also Greene (2003, p. 210) for further details. ■

Appendix C. Further results of the empirical exercise

Table 10 - ADF Unit Root test

	USA	CAN	JPN	UK	GER	ITA	FRA
C_t	-0.99	-2.19	-1.13	-1.17	-2.43	-2.04	-1.32
ΔC_t	-5.44 (**)	-14.02 (**)	-16.65 (**)	-15.76 (**)	-15.25 (**)	-9.78 (**)	-16.18 (**)
GNI_t^*	-1.51	-1.23	-1.85	-1.34	-1.15	-0.52	-2.14
ΔGNI_t^*	-6.08 (**)	-16.74 (**)	-14.58 (**)	-20.10 (**)	-12.01 (**)	-5.58 (**)	-15.62 (**)
CA_t	-3.20 (*)	-2.58 (+)	-3.21 (*)	-2.82 (+)	-2.02	-1.87	-1.24
ΔZ_t	-14.29 (**)	-15.41 (**)	-14.68 (**)	-18.61 (**)	-13.35 (**)	-14.11 (**)	-17.32 (**)

Notes: a) $GNI_t^* \equiv Y_t + rB_t - I_t - G_t$.

b) (**) indicates rejection of the null hypothesis of unit root at 1% level; (*) at 5% level and (+) at 10% level

c) $CA_t = Y_t + rB_t - I_t - G_t - \theta C_t$

d) The USA, CAN, JPN, UK and GER series range from 1960:1-2007:1, whereas ITA and FRA range from 1980:1-2007:1.

Table 11 - Johansen´s Cointegration Test

The cointegration vector between $(Y_t + rB_t - I_t - G_t)$ and C_t is given by $(1, -\theta)$

		Ho:	p=0	Ho:	p≤1
Country	θ	trace stat .	$\lambda_{máx.}$ stat.	trace stat.	$\lambda_{máx.}$ stat.
USA	0.895 (0.018)	13.72	10.63	3.09	3.09
CAN	1.024 (0.049)	9.54	8.41	1.13	1.13
JPN	1.105 (0.016)	15.06	12.83	2.23	2.23
UK	0.921 (0.026)	13.59	10.89	2.70	2.70
GER	1.114 (0.077)	5.61	4.05	1.56	1.56
ITA	1.120 (0.058)	13.16	9.23	3.93 (*)	3.93 (*)
FRA	1.267 (0.149)	4.64	4.60	0.03	0.03

Notes: a) p is the number of cointegrating relations; (*) indicates rejection of Ho at 5% level.

b) In column θ the standard deviation is presented in parentheses.

c) The USA, CAN, JPN, UK and GER series range from 1960:1-2007:1, whereas ITA and FRA range from 1980:1-2007:1.

Table 12 - Comparison of θ with the literature

Country	IG	Ghosh (95)	Agénor et al.(99)
USA	0.895	0.994	-
CAN	1.024	0.96	-
JPN	1.105	1.04	-
UK	0.921	0.98	-
GER	1.114	1.08	-
ITA	1.120	-	-
FRA	1.267	-	0.982

Note: IG means Issler & Gaglianone (our results).

Table 13 - Wald test (p-value) - Per capita time series, in 2000 U.S. dollars

Country	OLS	SUR	FB.1	FB.2
$USA_{(1)}$	0.1369	0.0201 (*)	0.0351 (*)	0.0586
$CAN_{(1)}$	0.3238	0.1430	0.0620	0.0734
$JPN_{(1)}$	0.0042 (**)	0.0001 (**)	0.0002 (**)	0.0004 (**)
$UK_{(1)}$	0.1685	0.1049	0.7692	0.7088
$GER_{(1)}$	0.1363	0.2351	0.6118	0.6690
$ITA_{(1)}$	0.1596	0.0301 (*)	0.9850	0.9652
$FRA_{(1)}$	0.0078 (**)	0.0000 (**)	0.0149 (*)	0.0139 (*)
$USA_{(2)}$	0.0256 (*)	0.0039 (**)	0.0120 (*)	0.0158 (*)
$CAN_{(2)}$	0.1932	0.0531	0.0509	0.2274
$JPN_{(2)}$	0.0015 (**)	0.0001 (**)	0.0000 (**)	0.0019 (**)
$UK_{(2)}$	0.0078 (**)	0.0001 (**)	0.5571	0.2576
$GER_{(2)}$	0.0174 (*)	0.0096 (**)	0.4876	0.1167

Notes: a) Ho: $(CA_t^* - \mu_{CA}^*) = (CA_t - \mu_{CA})$

b) (**) means rejection at 1% level, (*) at 5% level;

c) $USA_{(1)}$ indicates sample 1 (1980:1-2007:1), and $USA_{(2)}$ means sample 2 (1960:1-2007:1);

d) FB.1 means Fuller & Battese (1974) error decomposition, with a unique global shock, whereas FB.2 considers distinct global shocks for CA_t and ΔZ_t equations.

Table 14 - Wald test (p-value) - Time series in U.S. dollars (not per capita)

Country	OLS	SUR	FB.1	FB.2
$USA_{(1)}$	0.1400	0.0207 (*)	0.0340 (*)	0.0394 (*)
$CAN_{(1)}$	0.2402	0.0312 (*)	0.8692	0.8766
$JPN_{(1)}$	0.0291 (*)	0.0205 (*)	0.0032 (**)	0.0051 (**)
$UK_{(1)}$	0.1796	0.1550	0.9970	0.9979
$GER_{(1)}$	0.0480 (*)	0.0097 (**)	0.6667	0.6898
$ITA_{(1)}$	0.1425	0.0789	0.9771	0.9594
$FRA_{(1)}$	0.0295 (*)	0.0000 (**)	0.7957	0.7525
$USA_{(2)}$	0.0291 (*)	0.0069 (**)	0.0000 (**)	0.0000 (**)
$CAN_{(2)}$	0.1499	0.0123 (*)	0.8288	0.8361
$JPN_{(2)}$	0.0201 (*)	0.0160 (*)	0.0363 (*)	0.3019
$UK_{(2)}$	0.0139 (*)	0.0001 (**)	0.8727	0.6861
$GER_{(2)}$	0.0143 (*)	0.0078 (**)	0.7161	0.3260

Notes: a) Ho: $(CA_t^* - \mu_{CA}^*) = (CA_t - \mu_{CA})$

b) (**) means rejection at 1% level, (*) at 5% level;

c) $USA_{(1)}$ indicates sample 1 (1980:1-2007:1), and $USA_{(2)}$ means sample 2 (1960:1-2007:1);

d) FB.1 means Fuller & Battese (1974) error decomposition, with a unique global shock, whereas FB.2 considers distinct global shocks for CA_t and ΔZ_t equations.

Appendix D. Some details of the Monte Carlo simulation

Generating a covariance-stationary VAR

An initial idea to design the Monte Carlo experiment could consist on constructing the companion matrix, sorting it values from uniform distributions, in order to satisfy the restrictions imposed by the null hypothesis, and then verifying whether or not the eigenvalues of the companion matrix all lie inside the unit circle. However, this strategy could lead to a wide spectrum of search for adequate values for the companion matrix. This way, we propose an analytical solution to generate a covariance-stationary VAR, based initially on the choice of the eigenvalues, and then on the generation of the respective companion matrix. According to Hamilton (1994, page 259), if the eigenvalues of the companion matrix (F) all lie inside the unit circle, then the VAR turns out to be covariance-stationary. The eigenvalue vector (λ) of the companion matrix (F) for a VAR(p) is obtained from the following equation:

$$|F - \lambda I| = 0, \tag{39}$$

where I is the identity matrix. Two important properties of the eigenvalue vector, presented in Simon&Blume (1994, page 599), are reproduced below:

$$trace(F) = \sum_{i=1}^{n} \lambda_i \quad ; \quad \det(F) = \prod_{i=1}^{n} \lambda_i, \tag{40}$$

where n is the number of eigenvalues (equal to $2p$). The companion matrix for a VAR(1) is given by

$$F = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \tag{41}$$

The restrictions imposed to a VAR(1) to consider the optimal current account equal to the observed one (i.e., a true null hypothesis) are given by $c = a$; and $d = b + (1 + r)$. This way, in order to impose a false null hypothesis into the model we multiply the restrictions above by a gamma factor, resulting in the following companion matrix

$$F = \begin{bmatrix} a & b \\ \gamma a & \gamma(b + (1 + r)) \end{bmatrix}. \tag{42}$$

It should be mentioned that setting the gamma factor equal to unity we consider a true Ho, but imposing gamma less than unity we generate a false null hypothesis. This way, the eigenvalues of the companion matrix are given by

$$\det(F) = \gamma a(b + 1 + r) - \gamma ab = \lambda_1 \lambda_2 \quad \therefore \quad a = \frac{\lambda_1 \lambda_2}{\gamma(1 + r)} \tag{43}$$

$$trace(F) = a + \gamma(b + (1 + r)) = \lambda_1 + \lambda_2 \quad \therefore \quad b = \frac{\lambda_1 + \lambda_2 - a - \gamma(1 + r)}{\gamma}. \tag{44}$$

Therefore, to construct a covariance-stationary VAR(1) we sort from a uniform distribution (-1;1) the values of λ_1 and λ_2 and calculate the parameters a and b from the equations above. In a VAR(2) case, the companion matrix (considering the restrictions implied by the null hypothesis) is given by

$$F = \begin{bmatrix} a_1 & a_2 & b_1 & b_2 \\ 1 & 0 & 0 & 0 \\ \gamma a_1 & \gamma a_2 & \gamma(b_1 + (1+r)) & \gamma b_2 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{45}$$

The eigenvalue vector is obtained from

$$|F - \lambda I| = \begin{vmatrix} a_1 - \lambda & a_2 & b_1 & b_2 \\ 1 & -\lambda & 0 & 0 \\ \gamma a_1 & \gamma a_2 & \gamma(b_1 + (1+r)) - \lambda & \gamma b_2 \\ 0 & 0 & 1 & -\lambda \end{vmatrix} = 0. \tag{46}$$

Since $\det(F) = 0$, at least one eigenvalue is null ($\lambda_1 = 0$). Thus, from the last equation, it follows that

$$|F - \lambda I| = \lambda^4 - \lambda^3(a_1 + \gamma b_1 + \gamma(1+r)) + \lambda^2(\gamma a_1(1+r) - a_2 - \gamma b_2) + \lambda\gamma a_2(1+r) = 0 \tag{47}$$

$$\therefore a_2 = \frac{\lambda(\gamma a_1(1+r) - \gamma b_2 - \lambda(a_1 + \gamma b_1 + \gamma(1+r)) + \lambda^2)}{\lambda - \gamma(1+r)}. \tag{48}$$

The equation above must be valid for all values of λ_i. In particular, one can construct a system of 2 equations, with the expression above, setting $\lambda = \lambda_2$ and $\lambda = \lambda_3$. This way, we can explicit b_2 as a function of a_1, b_1, r, λ_2 and λ_3, as it follows:

$$\begin{aligned} b_2 = {} & ((-\lambda_2\gamma^2 + \gamma^2 a_1 - \lambda_3\gamma^2)r^2 + (\lambda_2^2\gamma + 2\lambda_2\gamma\lambda_3 - \lambda_2\gamma^2 b_1 - \lambda_3\gamma^2 b_1 - a_1\lambda_2\gamma - 2\lambda_2\gamma^2 + 2\gamma^2 a_1 \\ & -a_1\lambda_3\gamma + \lambda_3^2\gamma - 2\lambda_3\gamma^2)r + 2\lambda_2\gamma\lambda_3 - \lambda_2^2\lambda_3 + \lambda_2\gamma b_1\lambda_3 + a_1\lambda_2\lambda_3 - \lambda_3^2\lambda_2 + \lambda_2^2\gamma - \lambda_2\gamma^2 b_1 \\ & -\lambda_3\gamma^2 b_1 - a_1\lambda_2\gamma + \lambda_3^2\gamma - \lambda_3\gamma^2 - a_1\lambda_3\gamma + \gamma^2 a_1 - \lambda_2\gamma^2)/((1+r)\gamma^2). \end{aligned} \tag{49}$$

On the other hand, the trace of the companion matrix is given by

$$trace(F) = a_1 + \gamma(b_1 + (1+r)) = \lambda_2 + \lambda_3 + \lambda_4 \tag{50}$$

$$\therefore b_1 = \frac{(\lambda_2 + \lambda_3 + \lambda_4) - a_1 - \gamma(1+r)}{\gamma}. \tag{51}$$

Therefore, to construct a covariance-stationary VAR(2) we set $\lambda_1 = 0$ and sort (independently) from uniform distributions (-1;1) the values of a_1, λ_2, λ_3 and λ_4. Then, we choose a gamma factor in order to simulate the false (or true) null hypothesis. Thus, we obtain b_1 by the last expression and calculate b_2 and a_2 from the previous equations. This way, we construct the companion matrix with all eigenvalues inside the unit circle. A VAR(p) can be constructed in a similar way by following the presented methodology.

Besides the proper construction of the companion matrix, another important issue of the Monte Carlo simulation is the generation of the residuals for the VAR. Since the current accounts for different countries are nowadays expected to be globally linked, we construct the residuals of the VAR based on two components: an idiosyncratic shock, and a global shock, which is assumed to be common among the considered countries. The construction of these residuals is detailed in next section.

Constructing Global Shocks

A key issue regarding the Monte Carlo experiment is the generation of the residuals for the VAR. Inspired by Glick & Rogoff (1995), which study the current account response to different productivity shocks in the G-7 with a structural model including global and country-specific shocks, and based on the 'common sense' that the current account fluctuations have become more closely linked across countries in the last decades, we decompose the residuals of the VAR into an idiosyncratic shock and a global (common) component among countries. Thus, the VAR(p) for a country i can be written as:

$$\begin{bmatrix} \Delta Z_t^i \\ CA_t^i \end{bmatrix} = \begin{bmatrix} a^i(L) & b^i(L) \\ c^i(L) & d^i(L) \end{bmatrix} \times \begin{bmatrix} \Delta Z_{t-1}^i \\ CA_{t-1}^i \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t}^i \\ \varepsilon_{2t}^i \end{bmatrix}, \tag{52}$$

where the residuals ε_t^i can be decomposed into an idiosyncratic shock (ν_t^i) and a global (common) component among countries (ξ_t), as it follows:

$$\begin{bmatrix} \varepsilon_{1t}^i \\ \varepsilon_{2t}^i \end{bmatrix} = \begin{bmatrix} \nu_{1t}^i \\ \nu_{2t}^i \end{bmatrix} + scale * \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \end{bmatrix}, \tag{53}$$

or in a reduced form: $\varepsilon_t^i = \nu_t^i + scale * \xi_t$. The parameter *scale* is used to measure the importance of the global shock into the residuals of a country i. The numerical procedure adopted to construct the residuals in the Monte Carlo experiment first drops (independently) from a normal standard distribution the ν_{1t}^i and ν_{2t}^i series of shocks, resulting on a I_2 covariance matrix:

$$\Sigma^i = Var(\nu_t^i) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{54}$$

However, this covariance matrix does not represent a bivariate shock, and we must transform the covariance matrix Σ^i into a covariance matrix Ω^i, in a framework of a bivariate normal distribution, with parameters as close as possible to the data of the countries.

$$\Omega^i = \begin{bmatrix} \sigma_{1i}^2 & r^i \\ r^i & \sigma_{2i}^2 \end{bmatrix} \tag{55}$$

Thus, we must find a symmetric matrix X to make the following transformation: $\Omega = X\Sigma X'$. In our case, Σ is a diagonal and symmetric matrix, as it follows:

$$\therefore \Omega = X\Sigma^{1/2}\Sigma^{1/2}X' = \Sigma^{1/2}XX\Sigma^{1/2} \therefore XX = \Sigma^{-1/2}\Omega\Sigma^{-1/2}. \tag{56}$$

To obtain X we must calculate the square root of the matrix (XX). Adopting the eigenvalue decomposition, according to Simon&Blume (1994),[27] one could rewrite the (XX) matrix as a function of the eigenvectors (V) and the eigenvalue matrix (D), as it follows: $XX = VDV^{-1} = VD^{1/2}D^{1/2}V^{-1} = (VD^{1/2}V^{-1})(VD^{1/2}V^{-1})$, where D is a diagonal matrix filled with the eigenvalues of (XX). This way, the matrix X can be obtained

[27] For further details see Simon&Blume (1994), pages 590-595, and page 866 of Ruud (2000).

by $(VD^{1/2}V^{-1})$. Finally, after calculating X, one could use the constructed vector of residuals ν_t^i to obtain the $\widetilde{\nu}_t^i$ residuals (following a bivariate normal distribution with covariance matrix Ω^i):

$$\underset{\text{Tx2}}{\widetilde{\nu}_t} = \underset{\text{Tx2}}{\nu_t} \underset{\text{2x2}}{X} \tag{57}$$

where $Var(\widetilde{\nu}_t^i) = \Omega^i$, $Var(\nu_t^i) = \Sigma^i$, and T is the number of observations. The next step is to construct the global shocks, common to all countries. The procedure adopted in the Monte Carlo experiment drops (independently) from a normal standard distribution the ξ_{1t} and ξ_{2t} series of shocks, resulting in a covariance matrix equal to an I_2 matrix. However, this covariance matrix also does not represent a bivariate shock, and must be transformed, in the same way presented above (adopting the eigenvalue decomposition), into a covariance matrix Γ, representing a bivariate normal distribution:

$$\Gamma = Var(\xi_t) = \begin{bmatrix} 1 & w \\ w & 1 \end{bmatrix}, \tag{58}$$

where the parameter w represents the covariance between the global shocks on ΔZ_t^i and CA_t^i equations.

References

[1] Agénor, P.R., Bismut, C., Cashin, P., McDermott, C.J., 1999. Consumption smoothing and the current account: evidence for France, 1970-1996. Journal of International Money and Finance 18, 1-12.

[2] Aguiar, M., Gopinath, G., 2006. Defaultable debt, interest rates and the current account. Journal of International Economics 69, 64– 83.

[3] Anderson, T. W., Hsiao, C., 1981. Estimation of Dynamic models with Error Components. Journal of the American Statistical Association 76, 598-606.

[4] Andrews, D.W.K., 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. Econometrica 59, 817-858.

[5] Baltagi, B. H., 1980. On seemingly unrelated regressions with error components. Econometrica 48(6), 1547-1551.

[6] Baltagi, B.H., 1981. Pooling: An experimental study of alternative testing and estimation procedures in a two-way error component model. Journal of Econometrics 17, 21-49.

[7] Baltagi, B.H., 2001. Econometric Analysis of Panel Data, John Wiley & Sons Ltd.

[8] Bergin, P.R., 2003. Putting the 'New Open Economy Macroeconomics' to a test. Journal of International Economics 60, 3–34.

[9] Bergin, P.R., Sheffrin, S.M., 2000. Interest rates, exchange rates and present value models of the current account. The Economic Journal 110, 535-558.

[10] Campbell, J.Y., Shiller, R., 1987. Cointegration and tests of present value models. Journal of Political Economy 95, 1062-1088.

[11] Canova, F., Dellas, H., 1993. Trade interdependence and the international business cycle. Journal of International Economics 34, 23-47.

[12] Canova, F., Marrinan, F., 1998. Sources and Propagation of Internation Output Cycles: Common Shocks or Transmission? Journal of International Economics 46, 133–166.

[13] Centoni, M., Cubadda, G., Hecq, A., 2003. Common Shocks, Common Dynamics, and the International Business Cycle. Available at http://www.personeel.unimaas.nl/A.Hecq/CCH.pdf

[14] Chinn, M. D., Prasad, E.S., 2003. Medium-term determinants of current accounts in industrial and developing countries: an empirical exploration. Journal of International Economics 59, 47–76.

[15] Driscoll, J., Kraay, A., 1995. Spatial Correlations in Panel Data. The World Bank, Policy Research Working Paper 1553.

[16] ________, 1998. Consistent covariance matrix estimation with spatially dependent panel data. Review of Economics and Statistics 80, 549–560.

[17] Fuller, W.A. and Battese, G.E., 1974. Estimation of Linear Models with Crossed-Error Structure. Journal of Econometrics 2, 67-78.

[18] Ghosh, A.R., 1995. International capital mobility amongst the major industrialised countries: too little or too much? The Economic Journal 105, 107-128.

[19] Ghosh, A.R., Ostry, J.D., 1995. The current account in developing countries: a perspective from the consumption smoothing approach. World Bank Economic Review 9, 305-333.

[20] _______, 1997. Macroeconomic uncertainty, precautionary saving, and the current account. Journal of Monetary Economics 40, 121-139.

[21] Glick, R., Rogoff, K., 1995. Global versus country-specific productivity shocks and the current account. Journal of Monetary Economics 35, 159-192.

[22] Greene, W.H., 1993. Econometric Analysis. Macmillan Publishing Company.

[23] Grohé, S.S., Uribe, M., 2003. Closing small open economy models. Journal of International Economics 61, 163–185.

[24] Gruber, J.W., 2000. A present value test of habits and the current account. Manuscript, Johns Hopkins University.

[25] Hamilton, J.D., 1994. Time series analysis. Princeton University Press.

[26] Hausman, J.A., 1978. Specification Tests in Econometrics. Econometrica, 46, 1251-1271.

[27] Holtz-Eakin, D., Newey, W., Rosen, H.S., 1988. Estimating Vector Autoregressions with Panel Data. Econometrica 56(6), 1371-1395.

[28] Hsiao, C., 2003. Analysis of Panel Data. Cambridge University Press.

[29] Hussein, K.A., Mello, L.R., 1999. International capital mobility in developing countries: theory and evidence. Journal of International Money and Finance 18, 367-381.

[30] İşcan, T.B., 2002. Present value tests of the current account with durables consumption. Journal of International Money and Finance 21, 385-412.

[31] King, R.G., Plosser, C.I., Stock, J.H., Watson, M.W., 1991. Stochastic Trends and Economic Fluctuations. American Economic Review 81, 819-840.

[32] Maddala, G. S., 1971. The Use of Variance Components Models in Pooling Cross Section and Time Series Data. Econometrica 39, 341-358.

[33] Mantalos, P., 2000. A graphical investigation of the size and power of the Granger-causality tests in integrated-cointegrated VAR systems. Studies in Nonlinear Dynamics and Econometrics, vol. 4 (1). The MIT Press.

[34] Nason, J.M., Rogers, J.H., 2006. The present-value model of the current account has been rejected: Round up the usual suspects. Journal of International Economics 68, 159–187.

[35] Newey, W.K., West, K.D., 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55, 703-708.

[36] Obstfeld, M., Rogoff, K., 1994. The intertemporal approach to the current account. National Bureau of Economic Research, working paper 4893.

[37] _______ , 1996. Foundations of International Macroeconomics. MIT Press, Cambridge.

[38] Otto, G., 1992. Testing a present-value model of the current account: evidence from the U.S. and Canadian time series. Journal of International Money and Finance 11, 414-430.

[39] Pesaran, M.H., Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. Journal of Econometrics 68, 79-113.

[40] Phillips, P.C.B., Moon, H.R., 1999. Linear Regression Limit Theory for Nonstationary Panel Data. Econometrica 67(5), 1057-1112.

[41] Ruud, P.A., 2000. An introduction to classical econometric theory. Oxford University Press.

[42] Sachs, J., 1982. The current account in the macroeconomic adjustment process. Scandinavian Journal of Economics 84, 147-159.

[43] SAS Institute Inc. 1979. SAS Technical Report S-106, TSCSREG: A SAS Procedure for the Analysis of Time-Series Cross-Section Data, Cary, NC: SAS Institute Inc.

[44] Sheffrin, S.M., Woo, W.T., 1990. Present value tests of an intertemporal model of the current account. Journal of International Economics 29, 237-253.

[45] Simon, C.P., Blume, L., 1994. Mathematics for economists. W.W. Norton & Company Inc.

[46] Wallace, T. D., Hussain, A., 1969. The Use of Error Components Models in Combining Cross Section with Time Series Data. Econometrica 37, 55-72.

[47] White, H., 1984. Asymptotic Theory for Econometricians. Academic Press, New York.

[48] Zellner, A., 1962. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias. Journal of the American Statistical Association 57, 348–368.

Chapter 2

Debt ceiling and fiscal sustainability in Brazil: a quantile autoregression approach[1]

Abstract

In this paper we investigate fiscal sustainability by using a quantile autoregression (QAR) model. We propose a novel methodology to separate periods of nonstationarity from stationary ones, allowing us to identify various trajectories of public debt that are compatible with fiscal sustainability. We use such trajectories to construct a debt ceiling, that is, the largest value of public debt that does not jeopardize long-run fiscal sustainability. We make an out-of-sample forecast of such a ceiling and show how it could be used by Policy makers interested in keeping the public debt on a sustainable path. We illustrate the applicability of our results using Brazilian data.

JEL Classification: C22, E60, H60.
Keywords: Fiscal Policy, Debt Ceiling, Quantile Autoregression.

[1] This article was jointly made with Luiz Renato Lima and Raquel M. Sampaio. We would like to thank Zhijie Xiao and Roger Koenker for their insightful advice, and the participants of Econometrics in Rio (July, 2006) where this paper was presented . We are also grateful to Co-Editor Lant Pritchett and two anonymous referees, for their helpful comments and suggestions, and seminar participants at EPGE-FGV (August, 2006), specially Antônio Galvão, Caio Ibsen, Carlos E. da Costa, Fernando de Holanda Barbosa, Luis H. B. Braido, Maria Cristina Terra and Renato Fragelli Cardoso.

1 Introduction

For decades, a lot of effort has been devoted to investigating whether long-lasting budget deficits represent a threat to public debt sustainability. Hamilton and Flavin (1986) was one of the first studies to address this question testing for the non-existence of a Ponzi scheme in public debt. They conducted a battery of tests using data from the period 1962-84 with the assumption of a fixed interest rate. Their results indicate that the government's intertemporal budget constraint holds. In a later work, Wilcox (1989) extends Hamilton and Flavin's work by allowing for stochastic variation in the real interest rate. His focus was on testing for the validity of the present-value borrowing constraint, which means that public debt will be sustainable in a dynamically efficient economy[2] if the discounted public debt is stationary with an unconditional mean equal to zero.

An important and common feature in the aforementioned studies is the underlying assumption that economic time series possess symmetric dynamics. In recent years, considerable research effort has been devoted to studying the effect of different fiscal regimes on long-run sustainability of the public debt. When the public debt possesses a nonlinear dynamic, it may be sustainable in the long-run but can present episodes of unsustainability in the short-run. Indeed, some studies have reported the existence of short-run fiscal imbalances. For instance, Sarno (2001) uses a smooth transition autoregressive (STAR) model to investigate the U.S. debt-GDP ratio and states that the difficulty encountered in the literature in detecting mean reversion of the debt process may be due to the linear hypothesis commonly adopted in the testing procedures. According to the author, the U.S. debt-GDP ratio is well characterized by a nonlinearly mean reverting process, and governments respond more to primary deficits (surpluses) when public debt is particularly high (low).

More recently, Davig (2005) uses a Markov-switching time series model to analyze the behavior of the discounted U.S. federal debt. The author uses an extended version of Hamilton and Flavin (1986) and Wilcox (1989) data and identifies two fiscal regimes: in the first one, the discounted federal debt is expanding, whereas, it is collapsing in the second one. He concludes that although the expanding regime is not sustainable, it does not pose a threat to the long-run sustainability of the discounted U.S. federal debt. Arestis et al. (2004) consider a threshold autoregressive model and, by using quarterly deficit data from the period 1947:2 to 2002:1, they find evidence that the U.S. budget deficit is sustainable in the long-run, but that fiscal authorities only intervene to reduce budget deficits when they reach a certain threshold, deemed to be unsustainable.

A common finding in the studies of Sarno (2001), Arestis et al. (2004), and Davig (2005) is that the presence of a nonlinear dynamic in public debt permits the existence of short episodes in which public debt exhibits a nonsustainable behavior. Such short-run behavior, however, does not pose a threat to long-run sustainability. Therefore, there could be three possible paths for public debt: (i) long-run sustainable paths with episodes of fiscal imbalances; (ii) long-run sustainable paths without episodes of fiscal imbalances and;

[2] Abel et al. (1996) provides evidence that the U.S. economy is dynamically efficient.

(iii) long-run unsustainable paths. How can we identify and separate each of the aforementioned paths? This paper addresses this question by proposing a novel measurement of debt ceiling that can be used to guide fiscal-policy managers in their task of keeping public debt sustainable in the long run.

The methodology developed in this paper is based on the so called quantile autoregressive (QAR) model, introduced by Koenker and Xiao (2002, 2004a, 2004b). The QAR approach provides a way to directly examine how past information affects the conditional distribution of a time series. This feature of the QAR model is fundamental to the methodology proposed in this paper since our measurement of debt ceiling ($\widetilde{D}_t$) will be nothing more than the upper conditional quantile of the public debt that satisfies the transversality condition of a no-Ponzi game. Compared to the QAR approach, other non-linear methods such as smooth transition autoregressive (STAR), threshold autoregressive (TAR) or Markov switching are not able to estimate conditional quantiles since they were originally proposed to estimate nonlinear models for conditional means (or variance).

The proposed measurement of debt ceiling has the following main feature: if public debt y_t has non-stationary behavior at time $t = t_A$, then $y_t > \widetilde{D}_t$ at $t = t_A$, otherwise $y_t \leq \widetilde{D}_t$. We also estimate $H \equiv \frac{1}{T}\sum_t I_t(y_t > \widetilde{D}_t)$, where $I(.)$ is an indicator function and T is the sample size, representing the percentage of periods in which public debt had an (local) unsustainable behavior. There are, therefore, two important issues we want to address in this paper. Firstly, how to identify $\widetilde{D}_t$ and, consequently, H? With this information in hand, the policy maker can evaluate whether a given fiscal policy is at risk, that is, if y_t is above $\widetilde{D}_t$, or whether it is sustainable but too austere, in the sense that y_t is too far below $\widetilde{D}_t$. Secondly, how to make multi-step-ahead forecasts of the debt ceiling? A decision maker (fiscal authority) can use such a forecast to decide whether or not to take some action against long-run unsustainable paths of public debt.

The methodology developed in this paper complements the study by Garcia and Rigobon (2004), which proposed a very attractive technique to study debt sustainability from a risk management perspective by using a Value at Risk (VaR) approach based on Monte Carlo simulations. However, in their article, the choice of the quantile needed to compute the "risky" threshold of sustainability for public debt was somewhat arbitrary. The methodology proposed in this paper goes beyond their approach by computing the exact quantile, the so-called critical quantile, that is used to separate sustainable fiscal policies from unsustainable ones. Therefore, our measurement of debt ceiling can be viewed as a more elaborated concept of VaR in the sense that it appropriately uses economic theory to identify the quantile needed to compute the "risky" threshold, rather than choosing it arbitrarily.

We illustrate the applicability of our debt ceiling measurement by using data from the Brazilian public debt. Fiscal stabilization in Latin American countries, and especially in Brazil, has received a lot of attention over the last decade. In effect, Issler and Lima (2000) showed that public debt sustainability in Brazil from 1947 to 1992 was achieved mostly through the usage of revenue from seigniorage. However, after the Brazilian stabilization plan in 1994, this source of revenue disappeared, leading fiscal authorities to propose tax increases in order to the run high primary surpluses needed to guarantee fiscal sustainability. The need for obtaining high primary surpluses possibly implied a shift to fiscal austerity and probably a cost in terms of foregone output and higher unemployment. Has the fiscal policy in Brazil been too austere or has it been just restrictive enough to avoid an excessive build up of debt? We answer these questions by using the measurement of debt ceiling developed in this paper.

This study is organized as follows: Section 2 describes the theoretical model for investigating public debt and the respective transversality condition to be tested, Section 3 presents the quantile autoregression model and a novel methodology to separate nonstationary observations from stationary ones. Section 4 describes debt ceiling on a QAR approach, Section 5 provides the empirical results for Brazilian public debt, and Section 6 summarizes the main conclusions.

2 Methodology

2.1 Theoretical Model

There is extensive literature on the government's intertemporal budget constraint. The general conclusion is that fiscal policy is sustainable if the government budget constraint holds in present value terms. In other words, the current debt should be offset by the sum of expected future discounted primary budget surpluses. The approaches used to analyze sustainability of fiscal policy consist in testing if the public debt and/or budget deficit is a stationary process.

The theoretical framework used here to investigate the sustainability of the Brazilian federal debt follows Uctum and Wickens (2000), which extends the results of Wilcox (1989) to a stochastic and time-varying discount rate, considering a discounted primary deficit that can be either strongly or weakly exogenous. According to the authors, a necessary and sufficient condition for sustainability is that the discounted debt-GDP ratio should be a stationary zero-mean process. As a starting point of the analysis, Uctum and Wickens (2000) investigate the one-period government intertemporal budget constraint, which can be written in nominal terms as

$$G_t - T_t + i_t B_{t-1} = \Delta B_t + \Delta M_t = -S_t, \tag{1}$$

where G = government expenditure, T = tax revenue, B = government debt at the end of period t, M = monetary base, S = total budget surplus, i = interest rate on government debt. Dividing each term of (2) by nominal GDP, one could obtain the budget constraint in terms of proportion of GDP

$$g_t - \tau_t + (i_t - \pi_t - \eta_t)b_{t-1} = \Delta b_t + \Delta m_t + (\pi_t + \eta_t)m_{t-1} = -s_t. \tag{2}$$

The variables g, τ, b, m, and s denote the ratio of the respective variables to nominal GDP, $\pi_t = (P_t - P_{t-1})/P_{t-1}$ and $\eta_t = (Y_t - Y_{t-1})/Y_{t-1}$, with P and Y standing for the price level and real GDP. This way, equation (2) can be rewritten as

$$d_t + \rho_t b_{t-1} = \Delta b_t, \tag{3}$$

where $d_t = g_t - \tau_t - \Delta m_t - (\pi_t + \eta_t)m_{t-1}$ is the primary government deficit expressed as a ratio to nominal GDP, and $\rho_t = i_t - \pi_t - \eta_t$ is the real ex-post interest rate adjusted for real output growth. According to the authors, if $\rho_t < 0$ for all t then equation (3) is a stable difference equation, which can therefore be solved backwards, implying that the debt-GDP ratio b_t will remain finite for any sequence of finite primary deficits d_t. It should be noted that for the constants ρ and d, the steady-state value of b is given by $-d/\rho$.

On the other hand, if $\rho_t > 0$ for all t, then the debt-GDP ratio will eventually explode for $d_t > 0$. Thus, primary surpluses are required to avoid this case (i.e. $d_t < 0$), and equation (3) must be solved forwards, in order to determine whether the sum of expected future discounted surpluses is sufficient to meet the current level of debt-GDP ratio. In addition, the authors rewrite (in ex-ante terms) the budget constraint for period $t+1$ as

$$b_t = E_t[(1+\rho_{t+1})^{-1}(b_{t+1} - d_{t+1})], \tag{4}$$

where b_t is known in period t, and expectations are taken based on information at time t. Equation (4) is solved forwards, resulting in the n-period intertemporal budget constraint

$$b_t = E_t \delta_{t,n} b_{t+n} - E_t \sum_{i=1}^{n} \delta_{t,i} d_{t+i}, \tag{5}$$

where $\delta_{t,n} = \prod_{s=1}^{n} (1+\rho_{t+s})^{-1}$ is the time-varying real discount factor n periods ahead, adjusted for real GDP growth rate. The discount factor $\delta_{t,n}$ can also be written as $\delta_{t,n} = a_{t+n}/a_t$, where $a_t = \prod_{i=1}^{t} (1+\rho_i)^{-1}$. The authors normalize $a_0 = 1$ and define $X_t = a_t b_t$ and $Z_t = a_t d_t$ as the discounted debt-GDP and primary deficit-GDP ratios respectively. This way, equation (5), representing the present-value borrowing constraint (PVBC), can be rewritten as

$$a_t b_t = E_t a_{t+n} b_{t+n} - E_t \sum_{i=1}^{n} a_{t+i} d_{t+i}, \tag{6}$$

or as

$$X_t = E_t X_{t+n} - E_t \sum_{i=1}^{n} Z_{t+i}. \tag{7}$$

The one-period budget constraint given by expression (3) can also be written in discounted terms, in the following way

$$b_{t-1} = (1+\rho_t)^{-1}(b_t - d_t) = (a_t/a_{t-1})(b_t - d_t), \tag{8}$$

$$\therefore X_{t-1} = a_{t-1} b_{t-1} = a_t b_t - a_t d_t = X_t - Z_t \therefore Z_t = \Delta X_t. \tag{9}$$

Hence, equation (4) can be expressed by

$$X_t = E_t(X_{t+1} - Z_{t+1}). \tag{10}$$

2.2 Sustainability for infinite horizon

According to Uctum and Wickens (2000), a necessary and sufficient condition for sustainability is that as n goes toward infinity, the expected value of the discounted debt-GDP ratio converges to zero. This condition is usually known in the literature as the transversality condition (or no-Ponzi-scheme condition), and can be summarized by

$$\lim_{n\to\infty} E_t X_{t+n} = 0. \tag{11}$$

This way, the current debt-GDP ratio is counterbalanced by the sum of current and expected future discounted surpluses, also expressed as a proportion of GDP, implying that the government's budget constraint is given (in present value terms) by

$$b_t = -\lim_{n\to\infty} E_t \sum_{i=1}^{n} \delta_{t,i} d_{t+i}, \tag{12}$$

or

$$X_t = -\lim_{n\to\infty} E_t \sum_{i=1}^{n} Z_{t+i}. \tag{13}$$

Uctum and Wickens (2000) show that the necessary and sufficient condition for the intertemporal budget constraint (13) to hold is that the discounted debt-GDP ratio (X_t) be a stationary zero-mean process. This way, if fiscal policy is currently (locally) unsustainable, then it will need to change in the future to guarantee (global) sustainability. In addition, the transversality condition requires the discounted debt-GDP ratio to converge to zero.

A starting point for investigating this condition arises from a graphical analysis of the discounted debt time series, which should be declining over the sample period. In this paper, we perform a formal test of the sustainability of the Brazilian federal debt, investigating the validity of the (necessary and sufficient) condition of stationarity with zero mean for the discounted debt-GDP ratio process. We will do so by using the quantile autoregression model which is briefly described in the next section.

3 The Quantile Autoregression Model

In a sequence of recent papers Koenker and Xiao (2002, 2004a, 2004b) introduced the so-called quantile autoregression (QAR) model. In this paper, we will show how one can separate nonstationary observations from stationary ones by using the QAR model. This result will have important implications on the literature of public-debt sustainability as shown in the next sections. For now, consider the following assumptions:

Assumption 1 let $\{U_t\}$ be a sequence of iid standard uniform random variables;

Assumption 2 Let $\alpha_i(U_t)$, $i = 0, ..., p$ be comonotonic random variables.[3]

We define the pth order autoregressive process as follows,

$$y_t = \alpha_0(U_t) + \alpha_1(U_t) y_{t-1} + ... + \alpha_p(U_t) y_{t-p}, \tag{14}$$

[3] According to Koenker (2006), two random variables $X, Y : \Omega \to \mathbb{R}$ are said to be comonotonic if there exists a third random variable $Z : \Omega \to \mathbb{R}$ and increasing functions f and g such that $X = f(Z)$ and $Y = g(Z)$. In our paper, $\beta_{i,t} = \alpha_i(U_t)$, $i = 0, 1, ..., p$ are comonotonic and $\alpha_i(\cdot)$ are, by definition, increasing functions. See our proofs in the Appendix to understand the crucial usefulness of this assumption.

where α_j's are unknown functions $[0,1] \to \mathbb{R}$ that we will want to estimate. We will refer to this model as the QAR(p) model. Given assumptions 1 and 2, the conditional quantile of y_t is given by

$$Q_{y_t}(\tau \mid \mathcal{F}_{t-1}) = \alpha_0(\tau) + \alpha_1(\tau) y_{t-1} + ... + \alpha_p(\tau) y_{t-p},$$

where $\mathcal{F}_{t-1} = (y_{t-1}, ..., y_{t-p})$ and τ is the quantile of U_t. In order to develop intuitive understanding of the QAR model, let us consider the following simple example

$$y_t = \alpha_0(U_t) + \alpha_1(U_t) y_{t-1}, \tag{15}$$

which is simply a QAR(1) model. It should be noted that the QAR model can play a useful role in expanding the territory between classical stationary linear time series and their unit root alternatives. To see this, suppose in our QAR(1) example that $\alpha_1(U_t) = U_t + 0.5$. In this case, if $0.5 \leq U_t < 1$ then the model generates y_t according to the nonstationary model, but for smaller realizations of U_t, we have mean reversion tendency. Thus, the model exhibits a form of asymmetric persistence in the sense that sequences of strongly positive innovations of the iid standard uniform random variable U_t tend to reinforce its nonstationary like behavior, while occasional smaller realizations induce mean reversion and thus undermine the persistency of the process. Therefore, it is possible to have locally nonstationary time series being globally stationary.[4]

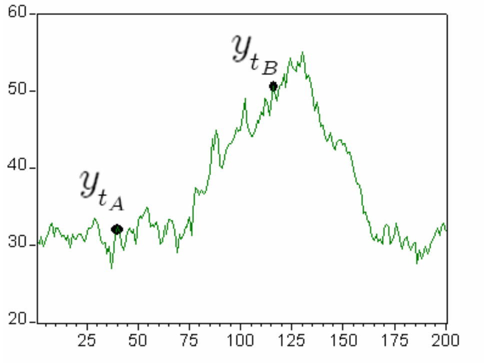
3.1 Identifying Nonstationary Observations

We continue our reasoning by considering again the QAR(1) model (15) with the same autoregressive coefficient $\alpha_1(U_t) = U_t + 0.5$. If at a given period $t = t_A$, $U_{t_A} = 0.2$, then $\alpha_1(U_{t_A}) = 0.7$ and the model will present a mean reversion tendency at $t = t_A$. However, if at $t = t_B$, $U_{t_B} = 0.5$, then $\alpha_1(U_{t_B}) = 1$, and y_t will have a local unit-root behavior. Suppose, for illustrative purposes, that this model can be represented by the stochastic process depicted in Figure 1, in which y_t has a mean reversion tendency around the period t_A. Now assume that for periods $t > t_A$, there is a sequence of strong realizations of U_t inducing the model to nonstationary behavior at period t_B.[5]

[4] See the Appendix for further details regarding the QAR model, including alternative representations, stationarity conditions, central limit theorem, estimation, autoregressive order choice, global stationarity, unconditional mean tests, and local analysis through the Koenker & Xiao (2004b) test.

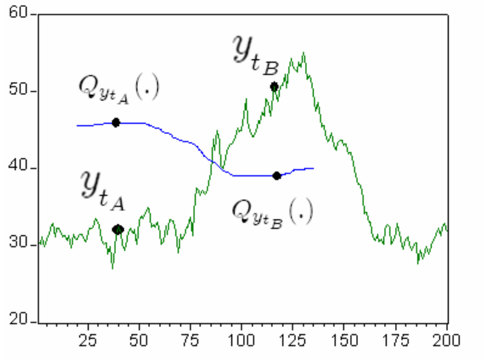
[5] The DGP used to construct this example is represented by the QAR(1) model $y_t = \alpha_1(U_t) y_{t-1}$ where $\{U_t\}$ is a sequence of iid standard uniform random variables, and the coefficients α_1 is a function on $[0,1]$, given by $\alpha_1(U_t) = \min\{1; \gamma_1 * U_t\}$, where $F : \mathbb{R} \to [0,1]$ is the standard normal cumulative distribution function. We set the parameter $\gamma_1 = 0.8$ for $t = (1, ..., 65)$; 10 for $t = (66..., 90)$; 5 for $t = (91, ..., 152)$ and 0.8 for $t = (153, ..., 200)$.

Figure 1 - Example of a QAR(1) model



A natural question that arises in this context is how to separate periods of stationarity from periods where y_t exhibits nonstationary behavior? In other words, is it possible to construct a function $Q_{y_t}(.)$ such that if y_t has a mean reversion tendency at time $t = t_A$ then $Q_{y_{t_A}}(.) \geq y_{t_A}$, but if y_t presents nonstationary behavior at time $t = t_B$ then $Q_{y_{t_B}}(.) < y_{t_B}$?

Figure 2 - Separating periods of nonstationarity



This is a theoretical question that we aim to answer in this paper by using the QAR approach. In order to separate observations of y_t that exhibit a unit-root behavior from other observations with stationary behavior, we will need the following definitions:

Definition 1 *Critical Quantile ($\tau_{crit.}$) is the largest quantile $\tau \in \Gamma = (0,1)$ such that $\alpha_{1,t}(\tau) = \sum_{i=1}^{p} \alpha_i(\tau) < 1$, where τ is the quantile of U_t*

Definition 2 *Critical Conditional Quantile of y_t: $Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1}) = \alpha_0(\tau_{crit.}) + \alpha_1(\tau_{crit.})\, y_{t-1} + ... + \alpha_p(\tau_{crit.})\, y_{t-p}$, where $\mathcal{F}_{t-1} = (y_{t-1}, ..., y_{t-p})$.*

The critical quantile $\tau_{crit.}$ can easily be identified by using the Koenker & Xiao (2004b) test for $H_0 : \alpha_{1,t}(\tau) = 1$ for selected quantiles $\tau \in \Gamma = (0,1)$, presented in the Appendix. The critical conditional quantile $Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1})$, is merely the τth conditional quantile function of y_t evaluated at $\tau = \tau_{crit.}$. Consider the additional assumption

Assumption 3 Let $\Omega = (t_1, t_2, ...t_T)$ be the set of all observations T. Assume that for the subset of time periods $\Upsilon \subset \Omega$, the time series y_t exhibits nonstationary behavior, i.e., unit root model. Now we can state proposition 1[6].

Proposition 1 *Consider the QAR(p) model (14) and Assumptions 1, 2 and 3. The critical conditional quantile of y_t will always be lower than y_t for all periods in which y_t exhibits a unit-root behavior, that is,* $Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1}) < y_t \;\; ; \;\; \forall t \in \Upsilon.$

Proof. See Appendix. ■

In order to clarify this result, suppose that all observations of y_t, $t = 1, ...T$, exhibit unit-root (stationary) behavior. In this case, the path of y_t would always be above (below) the path generated by $Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1})$. There may exist an intermediate case in which some observations of y_t exhibit unit-root behavior. In this case, the path of y_t would be above the path generated by $Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1})$ only at the periods where y_t has a unit root.

In addition, by just comparing both time series y_t and $Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1})$, one can compute the statistic H, which represents the percentage of periods in which y_t exhibits (local) nonstationary behavior.

Definition 3 *Let H be the relative frequency of nonstationary periods, that is,* $H \equiv \frac{1}{T}\sum_{t=1}^{T} I_t \left\{ y_t > Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1}) \right\}$, *where T is the sample size and I_t is an indicator function such that* $I_t = \begin{cases} 1 \;\; ; \;\; if \;\; y_t > Q_{y_t}\left(\tau_{crit.} \mid \mathcal{F}_{t-1}\right) \\ 0 \;\; ; \;\; otherwise \end{cases}$

In order to link the statistic H with the critical quantile, we can also state Proposition 2 :

Proposition 2 *If Assumptions 1 and 2 hold, then* $H = (1 - \tau_{crit.})$.

Proof. See Appendix. ■

These two propositions enable us to identify periods in which the time series y_t exhibits nonstationary (stationary) behavior. This methodology will be crucial to our analysis of fiscal sustainability as further described in section 4.

3.2 Out-of-sample forecast

In the previous sections, we showed how to identify the critical conditional quantile. Now, we show how to make multi-step-ahead forecasts for the critical conditional quantile. In order to do so, we first forecast y_t based on the simple idea of recursive generation of its conditional density, which is a quite novel approach introduced by Koenker and Xiao (2006)[7].

[6] A Monte Carlo experiment is presented in the Appendix to verify the result of Proposition 1 in finite samples. The simulation reveals that the critical conditional quantile indeed exhibits good behavior in finite samples, by correctly separating nonstationary periods from stationary ones.

[7] Koenker and Xiao presented this forecasting approach at the Econometrics in Rio conference, which took place at the economic department of the Getulio Vargas Foundation, Rio de Janeiro, Brazil.

Recall that T is the sample size and let s be the forecast horizon. Given an estimated QAR model $\widehat{Q}_{y_t}(\tau \mid \mathcal{F}_{t-1}) = x_t'\widehat{\alpha}(\tau)$ based on data $t = 1, ..., T$ we can forecast

$$\widehat{y}_{T+s} = \widetilde{x}'_{T+s}\widehat{\alpha}(U_{T+s}) \quad ; \text{ for } s = 1, ..., s_{\max} \tag{16}$$

where U_{T+s} ~iid Uniform $(0,1)$; $\widetilde{x}'_{T+s} = [1, \widetilde{y}_{T+s-1}, ..., \widetilde{y}_{T+s-p}]'$ and $\widetilde{y}_t = \begin{cases} y_t & \text{if } t \leq T \\ \widehat{y}_t & \text{if } t > T \end{cases}$. Conditional density forecasts can be made based on an "ensemble" of such forecast paths, i.e., a great number (k) of future trajectories of y_t enables us to construct the conditional density of y_t at each future period $T + s$.

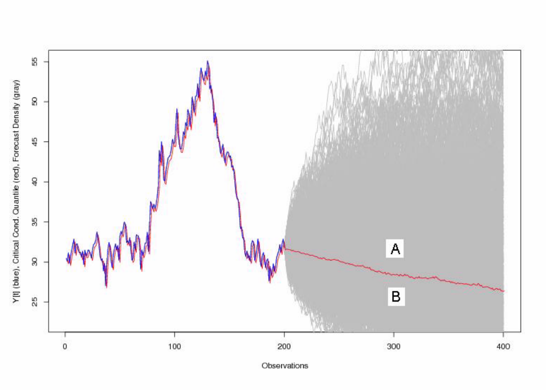
To better understand this idea, notice that $U_{T+s} \sim$ iid Uniform $(0,1)$. Hence, it is always possible to establish a 1:1 relationship between τ and a realization, u_{T+s}, of this iid standard uniform random variable U_{T+s}. Thus, for each realization of U_{T+s}, there is a 1:1 corresponding quantile $\tau = u_{T+s}$. Moreover, in estimating the conditional quantile function of y_t, $Q_{y_t}(\tau \mid \mathcal{F}_{t-1})$, one can find the estimated coefficients $\widehat{\alpha}_i(\tau)$ for each τ and, therefore, we can find $\widehat{\alpha}(U_{T+s})$ for any realization of U_{T+s}. We proceed by generating a sequence of realizations of U_{T+s} of size $s_{\max}$, that is, $\{u_{T+s}\}$, $s = 1, 2, ...s_{\max}$. This way, we can make an out-of-sample trajectory of y_t through equation (16).

If we repeat the above steps k times, then we will end up with an ensemble of forecast paths. We can now forecast the critical conditional quantile based on this ensemble of forecasts. In order words, for a given period $T + s$, $\widehat{Q}_{y_{T+s}}(\tau_{crit.} \mid \mathcal{F}_T) = \widehat{y}^{k^*}_{T+s}$ so that $\Pr\left(\widehat{y}^{k}_{T+s} \leq \widehat{y}^{k^*}_{T+s} \mid \mathcal{F}_T\right) = \tau_{crit}$. This way, we are able to generate the sequence $\{\widehat{Q}_{y_{T+s}}(\tau_{crit.} \mid \mathcal{F}_T)\}$ for $s = 1, .., s_{\max}$, which is nothing more than the forecast path of the critical conditional quantile. This methodology allows us to classify the future observations of the time series y_t into stationary and nonstationary ones.[8]

In order to clarify the idea of a multi-step-ahead forecast, consider again the QAR(1) model discussed in section 3.1. Thus, based on the estimated coefficients $\widehat{\alpha}_i(\tau)$ and the generation of k sequences of $U_{T+s} \sim$ iid Uniform of size $s_{\max}$, we can compute (see Figure 3) the conditional densities of y_{T+s} for the forecast horizons $s = 1, ..., s_{\max}$. In our example, we considered $k = 1,000$ trajectories and $s_{\max} = 200$ periods.

[8] See appendix for further details regarding the numerical procedure.

Figure 3 - Out-of-sample forecast of y_t



Notes: (a) The picture shows the forecast conditional densities of the mentioned QAR(1) model for k=1,000 trajectories.

(b) The red line represents the in-sample and out-of-sample forecasts of the critical conditional quantile.

Figure 3 summarizes the above discussion. The red line represents the forecast of the critical conditional quantile. We can see that the out-of-sample forecast of the critical conditional quantile splits the ensemble of forecasts into two regions, A and B. All the paths in region A are nonstationary whereas they are stationary in region B. As we will show in our empirical exercise in section 5, this separation has strong economic implications. Furthermore, the out-of-sample forecast of the critical conditional quantile apparently tends toward zero as long as the forecast horizon increases. In fact, as we will formally show in Proposition 3, if the time series process y_t is a zero-mean stationary process, then its critical conditional quantile will converge to zero at an infinite horizon.

4 Debt Ceiling and Fiscal Sustainability

Hereafter let y_t be the discounted debt-GDP ratio process (X_t) presented in section 2. Before introducing a "sustainability" concept, lets consider the following (testable) additional assumptions:

Assumption 4: The time series y_t is covariance stationary.

Assumption 5: The unconditional mean of y_t is zero, i.e., $\mu_y = 0$.

Notice that assumption 5 holds if we set $\alpha_0(U_t) = 0$ in Eq. (14). In this case, the out-of-sample forecast $\widehat{y}_{T+s} = \widetilde{x}'_{T+s}\widehat{\alpha}\,(U_{T+s})$ would be computed from a vector without intercept $\widetilde{x}'_{T+s} = [\widetilde{y}_{T+s-1}, ..., \widetilde{y}_{T+s-p}]$.

Hence, based on the study by Uctum&Wickens (2000), we adopt the following concept of public debt sustainability:

Definition 4 *A fiscal policy is "globally sustainable" if and only if the discounted debt-GDP ratio y_t is a stationary zero-mean process, that is, it satisfies assumptions 4 and 5.*

The previous assumptions denote that y_t is a stationary zero-mean process, which is a necessary and sufficient condition for global sustainability. If a fiscal policy is sustainable in the long run, there can still be local episodes of fiscal imbalances. How can we identify such local episodes and separate sustainable fiscal policies from unsustainable ones? In order to answer these questions, we define the concept of debt ceiling.

Definition 5 *Debt ceiling ($\widetilde{D}_t$) is equal to the critical conditional quantile when assumptions 1-5 hold.*

The above definition establishes that the debt ceiling is nothing other than the critical conditional quantile of the discounted debt-GDP ratio, $\widetilde{D}_t \equiv Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1})$. In order to clarify the concept of debt ceiling, suppose that all the observations of y_t, $t = 1, ...T$, exhibit sustainable behavior. In this case, they would always be below or on the path generated by $\widetilde{D}_t$. There may exist an intermediate case in which the public debt is still globally sustainable despite some episodes of local unsustainability. In this case, the path of y_t would be above the path generated by $\widetilde{D}_t$ only at the periods where y_t takes on unsustainable behavior. The proposed debt ceiling is a simple way to separate paths of public debt (fiscal policies) that are not sustainable from ones that satisfy the long-run transversality condition. This discussion is summarized in the following corollary.

Corollary 1 *Consider the QAR(p) model (14), where y_t now represents the discounted debt-GDP ratio process. If Assumptions 1 to 5 hold, then the respective Debt Ceiling ($\widetilde{D}_t$) will always be lower than y_t in all periods where y_t is nonsustainable, that is, $\widetilde{D}_t < y_t$, $\forall t \in \Upsilon$.*

Proof. See Appendix. ∎

Corollary 1 is an immediate consequence of Proposition 1 when the definitions 4-5 and the assumptions 4-5 are also considered. Based on Corollary 1, we have the nice result that $y_t > \widetilde{D}_t$ for all periods in which the public debt takes on an unsustainable dynamic. Moreover, given that y_t is, by (testable) assumptions, a stationary zero-mean process, by just comparing y_t and $\widetilde{D}_t$ one can also compute what we call "debt tolerance", that is, the percentage of episodes of local unsustainability that does not jeopardize long-run sustainability, that is:

$$H \equiv \frac{1}{T}\sum_{t=1}^{T} I_t \left\{y_t > \widetilde{D}_t\right\}, \tag{17}$$

where $I(.)$ is an indicator function and T is the sample size. Therefore, given a globally sustainable fiscal policy, H represents the percentage of violations of the transversality condition still compatible with long-run fiscal sustainability[9].

Regarding the out-of-sample forecast, the following Proposition guarantees that the forecast path of debt ceiling will go to zero as the forecast horizon goes to infinity. This is an expected result from the literature on public debt sustainability, since the transversality condition (or no-Ponzi-game condition) states that the forecast value of a sustainable (discounted) debt-GDP ratio must converge to zero.

[9] Reinhart et al. (2003) developed the concept of "debt intolerance" based on a historical analysis of external debt. They divided the countries into debtors' clubs and vulnerability regions, depending principally on a country's own history of default and high inflation.

Proposition 3 *If assumptions 1 to 5 hold, then the forecast path of the Debt Ceiling* $(\widehat{D}_{T+s})$ *will go to zero as the forecast horizon s goes to infinity, i.e.,* $\lim_{s\to\infty} \widehat{D}_{T+s} = 0$.

Proof. See Appendix. ■

In the next section, we show that the debt ceiling concept can also be seen as a more elaborat concept of Value at Risk.

4.1 Debt Ceiling and Value at Risk

In the financial literature, Value at Risk (VaR) is a measurement representing the worst expected loss of an asset or portfolio over a specific time interval, at a given confidence level. It is typically used by security houses or investment banks to measure the market risk of their asset portfolios.[10] The VaR_t can be defined as

$$\Pr(r_t \leq \mathrm{VaR}_t \mid \mathcal{F}_{t-1}) = \tau, \tag{18}$$

where r_t is the return on some financial asset, $\mathcal{F}_{t-1}$ is the information set available at time $t-1$, and $\tau \in (0,1)$ is the confidence level. From this definition, it is clear that finding a VaR_t is basically the same as finding a conditional quantile. Following Hafner & Linton (2006), it is straightforward to show that the estimation of a VaR_t is a natural application of the QAR model, that is

$$\Pr(y_t \leq Q_{y_t}(.) \mid \mathcal{F}_{t-1}) = \tau_{crit.}. \tag{19}$$

In our application of the QAR model, we estimate the exact conditional quantile that represents the limit of stationarity (our critical conditional quantile), which is used to define the debt ceiling, in accordance with the government's intertemporal budget constraint. Thus, our proposed debt ceiling is nothing more than a "qualified" Value at Risk, that is

$$\widetilde{D}_t \equiv Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1}) = \mathrm{VaR}_t.$$

It is important to note, however, that the proposed "qualified" VaR concept goes far beyond the financial applications, in which an "ad-hoc" value for $\tau_{crit.}$ is adopted (usually 1% or 5%).[11] In our approach, we identify the exact critical quantile that represents a threshold, $\tau_{crit.}$, according to a given theoretical economic model.

This is a novel approach in the literature of public debt sustainability, but it may have other applications in finance and macroeconomics. Garcia and Rigobon (2004) studied debt sustainability from a risk management perspective by using a Value at Risk (VaR) approach. The authors proposed a very attractive technique, based on Monte Carlo simulations, to compute "risk probabilities", i.e., probabilities that the simulated

[10] For instance, if a given portfolio has a 1 day VaR of $5 million (at a 95% confidence level), this implies, with a probability of 95%, that the value of its portfolio is expected to decrease by 5 million or less during 1 day.

[11] For instance, a bank capital requirements analysis usually fixes the critical quantile at 1%, whereas risk management models typically impose a confidence level of 5%.

debt-GDP ratio exceeds a given threshold deemed "risky". However, their choice of the quantile needed to compute the "risky" threshold of sustainability was somehow arbitrary (see figure 4 of Garcia and Rigobon, 2004). The methodology proposed in this paper complements their approach by computing the exact "risky" quantile, the so-called $\tau_{crit.}$, which enables us to properly separate nonsustainable paths of public debt from sustainable ones, instead of choosing an "ad-hoc" threshold of sustainability.[12]

5 Empirical Results

5.1 The Database

The methodology presented in this paper is applied to the analysis of the discounted Brazilian federal debt. All data are quarterly and are obtained from the Central Bank of Brazil (BCB), the Institute of Applied Economic Research (IPEA), and the Brazilian Institute of Geography and Statistics (IBGE). Our sample covers the period from 1976.I to 2005.I (117 observations). The undiscounted debt represents the series "Dívida Mobiliária Interna Federal fora do Banco Central", or federal domestic debt held by the public, in percentage of GDP.[13] The discounted debt is given by the undiscounted debt series multiplied by the stochastic discount factor. Bohn (2004) mentions that the debt-GDP ratio suggests a "more benign view" of fiscal policy than the nominal and real series.

The stochastic discount factor (a_t), as previously mentioned in the theoretical model, is generated from ρ_t (the real ex-post interest rate adjusted for real output growth), which depends on the inflation and nominal interest rates, and real output growth. The inflation rate (π_t) is measured in a standard approach by a general price index (IGP–DI), and the nominal interest rate (i_t) is measured by the over/selic interest rate (equivalent to the U.S. Fed funds rate). Regarding real output growth (η_t), we generate a quarterly series based on the quarterly GDP, which is released by IBGE, with seasonal adjustments made by the MA(12)[14] and X-11 methods.[15]

$$a_t = \prod_{i=0}^{t-1} \frac{1}{(1+\rho_i)} \quad ; \quad a_0 = 1 \tag{20}$$

$$(1+\rho_t) = \frac{(1+i_t)}{(1+\pi_t)(1+\eta_t)} \tag{21}$$

According to Uctum and Wickens (2000), there are two major issues that must be addressed when using government debt data: whether to measure debt at market value or at face value (at par), and how to measure

[12] Moreover, this paper presents a distribution-free approach to make out-of-sample forecasts of the debt ceiling. The same does not happen in Garcia and Rigobon (2004) since their simulations are based on the assumption of normal distribution innovations.

[13] Following Rocha (1997), we focused the analysis on the domestic debt, since the sustainability of external debt is guaranteed by current account surpluses, and not by fiscal surpluses or seigniorage. Despite the fact that the debt-GDP ratio is not high in comparison to other nations, its sharp increase in the last decade is very concerning.

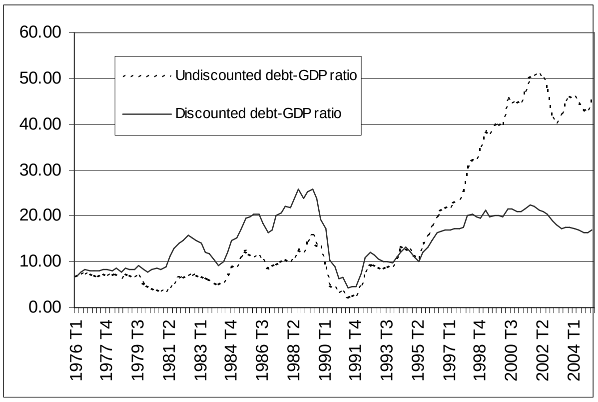
[14] Following Garcia and Rigobon (2004).

[15] Since the results based on these two techniques are very similar, we only report the MA(12) results.

the discount rate.[16] The authors state that the correct implementation of the government's intertemporal budget constraint requires the use of the discounted net market value of debt. However, the market value of debt is usually not available, and the debt is generally expressed at par. An estimate of the market value of debt is obtained by multiplying the face value by the implied market price $1/(1+p_t)$, where p_t is the yield on government debt. Some studies on the sustainability of the Brazilian public debt, such as Pastore (1995), Rocha (1997) and Giambiagi and Ronci (2004), used debt value at par, whereas Luporini (2000) uses market value. In our case, the analysis will only be conducted for the discounted debt at face value, since these two series, in our sample period, are very similar.

Figure 4 presents the undiscounted and discounted Brazilian federal debt-GDP ratio. A simply visual inspection of figure 4 suggests that the discounted debt seems to be stationary, despite the sharply increasing path of the undiscounted series in the 1990s. The formal evidence on sustainability of the Brazilian public debt is investigated in the following sections.

Figure 4 - Brazilian federal debt (% GDP)



Note: Undiscounted debt corresponds to the federal domestic debt held by the public, in percentage of GDP.

5.2 Autoregressive Order Choice

We first determine the autoregressive order of the QAR(p) model (14) using the Kolmogorov-Smirnov test based on LR statistics, following Koenker & Machado (1999). We start estimating the quantile regression below with $p = p_{\max} = 3$, that is:

$$Q_{y_t}(\tau \mid y_{t-1}, ..., y_{t-p}) = \alpha_0(\tau) + \alpha_1(\tau)\, y_{t-1} + \alpha_2(\tau)\, y_{t-2} + \alpha_3(\tau)\, y_{t-3}.$$

The index set used for quantiles is $\tau \in \Gamma = [0.1, 0.9]$ with steps of 0.005. Next, we test if the third order covariate is relevant in our model, i.e., we considered the null hypothesis:

$$H_0 : \alpha_3(\tau) = 0, \quad \textit{for all } \tau \in \Gamma.$$

[16] According to Giambiagi and Ronci (2004), one should ideally use net-of-taxes real rate of interest. However, net-of-tax yield is a difficult task since tax rates vary according to security holder, and there is limited information on its identity.

The results are reported in Table 1. Using critical values obtained in Andrews (1993), we can infer that the autoregressive variable y_{t-3} can be excluded from our econometric model.

Table 1: Choice of the autoregressive order

excluded variable	$\sup_{\tau\in\Gamma} L_n(\tau)$ estimate	5% critical value	10% critical value	H_0	Result
y_{t-3}	3.989623	9.31	7.36	$\alpha_3(\tau)=0$	*do not reject*
y_{t-2}	23.79831	9.31	7.36	$\alpha_2(\tau)=0$	*reject*

Since the third order is not relevant, we proceed by analyzing if the second order covariate is relevant.[17] Thus, we considered the null hypothesis:

$$H_0 : \alpha_2(\tau) = 0, \quad for\ all\ \tau \in \Gamma,$$

whose results are also presented in Table 1. Indeed, we verify that the second autoregressive variable cannot be excluded. Thus, the optimal choice of lag length in our model is $p = 2$ and this order will be used in the subsequent estimation and hypothesis tests presented in this paper. In summary, our econometric model will be:

$$y_t = \alpha_0(U_t) + \alpha_1(U_t)\, y_{t-1} + \alpha_2(U_t)\, y_{t-2}, \tag{22}$$

and the associated ADF formulation is:[18]

$$y_t = \mu_0 + \alpha_{1,t} y_{t-1} + \alpha_{2,t}\Delta y_{t-1} + u_t, \tag{23}$$

where

$$\begin{aligned} \alpha_{1,t} &= \sum_{i=1}^{2} \alpha_i(U_t) \\ \alpha_{2,t} &= -\alpha_2(U_t), \\ u_t &= \alpha_0(U_t) - \mu_0. \end{aligned}$$

5.3 Global sustainability

The concept of global sustainability used in this paper states that local episodes of fiscal imbalances must be offset by periods of fiscal responsibility, so that the PVBC condition holds in the long-run. Recall from section 2 that the necessary and sufficient condition for the intertemporal budget constraint (13) to hold is that the discounted debt-GDP ratio, represented by y_t, must be a stationary zero-mean process. If this happens, then the Brazilian federal debt will be globally sustainable.

[17] As usual, we performed the test for exclusion of y_{t-2} with same sample size used to test the exclusion of y_{t-3}.

[18] For the sake of completion, we carried out the same tests in the ADF form. As expected, the Kolmogorov-Smirnov based on LR statistics estimates were exactly the same as the estimates reported in Table 2.

In order to test for global stationarity, we need to test the null hypothesis H_0: $\alpha_{1,t}= 1$ in Eq. (23). If such a null hypothesis is rejected against the alternative H_1: $\alpha_{1,t}< 1$, then we say that the Brazilian federal debt is globally stationary. We test H_0: $\alpha_{1,t}= 1$ by using the so-called Quantile Komogorov-Smirnoff (QKS) test proposed by Konker and Xiao (2004). The computational details on the QKS test statistic are described in the appendix. The critical values used in the QKS test are computed by the residual-based block (RBB) bootstrap recently proposed by Paparoditis and Politis (2003). Therefore, the critical values will ultimately depend on the block length arbitrarily chosen by the user.[19] Table 2 reports the statistics and critical values for eight different block lengths, b, arbitrarily chosen. We considered 10,000 bootstrap replications.

Table 2: Results for the global stationarity test

Block length b	QKS	5% critical value	10% critical value	H_0: $\alpha_{1,t}= 1$
12	13.2753601	14.0261732	11.9415994	reject at 10%
14	13.2753601	13.8578960	11.91955628	reject at 10%
16	13.2753601	14.6971102	12.18249493	reject at 10%
18	13.2753601	14.2410137	12.00697801	reject at 10%
20	13.2753601	15.4325526	12.76201548	reject at 10%
22	13.2753601	13.7142703	11.36398838	reject at 10%
24	13.2753601	12.8470731	11.12157297	reject at 5%
26	13.2753601	12.3618035	10.87127688	reject at 5%

There is evidence that the discounted debt is not a unit root process, with a significance level of 10% for almost all values of b (except for b = 24 and 26, where we reject the unit root null at a significance level of 5%). Overall, the results in Table 2 suggest that, at worst, the discounted Brazilian debt is globally stationary at 10% of significance.

We now test the null hypothesis that y_t has zero unconditional mean, i.e., $H_0 : \mu_y = 0$. We conduct a t-test for the unconditional mean and use the NBB resampling method with 10,000 replications to compute 5% critical values. Table 3 reports the t-statistic for the discounted public debt series. The reported results suggest that the unconditional mean of the autoregressive process is not statistically different from zero. The result of the test depends on the block length used to compute the bootstrap sample. The results in Table3

[19] The fundamental issue of the RBB bootstrap is its ability to simulate the weak dependence appearing in the original data series by separating the residuals in blocks. For more details, see Lima and Sampaio (2005).

proved to be robust to various values of the block length (b).

Table 3 : Results for the unconditional mean test

Block length b	t	2.5% critical value	97.5% critical value	Ho: intercept=0
12	28.9146968	17.8434187	32.8290331	do not reject at 5%
14	28.9146968	19.5784337	34.5324403	do not reject at 5%
16	28.9146968	21.241210	35.5579257	do not reject at 5%
18	28.9146968	22.7204141	36.7304168	do not reject at 5%
20	28.9146968	24.1544067	38.4237097	do not reject at 5%
22	28.9146968	25.4286859	39.7284730	do not reject at 5%
24	28.9146968	26.7328681	40.4919047	do not reject at 5%
26	28.9146968	28.0966684	41.7003689	do not reject at 5%

Putting it all together, the discounted Brazilian federal debt is indeed globally sustainable. This result is in accordance with many previous studies, such as in Pastore (1995), Rocha (1997), and Issler and Lima (2000), suggesting the sustainability of the Brazilian public debt.

5.4 Local Sustainability Test

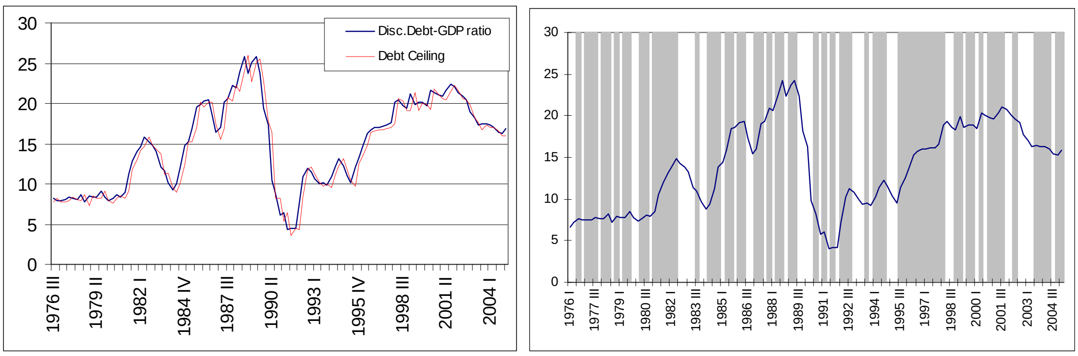
Given that the Brazilian public debt is a stationary zero mean process, we can now proceed to the "local" analysis by using the Koenker & Xiao (2004b) test. In order to identify the debt ceiling of the Brazilian public debt, we need to test the null hypothesis $H_0 : \alpha_1(\tau) = 1$ at various quantiles by using the t-ratio test $t_n(\tau)$ proposed by Koenker and Xiao (2004.b), with the zero-mean restriction imposed in the ADF representation of Eq. (22) . Table 4 reports the results. The second column displays the estimate of the autoregressive term at each decile. Note that, in accordance with our theoretical model, $\widehat{\alpha}_1(\tau)$ is monotonic increasing in τ, and it is close to unity when we move towards upper quantiles. Table 4 shows that the null hypothesis $H_0 : \alpha_1(\tau) = 1$ is rejected against the alternative hypothesis $H_1 : \alpha_1(\tau) < 1$ for $\tau \in [0.1; 0.4]$. The critical values were obtained by interpolation of the critical values extracted from Hansen (1995, page 1155). The last column summarizes the local sustainability analysis.

Table 4 : Koenker-Xiao test

τ	$\widehat{\alpha}_1(\tau)$	$t_n(\tau)$	δ^2	H_0: $\alpha_1(\tau)=1$	Sustainability
0.10	0.8955590	-5.89314	0.12338216	reject	OK
0.20	0.9547887	-3.58005	0.09245611	reject	OK
0.30	0.9671567	-4.85814	0.28034278	reject	OK
0.40	0.9810935	-3.53123	0.16280786	reject	OK
0.50	0.9963589	-0.49040	0.13264691	do not reject	-
0.60	1.0093934	1.05544	0.19937313	do not reject	-
0.70	1.0339169	2.81045	0.18691883	do not reject	-
0.80	1.0694750	5.68227	0.04007214	do not reject	-
0.90	1.0948026	6.29170	0.02279413	do not reject	-

Table 4 shows that the critical quantile found using Brazilian public-debt data is equal to 0.40 ($\tau_{crit.} = 0.40$). Consequently, the debt ceiling of the Brazilian debt-GDP ratio corresponds to the path generated by the fourth conditional decile, that is, $\widetilde{D}_t = Q_{y_t}(0.40 \mid y_{t-1}, ..., y_{t-p})$. Hence, according to corollary 1 in this paper, if a given fiscal policy yielding a path of the (discounted) debt-GDP ratio above $Q_{y_t}(0.40 \mid y_{t-1}, ..., y_{t-p})$ were to persist forever, then such a fiscal policy would not be sustainable in the long run. Figure 5 displays the in-sample path of the debt-ceiling which is nothing more than the in-sample forecast of the 0.4th conditional decile function.

Figure 5 - Debt ceiling ($\widetilde{D}_t$) and discounted debt-GDP ratio (y_t)



Note: The debt ceiling series is constructed through in-sample forecast of the 0.4th conditional quantile, given by the ADF formulation: $\widetilde{D}_t = \widehat{\alpha}_1(0.4)\, y_{t-1} + \widehat{\alpha}_2(0.4)\, \Delta y_{t-1}$

The gray bar in Figure 5 indicates episodes in which the public debt presented unsustainable behavior. Recall that Tables 2 and 3 show that the discounted debt-GDP ratio in Brazil is globally sustainable. It means that despite the many episodes of fiscal imbalances exhibited in Figure 5 by the gray bars, there were other episodes of fiscal adjustments (white bars) that were enough to guarantee global sustainability of the Brazilian debt. These episodes of fiscal imbalances were triggered by external shocks, such as oil price shocks in the 70s, and the sequence of financial crises in the 80s and 90s. In the domestic scenario, some recent macroeconomic shocks, such as the exchange rate fluctuation in 1999, and the political uncertainty related to the presidential elections of 2002, are also related to periods of local unsustainability of Brazilian debt.

In sum, the results displayed by Figure 5 suggest that the Brazilian authorities are able to intervene through deficit cuts when debt has reached high levels. However, as suggested in Issler and Lima (2000), their mechanism of intervention is never based on spending cuts: it is either based on increases in the tax burden or on the usage of seigniorage revenue.

Table 5 gives us a historical perspective of the Brazilian public debt solvency. The overall result of Table 5 reveals that the debt tolerance $\widehat{H} = 0.60$, i.e., the percentage of episodes in our sample period in which the discounted debt-GDP ratio was above its debt ceiling ($y_t > \widetilde{D}_t$) was 60%, which is perfectly compatible with Proposition 2, since we have found $\tau_{crit.} = 0.40$.[20] Furthermore, due to the nonlinear dynamics of y_t, it is possible to identify different fiscal regimes by estimating, for each historical period, the respective statistic H. Indeed, our estimates for the fiscal policy by the end of the military regime suggest that for 59% of this period the public debt was above the debt ceiling, which is an amount slightly below the theoretical value for the debt tolerance H. As for the beginning of the new republic, in the Sarney administration (1985.II-1990.I), the fiscal policy implemented during the period was not sustainable 55% of the time, which is lower than the debt tolerance of 60%. However, we should point out that seigniorage revenue played a crucial role in balancing the public budget in that period.

Table 5: Quarters during which the discounted public debt-GDP ratio is larger than the 0.4th conditional quantile forecast ($y_t > \widetilde{D}_t$)

	number of quarters (a)	total of quarters (b)	H = (a) / (b)
End of Military Regime (1976.I-1985.I)	22	37	0.59
Sarney administration (1985.II-1990.I)	11	20	0.55
Collor and Franco administration (1990.II-1994.IV)	10	19	0.53
First Cardoso administration (1995.I-1998.IV)	12	16	0.75
Second Cardoso administration (1999.I-2002.IV)	9	16	0.56
Lula administration (2003.I-2005.I)	6	9	0.67
Total sample (1976.I-2005.I)	70	117	0.60

Regarding the Collor and Franco administration (1990.II-1994.IV), it is important to notice that the fiscal stabilization plan launched in the middle of March 1990 was responsible for the sharp decrease observed in public debt stock, since around 80% of the money stock was "frozen" (M4=M1+all other financial assets).[21] As a result, the percentage of periods in which the public debt moved above its debt ceiling was only 53%.

[20] The Brazilian debt is globally sustainable despite the fact that 60% of its observations exhibit (local) unsustainable behavior. This finding results from the combination of the global stationarity and unconditional mean tests with local investigation in a selected range of quantiles, based on the Koenker & Xiao (2004) test.

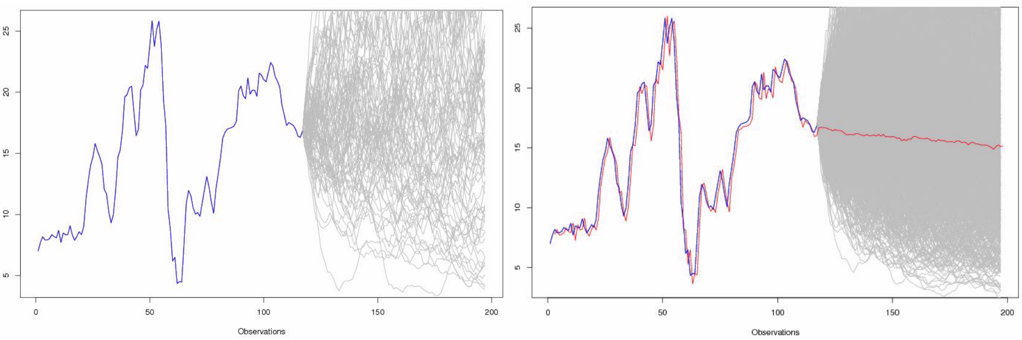
[21] See Rocha (1997).

Notice, however, that such a number should be analyzed with some caution since the Brazilian Supreme Court decided that the majority part of this "unpaid" debt had to be repaid in the Cardoso government under the denomination of hidden liabilities (skeletons). Indeed, regarding Cardoso's first term (1995.I-1998.IV), the episodes of fiscal unsustainability was equal to 75%, well above the 60% debt tolerance. The elevation of the debt-GDP ratio was mainly due to the recognition of skeletons of around 10% of GDP. However, despite the sharp increase in the debt, the recognition of skeletons improved the fiscal statistics, providing greater transparency and accuracy in Brazil's fiscal position.

Table 5 shows an improvement of the Brazilian fiscal position in the second term of President Cardoso. This improvement occurred despite the significant real exchange rate depreciation starting in 1999.I,[22] which provoked a considerable increase in debt because most Brazilian bonds at that time were indexed to hard currencies. Since government spending did not stop rising in Cardoso's second term, most of the fiscal effort was based on the fact that tax revenue increased much faster than government spending. More recently, regarding President Lula's administration, it should be noticed that despite the fiscal effort to keep discounted debt on a sustainable path, the majority of the observations are beyond the debt ceiling. Therefore, we find that the fiscal policy in effect since the beginning of 2003 has not been austere enough to guarantee long-run sustainability.

Next, we present the out-of-sample forecasts of the Brazilian public debt, based on the methodology of recursive generation of conditional densities of y_t, previously described in section 3.2. The out-of-sample forecasts were constructed with a maximum forecast horizon $s_{\max} = 80$ periods (or 20 years), with 1,000 trajectories for the y_t process:

Figure 6 - Out–of-sample forecast of Brazilian debt



Notes: (a) The pictures respectively show the out-of-sample forecasts for 100 and 1,000 trajectories.

(b) The right picture exhibits (with the red line) the in-sample and out-of-sample forecast of the critical conditional quantile.

The red line, representing the forecast debt ceiling, which is the upper trajectory that satisfies the transversality condition of no-Ponzi scheme. Notice that it is indeed decreasing, in accordance with Proposition 3, which states that it must converge to zero in the long run. A decision maker will use the debt ceiling forecast to decide whether or not to take some action. For example, if the future values of public debt are

[22] Real exchange rate adjustment has occurred under the new floating exchange regime.

above its predicted ceiling, then the fiscal authorities may decide to cut expenditure or increase tax revenue to bring public debt back to its sustainable path. Based on the information available up to time T, one can consider the following additional statistic:

Definition 6 *Future percentage of violations* $H^* \equiv \frac{1}{s_{\max}} \sum_{t=T+1}^{T+s_{\max}} I_t^*$*, where* I_t^* *is an indicator function, for* $t = T + s$ *and* $s = 1, ..., s_{\max}$*, such that* $I_t^* = \begin{cases} 1 & ; \text{ if } \widehat{y}_t > \widetilde{D}_t \\ 0 & ; \text{ otherwise} \end{cases}$.

Based on the above definition,we could classify the future paths of the public debt into three different categories:

(i) Globally sustainable fiscal policies: those trajectories always below the "red line", i.e., $H^* = 0$;

(ii) Unsustainable fiscal policies: those paths always above the red line, or with a percentage of violations above 60%.i.e., $H^* > 0.6$.

(ii) Globally sustainable fiscal policies but with some local unsustainable episodes: those trajectories with percentage of violations below 60%, i.e., $H^* \leq 0.6$;

Therefore, a decision maker (fiscal authority) may decide to intervene in the path of public debt (by increasing budget surplus) if the percentage of violations (H^*) during, say, the next four quarters is larger than 60%. Since our sample ends in 2005.I, and (by now) new observations have become available, we can compare them to the forecast debt ceiling. Notice that the actual undiscounted debt-GDP ratio for the periods 2005.II, 2005.III, 2005.IV, and 2006.I was respectively 47.21%,48.95%, 49.53% and 50.76%. However, the predicted debt ceiling for the same period was 42.83%, 42.78%, 42.75%, and 42.51%, respectively. Hence, for the 4 quarters considered, the number of violations was 100%, that is, $H^* = 1$. Therefore, the out-of-sample forecast based analysis reveals that the more recent dynamic of the Brazilian public debt is not sustainable and additional fiscal efforts are needed to bring the debt-GDP ratio back to values below the debt ceiling.

It is important to mention that other decision-making parameters might also be considered by the fiscal authority. For example, the government might have to decide today (at the time that the forecast is made) how many expenditure cuts or tax revenue increases should occur in the next four quarters in order to guarantee that the public debt would be lower than its forecast ceiling. Another interesting application is to define $z_t = 1$ if $y_t > \widetilde{D}_t$ and $z_t = 0$, otherwise. Hence, we could estimate $\widehat{\pi}_t^i = prob(y_t > \widetilde{D}_t)$ according to some economic model "i" and use the Kuipers Score to evaluate such probability forecasts (See Granger and Pesaran ,1999, for further details). We did not consider either of these techniques in this paper, but we recognize that they can easily be employed to study other aspects of public debt sustainability, such as the determinants of local fiscal imbalances.

Since additional fiscal effort is needed, it is relevant to understand how long-run fiscal sustainability has normally been reached in Brazil. Issler and Lima (2000) show that from 1947 to 1994, the public budget in Brazil was balanced through seigniorage revenue with no reduction in government spending. After the Real plan, seigniorage revenue disappeared, leading the Brazilian government to correct fiscal imbalances through tax increases. Indeed, the tax burden in Brazil is already 38% of GDP, meaning that Brazilians are now the most heavily taxed citizens in Latin America with almost no counterpart in public goods. Hence, it would be ideal if the aforementioned fiscal goal of raising the primary surplus were to be achieved through expenditure cuts. It turns out, however, that in the last four quarters of the Lula administration, the GDP growth rate has been very low and government spending has increased by 14% (year to date), while tax revenue has increased by only 11% (year to date).

If public expenditures keep rising faster than tax revenue, we might expect that the fiscal position in Brazil will worsen in the near future. Notice, however, that a new presidential term will start in January, 2007. Based on the fact that popularity concerns[23] (political constraints) are partially eliminated at the beginning of a new term, we could expect that a fiscal policy based on expenditure cuts through the reduction of interest rate payments is perfectly viable in Brazil as long as the market believes that the new government is able to implement a reform agenda that would increase the productivity of the Brazilian economy in the long run. Such an agenda should include changes in job-market legislation, the social security system, the educational system, and simplification of the bureaucracy, among other changes needed to increase the productivity of the Brazilian economy.[24] Without such reforms, it will be hard for the Brazilian fiscal authorities to convince the market that they will be able to bring the debt-GDP ratio back to its sustainable path, unless, of course, they decide to resort to seigniorage revenue.

Table 6: Out-of-sample forecast of Brazilian debt (% GDP)

Periods	Debt ceiling (discounted debt)	Debt ceiling (undiscounted debt)	Observed debt
2005.II	15.71	42.83	47.21
2005.III	15.69	42.78	48.95
2005.IV	15.68	42.75	49.53
2006.I	15.59	42.51	50.76
2006.II	15.54	42.37	-
2006.III	15.43	42.06	-

Note: The stochastic discount factor used to transform the discounted Debt Ceiling (% GDP) into the undiscounted value is the same one used in the last sample point, that is, 2005.I. However, there are other ways to deal with future values of the stochastic discount factor. For example, one could use the market expectations for inflation, output growth and interest rate, published by the Central Bank of Brazil.

[23] The existence of delayed stabilization in Brazil was recently reported by Lima and Simonassi (2005) who investigated whether the Brazilian public debt is sustainable in the long run by considering threshold effects on the Brazilian budget deficit. They show that popularity concerns (political constraints) taking place in the end of the presidential term are the main reason for the existence of delays in the fiscal stabilization in Brazil.

[24] It is important to notice that government intervention through deficit cuts might not necessarily be incompatible with the minimization of output and employment loss. Indeed, Giavazzi and Pagano (1990) found empirical evidence, for some European Countries, in favour of an "expansionary expectational effect" of a fiscal consolidation.

6 Conclusions

After the fiscal stabilization plan of 1994, the Brazilian government was no longer able to use seigniorage as a (major) source of revenue. In order to avoid an excessive build up of debt and consequent pressure on monetary policy, fiscal authorities had to adopt restrictive fiscal policies. The fiscal austerity led to very low rates of growth for the Brazilian economy, with negative impact on employment. Some politicians have constantly argued that Brazil's primary budget surplus in Brazil is too large and, therefore, should be reduced to allow for an increase in public spending on infrastructure, education and health services. They claim that fiscal policy would still be sustainable (without the necessity to use seigniorage) with lower budget surpluses.

Running lower budget surpluses without resorting to seigniorage revenue would ultimately lead to an increase in public debt. In this paper, we attempted to answer the following question: how austere should fiscal policy be to guarantee long-run sustainability? By using a fresh econometric model, we showed that: (i) contrary to the opinions of many politicians, Brazilian public debt is not currently low enough to guarantee long-run sustainability and, therefore, the budget surplus should rise rather than be decreased. In other words, we found that the debt-GDP ratio has moved beyond its ceiling during the majority of quarters in the last two years; (ii) in the absence of shocks, the Brazilian government would have to reduce the debt-GDP ratio during the next quarters to guarantee long-run fiscal sustainability and; (iii) despite occasional periods in which the Brazilian public debt moved beyond its sustainability ceiling, our historical analysis reveals that public debt in Brazil has been globally sustainable, suggesting that Brazilian government authorities have reacted to high levels of public debt, mainly through increases in the tax burden or seigniorage revenue.

Issler and Lima (2000) concluded their article with a brief reflection on the solvency of the Brazilian public debt. They suggested that, for exogenous expenditures, as they verified in the sample from 1947-1992, there would be just two polar forms for restoring long-run sustainability in Brazil: tax increases or increases of seigniorage revenue. Since the overall tax burden has risen almost twofold and already reached 38% of GDP,[25] it seems that Brazilian fiscal authorities did opt to balance the budget via tax increases. With such a tax burden, Brazilians are now the most heavily taxed citizens in Latin America and, therefore, may start penalizing politicians who propose additional tax increases. Hence, the aforementioned fiscal goal of raising the primary surplus will probably have to be achieved through expenditure cuts or increases in seigniorage revenue. In the second case, inflation will increase again, a price Brazilians may be willing to pay for tax relief. As in Issler and Lima (2000), we all hope that expenditures will cease to be "exogenous" in Brazil.

Despite the process of institutional transformations and the recent austere fiscal policy adopted in Brazil with the implementation of a target for the budget surplus, Brazil has an unfortunate history of serious difficulties in balancing its public budget. Therefore, it appears that the construction of indebtedness targets for Brazil is necessary, to provide a benchmark to guide fiscal authorities in their task of keeping the public

[25] In the first semester of 2006.

debt on a sustainable path. The measure of debt ceiling introduced in this paper aims to contribute in this direction, developing a "debt-warning system" that helps the macroeconomist to identify "dangerous" debt paths, deemed to be unsustainable.

Appendix A. Inference Methods of the QAR model

Other Representations and Regularity Conditions of the QAR(p) Model

We define the pth order autoregressive process as follows,

$$y_t = \alpha_0 (U_t) + \alpha_1 (U_t) y_{t-1} + ... + \alpha_p (U_t) y_{t-p}$$

where α_j's are unknown functions $[0, 1] \to \mathbb{R}$ that we will want to estimate. We will refer to this model as the QAR(p) model.[26]

In order to investigate stationarity of the y_t process, we initially rewrite the QAR(p) model in a vector QAR(1) representation, as follows

$$Y_t = \mu + A_t Y_{t-1} + V_t$$

where $Y_t = [y_t, ..., y_{t-p+1}]'$; $\mu = \begin{bmatrix} \mu_0 \\ 0_{p-1} \end{bmatrix}$; $A_t = \begin{bmatrix} a_t & \alpha_p (U_t) \\ I_{p-1} & 0_{p-1} \end{bmatrix}$; $V_t = \begin{bmatrix} u_t \\ 0_{p-1} \end{bmatrix}$; $a_t = [\alpha_1 (U_t), ..., \alpha_{p-1} (U_t)]$ and $u_t = \alpha_0 (U_t) - \mu_0$.

Then, lets assume the following conditions:

C.1 $\{u_t\}$ is *iid* with mean 0 and variance $\sigma^2 < \infty$. The CDF of u_t, F, has a continuous density f with $f(u) > 0$ on $U = \{u : 0 < F(u) < 1\}$.

C.2 Eigenvalues of $\Omega_A = E(A_t \otimes A_t\}$ have moduli less than unity.

Koenker & Xiao (2004b) state that under conditions C.1 and C.2, the QAR(p) process y_t is covariance stationary and satisfies a central limit theorem

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left(y_t - \mu_y\right) \Rightarrow N\left(0, \omega_y^2\right) \tag{24}$$

with

$$\begin{aligned} \mu_y &= \frac{\mu_0}{1 - \sum_{j=1}^{p} \beta_j} \qquad (25) \\ \beta_j &= E\left(\alpha_j (U_t)\right) \quad , \quad j = 1, ..., p \\ \omega_y^2 &= \lim \frac{1}{n} E\left[\sum_{t=1}^{n} \left(y_t - \mu_y\right)\right]^2 \end{aligned}$$

[26] More on regularity conditions underlying model (14) are found in Koenker and Xiao (2004a) as well as in the appendix of this paper.

The QAR(p) model (14) can be reformulated in a more conventional random coefficient notation as

$$y_t = \mu_0 + \beta_{1,t} y_{t-1} + ... + \beta_{p,t} y_{t-p} + u_t, \tag{26}$$

where

$$\begin{aligned} \mu_0 &= E\alpha_0 (U_t), \\ u_t &= \alpha_0 (U_t) - \mu_0, \\ \beta_{j,t} &= \alpha_j (U_t), \quad j = 1, ..., p. \end{aligned}$$

Thus, $\{u_t\}$ is an iid sequence of random variables with distribution $F(\cdot) = \alpha_0^{-1} (\cdot + \mu_0)$, and the $\beta_{j,t}$ coefficients are functions of this u_t innovation random variable. An alternative form of model (26) widely used in economic applications is the ADF (augmented Dickey-Fuller) representation (27). According to Koenker & Xiao (2004b), in the ADF formulation the first order autoregressive coefficient plays an important role in measuring persistency in economic and financial time series, which in our case will be crucial to determining the sustainability of public debt.

$$y_t = \mu_0 + \alpha_{1,t} y_{t-1} + \sum_{j=1}^{p-1} \alpha_{j+1,t} \Delta y_{t-j} + u_t, \tag{27}$$

where, corresponding to (14),

$$\begin{aligned} \alpha_{1,t} &= \sum_{i=1}^{p} \alpha_i(U_t), \\ \alpha_{j+1,t} &= -\sum_{i=j}^{p} \alpha_i(U_t), \quad j = 1, ..., p. \end{aligned}$$

Under regularity conditions, if $\alpha_{1,t} = 1$, y_t contains a unit root and is persistent; and if $|\alpha_{1,t}| < 1$, y_t is stationary. Notice that equations (14), (26) and (27) are equivalent representations of our econometric model. Each representation is convenient for inference analysis.

Appendix B. Proofs of Propositions

Proof of Proposition 1. Consider the ADF representation of the QAR(p) model (27). The existence and uniqueness of the critical conditional quantile is proven by the following simple argument:

For $\forall t \in \Upsilon$, let u_t be a realization of the iid uniform random variable U_t such that $\alpha_{1,t} = \sum_{i=1}^{p} \alpha_i(u_t) = 1$, and $\alpha_{j+1,t} = -\sum_{i=j+1}^{p} \alpha_i(u_t) = \alpha_{j+1} \quad j = 1, ..., p$. By assumptions 1 and 2, $\alpha_i (u_t)$, $i = 0, ...p$, are increasing functions in u_t.Since the sum of monotone increasing functions is itself a monotone increasing function, it follows that $\alpha_{1,t}(u_t)$ and $\alpha_{j+1,t}(u_t)$ are monotone increasing. Assumptions 1 and 2 guarantee that $Q_{\alpha_i(U_t)} = \alpha_i(Q_{U_t}) = \alpha_i(\tau)$, which is an increasing function in τ. Moreover, comonotonicity guarantees that $Q_{\sum_{i=1}^{p} \alpha_i(U_t)} = \sum_{i=1}^{p} Q_{\alpha_i(U_t)} = \sum_{i=1}^{p} \alpha_i (Q_{U_t}) = \sum_{i=1}^{p} \alpha_i (\tau)$. This implies that $\alpha_{1,t}(\tau)$ and $\alpha_{j+1,t}(\tau)$

are monotone increasing in τ. Thus, assumptions 1 and 2 guarantee that the conditional quantile function of y_t is monotone increasing in τ.

Given assumption 1, we know that $u_t \in (0,1)$. Based on the above argument, it is always possible to find a unique quantile τ^* such that $\alpha_{1,t}(\tau^*) = \sum_{i=1}^{p} \alpha_i(\tau^*) = 1$ and $\alpha_{j+1,t}(\tau^*) = -\sum_{i=j+1}^{p} \alpha_i(\tau^*) = \alpha_{j+1}$. This suggests the nice result that

$$Q_{y_t}(\tau^* \mid \mathcal{F}_{t-1}) = y_t, \ \forall t \in \Upsilon$$

that is, the trajectory of the conditional quantile function $Q_{y_t}(\tau^* \mid \mathcal{F}_{t-1})$ will hit the points in which the time series process y_t has a unit root behavior.

Now recall that the critical quantile $\tau_{crit.}$ is the largest quantile τ such that $\alpha_{1,t}(\tau) < 1$. Define $\widetilde{\Gamma} \subseteq \Gamma = (0;1)$ as the subset of quantiles so that $\alpha_{1,t}(\tau) < 1$. Hence, based on the fact that $\alpha_{1,t}(\tau)$ is monotone increasing in τ, it follows that the critical quantile is

$$\tau_{crit.} = \sup \widetilde{\Gamma}$$

Thus, $\tau_{crit.} < \tau^*$ by definition and, since $Q_{y_t}(\tau \mid \mathcal{F}_{t-1})$ is monotone increasing in τ, we must have that $Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1})$ must lie below y_t in all periods where y_t is nonstationary, that is, $Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1}) < y_t = Q_{y_t}(\tau^* \mid \mathcal{F}_{t-1})$, $\forall t \in \Upsilon$ ■

Proof of Corollary 1. Based assumptions 4 and 5, we have that the public debt process y_{t_i} for $\forall t_i \in \Upsilon$ is now represented by

$$y_{t_i} = y_{t_i - 1} + \sum_{j=1}^{p-1} \alpha^*_{j+1,t_i} \Delta y_{t_i - j} \ \ ; \ \ i = 1, ..., N$$

which is the same y_{t_i} process discussed in Proposition 1, but without intercept. In the same manner, the conditional quantile function can be written, by using the ADF formulation, as

$$Q_{y_t}(\tau \mid \mathcal{F}_{t-1}) = \alpha_{1,t}(\tau) y_{t-1} + \sum_{j=1}^{p-1} \alpha_{j+1,t}(\tau) \Delta y_{t-j}$$

This way, the proof of Corollary 1 is achieved in a straightforward manner, by just following the proof of Proposition 1 considering no intercept in the stochastic process y_t, given that the local analysis of public debt depends on the zero-mean process assumption (or global sustainability for public debt). ■

Proof of Proposition 2. By definition, we have that $H \equiv \frac{1}{T} \sum_t I_t \left\{ y_t > Q_{y_t}(\tau_{crit.} | \mathcal{F}_{t-1}) \right\}$, where $I_t(.)$ is an indicator function and T is the sample size. By Assumption 3, we can rewrite this expression as $H = \frac{1}{T} \left(\sum_{t \in \Upsilon} I_t \left\{ y_t > Q_{y_t}(\tau_{crit.} | \mathcal{F}_{t-1}) \right\} + \sum_{t \in (\Omega / \Upsilon)} I_t \left\{ y_t > Q_{y_t}(\tau_{crit.} | \mathcal{F}_{t-1}) \right\} \right) = \frac{(N+0)}{T}$, based on Proposition 1. On the other hand, we can state the critical quantile as $\tau_{crit.} = \Pr\left(y_t < Q_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1}) \right) = \Pr\left(t \in [\Omega / \Upsilon] \mid \mathcal{F}_{t-1} \right)$ $= \Pr\left(t \in \Omega \mid \mathcal{F}_{t-1} \right) - \Pr\left(t \in \Upsilon \mid \mathcal{F}_{t-1} \right) = 1 - \frac{N}{T} = 1 - H$ ■

Proof of Proposition 3. Notice that for each realization of U_{T+s}, there is a 1:1 corresponding quantile $\tau = u_{T+s}$. Hence, let u_{crit} be the largest realization of U_{T+s} so that $\alpha_{1,t} = \sum_{i=1}^{p} \alpha_i(U_{T+s}) < 1$, which guarantees stationarity whenever the realization u_{crit} takes place. By proposition 1, there exist a critical quantile $\tau_{crit.} = u_{crit}$, and its corresponding conditional critical quantile $\widehat{D}_{T+s} = \widehat{Q}_{y_{T+s}}(\tau_{crit.} \mid \mathcal{F}_T)$, so that $\widehat{y}_{T+s} = \widehat{D}_{T+s}$ whenever the realization u_{crit} takes place. Given that the process y_t has zero mean (no intercept) and $\widehat{y}_{T+s}$ is a forecasted path of y_t, it follows that $\lim_{s\to\infty} \widehat{D}_{T+s} = 0$. ■

Appendix C. Estimation and Hypothesis Testing

Provided that the right hand side of (14) is monotone increasing in U_t, it follows that the τth conditional quantile function of y_t can be written as

$$Q_{y_t}(\tau \mid y_{t-1}, ..., y_{t-p}) = \alpha_0(\tau) + \alpha_1(\tau)\, y_{t-1} + ... + \alpha_p(\tau)\, y_{t-p}, \tag{28}$$

or somewhat more compactly as

$$Q_{y_t}(\tau \mid y_{t-1}, ..., y_{t-p}) = x_t' \alpha(\tau),$$

where $x_t' = (1, y_{t-1}, ..., y_{t-p})'$. The transition from (14) to (28) is an immediate consequence of the fact that for any monotone increasing function g and a standard uniform random variable, U, we have:

$$Q_{g(U)}(\tau) = g(Q_U(\tau)) = g(\tau),$$

where $Q_U(\tau) = \tau$ is the quantile function of U_t. Analogous to quantile estimation, quantile autoregression estimation involves the solution to the problem

$$\min_{\{\alpha \in R^{p+1}\}} \sum_{t=1}^{n} \rho_\tau(y_t - x_t'\alpha), \tag{29}$$

where ρ_τ is defined as in Koenker and Basset (1978):

$$\rho_\tau(u) = \begin{cases} \tau u, u \geq 0 \\ (\tau - 1)\, u, u < 0 \end{cases}.$$

The quantile regression method is robust in distributional assumptions, a property that is inherited from the robustness of the ordinary sample quantiles. Moreover, in quantile regression, it is not the magnitude of the dependent variable that matters but its position relative to the estimated hyperplane. As a result, the estimated coefficients are less sensitive to outlier observations than, for example, the OLS estimator. This superiority over OLS estimator is common to any M-estimator.[27]

Autoregressive Order Choice

Equation (14) gives our pth order quantile autoregression model. We now discuss how to choose the optimal lag length p. We follow Koenker and Machado (1999) in testing for the null hypothesis of exclusion for the pth control variable.τ

$$H_0 : \alpha_p(\tau) = 0, \quad \text{for all } \tau \in \Gamma, \tag{30}$$

[27] The quantile estimator is (in fact) an M-estimator.

and some index set $\Gamma \subset (0,1)$. Let $\widehat{\alpha}(\tau)$ denote the minimizer of

$$\widehat{V}(\tau) = \min_{\{\alpha \in \mathbb{R}^{p+1}\}} \sum \rho_\tau (y_t - x_t' \alpha),$$

where $x_t' = (1, y_{t-1}, y_{t-2}, ..., y_{t-p})'$and $\widetilde{\alpha}(\tau)$ denotes the minimizer for the corresponding constrained problem without the pth autoregressive variable, with

$$\widetilde{V}(\tau) = \min_{\{\alpha \in \mathbb{R}^{p}\}} \sum \rho_\tau (y_t - x_{1t}' \alpha),$$

where $x_{1t}' = \left(1, y_{t-1}, y_{t-2}, ..., y_{t-(p-1)}\right)'$. Thus, $\widehat{\alpha}(\tau)$ and $\widetilde{\alpha}(\tau)$ denote the unrestricted and restricted quantile regression estimates. Koenker and Machado (1999) state that we can test the null hypothesis (30) using a related version of the Likelihood process for a quantile regression with respect to several quantiles. Suppose that the $\{u_t\}$ are iid but drawn from some distribution, say, F, and satisfying some regularity conditions. The LR statistics at a fixed quantile is derived as follows:

$$L_n(\tau) = \frac{2\left(\widetilde{V}(\tau) - \widehat{V}(\tau)\right)}{\tau(1-\tau)s(\tau)}, \tag{31}$$

where $s(\tau)$ is the sparsity function

$$s(\tau) = \frac{1}{f(F^{-1}(\tau))}.$$

The sparsity function, also termed the quantile-density function, plays the role of a nuisance parameter. We want to carry out a joint test about the significance of the pth autoregressive coefficient with respect to a set of quantiles Γ (not only at fixed quantile). Koenker and Machado (1999) suggest using the Kolmogorov-Smirnov type statistics for the joint test:

$$\sup_{\tau \in \Gamma} L_n(\tau),$$

and show that under the null hypothesis (30):

$$\sup_{\tau \in \Gamma} L_n(\tau) \rightsquigarrow \sup_{\tau \in \Gamma} Q_1^2(\tau),$$

where $Q_1(\cdot)$ is a Bessel process of order 1. Critical values for $supQ_q^2(\cdot)$ are extensively tabled in Andrews (1993).

Global Stationarity

Given the choice of the optimal lag length p, one must check for global stationarity of the y_t process, in order to verify whether conditions C.1 and C.2 described in section 3 indeed hold, and y_t is covariance stationary in the sense of Koenker & Xiao (2004b). An approach for testing the unit root property is to examine it over a range of quantiles $\tau \in \Gamma$, instead of focusing only on a selected quantile. We may, then, construct a Kolmogorov-Smirnov (KS) type test based on the regression quantile process for $\tau \in \Gamma$. We considered $\tau \in \Gamma = [0.1, 0.9]$ with steps of 0.005. Koenker and Xiao (2004b) proposed the following quantile regression based statistics for testing the null hypothesis of a unit root:

$$QKS = \sup_{\tau \in \Gamma} \mid U_n(\tau) \mid, \tag{32}$$

where $U_n(\tau)$ is the coefficient based statistics given by:

$$U_n(\tau) = n(\widehat{\alpha}_1(\tau) - 1).$$

Koenker and Xiao (2004b) suggest the approximation of the limiting distribution of (32) under the null hypothesis by using the autoregressive bootstrap (ARB). In this paper, we approximate the distribution under the null using the residual based block bootstrap procedure (RBB). The advantages of the RBB over ARB are documented in Lima and Sampaio (2005).

Unconditional Mean Test

In order to test whether or not the unconditional mean of the process is zero, we recall that the following null hypotheses are equivalent:

$$\begin{aligned} H_0 &: \mu_y = 0 \\ H_0' &: \mu_0 = 0 \end{aligned}$$

Consider the pth order quantile autoregressive process given by

$$\begin{aligned} y_t &= \alpha_0(U_t) + \alpha_1(U_t)y_{t-1} + ... + \alpha_p(U_t)y_{t-p} \\ &= \mu_0 + \beta_{1,t}y_{t-1} + ... + \beta_{p,t}y_{t-p} + u_t \end{aligned}$$

where $u_t = \alpha_0(U_t) - \mu_0$. Now note that the τth conditional quantile function of y_t is given by

$$Q_{y_t}(\tau \mid y_{t-1}, ..., y_{t-p}) = \alpha_0(\tau) + \alpha_1(\tau)y_{t-1} + ... + \alpha_p(\tau)y_{t-p},$$

and it does not allow us to identify the intercept coefficient μ_0, since $Q_u(\tau) = \alpha_0(\tau) - \mu_0$, where $\tau = Q_U(\tau)$ is the quantile function of U. Thus, the next natural attempt would be to ignore the existence of an asymmetric dynamic and estimate a symmetric regression (constant coefficient model)

$$y_t = \mu_0 + \beta_1 y_{t-1} + ... + \beta_p y_{t-p} + v_t. \tag{33}$$

The null hypothesis H_0' could be tested using conventional t-statistics

$$t = \frac{\widehat{\mu_0}}{\widehat{SE(\mu_0)}}.$$

However, in omitting asymmetries, the new error term v_t is no longer an iid sequence, i.e.,

$$v_t = \left(\beta_{1,t} - \beta_1\right)y_{t-1} + ... + \left(\beta_{p,t} - \beta_p\right)y_{t-p} + u_t,$$

which *invalidates* the conventional t-statistics type test. Putting that aside, we decided to directly test the null hypothesis $H_0 : \mu_y = 0$ using a resampling method for dependent data according to Carlstein (1986), named Nonoverlapping Block Bootstrap (NBB). The key feature of this bootstrap method is that its blocking rule is based on nonoverlapped segments of the data, making it able to simulate the weak dependence in the original series, y_t. Further details regarding NBB bootstrap are available in Lahiri (2003).

The Koenker-Xiao Test

In this section, we introduce the Koenker-Xiao test, which is used to test the null hypothesis $H_0 : \alpha_1(\tau) = 1$, for a given $\tau \in (0,1)$. We express the null hypothesis in the ADF representation (16) as:

$$H_0 : \alpha_1 (\tau) = 1, \text{ for selected quantiles } \tau \in (0,1).$$

In order to test such a hypothesis, Koenker and Xiao (2004b) proposed a statistic similar to the conventional augmented Dick-Fuller (ADF) t-ratio statistic. The t_n statistic is the quantile autoregression counterpart of the ADF t-ratio test for a unit root and is given by:

$$t_n (\tau) = \frac{f\,(\widehat{F^{-1}(\tau)})}{\sqrt{\tau (1-\tau)}} \left(Y_{-1}^{\intercal} P_X Y_{-1}\right)^{\frac{1}{2}} (\widehat{\alpha}_1 (\tau) - 1),$$

where, $f\,(\widehat{F^{-1}(\tau)})$ is a consistent estimator of $f\left(F^{-1}(\tau)\right)$, Y_{-1} is a vector of lagged dependent variables (y_{t-1}) and P_X is the projection matrix onto the space orthogonal to $X = (1, \Delta y_{t-1}, ..., \Delta y_{t-p+1})$. Koenker and Xiao (2004b) show that the limiting distribution of $t_n(\tau)$ can be written as:

$$t_n (\tau) \Rightarrow \delta \left(\int_0^1 \underline{W}_1^2\right)^{-\frac{1}{2}} \int_0^1 \underline{W}_1 dW_1 + \sqrt{1-\delta^2} N(0,1),$$

where $\underline{W}_1 (r) = W_1 (r) - \int_0^1 W_1 (s)\, ds$ and $W_1 (r)$ is a standard Brownian Motion. Thus, the limiting distribution of $t_n(\tau)$ is nonstandard and depends on parameter δ given by:

$$\delta = \delta(\tau) = \frac{\sigma_{\omega\psi}(\tau)}{\sigma_\omega^2},$$

and can be consistently estimated (see Koenker and Xiao, 2004b, for more details). Critical values for the statistic $t_n(\tau)$ are provided by Hansen (1995, page 1155) for values of δ^2 in steps of 0.1. For intermediate values of δ^2, Hansen suggests obtaining critical values by interpolation.

Appendix D. Monte Carlo Simulation

A Monte Carlo simulation is designed to investigate the finite sample performance of the result shown in Proposition 1, that is, if the critical conditional quantile is able to separate nonstationary points from stationary ones. One of the critical issues regarding this experiment is the Data-Generating Process (DGP), which will be represented by the following QAR(1) model

$$y_t = \alpha_0 (U_t) + \alpha_1 (U_t)\, y_{t-1}, \tag{34}$$

where $\{U_t\}$ is a sequence of iid standard uniform random variables, and the coefficients α_0 and α_1 are functions on $[0,1]$, given by $\alpha_0 (U_t) = F^{-1}(U_t)$, where $F : \mathbb{R} \to [0,1]$ is the standard normal cumulative distribution function, and $\alpha_1 (U_t) = \min\{1 \;\; ; \;\; \gamma_0 + \gamma_1 U_t\}$ with $\gamma_0 \in (0,1)$ and $\gamma_1 > 0$.

In our case, we initially assume $\gamma_0 = 0.7$ and $\gamma_1 = 0.4$ in order to limit the variance of α_1. If $U_t > \frac{(1-0.7)}{0.4} = 0.75$, then the model generates y_t according to the unit root model, but for smaller realizations of U_t we have a mean reversion tendency. In other words, we expect that 25% of the U_t realizations will induce a unit root behavior. We also consider the case $\gamma_0 = 0.8$, which leads to 50% of the realizations of U_t generating a unit root model.

In our experiment, we construct 10,000 replications of $\{y_t\}$ with 100 or 300 observations. We adopt a hybrid solution for this experiment using R and Ox environments, since the proposed simulation is extremely computational intensive. Ox is much faster than R in large computations. On the other hand, R language is more interactive and user-friendly than Ox, and the QAR model must be estimated in R, since its package for quantile regressions (quantreg) is more complete and updated than the Ox package. The main steps of the algorithm used in the Monte Carlo simulation are as follows:[28]

a) Initialization of the R code (setting parameters γ_0, γ_1)

b) Generation of one DGP

- b.1) R code calls Ox code informing the input parameters
- b.2) Ox code generates one DGP y_t
- b.3) R code imports the data generated by Ox code

c) Calculation of the optimal lag length (p) for the QAR(p) model

d) Estimation of the coefficients for the QAR(p) model

e) Testing for local unit root in all quantiles

f) Search for the critical quantile

g) Generation of the conditional quantiles

h) Computation of the Debt Ceiling

i) Saving of the results for this DGP

j) Repeat steps (b) to (i) for 10,000 replications

Therefore, we proceed as follows: Ox code initially generates the time series y_t and then returns these data to R, which estimates the QAR(p) model, computes the descriptive statistics and saves the results in a text file. Once the Ox code generates the $\{y_t\}$ process, the optimal lag length of the QAR(p) model is chosen based on the Koenker and Machado (1999) procedure. This way, the coefficients are estimated for all quantiles and a local unit root test is conducted in order to find the critical quantile $\tau_{crit.}$, i.e., the last quantile associated with an autoregressive coefficient, which still represents a mean reversion tendency (or in other words, where the null $H_0 : \alpha_1(\tau) = 1$ is still rejected, according to the Koenker-Xiao test for unit root). Furthermore, the R code generates the conditional quantiles, including the critical quantile, according to the following ADF formulation

$$\widehat{Q}_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1}) = \widehat{\alpha}_0(\tau_{crit.}) + \widehat{\alpha}_1(\tau_{crit.})\, y_{t-1} + \sum_{j=1}^{p-1} \widehat{\alpha}_{j+1}(\tau_{crit.})\, \Delta y_{t-j}. \tag{35}$$

Based on the critical conditional quantile, one can verify if the adopted QAR(p) model for a finite sample is able to correctly identify the stationarity limit, by comparing the $\{y_t\}$ process with $\widehat{Q}_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1})$

[28] Both R and Ox codes are available from the authors upon request.

for observations where the DGP imposes a unit root model. To investigate this issue carefully, lets initially define (for a given replication i) the following dummy variables W_t^i and Z_t^i:

$$W_t^i = \left\{ \begin{array}{l} 1 \;\; ; \;\; \text{if } \alpha_1 (U_t) = 1 \\ 0 \;\; ; \;\; \text{otherwise} \end{array} \right\}, \tag{36}$$

$$Z_t^i = \left\{ \begin{array}{l} 1 \;\; ; \;\; \text{if } y_t > \widehat{Q}_{y_t}(\tau_{crit.} \mid \mathcal{F}_{t-1}) \;\; \text{and} \;\; \alpha_1 (U_t) = 1 \\ 0 \;\; ; \;\; \text{otherwise} \end{array} \right\}. \tag{37}$$

Thus, the W_t^i variable indicates observations with an autoregressive coefficient equal to unity, according to the DGP, and Z_t^i reveals observations associated with a unit root behavior and, at the same time, where the generated y_t time series is above the critical conditional quantile. Note that $\frac{1}{T}\sum\limits_{t=1}^{T} Z_t^i = H^i$, which is exactly the H statistic, presented in definition 3, computed for replication i. Therefore, one can compute the ratio R^i as follows:

$$R^i \equiv \frac{\frac{1}{T}\sum\limits_{t=1}^{T} Z_t^i}{\frac{1}{T}\sum\limits_{t=1}^{T} W_t^i}. \tag{38}$$

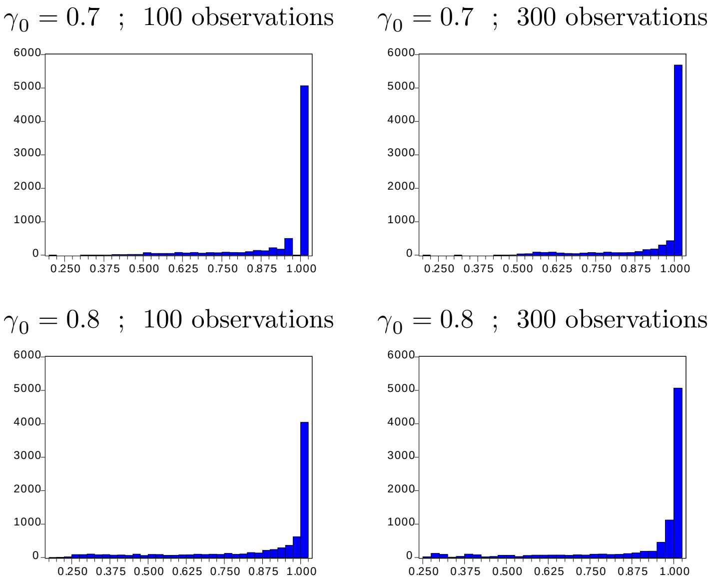
One should expect the ratio R^i to be as close to unity as possible, since in the QAR(p) model all observations of the y_t process associated with a unit root model must be above the critical quantile, according to Proposition 1.

Our simulation computes the R^i statistic for each replication i and summarizes the results in the following histograms, where the frequency of R^i is plotted for the set of 10,000 replications.[29] It is worth mentioning that only the replications in which the null hypothesis of a local unit root for the y_t process can not be rejected are displayed in the following histograms. In other words, we select among the 10,000 replications only those representing a stochastic process y_t containing at least one quantile with a local unit root, i.e., the null $H_0 : \alpha_1 (\tau) = 1$ is not rejected for (at least) one quantile $\tau' \in (0, 1)$.

Since U_t follows a standard uniform distribution and $\alpha_1 (U_t) = \min\{1 \;\; ; \;\; \gamma_0 + \gamma_1 U_t\}$ it is possible that for a given replication j the stochastic process $\{y_t^j\}$ has no local unit root, i.e., $\alpha_1(\tau) < 1; \forall \tau \in (0, 1)$. In fact, these cases occur for lower values of γ_0 and γ_1, but since they are not the object of our investigation we decided to not consider them in our analysis.

[29] Each vertical bar graph represents the frequency distribution of R^i, in which the height of the bar is proportional to the frequency within each class interval.

Figure 7 - Histograms with the frequency of R^i



Note: Total of i=10,000 replications, excluding those with no local unit root.

Table 7 - Summary of the Monte Carlo results

Parameter γ_0	T	T^*	mean (R_i)	median (R_i)	std.dev.(R_i)
$\gamma_0 = 0.7$	100	25	0.935	1.000	0.131
	300	75	0.951	1.000	0.112
$\gamma_0 = 0.8$	100	50	0.866	0.982	0.213
	300	150	0.909	1.000	0.182

Notes: (a) Total of i=10,000 replications for each simulation, excluding those with no local unit root;

(b) T is the total number of observations and T^* is the expected number of observations,

across the 10,000 replications, associated with a unit root model.

According to the Monte Carlo experiment, we found that the result of Proposition 1 indeed exhibits a good performance in the finite sample investigation. As long as the number of observations T increases (for a given parameter γ_0), the empirical distribution of R^i approaches the unity value, with a respective decreasing standard deviation, as we already expected. In our simulations, the distribution of R^i for $\gamma_0 = 0.7$ is more concentrated than the respective distribution for $\gamma_0 = 0.8$, since the DGP for $\gamma_0 = 0.8$ induces a larger expected number T^* of realizations of U_t generating a unit root model. In this case, for $\gamma_0 = 0.7$ and $T = 100$ observations, we found that (on the average) the QAR model imposes 93,5% of observations of the y_t process associated with a unit root model (T^*) correctly above the estimated critical quantile.

Appendix E. Out-of-sample Forecast: generation of a discrete uniform random variable

In practical terms, the numerical procedure described in the construction of the out-of-sample forecast of y_t must be implemented by an algorithm considering a perfect match between the discrete set of quantiles $\tau \in \Lambda = [0.1, ..., 0.9]$ and a discrete support of the U_t random variable. Firstly, we must choose the number of elements n for the grid Λ of quantiles and, then, estimate the QAR model to generate the set of coefficients $\widehat{\alpha}_i(\tau)$ for all $\tau \in \Lambda$. The discrete set of quantiles Λ, containing n elements, is defined by

$$\tau \in \Lambda \equiv [0.1, 0.1 + \tau_{step}, 0.1 + 2\tau_{step}, ..., 0.9 - \tau_{step}, 0.9], \tag{39}$$

where $\tau_{step} = (0.9 - 0.1)/(n-1)$. In addition, one must ensure that the dropping of the discrete version of the random variable U_t, defined as $\widetilde{U}_t$, is made based on the same set Λ, in order to guarantee that, for every realization of U_t, the algorithm correctly calculates the respective $\widetilde{U}_t$, in order to find an associated quantile τ and, therefore, an estimated coefficient $\widehat{\alpha}_i\left(\tau = \widetilde{U}_t\right)$.

This way, a perfect 1:1 mapping between τ and U_t depends on the random variable $\widetilde{U}_t$, which can be obtained from the realization of the continuous random variable U_t, in the following way: Assume that U_t belongs to the continuous set $[0.1, 0.9]$. If we define $\widetilde{U}_t$ as follows, we can guarantee that indeed $\widetilde{U}_t$ belongs to the same discrete set Λ of quantiles:

$$\widetilde{U}_t \equiv 0.1 + \tau_{step} * round\{\frac{(U_t - 0.1)}{\tau_{step}}\}, \tag{40}$$

where the $round(.)$ function approximates its argument to the nearest integer value.[30]

References

[1] Alesina, A., Drazen, A., 1991. Why are Stabilization Delayed? American Economic Review 81(5), 1170-1188.

[2] Andrews, D.W.K., 1993. Tests for Parameter Instability and Structural Change with Unknown Change Point. Econometrica 61, 821-856.

[3] Arestis, P., Cipollini, A., Fattouh, B., 2004. Threshold Effects in the U.S. Budget Deficit. Economic Inquiry 42(2), 214-222.

[30] Lets present a simple example for some grid Λ_1, with $n = 9 \therefore \tau_{step} = 0.1 \therefore \Lambda_1 = [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]$. Now, assume the realized value $U_t = 0.376815$. By definition, we have $\widetilde{U}_t = 0.1 + 0.1 * round\{\frac{(0.376815-0.1)}{0.1}\} = 0.4 \in \Lambda_1$. Therefore, for a given drop of U_t from a continuous range set $[0.1, 0.9]$ we constructed its respective discrete version $\widetilde{U}_t$ that indeed belongs to the considered grid Λ_1 of quantiles. Now consider another example, with a different grid Λ_2 by assuming $n = 3 \therefore \tau_{step} = 0.4$ and $\Lambda_2 = [0.1, 0.5, 0.9]$. If one sorts again $U_t = 0.376815$, then $\widetilde{U}_t = 0.1 + 0.4 * round\{\frac{(0.376815-0.1)}{0.4}\} = 0.5 \in \Lambda_2$. On the other hand, if $U_t = 0.68$, then $\widetilde{U}_t = 0.5 \in \Lambda_2$, but if $U_t = 0.72$, then $\widetilde{U}_t = 0.9 \in \Lambda_2$.

[4] Bohn, H., 2004. The sustainability of Fiscal Policy in the United States. Department of Economics, University of California at Santa Barbara (November 2004).

[5] Carlstein, E., 1986. The Use of Subseries Methods for Estimating the Variance of a General Statistic from a Stationary Time Series. The Annals of Statistics 14, 1171-1179.

[6] Davig, T., 2005. Periodically Expanding discounted public debt : A Threat to Fiscal Policy Sustainability? Journal of Applied Econometrics 20(7), 829-840.

[7] Garcia, M., Rigobon, R., 2004. A Risk Management Approach to Emerging Market's Sovereign Debt Sustainability with an Application to Brazilian Data. NBER Working Paper 10336.

[8] Giambiagi, F., Ronci, M., 2004. Fiscal Policy and Debt Sustainability: Cardoso's Brazil, 1995-2002. IMF Working Paper 04/156.

[9] Giavazzi, F., Pagano, M., 1990. Can Severe Fiscal Contractions be Expansionary? Tales of Two Small European Countries. NBER Working Papers 3372.

[10] Granger, C., Pesaran, H., 2000. Economic and Statistical Measures of Forecast Accuracy. Journal of Forecasting 19(7), 537-560.

[11] Hafner, C.M., Linton, O.B., 2006. Discussion of Quantile Autoregression by Koenker and Xiao, Available at http://personal.lse.ac.uk/lintono/downloads/disckxfin.pdf.

[12] Hakkio, C.S., Rush, M., 1991. Is the Budget Deficit Too Large? Economic Inquiry 24, 429-445.

[13] Hamilton, J.D., Flavin, M.A., 1986. On the Limitations of Government Borrowing: A Framework for Empirical Testing. American Economic Review 76, 808-819.

[14] Hansen, B., 1995. Rethinking the univariate approach to unit root tests: How to use Covariates to increase power. Econometric Theory 8, 1148-1171.

[15] Haug, A.A., 1991. Cointegration and Government Borrowing Constraints: Evidence for the United States. Journal of Business and Economic Statistics 9(1), 97-111.

[16] Issler, J.V., Lima, L.R., 2000. Public Debt Sustainability and Endogenous Seigniorage in Brazil: Time Series Evidence from 1947-1992. Journal of Development Economics 62, 131-147.

[17] Koenker, R., Bassett, G., 1978. Regression Quantiles. Econometrica 46, 33-49.

[18] Koenker, R., Bassett, G., 1982. Tests of Linear Hypotheses and l_1 Estimation. Econometrica 50, 1577-1584.

[19] Koenker, R., Machado, J.A.F., 1999. Goodness of Fit and Related Inference Processes for Quantile Regression. Journal of the American Statistical Association 94(448), 1296-1310.

[20] Koenker, R., Xiao, Z., 2002. Inference on the Quantile Regression Process. Econometrica 70, 1583-1612.

[21] Koenker, R., Xiao, Z., 2004a. Quantile Autoregression. Working paper. University of Illinois at Urbana-Champaign. Available at: http://www.econ.uiuc.edu/ ~roger/research/qar/qar9.pdf.

[22] Koenker, R., Xiao, Z., 2004b. Unit Root Quantile Autoregression Inference. Forthcoming Journal of the American Statistical Association.

[23] Lahiri, S.N., 2003. Resampling Methods for Dependent Data. Springer Series in Statistics.

[24] Lima, L.R., Sampaio, R.M.B., 2005. The Asymmetric Behavior of the U.S. Public Debt. Working Paper, Getulio Vargas Foundation, EPGE 593.

[25] Lima, L.R., Simonassi, A., 2005. Dinâmica Não-Linear e Sustentabilidade da Dívida Pública Brasileira. Pesquisa e Planejamento Econômico 35 (2), 1-20.

[26] Luporini, V., 2000. Sustainability of the Brazilian Fiscal Policy and Central Bank Independence. Revista Brasileira de Economia 54(2), 201-226.

[27] Paparoditis, E., Politis, D.N., 2003. Residual-Based Block Bootstrap for Unit Root Testing. Econometrica 71(3), 813-855.

[28] Pastore, A.C., 1995. Déficit Público, a Sustentabilidade do Crescimento das Dívidas Interna e Externa, Senhoriagem e Inflação: Uma Análise do Regime Monetário Brasileiro. Revista de Econometria 14(2), 177-234.

[29] Reinhart, C.M., Rogoff, K.S., Savastano, M.A., 2003. Debt Intolerance. Brookings Papers on Economic Activity, 1:2003.

[30] Rocha, F., 1997. Long-Run limits on the Brazilian Government Debt. Revista Brasileira de Economia 51(4), 447-470.

[31] Rocha, F., 2001. Is There any Rationale to The Brazilian Fiscal Policy? Revista Brasileira de Economia 55(3), 315-331.

[32] Sarno, L., 2001. The behavior of US public debt: a nonlinear perspective. Economics Letters 74(1), 119-125.

[33] Trehan, B., Walsh, C.E., 1988. Common Trends, the Government's Budget Constraint and Revenue Smoothing. Journal of Economic Dynamics and Control 12, 425-444.

[34] Uctum, M., Wickens, M., 2000. Debt and deficit ceilings, and sustainability of fiscal policies: an intertemporal analysis. Oxford Bulletin of Economics and Statistics 62 (2), 197-222.

[35] Wilcox, D., 1989. The Sustainability of Government Deficits: Implications of the Present-Value Borrowing Constraint. Journal of Money, Credit and Banking 3, 291-306.

Chapter 3

Evaluating Value-at-Risk models via quantile regressions[1]

Abstract

We propose an alternative backtest to evaluate the performance of Value-at-Risk (VaR) models. The presented methodology allows us to directly test the performance of many competing VaR models, as well as identify periods of an increased risk exposure based on a quantile regression model (Koenker & Xiao, 2002). Quantile regressions provide us an appropriate environment to investigate VaR models, since they can naturally be viewed as a conditional quantile function of a given return series. A Monte Carlo simulation is presented, revealing that our proposed test might exhibit more power in comparison to other backtests presented in the literature. Finally, an empirical exercise is conducted for daily S&P500 series in order to explore the practical relevance of our methodology by evaluating five competing VaRs through four different backtests.

JEL Classification: C12, C14, C52, G11
Keywords: Value-at-Risk, Backtesting, Quantile Regression

[1] This article was jointly made with Luiz Renato Lima and Oliver Linton. We would like to thank Qiwei Yao and Myung Seo for their helpful comments and suggestions, and the seminar participants of the Joint PhD afternoon - LSE Economics and Statistics departments, where this paper was first presented (May 2007). We are also grateful to the participants of the 2nd AlBan Conference - Grenoble, and all the participants of the Empirical Finance for Central Banks course at CCBS - Bank of England, specially Ibrahim Stevens, Francoise Ben Zur, David Delgado and Sarat Dhal.

1 Introduction

Several large scale crashes and financial losses in the previous decades, such as the "Black Monday" in 1987 (with a 23% drop in value of U.S. stocks, equivalent to $1trillion lost in one day), or the Asian turmoil of 1997 and Russian default (leading to a near failure of LTCM) in 1998 and, more recently, the World Trade Center attack in 2001, freezing the financial market for six days, with U.S. stock market losses of $1.7 trillion, have brought risk management of financial institutions to the forefront of internal control and regulatory debate. Value-at-Risk (hereafter, VaR) models arose as a subject for both regulators and investors concerned with large crashes and the respective adequacy of capital to meet such risk.

In fact, VaR is a statistical risk measure of potential losses, and summarizes in a single number the worst loss over a target horizon that will not be exceeded with a given level of confidence. Despite several other competing risk measures proposed in the literature, VaR has effectively become a cornerstone of internal risk management systems in financial institutions, following the success of the J.P. Morgan RiskMetrics system, and nowadays form the basis of the determination of market risk capital, since the 1996 Amendment of the Basel Accord.

A crucial question that arises in this context is how to evaluate the performance of a VaR model? When several risk forecasts are available, it is desirable to have formal testing procedures for comparison, which do not necessarily require knowledge of the underlying model, or, if the model is known, which do no restrict attention to a specific estimation procedure. The literature has proposed several tests (also known as "backtests"), such as Kupiec (1995), Christoffersen (1998) and Engle and Manganelli (2004), mainly focused on a hit sequence from which statistical properties are derived and further tested.

In this article, we go a step further by arguing that these backtests may provide only necessary but not sufficient conditions to test whether or not a given VaR measure is properly specified. In fact, by investigating solely a violation sequence, one might ignore an important piece of information contained in the magnitude of violations. In this sense, we propose an alternative backtest based on a quantile regression framework that can properly account for it. It is natural to evaluate a VaR model by a quantile regression method due to its capability of conditional distribution exploration with distribution-free assumption, also allowing for serial correlation and conditional heteroskedasticity. Furthermore, Value-at-Risk models are nothing else than conditional quantiles functions, as will be further explored throughout this paper.

There are a variety of approaches to estimate conditional quantiles in general and Value-at-Risk in particular. A short list includes Koenker and Zhao (1996), Danielsson and de Vries (1997), Embrechts et al. (1997, 1999), Chernozhukov and Umantsev (2001), Christoffersen et al. (2001). For instance, Koenker and Zhao (1996) provide a discussion about conditional quantile estimation and inference under Engle's (1982) ARCH models, whereas Hafner and Linton (2006) show that a QAR(p) process can be represented by a semi-strong ARCH(p) process, and the GARCH(1,1) can be nested by a QAR process extended to infinite order.

Quantile regressions can also be used to construct VaR measures without imposing a parametric distribution or the iid assumption: Chen (2001) discusses a multiperiod VaR model based on quantile regressions, and Wu and Xiao (2002) present an ARCH quantile regression approach to estimate VaR and left-tail measures. Surprisingly, however, little empirical work has been done by using quantile regressions to evaluate competing VaR models (e.g., Engle and Manganelli (2004) and Giacomini and Komunjer (2005)).

This way, the main objective of this paper is to provide a backtest based on quantile regressions that allows us to formally evaluate (through a standard Wald statistic) the performance of a VaR model, and also permits one to identify periods of an increased risk exposure, which we believe to be a novelty in the literature. The test statistic is derived from a Mincer-Zarnowitz (1969) type-regression considered in a quantile environment.

The proposed test is quite simple to be computed and can be carried out using software available for conventional quantile regression, and also presents the advantage of making "full use of information", in the sense that takes into account the magnitudes of model violations, rather than simply checking whether the violation series follows an iid sequence. In addition, our methodology is applicable even when the VaR does not come from a conditional volatility model.

The practical relevance of our theoretical results are documented by a small Monte Carlo simulation, in which the quantile regression test seems to have more power in comparison to other backtests, previously documented in the literature as exhibiting low power to detect misspecified VaRs in finite samples (see Kupiec (1995), Pritsker (2001) and Campbell (2005)). The increased power might be due to the quantile framework, which provides an adequate null hypothesis, in comparison to other backtests, which is also an issue addressed in this paper.

This study is organized as follows: Section 2 defines Value-at-Risk, presents a quantile regression-based hypothesis test to evaluate VaRs and describes other backtests suggested in the literature. Section 3 shows the Monte Carlo simulation comparing the size and power of the competing backtests. Section 4 provides an empirical exercise based on daily S&P500 series, and Section 5 summarizes our conclusions.

2 The econometric model and assumptions

Value-at-Risk models were developed in response to the financial disasters of the early 90s, and have become a standard measure of market risk, which is increasingly used by financial and non-financial firms as well. VaR models have also been sanctioned for determining market risk capital requirements for financial institutions through the 1996 Market Risk Amendment to the Basle Accord.[2]

According to Jorion (2007), Mr. Till Guldimann is the creator of the term "Value-at-Risk", while head of global research at J.P. Morgan in the late 80s. The introduction of the VaR concept through the RiskMetrics

[2] See Appendix B for further details.

methodology has collapsed the entire distribution of the portfolio returns into a single number, which investors have found very useful and easily interpreted as a measure of market risk. Generally speaking, Value-at-Risk can be interpreted as the amount lost on a portfolio, with a given small probability, over a fixed time period.

Jorion (2007) also argues that a VaR summarizes the worst loss (or the highest gain) of a portfolio over a target horizon that will not be exceeded with a given level of confidence. The author formally defines VaR as the quantile of the projected distribution of gains and losses over the target horizon. If $\tau^* \in (0;1)$ is the selected tail level of the mentioned distribution, the respective VaR is implicitly defined by the following expression:

$$\Pr\left[R_t \leq V_t | \mathcal{F}_{t-1}\right] = \tau^*, \tag{1}$$

where $\mathcal{F}_{t-1}$ is the information set available at time $t-1$, R_t is the return series and V_t is the respective VaR. From this definition, it is clear that finding a VaR is essentially the same as finding a $(100 * \tau)\%$ conditional quantile. Note that, for convention, the VaR is defined for the right tail of the distribution, which is assumed without loss of generality, since our methodology can easily be adapted to investigate the left tail.[3] In this case, the VaR would be defined by $\Pr\left[R_t \leq -V_t | \mathcal{F}_{t-1}\right] = \tau^*$. Note that the sign is changed to avoid a negative number in the V_t time series, since the VaR is usually reported by risk managers as a positive number.

Following the idea of Christoffersen et al. (2001), one can think of generating a VaR measure as the outcome of a quantile regression, treating volatility as a regressor. For instance, from a regression of the form: $y_t = \alpha_0\left(U_t\right) + \alpha_1\left(U_t\right)\sigma_t^2$, where σ_t^2 is the conditional volatility of y_t, it follows that $Q_{y_t}\left(\tau \mid \mathcal{F}_{t-1}\right) = \alpha_0\left(\tau\right) + \alpha_1\left(\tau\right)\sigma_t^2$, which implies that the conditional quantile[4] is some linear function of volatility. In this sense, Engle and Patton (2001) argue that a volatility model is typically used to forecast the absolute magnitude of returns, but it may also be used to predict quantiles.

In this paper, we adapt the model suggested by Christoffersen et al. (2001) to investigate the accuracy of a given VaR model. In other words, instead of using the conditional volatility as a regressor, we simply use V_t in place of σ_t^2 in the above model, where V_t is the VaR measure of interest. That is exactly the idea we next explore, where a convenient hypothesis test is formally derived to evaluate VaR models. In sum, we consider the following model

$$\begin{aligned} R_t &= \alpha_0(U_t) + \alpha_1(U_t)V_t \qquad (2)\\ &= x_t'\beta(U_t), \qquad (3) \end{aligned}$$

[3] According to Nankervis et al. (2006), it is usual that VaR is separately computed for the left and right tails of the distribution depending on the position of the risk managers or traders. For traders with a long position (when they buy and hold a traded asset), the risk comes from a drop in the price of the asset, while traders with a short position (who first borrow the asset and subsequently sell it in the market) lose money when the price increases. Due to the existence of leverage effects, a well-known stylized fact in financial asset returns, models that allow positive and negative returns to have different impacts on volatility are required to compute and distinguish the VaR for the long and short positions.

[4] Where the quantile function of a given random variable z_t is defined as the reciprocal of its cumulative distributive function F_z, i.e., $Q_{z_t}(\tau) = F_z^{-1}(\tau) = \inf\{z : F(z) \geq \tau\}$.

where U_t is an iid standard uniform random variable, $U_t \sim U(0,1)$, the functions $\alpha_i(U_t)$, $i = 0, 1$ are assumed to be comonotonic, $\beta(U_t) = [\alpha_0(U_t); \alpha_1(U_t)]'$ and $x_t' = [1, V_t]$. Notice that Eq. (2) can be re-written as

$$R_t = \varphi_t + \epsilon_t, \tag{4}$$

where $\varphi_t = \alpha_0 + \alpha_1(U_t)V_t$, $\epsilon_t = \alpha_0(U_t) - \alpha_0$ is an iid random variable and $\alpha_0 = E[\alpha_0(U_t)]$. An important feature of (4) is that the conditional mean is affected by the VaR, which was computed using information available up to period $t-1$. Since the value at risk V_t is nothing else than the conditional quantile of R_t, then the above model can be seen as a quantile-in-mean model. Indeed, if we allow V_t to be equal to the conditional variance of R_t, then the above model becomes a particular case of the so-called ARCH-in-mean model introduced by Engle (1987).

Following Koenker and Xiao (2002), we will assume that the returns $\{R_t\}$ are, conditional on $\mathcal{F}_{t-1}$, independent with linear conditional quantile functions given by (5). Since V_t is already available at the end of period $t-1$, before the realization of R_t at time t, then we can compute the conditional quantile of R_t as follows:

$$Q_{R_t}(\tau \mid \mathcal{F}_{t-1}) = \alpha_0(\tau) + \alpha_1(\tau) V_t. \tag{5}$$

Now, recall the definition of Value-at-Risk (V_t), in which the conditional probability of a return R_t to be lower than V_t, over the target horizon, is equal to $\tau^* \in (0,1)$, i.e., $\Pr(R_t \leq V_t \mid \mathcal{F}_{t-1}) = \tau^*$. From this definition, it is clear that finding a VaR is exactly the same as finding a conditional quantile function. In fact, from our quantile regression methodology, we also have that $\Pr(R_t \leq Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) \mid \mathcal{F}_{t-1}) = \tau^*$. Thus, by considering that the VaR model's true level of coverage is τ^*, it follows that V_t must coincide with the related conditional quantile function of R_t at the same level τ^*. Therefore, a natural way to test for the overall performance of a VaR model is to test the null hypothesis

$$Ho : \begin{cases} \alpha_0(\tau^*) = 0 \\ \alpha_1(\tau^*) = 1 \end{cases}. \tag{6}$$

This hypothesis can be presented in a classical formulation as $Ho : R\theta(\tau^*) = r$, for the fixed quantile $\tau = \tau^*$, where R is a 2×2 identity matrix; $\beta(\tau^*) = [\alpha_0(\tau^*); \alpha_1(\tau^*)]'$ and $r = [0; 1]$. Note that, due to the simplicity of our restrictions, the later null hypothesis can still be reformulated as $Ho : \theta(\tau^*) = 0$, where $\theta(\tau^*) = [\alpha_0(\tau^*); (\alpha_1(\tau^*) - 1)]'$.[5]

The first issue to implement such a hypothesis test is to construct the confidence intervals for the estimated coefficients $\widehat{\theta}(\tau^*)$. Following Koenker (2005, p.74) the method used in this paper to compute the covariance matrix of the estimated coefficients takes the form of a Huber (1967) sandwich:[6]

$$\sqrt{T}(\widehat{\theta}(\tau^*) - \theta(\tau^*)) \xrightarrow{d} N(0, \tau^*(1-\tau^*)H_{\tau^*}^{-1}JH_{\tau^*}^{-1}) = N(0, \Lambda_{\tau^*}), \tag{7}$$

[5] Recall that our focus is to test a VaR on the right tail of the distribution of returns. In order to investigate a VaR for the left tail, one must consider the modified null hypothesis: $\widetilde{\theta}(\tau^*) = 0$, in which $\widetilde{\theta}(\tau^*) \equiv [\alpha_0(\tau^*); (\alpha_1(\tau^*) + 1)]'$.

[6] A technical issue on the estimation process emerges from the fact that the objective function is not differentiable with respect to parameters at interested quantiles. The discontinuity in first order condition of the corresponding objective function makes the derivation of asymptotics of quantile regression estimators quite difficult, since conventional techniques (based on Taylor expansion) are no more applicable. The argument of stochastic uniform continuity, called stochastic equicontinuity, is one of the solutions for deriving

where $J = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} x_t x_t'$ and $H_{\tau^*} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} x_t x_t' [f_t(Q_{R_t}(\tau^*|x_t)]$ under the quantile regression model $Q_{R_t}(\tau \mid x_t) = x_t'\theta(\tau)$. The term $f_t(Q_{R_t}(\tau^*|x_t)$ represents the conditional density of R_t evaluated at the quantile τ^*. Given that we are able to compute the covariance matrix of the estimated $\widehat{\theta}(\tau)$ coefficients, we can now construct our hypothesis test to verify the performance of the Value-at-Risk model based on quantile regressions (hereafter, VQR test).

Definition 1: Let our test statistic be defined under the null by $\zeta_{VQR} = T[\widehat{\theta}(\tau^*)'(\tau^*(1-\tau^*)H_\tau^{-1}JH_\tau^{-1})^{-1}\widehat{\theta}(\tau^*)]$.

In addition, consider the following assumptions:

Assumption 1: Let $x_t \geq 0$ be measurable with respect to $\mathcal{F}_{t-1}$ and $z_t \equiv \{R_t; x_t\}$ be a strictly stationary process;

Assumption 2: (Density) Let $\{R_t\}$ have distribution functions F_t, with continuous Lebesgue densities f_t uniformly bounded away from 0 and ∞ at the points $Q_{R_t}(\tau \mid x_t) = F_{R_t}^{-1}(\tau \mid x_t)$;

Assumption 3: (Design) There exist positive definite matrices J and H_τ, such that $J = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} x_t x_t'$ and $H_\tau = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} x_t x_t' [f_t(Q_{R_t}(\tau \mid x_t))]$;

Assumption 4: $\max_{i=1,\ldots,T} \|x_i\| / \sqrt{T} \to 0$.

The following Proposition, which is merely an application of Hendricks and Koenker (1992) and Koenker (2005, Theorem 4.1), by considering a fixed quantile τ^*, summarizes our VQR test, designed to check whether the Value-at-Risk model (V_t) equals the respective conditional quantile function of R_t (at quantile τ^*), obtained from (2), i.e., $Ho : V_t = Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$.

Proposition 1 *(VQR test) Consider the quantile regression (2). Under the null hypothesis (6), if assumptions (1)-(4) hold, then, the test statistic ζ_{VQR} is asymptotically chi-squared distributed with two degrees of freedom.*

Proof. See Appendix. ■

Remark 1: The Wald statistic is often adopted in joint tests for quantile regressions, such as Machado and Mata (2004), and Proposition 1 is a special case of a general linear hypothesis test, which was properly adapted to our setup. According to Schulze (2004), a more general Wald statistic is given by $W = T(R\widehat{\varphi} -$

the asymptotics from the non-differentiable objective function, revalidating the conventional techniques under nonstandard conditions. This idea was pioneering illustrated by Huber (1967) in discussion of deriving the asymptotics of maximum likelihood estimators with iid random variables under nonstandard conditions. The main idea is to make the discontinuous first order conditions asymptotically and uniformly continuous by stochastic equicontinuity argument, i.e., by approximating it through a uniformly continuous function. After justifying stochastic equicontinuity, all conventional techniques for deriving asymptotics are again applicable. See Chen (2001) for further details.

$r)'(R\widehat{\Omega}R')^{-1}(R\widehat{\varphi}-r)$, where $\widehat{\varphi}=[\widehat{\alpha}(\tau_1);...;\widehat{\alpha}(\tau_m)]'$ and $\widehat{\Omega}$ is the estimated asymptotic joint matrix of the estimated coefficients considering a full range of quantiles $\tau \in [\tau_1;...;\tau_m]$.[7] Note that this formulation includes a wide variety of testing situations. However, since we are only focused on testing the estimated coefficients $\widehat{\alpha}(\tau)$ for the specific quantile $\tau = \tau^*$, we adopted the simplified version of the Wald statistic presented in Proposition 1.

Remark 2: Assumption (1) together with comonotonicity of $\alpha_i(U_t)$, $i = 0, 1$ guarantee the monotonic property of the conditional quantiles. We recall the comment of Robinson (2006), in which the author argues that comonotonicity may not be sufficient to ensure monotonic conditional quantiles, in cases where x_t can assume negative values. Assumption (2) relaxes iid in the sense that allows for non-identical distributions. Bounding the quantile function estimator away from 0 and ∞ is necessary to avoid technical complications. Assumptions (2)-(4) are quite standard in the quantile regression literature (e.g., Koenker and Machado (1999) and Koenker and Xiao (2002)) and familiar throughout the literature on M-estimators for regression models, and are crucial to claim the CLT of Koenker (2005, Theorem 4.1).

Remark 3: Under the null hypothesis it follows that $V_t = Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$, but under the alternative hypothesis the randomness nature of V_t, captured in our model by the estimated coefficients $\widehat{\theta}(\tau^*) \neq 0$, can be represented by $V_t = Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) + \eta_t$, where η_t represents the measurement error of the VaR on estimating the latent variable $Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$. Note that assumptions (1)-(4) are easily satisfied under the null and the alternative hypotheses. In particular, note that assumption (4) under H_1 implies that also η_t is bounded.

Remark 4: According to Giacomini and Komunjer (2005), when several forecasts of the same variable are available, it is desirable to have formal testing procedures, which do not necessarily require knowledge of the underlying model, or, if the model is known, which do no restrict attention to a specific estimation procedure. Note that assumptions (1)-(4) do not restrict our methodology to those cases in which V_t is constructed from a conditional volatility model, but instead allow for several cases, such as a Pareto or a Cauchy distribution of returns (in which the mean and variance do not even exist), frequently used in the EVT literature (see McNeil & Frey (2000) and Huisman et al. (2001)). In fact, our proposed methodology can be applied to a broad number of situations, such as:

(i) The model used to construct V_t is known. For instance, a risk manager trying to construct a reliable VaR measure. In such a case, it is possible that: (ia) V_t is generated from a conditional volatility model, e.g., $V_t = g(\widehat{\sigma}_t^2)$, where $g()$ is some function of the estimated conditional variance $\widehat{\sigma}_t^2$, say from a GARCH model; or (ib) V_t is directly generated, for instance, from a CAViaR model[8] or an ARCH-quantile method;[9]

(ii) V_t is generated from an unknown model, and the only information available is $\{R_t; V_t\}$. In this case, we are still able to apply Proposition 1 as long as assumptions (1)-(4) hold. This might be the case described

[7] See Koenker & Basset (1978, 1982 a, b); Koenker & Portnoy (1999) and Koenker (2005, p. 76) for further details.

[8] See section 2.2 for more details regarding the CAViaR model.

[9] A quantile regression model that allows for ARCH effect. See Koenker & Zhao (1996) and Wu & Xiao (2002) for further details.

in Berkowitz and O'Brien (2002), in which a regulator investigates the VaR measure reported by a supervised financial institution (see appendix B for further details);

(iii) V_t is generated from an unknown model, but besides $\{R_t; V_t\}$ a confidence interval of V_t is also reported. Suppose that a sequence $\{R_t; V_t; \underline{V}_t; \overline{V}_t\}$ is known, in which $\Pr\left[\underline{V}_t < V_t < \overline{V}_t \mid \mathcal{F}_{t-1}\right] = \delta$, where $[\underline{V}_t; \overline{V}_t]$ are respectively lower and upper bounds of V_t, generated (for instance) from a bootstrap procedure, with a confidence level δ. One could use this additional information to investigate the considered VaR by making a connection between the confidence interval of V_t and the previously mentioned measurement error η_t. The details of this route remain an issue to be further explored.

In next section, we provide an additional framework that might be useful for those interested in improving the performance of a rejected VaR as well as choosing the best model among competing measures.

2.1 Periods of risk exposure

The conditional coverage literature (e.g., Christoffersen (1998)) is concerned with the adequacy of the VaR model, in respect to the existence of clustered violations. In this section, we will take a different route to analyze the conditional behavior of a VaR measure. According to Engle and Manganelli (2004), a good Value-at-Risk model should produce a sequence of unbiased and uncorrelated hits H_t, and any noise introduced into the Value-at-Risk measure would change the conditional probability of a hit, vis-à-vis the related VaR. Given that our study is entirely based on a quantile framework, besides the VQR test, we are also able to identify the exact periods in which the VaR produces an increased risk exposure in respect to its nominal level τ^*, which is quite a novelty in the literature. To do so, let us first introduce some notation:

Definition 2: $W_t \equiv \{\widetilde{\tau} \in [0;1] \mid V_t = \widehat{Q}_{R_t}(\widetilde{\tau} \mid \mathcal{F}_{t-1})\}$, representing the "fitted quantile" of the VaR measure at period t given the regression model (2).

In other words, W_t is obtained by comparing V_t with a full range of estimated conditional quantiles evaluated at $\tau \in [0;1]$. Note that W_t enables us to conduct a local analysis, whereas the proposed VQR test is designed for a global evaluation based on the whole sample. It is worth mentioning that, based on our assumptions, $Q_{R_t}(\tau \mid \mathcal{F}_{t-1})$ is monotone increasing in τ, and W_t by definition is equivalent to a quantile level, i.e., $W_t > \tau^* \Leftrightarrow Q_{R_t}(W_t \mid \mathcal{F}_{t-1}) > Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$. Also note that W_t should (ideally) be as close as possible to τ^* for all t given that the VaR is computed for the fixed level τ^*. However, due to modeling procedures (i.e., in practice), it might be different from τ^*, suggesting that V_t could not belong to the proper conditional quantile of interest.

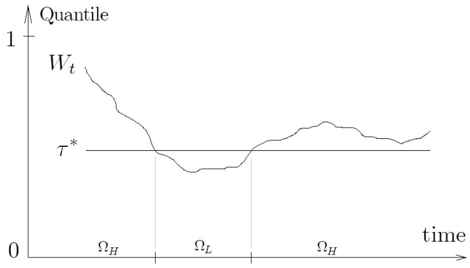
Now consider the set of all observations $\Omega = 1, ..., T$, in which T is the sample size, and define the following partitions of Ω:

Definition 3: $\Omega_H \equiv \{t \in \Omega \mid W_t \geqslant \tau^*\}$, representing the periods in which the VaR belongs to a quantile above the level of interest τ^* (indicating a conservative model);

Definition 4: $\Omega_L \equiv \{t \in \Omega \mid W_t < \tau^*\}$, representing the periods in which the VaR is below the nominal τ^* level and, thus, underestimate the risk in comparison to τ^*.

Since we partitioned the set of periods into two categories, i.e. $\Omega = \Omega_H + \Omega_L$, we can now properly identify the so-called periods of "risk exposure" Ω_L. Let us summarize the previous concepts through the following schematic graph:

Figure 1 - Periods of risk exposure



It should be mentioned that a VaR model that exhibits a good performance in the VQR test (i.e., in which Ho is not rejected) is expected to exhibit W_t as close as possible to τ^*, fluctuating around τ^*, in which periods of W_t below τ^* are balanced by periods above this threshold. On the other hand, a VaR model rejected by the VQR test should present a W_t series detached from τ^*, revealing the periods in which the model is conservative or underestimate risk. This additional information can be extremely useful to improve the performance of the underlying Value-at-Risk model, since the periods of risk exposure are now easily revealed.

Another important issue regarding model analysis is the choice of competing VaRs. Instead of only checking the performance of a single model, one might be interested in ranking several VaR measures (see Giacomini and Komunjer, 2005). Although this is not the main objective of this paper, we outline a simple nonparametric procedure, inspired by Lopez (1999), in which a loss function is used to measure the "conditional coverage distance" of a VaR from its nominal benchmark τ^*. According to the author, a numerical score could reflect regulatory concerns and provide a measure of relative performance to compare competing VaR models across time and institutions.

The generic loss function suggested by Lopez (1999) is given by $C(R_t; V_t) = \sum_{t=1}^{T} C_t(R_t; V_t)$, where $C_t(.) = \begin{cases} f(R_t; V_t) & ; \text{ if } R_t > V_t \\ g(R_t; V_t) & ; \text{ if } R_t \leq V_t \end{cases}$. Accurate VaR estimates are expected to generate lower numerical scores. Once the f and g functions are defined, the loss function can be constructed and used to evaluate the performance of a set of VaR models. Among several different specifications, Lopez (1999) suggests adopting $f(R_t; V_t) = 1 + (R_t - V_t)^2$ and $g(R_t; V_t) = 0$. An interesting advantage of this specification is to consider the magnitude of violations, since the magnitude as well as the number of violations is a serious matter of concern to regulators and risk managers. In addition, loss functions may be more suited to discriminate among competing VaR models than deciding for the accuracy of a single VaR model.

In this paper, we adapt the previous approach to our setup in order to rank competing VaRs. Let's first define C_t as the "empirical distance" of V_t in respect to the conditional quantile function $Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$, based on the W_t series:

$$C_t \equiv |W_t - \tau^*| \,. \tag{8}$$

Now, define the loss function $L(V_t)$ that summarizes the distances C_t, and assigns weights $[\gamma_1; \gamma_2]$ to each distance, according to the indicator function $I_t = \begin{cases} 1 \ ; \ \text{if } W_t > \tau^* \\ 0 \ ; \ \text{if } W_t \leq \tau^* \end{cases}$:

$$L(V_t) \equiv \frac{1}{T} \sum_{t=1}^{T} C_t * (\gamma_1 I_t + \gamma_2 (1 - I_t)). \tag{9}$$

This loss function is very convenient, since it is very easy to be computed and non-parametrically provides us an empirical way to rank a set of VaRs. Thus, the choice among $i = 1, ..., n$ competing models could be based on the minimization of the proposed loss function, by simply choosing $i = \underset{i=1,...,n}{\arg\min}[L(V_t^i)]$. In addition, note that by setting $\gamma_1 < \gamma_2$ one could penalize more the periods of risk exposure than those periods in which the "fitted quantile" W_t is above τ^*. Therefore, an asymmetric evaluation is allowed by this framework, in which the choice of weights $[\gamma_1; \gamma_2]$ could also be driven by the risk aversion degree of the regulator (or the risk manager).[10] Despite its simplicity, the descriptive statistic $L(.)$ might be useful to illustrate model comparison in our empirical exercise. It is worth mentioning that the backtest literature is mainly focused just on the signal I_t (as detailed in next section), whereas in this paper we try to go a step further by also considering the valuable information contained in the magnitude of the VaR violations.

2.2 Other Backtests

Generally speaking, backtesting a VaR model means checking whether the realized daily returns are consistent with the corresponding VaR produced by an internal model of a financial institution, over an extended period of time. Crouhy et al. (2001) argue that backtests provide a key check of how accurate and robust models are, by considering ex-ante risk measure forecasts and comparing it to ex-post realized returns. The authors also state: *"Backtesting is a powerful process with which to validate the predictive power of a VaR model ... and in effect, a self-assessment mechanism that offers a framework for continuously improving and refining risk modeling techniques"*.

According to Hull (2005), it represents a way to test how well VaR estimates would have performed in the past, i.e., how often was the actual 1-day (or 10-day) loss greater than the 95% (or 99%) VaR measure. Before presenting some commonly discussed backtests, let's initially recall that R_t is the observed returns and that V_t is the respective 1-day VaR defined for a quantile level τ^*. Now define a violation sequence[11] by the following indicator function:

$$H_t = \begin{cases} 1 \ ; \ \text{if } R_t > V_t \\ 0 \ ; \ \text{if } R_t \leq V_t \end{cases}, \tag{10}$$

[10] A more general setup could consider the weights as functions of the distance C_t, i.e., $\gamma_i = \gamma_i(C_t)$.

[11] Also called in the literature by "exception" or "hit sequence".

and compute the number of violations $N = \sum_{t=1}^{T} H_t$. Based on these definitions, we now present some backtests usually mentioned in the literature to identify misspecified VaR models (see Dowd (2005) and Jorion (2007) for a detailed description):

(i) Kupiec (1995): Some of the earliest proposed VaR backtests is due to Kupiec (1995), which proposes a nonparametric test based on the proportion of exceptions. Assume a sample size of T observations and a number of violations of N. The objective of the test is to know whether or not $\widehat{p} \equiv N/T$ is statistically equal to τ^* (the VaR confidence level).

$$Ho : p = E(H_t) = \tau^*. \tag{11}$$

The probability of observing N violations over a sample size of T is driven by a Binomial distribution and the null hypothesis Ho: $p = \tau^*$ can be verified through a LR test of the form (also known as the unconditional coverage test):

$$LR_{uc} = 2 \ln \left(\frac{\widehat{p}^N (1 - \widehat{p})^{T-N}}{\tau^{*N} (1 - \tau^*)^{T-N}} \right), \tag{12}$$

which follows (under the null) a chi-squared distribution with one degree of freedom. It also should be mentioned that this test is uniformly most powerful (UMP) test for a given T. However, Kupiec (1995) finds that the power of his test is generally poor in finite samples, and the test becomes more powerful only when the number of observations is very large.[12]

(ii) Christoffersen (1998): The unconditional coverage property does not give any information about the temporal dependence of violations, and the Kupiec (1995) test ignores conditioning coverage, since violations could cluster over time, which should also invalidate a VaR model. In this sense, Christoffersen (1998) extends the previous LR statistic to specify that the hit sequence should also be independent over time. The author argues that we should not be able to predict whether the VaR will be violated, since if we could predict it, then, that information could be used to construct a better risk model. The proposed test statistic is based on the mentioned hit sequence H_t, and on T_{ij} that is defined as the number of days in which a state j occurred in one day, while it was at state i the previous day. The test statistic also depends on π_i, which is defined as the probability of observing a violation, conditional on state i the previous day. It is also assumed that the hit sequence follows a first order Markov sequence with transition matrix given by

$$\Pi = \begin{array}{c} \text{Previous day} \\ \left[\begin{array}{cc} 1 - \pi_0 & 1 - \pi_1 \\ \pi_0 & \pi_1 \end{array} \right] \end{array} \begin{array}{l} \\ \text{current day (violation)} \\ \text{no violation} \end{array} \tag{13}$$

Note that under the null hypothesis of independence, we have that $\pi = \pi_0 = \pi_1 = (T_{01} + T_{11})/T$, and the following LR statistic can, thus, be constructed:

$$LR_{ind.} = 2 \ln \left(\frac{(1 - \pi_0)^{T_{00}} \pi_0^{T_{01}} (1 - \pi_1)^{T_{10}} \pi_1^{T_{11}}}{(1 - \pi)^{(T_{00} + T_{10})} \pi^{(T_{01} + T_{11})}} \right). \tag{14}$$

[12] According to Kupiec (1995), it would require more than six violations during a one-year period (250 trading days) to conclude that the model is misspecified. See Crouhy et al. (2001) for further details.

The joint test, also known as "conditional coverage test", includes unconditional coverage and independence properties, and is simply given by $LR_{cc} = LR_{uc} + LR_{ind.}$; where each component follows a chi-squared distribution with one degree of freedom, and the joint statistic LR_{cc} is asymptotically distributed as $\chi^2_{(2)}$. An interesting feature of this test is that a rejection of the conditional coverage may suggest the need for improvements on the VaR model, in order to eliminate the clustering behavior. On the other hand, the proposed test has a restrictive feature, since it only takes into account the autocorrelation of order 1 in the hit sequence.

(iii) Engle and Manganelli (2004): The Conditional Autoregressive Value-at-Risk by Regression Quantiles (CAViaR) model is proposed by the authors, which define $V_t(\tau)$ as the solution to $\Pr[R_t < -V_t(\tau) \mid \mathcal{F}_{t-1}) = \tau$; and describe the (generic) specification: $V_t(\tau) = \beta_0 + \sum_{i=1}^{q} \beta_i V_{t-i}(\tau) + \sum_{j=1}^{r} \gamma_j x_{t-j}$; where $[\beta_i; \gamma_j]$ are unknown parameters to be estimated and x_t is a generic vector of time t observable variables. The CAViaR approach directly models the return quantile rather than specifying a complete data generating process. The authors define various dynamic models for V_t itself, including the adaptative model: $V_t(\tau) = V_{t-1}(\tau) + \beta\,[\mathbf{1}(R_{t-1} \leq -V_{t-1}) - \tau]$; symmetric absolute value: $V_t(\tau) = \beta_0 + \beta_1 V_{t-1}(\tau) + \beta_2\,|R_{t-1}|$; asymmetric slope: $V_t(\tau) = \beta_0 + \beta_1 V_{t-1}(\tau) + \beta_2 R^{+}_{t-1} + \beta_3 R^{-}_{t-1}$; in which R_t is the return series. The estimation of the CAViaR model uses standard quantile regression techniques, but the model is only possible in GARCH special cases. How to proper simulate the model is another issue to be further explored.

Besides proposing the CAViaR class of models to directly estimate the conditional quantile process, Engle and Manganelli (2004) also suggested a specification test, also known as the dynamic conditional quantile (DQ) test, which involves running the following regression

$$DQ_{oos} = (Hit_t' X_t [X_t' X_t]^{-1} X_t' Hit_t)/(\tau(1-\tau)), \tag{15}$$

where $X_t = [c, V_t(\tau), Z_t]$; Z_t denotes lagged Hit_t; $Hit_t = I_t(\tau) - \tau$; and $I_t(\tau) = \begin{cases} 1 \;;\; \text{if } R_t < V_t(\tau) \\ 0 \;;\; \text{if } R_t \geq V_t(\tau) \end{cases}$. The null hypothesis is the independence between Hit_t and X_t. Under the null, the proposed metric to evaluate one-step-ahead forecasts (DQ_{out}) follows a χ^2_q, in which $q = rank(X_t)$. Note that the DQ test can be used to evaluate the performance of any type of VaR methodology (and not only the CAViaR family).

(iv) Berkowitz et al. (2006): These authors recently proposed a unified approach to a VaR assessment, based on the fact that the unconditional coverage and independence hypotheses are nothing but consequences of the martingale difference hypothesis of the Hit_t process, i.e., $E(I_t(\tau) - \tau \mid \mathcal{F}_{t-1}) = 0$, where $I(\tau)$ is defined as above. Based on a Ljung-Box type-test, they consider the nullity of the first K autocorrelations of the Hit_t process, instead of only considering the autocorrelation of order one, as done by Christoffersen (1998).

Several other related procedures are also well documented in the literature, such as the nonparametric test of Crnkovic and Drachman (1997),[13] the duration approach of Christoffersen and Pelletier (2004),[14] and the encompassing test of Giacomini and Komunjer (2005).[15] However, as shown in several simulation exercises (e.g., Kupiec (1995), Pritsker (2001) and Campbell (2005)), backtests generally have low power and are, thus, prone to misclassifying inaccurate VaR estimates as "acceptably accurate". The standard backtests often lack power, especially when the VaR confidence level is high and the number of observations is low, which has lead to a search for improved tests. Since the original Kupiec (1995) test is the most powerful among its class, more effective backtests would have to focus on different hypothesis or use more information, according to Jorion (2007). In next section, we exactly address this topic by comparing our proposed quantile regression-based backtest with some previously mentioned procedures.

2.3 Nested null hypotheses

In this section, we construct a parallel of our setup with some backtests. Since the main concern of the backtest literature is to evaluate the VaR accuracy, we pose a relevant question in this context: What do we really want to test? Given that a VaR measure is implicitly defined by Property 1 (hereafter, P1, reproduced below), the core issue of a backtest should be to verify whether it (in fact) is true. As we next show, the quantile regression framework provides a natural way to investigate the performance of a VaR model, and the proposed VQR test consists on a sufficient condition for P1. We also show that our considered null hypothesis implies some null hypotheses used in the literature to construct backtests, which are only necessary conditions for P1. To do so, firstly recall that a "violation" sequence is here defined by the following indicator (hit) function: $H_t = \begin{cases} 1 \;\; ; \text{ if } R_t > V_t \\ 0 \;\; ; \text{ if } R_t \leq V_t \end{cases}$, and secondly consider $\tau^{**} = (1 - \tau^*)$, in order to properly compare our VQR test, originally constructed for the right tail of the returns distribution, with the other considered backtests based on the left tail.

Property 1: $\Pr\left[R_t \leq V_t | \mathcal{F}_{t-1}\right] = \tau^*$;

Statement 1: (Null hypothesis of the VQR test) $V_t = Q_{R_t}\left(\tau^* \mid \mathcal{F}_{t-1}\right)$;

Statement 2: Berkowitz et al. (2006): $E(H_t - \tau^{**} \mid \mathcal{F}_{t-1}) = 0$;

Statement 3: Christoffersen (1998): $E[(H_t - \tau^{**})(H_{t-1} - \tau^{**})] = 0$;

Statement 4: Kupiec (1995): $E(H_t) = \tau^{**}$.

[13] The authors use a Kuiper's statistic, based not only on a selected quantile but focused on the entire forecasted distribution.

[14] The key idea is that if the VaR model is correctly specified for a given coverage rate p, then, the conditional expected duration between violations should be a constant $1/p$ days. The independence test based on duration allows one to consider wider dependences than those chosen under the Markov chain hypothesis. However, the core idea remains unchanged, and consists in putting the conditional coverage hypothesis to the test, still ignoring the magnitude of violations.

[15] A conditional quantile encompassing test is provided based on GMM estimation, with the focus on relative model evaluation, which involves comparing the performance of competing VaR models, and choosing the one that performs the best. The encompassing approach also gives a theoretical basis for quantile forecast combination, in cases when neither forecast encompasses its competitor. As a by-product, their framework also provides a link to the conditional coverage test of Christoffersen (1998).

Proposition 2 *(Nested Hypotheses) Consider Property 1 and statements S1-S4. If assumptions (1)–(4) hold for the regression (2), then, it follows that:*

(i) $P1 \Leftrightarrow S1$

(ii) $S1 \Rightarrow S2, S3, S4$

(iii) $S2, S3, S4 \nRightarrow S1$

Proof. See Appendix. ∎

Our proposed test aims to break down the paradigm of the hit sequence in the backtest literature, which investigates the accuracy of a VaR measure basically through the behavior of its hit sequence. Proposition 2 states an important result that the VQR test is a necessary and sufficient condition to verify Property 1. In addition, it also shows that our null hypothesis implies some null hypothesis of the backtest literature, but the reverse does not hold. In other words, assumption S1 is a sufficient condition for Property 1, whereas statements S2-S4 are only necessary conditions.

A small Monte Carlo simulation is conducted in next section, in order to verify the size and power of the VQR test in finite samples. Overall, the quantile regression test seems to have relatively more power in comparison to other backtests, previously documented in the literature as exhibiting low power to detect poor VaR models. These results are consistent with previous findings in Kupiec (1995), Pritsker (2001) and Campbell (2005). The increased power might be due to the quantile framework, which can provide an adequate null hypothesis as stated in Proposition 2.

3 Monte Carlo simulation

A small simulation experiment is conducted in order to investigate the finite sample properties of the VQR test, in comparison to other tests presented in the literature, such as the unconditional coverage test of Kupiec (1995), the conditional coverage test of Christoffersen (1998), and the out-of-sample DQ test of Engle and Manganelli (2004). To do so, we use two Data-Generating Processes: DGP1 is the RiskMetrics[16], and DGP2 is the GARCH(1,1) model $\sigma_t^2 = 0.02 + 0.05y_{t-1}^2 + 0.93\sigma_{t-1}^2$. In addition, we assume (in both DGPs) that $y_t = \sigma_t \varepsilon_t$, in which $\varepsilon_t \sim N(0,1)$. For each DGP, we generate $T + 2,000$ observations, discarding the first $2,000$ observations. Then, a total amount of $i = 5,000$ replications of the $\{y_t\}_{t=1}^T$ process is considered for each DGP.

We follow here the same computational strategy of Lima & Neri (2006), in which a hybrid solution using R and Ox environments is adopted, since the proposed simulation is extremely computational intensive. Ox is much faster than R in large computations, and also makes use of the package G@RCH 4.2 (see Laurent & Peters, 2006), which easily allows us to generate GARCH specifications. On the other hand, the R language is more interactive and user-friendly than Ox and the VQR test must in fact be conducted in R, since its

[16] Recall that RiskMetrics is just an integrated GARCH(1,1) model with the autoregressive parameter set to 0.94, i.e., $\sigma_t^2 = c + 0.06y_{t-1}^2 + 0.94\sigma_{t-1}^2$. In our simulation, we set $c = 0.02$.

package for quantile regressions (quantreg) is more complete and updated than the Ox package. Therefore, we proceed as follows: an Ox code initially generates the time series y_t for each DGP, and save all the replications in the hard disk. Next, an R code computes the four considered backtests for all replications and saves the final results in a text file.

For the size investigation, in order to generate data that supports the null hypothesis, we compute the respective VaR at the (standard normal) quantile τ^*. In other words, the VaR for $\tau^* = 95\%$ is given by $V_t = 1.64 * \sigma_t$, and for $\tau^* = 99\%$ is computed by $V_t = 2.33 * \sigma_t$. The empirical sizes for $T = \{250; 500; 1,000\}$ and the quantile levels $\tau^* = 95\%$ or 99% are presented in next table (for a nominal size of 5%):

Table 1 - Size investigation (T=250)

	$\tau^* = 95\%$		$\tau^* = 99\%$	
	DGP1	DGP2	DGP1	DGP2
ζ_{VQR}	0.0705	0.0591	0.1801	0.1851
ζ_{Kupiec}	0.0101	0.0069	0.0079	0.0007
$\zeta_{\text{Christ.}}$	0.1073	0.1053	0.0215	0.0175
ζ_{DQ}	0.0739	0.0429	0.0806	0.0931
T=500				
ζ_{VQR}	0.0632	0.0545	0.1114	0.1299
ζ_{Kupiec}	0.0088	0.0084	0.0242	0.0198
$\zeta_{\text{Christ.}}$	0.0960	0.1089	0.0325	0.0267
ζ_{DQ}	0.0592	0.0577	0.0762	0.0781
T=1,000				
ζ_{VQR}	0.0541	0.0513	0.0950	0.0991
ζ_{Kupiec}	0.0247	0.0192	0.0368	0.0254
$\zeta_{\text{Christ.}}$	0.0986	0.0920	0.0374	0.0332
ζ_{DQ}	0.0562	0.0517	0.0801	0.0855

Note: The values above represent the percentage of p-values below the nominal level of significance $\alpha = 5\%$.

Firstly, note that for $T = 250$, the VQR and DQ backtests exhibit relatively good sizes for $\tau^* = 95\%$. On the other hand, for $\tau^* = 99\%$ the results are slightly distorted: the DQ and VQR tests tend to over-reject the VaR model, whereas the Kupiec and Christoffersen backtests tend to under-reject it. The main reason is that, for $T = 250$ only a small number of observations is expected at the extreme quantiles, which is a serious problem for all backtests, and might also affect the QR estimation.

The increase of the sample size T can give us some flavor of the asymptotic behavior in the size investigation. Recall that each backtest is constructed to investigate different null hypotheses, which might partially explain the results presented in Table 1. In addition, note that an increase of the sample size produces the following effect in our simulation: As long as T increases, the estimation of the extreme quantiles becomes more precise, leading to a better estimation of the quantile density function evaluated at those quantiles. As a result, the empirical size of the proposed test tends to approach its nominal size (5%) as T goes to infinity.

Despite the relatively large sample size when $T = 1,000$, note that for $\tau^* = 99\%$ one should expect only 10 observations of y_t above the VaR measure, which could seriously influence the performance of any backtest. However, the small sample size should not be viewed anymore as a restriction, given that nowadays it is common to deal with intra-day data, and even for daily frequency, a sample size of $T = 1,000$ only requires four years of database.

Moreover, backtesting involves balancing two types of errors and dealing with the tradeoff between rejecting a correct model versus accepting a misspecified one. According to Christoffersen (2003, p.186), in risk management, may be very costly if the test fail to reject an incorrect model. Jorion (2007), in the same line, says that one would want a framework that has high power of rejecting an incorrect model. Therefore, if one is more concerned with discarding a poor VaR model (and, thus, the power of the tests), instead of validating a good VaR specification, the numerical results might be favorable to the VQR approach, as we shall next see.

In the power analysis, we conduct the investigation along three main directions: the sample size $T = \{250; 500; 1,000\}$, the quantile level $\tau^* = 95\%$ or 99%; and finally the set of alternative hypothesis. The first set of H_1 (here, so-called method 1) considers a sequence of DGPs based on a GARCH(1,1), with coefficients: $c = 0.02$; $\alpha = 0.06 - \phi/20$; $\beta = 0.94 - \phi/2$, and a Gamma (a, b) distribution with parameters $a = 200e^{-5\phi}; b = 5$. We control the "degree of misspecification" through the parameter $\phi \in [0, 1]$, which ranges from 0 to 1 with increments of 0.1. Then, in order to replicate a realistic situation, a VaR is estimated for each DGP via a RiskMetrics model with normal distribution. Note that when $\phi = 0$ we are under the null hypothesis,[17] but as long as we increase ϕ the alternative hypothesis is simulated.

The second approach for H_1 (method 2) is constructed as a complementary exercise, in which we now fix the DGP and then generate a sequence of VaRs. The idea is based on Engle and Manganelli (2004), which argue that: *"any noise introduced into the quantile estimate will change the conditional probability of a hit given the estimate itself"*. To do so, we initially generate y_t and σ_t^2 according to DGP2. Then, we construct a sequence of VaRs in the following way: $V_t(\phi) \equiv Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) + \phi\eta_t$, where $Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$ comes from the DGP2; $\eta_t \sim iid\ N(0, 1)$; $\phi \in [0, 1]$ ranges from 0 to 1. Note that the "degree of misspecification" is (again) given by ϕ, in which $V_t(\phi = 0)$ satisfies H_0, but as long as the ϕ parameter is augmented we expect to generate quite poor VaR measures due to the additional white noise η_t.

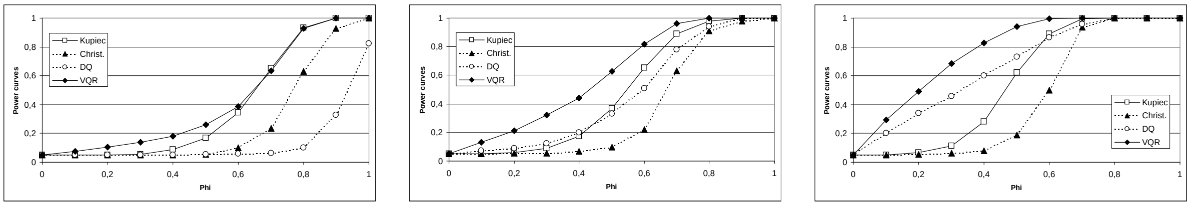
A final simulation for the power analysis is given by method 3, in which the DGP2 is used to generate y_t and σ_t^2, but the sequence of VaRs is now constructed from a normal-GARCH (1,1) specification with the following coefficients: $c = 0.02$; $\alpha = 0.05 + \phi/5$; $\beta = 0.93 - \phi/5$. This way, when $\phi = 0$, we are under the null hypothesis, and the model used for the VaR is compatible with the adopted DGP. However, as long as the

[17] Recall that a Gamma (a, b) distribution tends to a normal distribution as long as $a \to \infty$.

ϕ parameter increases, the constructed VaR measure will come from an increasingly misspecified volatility model. We next present the results for the power investigation, which are already corrected[18] for the size distortions shown in Table 1, i.e., size-adjusted power results (for a nominal size of 5%).

Figure 2a - Size-corrected Power Curves - Method 1 ($\tau^* = 95\%$)

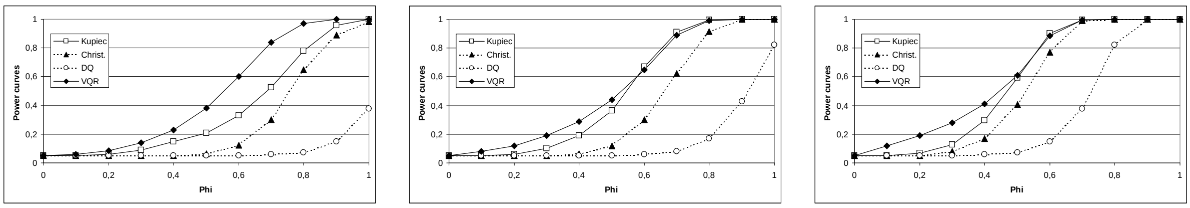
$T = 250$ $T = 500$ $T = 1,000$



Notes: Nominal level of significance is $\alpha = 5\%$. The results for methods 2 and 3 are presented in appendix.

Figure 2b - Size-corrected Power Curves - Method 1 ($\tau^* = 99\%$)

$T = 250$ $T = 500$ $T = 1,000$



Notes: Nominal level of significance is $\alpha = 5\%$. The results for methods 2 and 3 are presented in appendix.

The previous plots reveal meaningful differences among the considered tests. A very nice result is that in the smallest sample size case ($T = 250$, with $\tau^* = 99\%$), our test is indeed the most powerful among the considered backtests. Note that as long as T increases, all curves becomes closer to the origin, increasing the chance of detecting a misspecified model, which is a natural response since greater sample sizes lead to smaller variances. An important remark is that, in the same line of Engle and Manganelli (2004), one could include other (exogenous or lagged) variables in $\mathcal{F}_{t-1}$ and, thus, in the quantile regression (2), in order to still further increase the power of the VQR test in different directions.

The results for methods 2 and 3 are presented in appendix. The DQ curve exhibits the best shape in method 2, whereas the Christoffersen (1998) and Kupiec (1995) tests are relatively better than the VQR test for $\tau^* = 99\%$, but the VQR shows again a superior behavior for $\tau^* = 95\%$. Regarding method 3, the VQR test exhibits a good performance, beating the other backtests in almost all situations.

Previously results in the literature have already suggested that the Kupiec (1995) test might exhibit low power against poor VaR methodologies: Kupiec (1995) itself describes how his test has a limited ability to distinguish among alternative hypotheses and thus has low power in samples of size $T = 250$. See also

[18] "Size-corrected power" is just power using the critical values that would have yielded correct size under the null hypothesis.

Pritsker (2001), Campbell (2005) and Giacomini & Komunjer (2005).[19] In fact, our results reconfirm these earlier findings, and also suggest that the VQR test might be more powerful under some directions of the alternative hypothesis.

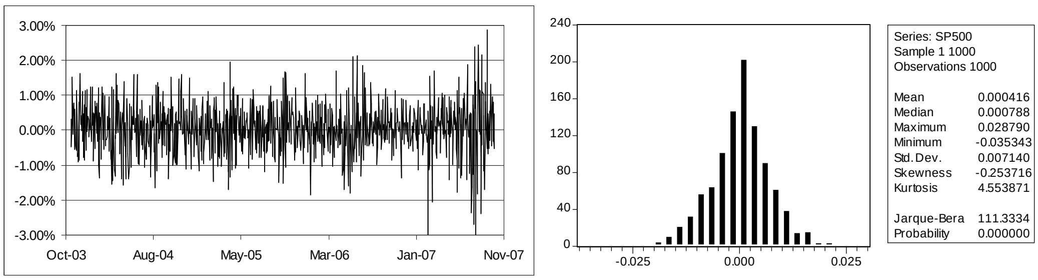
Besides the sample size, another reason to support the simulation results is given by Proposition 2, in which the null hypothesis of the VQR test seems to be a sufficient condition for the validity of Property 1, whereas the Kupiec (1995) test is only a necessary condition. Note that the unconditional coverage test of Kupiec (1995) is a LR test, which is uniformly most powerful for a given sample size. However, the related low power of this test in small samples is due to its inappropriate null hypothesis regarding Property 1. What we really want to test? Recall that an ideal VaR model should be well represented by Property 1. Yet another reason for the reported lack of power is the choice of a high confidence level (99%) that generates too few exceptions for a reliable backtest. Thus, simply changing the VaR quantile level from 99% to 95% sharply reduces the probability of accepting a misspecified model.[20]

4 Empirical exercise

4.1 Data

In this section, we explore the empirical relevance of the theoretical results previously derived. This is done by evaluating and comparing five different VaR models, based on the VQR test and other competing procedures commonly presented in the backtest literature. To do so, we investigate the daily returns of S&P500 over the last 4 years, with an amount of $T = 1,000$ observations, depicted in the following figure:[21]

Figure 3 - S&P500 daily returns (%)



Notes: a) The sample covers the period from 23/10/2003 until 12/10/2007;

b) Source: Yahoo!Finance.

Note from the graph and the summary statistics the presence of common stylized facts about financial data (e.g., volatility clustering; mean reverting; skewed distribution; kurtosis > 3; and non-normality. See Engle and Patton (2001) for further details). In addition, an analysis of the correlogram of the returns (not

[19] According to the authors, the unconditional coverage test of Kupiec (1995) assumes away parameter estimation uncertainty and, as we already discussed, only investigates the hit sequence instead of the magnitude of the violations.

[20] This could explain why some banks prefer to choose $\tau^* = 0.95$, in order to be able to observe sufficient number of observations to validate the internal model. See Jorion (2007, p. 147) for further details.

[21] We take the log-difference of the value of the S&P500 index in order to convert the data into returns.

reported) indicates only weak dependence in the mean. In this sense, a detailed analysis over a full range of quantiles could still be conducted based on the "quantilograms" of Linton and Whang (2007), which propose a diagnostic tool for directional predictability, by measuring nonlinear dependence based on the correlogram of the quantile hits. The authors provide a method to compute the correlogram of the quantile hits, so-called the "quantilogram", and to display this along with pointwise confidence bands, resulting in additional information in respect to the standard correlogram.

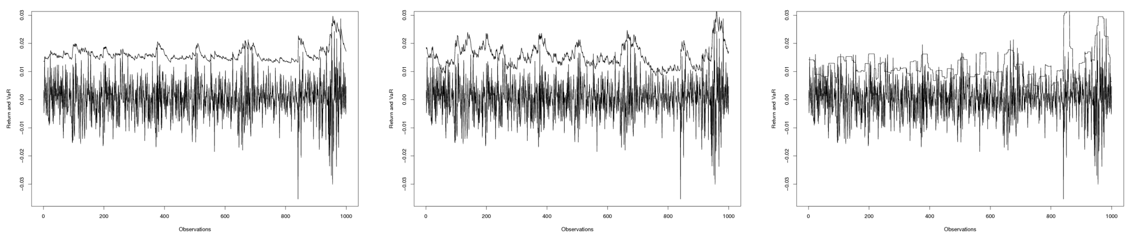
It is worth mentioning that Linton and Whang (2007) apply their methods to S&P500 stock index return data, from 1955 to 2002, and the empirical results suggest some directional predictability in daily returns, especially at the extreme lower quantiles. In addition, there is not much individual evidence of predictability in the median, which is similar to evidence at the mean using standard correlogram. In other words, extreme losses in one period are likely to be succeeded by large losses in the next period. This way, a good VaR measure should be able to capture this kind of dynamics. Note that our proposed framework is related to the quantilogram approach, in the sense that we also make use of additional information by investigating the magnitude of hits, and not only the hit sequence, but here we solely focus on model evaluation.

The five Value-at-Risk models adopted in our evaluation procedure are the following: Rolling Window (1 and 3 months), GARCH (1,1), RiskMetrics (hereafter, RM) and CAViaR. In the first two approaches, the last 30 (and 90) days of data are used to calculate the conditional variance (σ_t^2), based on a moving average of past observations. The third and fourth approaches are nothing else than conditional volatility models based on a GARCH (1,1) model,[22] since RiskMetrics is just an integrated GARCH(1,1) model with the autoregressive parameter set to 0.94. The respective VaR measures of these first four volatility models are, then, constructed by a linear function of σ_t (assuming normality). For instance, the Value-at-Risk for $\tau^* = 99\%$ is given by $V_t = 2.33 * \sigma_t$. Regarding the CAViaR model, we considered the asymmetric slope model: $V_t(\tau) = \beta_0 + \beta_1 V_{t-1}(\tau) + \beta_2 R_{t-1}^{+} + \beta_3 R_{t-1}^{-}$; in which R_t is the return series.

Practice generally shows that these various models lead to widely different VaR levels for the same considered return series, leading us to the crucial issue of model comparison and hypothesis testing. The Rolling Window method (also called Historical Simulation, hereafter, HS) has serious drawbacks and is expected to generate poor VaR measures, since it ignores the dynamic ordering of observations, and volatility measures look like "plateaus", due to the so-called "ghost effect". On the other hand, as shown by Christoffersen et al. (2001), apud Giacomini and Komunjer (2005), the GARCH-VaR model is the only VaR measure, among several alternatives considered by the authors, which passes the Christoffersen's (1998) conditional coverage test. The JP Morgan's RiskMetrics-VaR model is chosen as a benchmark model commonly used by practitioners. Finally, Engle and Manganelli (2004) show that the "asymmetric absolute value" and "asymmetric slope" models are the best CAViaR specifications for the S&P500 data.

[22] The following GARCH (1,1) model was estimated through EViews: $\sigma_t^2 = 2.44E - 06 + 0.049535 y_{t-1}^2 + 0.901294 \sigma_{t-1}^2$.

Figure 4 - S&P500 daily returns (R_t) and VaR (99%) V_t
GARCH(1,1), CAViaR and Rolling Window (1 month)



4.2 Results

Based on the quantile regression framework, we are now able to construct the VQR test for the five considered VaRs. The main results are summarized in the following table:

Table 2 - Results of the VQR test ($\tau^* = 99\%$)
$Ho : V_t = Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$

	CAViaR	GARCH	RM	HS1m	HS3m
$\widehat{\alpha}_0(\tau^*)$	0.00205 (0.00232)	-0.00594 (0.00976)	-0.00267 (0.00202)	0.00677 (0.00252)	0.00298 (0.00255)
$\widehat{\alpha}_1(\tau^*)$	0.83323 (0.19955)	1.39269 (0.63097)	1.16941 (0.08170)	0.80103 (0.22783)	0.91397 (0.19632)
ζ_{VQR}	0.81351	0.39766	15.23240	31.94366	11.87233
p-value	0.66581	0.81968	0.00049	1.15e-07	0.00264

Note: a) Standard error in parentheses.

As already expected, the rolling window models are all rejected, whereas the GARCH(1,1) and CAViaR models do not fail at the VQR test, which is a result perfectly in line with the literature (e.g., Christoffersen et al. (2001) and Giacomini and Komunjer (2005)). In addition, the RiskMetrics-VaR is rejected for $\tau^* = 99\%$. It should be mentioned that violations that are clustered in time are more likely to occur in a VaR model obtained from a rolling window procedure, which increases the number of scenarios for our backtest evaluation. We now present the results of other backtests often used in the literature for VaR evaluation:

Table 3 - Backtests comparison ($\tau^* = 99\%$)

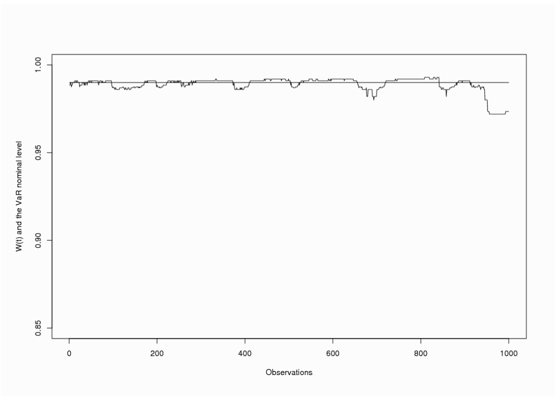
	CAViaR	GARCH	RM	HS1m	HS3m
% of hits	0.9	1.2	1.1	5.6	2.5
ζ_{Kupiec}	0.74884	0.53556	0.75198	0.00000 (**)	0.00000 (**)
$\zeta_{\text{Christ.}}$	0.87539	0.71333	0.84163	0.00000 (**)	0.00017 (**)
ζ_{DQ}	0.97173	0.94656	0.13848	0.00000 (**)	0.00000 (**)
ζ_{VQR}	0.66581	0.81968	0.00049 (**)	0.00000 (**)	0.00264 (**)

Notes: P-values are shown in the ζ's rows; (**) means rejection at 1%.

Note that the GARCH(1,1)-VaR model provides a quite good VaR measure, according to all considered backtests, despite its simplicity and the assumption of normality. Overall, the results are similar to those obtained from the VQR test, excepting the RiskMetrics model. The results of Table 3 indicate that RiskMetrics is only rejected by the VQR test, which is compatible[23] with the previous results of the Monte Carlo simulation (see Figure 2b, $T = 1,000$). In other words, our methodology is able to reject more VaR models in comparison to other backtests, which might be a major advantage of our approach. In fact, recall that the VQR test has more power in some directions of the alternative hypothesis, as described in the power investigation of section 3. The main reason could be that the other backtests are all based on a hit sequence, ignoring the respective magnitude of violations, which is properly considered in the quantile regression setup.

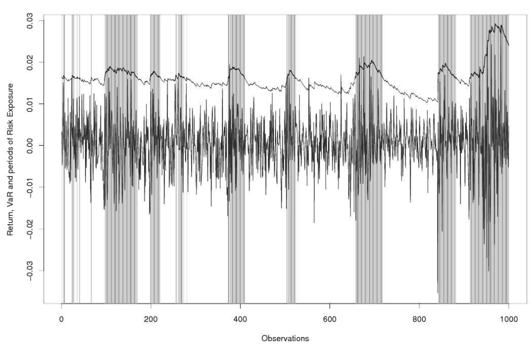
As a result of our proposed methodology, we are also able to construct the W_t series, described in section 2.1, in order to reveal the periods of risk exposure. Recall that whenever W_t is below the benchmark level τ^*, the VaR model increases the risk exposure by underestimating the related conditional quantile of returns, since (ideally) W_t should be as close as possible to τ^*. To illustrate the methodology, the estimated W_t series as well as the periods of risk exposure for the RiskMetrics-VaR(99%) model are depicted in Figures 5 and 6, where the gray bars indicate periods in which $W_t < \tau^*$.

Figure 5 - W_t (RiskMetrics-VaR 99%)



Notes: a) The black series is the computed W_t; $W_t \equiv \{\widetilde{\tau} \in [0;1] \mid V_t = \widehat{Q}_{R_t}(\widetilde{\tau} \mid \mathcal{F}_{t-1})\}$

Figure 6 - R_t and V_t (RiskMetrics-VaR 99%)



Note: Gray bars indicate $W_t < \tau^*$;

In other words, gray bars suggest periods in which the VaR measure underestimates the risk exposure. Since the RiskMetrics-VaR(99%) model is rejected by the VQR test, the risk exposure periods could be very useful for risk managers interested in improving the accuracy of the underlying model. For instance, a visual inspection on figure 6 indicates that the RiskMetrics model usually underestimates (gray bars) the degree of risk for high volatility periods. Therefore, we are able to unmask the bad performance of the RiskMetrics model in our empirical exercise based on a local behavior analysis, which brings some additional (and important) information to the backtest investigation by exposing some "reasons of rejection". Note

[23] Also note that the rank of the DQ test in the power curves of Figure 2b ($T = 1,000$) is not exactly the same as the rank of p-values in Table 3 (RM column). One possible explanation is that the power curves are size-adjusted, and the DQ test (see Table 1, $\tau^* = 99\%$, $T = 1,000$) is oversized whereas the Kupiec and Christoffersen backtests are undersized.

that this local behavior investigation could only be conducted through our proposed quantile regression methodology, which we believe to be a novelty in the backtest literature.

Other relevant issue regarding VaR evaluation is the comparison among several competing models. Although it is not the main objective of this paper, we outline (for the sake of completion of our empirical exercise) a simple nonparametric decision rule for model selection and apply it to our empirical exercise (see Giacomini and Komunjer (2005) for a detailed discussion of model comparison). We are, thus, concerned with relative evaluation, which involves comparing the performance of competing models and choosing the one that performs the best according to our suggested criterion of section 2.1. The main results are next summarized:

Table 4 - Loss Function $L(V_t)$ for $\tau^* = 99\%$

CAViaR	GARCH	RM	HS1m	HS3m
0.00320	0.00497	0.00560	0.06843	0.02099

Notes: a) Recall that $C_t \equiv |W_t - \tau^*|$; $I_t = \begin{cases} 1 \ ; \ \text{if } W_t > \tau^* \\ 0 \ ; \ \text{if } W_t \leq \tau^* \end{cases}$;

and $L(V_t) \equiv \frac{1}{T}\sum_{t=1}^{T} C_t * (\gamma_1 I_t + \gamma_2(1 - I_t))$;

b) We adopted $\gamma_1 = 1.0$ and $\gamma_2 = 1.5$.

Based on this procedure, one should choose the model in which W_t best tracks the desired τ^* level, according to the asymmetric weights γ_1 and γ_2.In our exercise, the CAViaR model exhibits the best performance (i.e., lowest value of $L(V_t)$), which is a natural result, given that it is exactly designed to produce Value-at-Risk measures, whereas the other discussed VaRs are only obtained from conditional volatility models together with the assumption of normality. Therefore, the proposed methodology to identify periods of risk exposure could be used to increase the performance of a poor VaR model, whereas, the suggested $L(V_t)$ distance could be applied to rank and select among competing models.

5 Conclusions

Backtesting could prove very helpful in assessing Value-at-Risk models and is nowadays a key component for both regulators and risk managers. Since the first procedures suggested by Kupiec (1995) and Christoffersen (1998), a lot of research has been done in the search for adequate methodologies to assess and help improve the performance of VaRs, which (preferable) do not require the knowledge of the underlying model.

As noted by the Basle Committee (1996), the magnitude as well as the number of exceptions of a VaR model is a matter of concern. The so-called "conditional coverage" tests indirectly investigate the VaR accuracy, based on a "filtering" of a serially correlated and heteroskedastic time series (V_t) into a serially independent sequence of indicator functions (hit sequence H_t). Thus, the standard procedure in the literature is to verify whether the hit sequence is iid. However, an important piece of information might be lost in that process: not only is the sequence of past hits that matters, but also the magnitude of H_t is of vital importance, since the conditional distribution of returns is dynamically updated. This issue is also discussed by Campbell (2005), which states that the reported quantile provides a quantitative and continuous measure of the magnitude of realized profits and losses, while the hit indicator only signals whether a particular threshold was exceeded. In this sense, the author suggests that quantile tests can provide additional power to detect an inaccurate risk model.

That is exactly the objective of this paper: to provide a VaR-backtest fully based on a quantile regression framework. Our proposed methodology enables us to: (i) formally conduct a Wald-type hypothesis test to evaluate the performance of VaR; and (ii) identify periods of an increased risk exposure. We illustrate the usefulness of our setup through an empirical exercise with daily S&P500 returns, in which we constructed five competing VaR models and evaluate them through our proposed test (and through other three backtests). In addition, we also suggest a simple nonparametric procedure to rank the competing models.

Since a Value-at-Risk model is implicitly defined as a conditional quantile function, the quantile approach provides a natural environment to study and investigate VaRs. One of the advantages of our approach is the increased power of the suggested quantile regression-backtest in comparison to some established backtests in the literature, as suggested by a small Monte Carlo simulation. Perhaps most importantly, our backtest is applicable under a wide variety of structures, since it does not depend on the underlying VaR model, covering either cases where the VaR comes from a conditional volatility model, or it is directly constructed (e.g., CAViaR or ARCH-quantile methods) without relying on a conditional volatility model. We also introduce a main innovation: based on the quantile estimation, one can also identify periods in which the VaR model might increase the risk exposure, which is a key issue to improve the risk model, and probably a novelty in the literature. A final advantage is that our approach can easily be computed through standard quantile regression softwares.

Although the proposed methodology have several appealing properties, it should be viewed as complementary rather than competing with the existing approaches, due to the limitations of the quantile regression technique discussed along this paper. Furthermore, several important topics remain for future research, such as: (i) time aggregation: how to compute and properly evaluate a 10-day regulatory VaR? Risk models constructed through QAR (Quantile Autoregressive) technique can be quite promising due to the possibility of recursively generation of multiperiod density forecast (see Koenker and Xiao (2006b)); (ii) Our randomness approach of VaR also deserves an extended treatment and leaves room for weaker conditions; (iii) multivariate VaR: although the extension of the analysis for the multivariate quantile regression is not straightforward, several proposals have already been suggested in the literature (e.g., Chaudhuri (1996) and Laine (2001)); (iv) inclusion of other variables to increase the power of VQR test in other directions; (v) improvement of the BIS formula for market required capital; among many others.

According to the Basel Committee (2006), new approaches to backtesting are still being developed and discussed within the broader risk management community. At present, different banks perform different types of backtesting comparisons, and the standards of interpretation also differ somewhat across banks. Active efforts to improve and refine the methods currently in use are underway, with the goal of distinguishing more sharply between accurate and inaccurate risk models. We aim to contribute to the current debate by providing a quantile technique that can be useful as a valuable diagnostic tool, as well as a mean to search for possible model improvements.

Appendix A. Proofs of Propositions

Proof of Proposition 1. By Assumption (1), $\alpha_i(U_t)$ are increasing functions of the iid standard uniform random variable U_t and, thus, $Q_{\alpha_i(U_t)} = \alpha_i(Q_{U_t}) = \alpha_i(\tau)$, since for any monotone increasing function g and a standard uniform random variable, U, we have $Q_{g(U)}(\tau) = g(Q_U(\tau)) = g(\tau)$, where $Q_U(\tau) = \tau$ is the quantile function of U_t. By comonotonicity, we have that $Q_{\sum_{i=1}^{p}\alpha_i(U_t)} = \sum_{i=1}^{p} Q_{\alpha_i(U_t)}$. This way, by also considering assumption (1), we guarantee that the conditional quantile function is monotone increasing in τ, which is a crucial property of Value-at-Risk models. In other words, we have that $Q_{R_t}(\tau_1 \mid \mathcal{F}_{t-1}) < Q_{R_t}(\tau_2 \mid \mathcal{F}_{t-1})$ for all $\tau_1 < \tau_2 \in (0;1)$. Assumptions (2)-(4) are regularity conditions necessary to define the asymptotic covariance matrix, and a continuous conditional quantile function, needed for the CLT (7) of Koenker (2005, Theorem 4.1). A sketch of the proof of this CLT, via a Bahadur representation, is also presented in Hendricks and Koenker (1992, Appendix). Given that we established the conditions for the CLT (7), our proof is concluded by using standard results on quadratic forms: For a given random variable $z \sim N(\mu, \Sigma)$ it follows that $(z-\mu)'\Sigma^{-1}(z-\mu) \sim \chi^2_r$ where $r = rank(\Sigma)$. See Johnson and Kotz (1970, p. 150) and White (1984, Theorem 4.31) for further details. ■

Lemma 1 *Consider two independent random variables X and Y. If X has a continuous pdf and $Y \sim N(0,1)$, then,* $\Pr(X > y \cap Y = y) = \Pr(X > 0)$.

Proof. Initially define the following events $(A) : X > 0$; $(B) : Y > 0$; $(C) : X > Y$. Thus, our objective is to show that $\Pr(C) = \Pr(A)$. Firstly, note that $\Pr(C) = \Pr(A \cap C) + \Pr(A^c \cap C)$ and $\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B^c)$. Moreover, $\Pr(C) = [\Pr(A \cap B \cap C) + \Pr(A \cap B^c \cap C)] + [\Pr(A^c \cap B \cap C) + \Pr(A^c \cap B^c \cap C)]$ and $\Pr(A) = [\Pr(A \cap B \cap C) + \Pr(A \cap B \cap C^c)] + [\Pr(A \cap B^c \cap C) + \Pr(A \cap B^c \cap C^c)]$. This way, $\Pr(C) - \Pr(A) = \Pr(A^c \cap B \cap C) + \Pr(A^c \cap B^c \cap C) - \Pr(A \cap B \cap C^c) - \Pr(A \cap B^c \cap C^c)$. Since $\Pr(A^c \cap B \cap C) = \Pr(A \cap B^c \cap C^c) = 0$, by construction, it follows that $\Pr(C) - \Pr(A) = \Pr(A^c \cap B^c \cap C) - \Pr(A \cap B \cap C^c)$. However, since Y has zero mean with a symmetric pdf, it follows that $\Pr(Y > x) = \Pr(Y < -x)$, where $x \in \mathbb{R}^+$. In other words, for any $X = x \in \mathbb{R}$ we have that $\Pr(Y > X \cap X, Y > 0) = \Pr(Y < X \cap X, Y < 0)$. Therefore, $\Pr(C) - \Pr(A) = 0$. ■

Proof of Proposition 2. (i) $P1 \Leftrightarrow S1$ Assume that the nominal quantile level of the VaR model is τ^*, i.e., $\Pr[R_t \leq V_t | \mathcal{F}_{t-1}] = \tau^*$. If assumptions (1)-(4) hold, then, it follows that $Q_{R_t}(\tau \mid \mathcal{F}_{t-1}) = \inf\{R_t : F(R_t \mid \mathcal{F}_{t-1}) \geq \tau\}$ and, thus, $\Pr(R_t \leq Q_{R_t}(\tau \mid \mathcal{F}_{t-1}) \mid \mathcal{F}_{t-1}) = \tau$. In particular, for $\tau = \tau^*$, we have that $\Pr(R_t \leq Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) \mid \mathcal{F}_{t-1}) = \tau^*$. Therefore, it follows that $\tau^* = \Pr(R_t \leq V_t \mid \mathcal{F}_{t-1}) = \Pr(R_t \leq Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) \mid \mathcal{F}_{t-1}) \Leftrightarrow V_t = Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$.

(iia) $S1 \Rightarrow S2$ From the definition of H_t, it follows that $E(H_t \mid \mathcal{F}_{t-1}) = 1{*}\Pr(R_t > V_t \mid \mathcal{F}_{t-1}) + 0{*}\Pr(R_t \leq V_t \mid \mathcal{F}_{t-1}) = \Pr(R_t > V_t \mid \mathcal{F}_{t-1}) = \Pr(R_t > Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) \mid \mathcal{F}_{t-1})$, where the last equality is due to S1. This way, $E(H_t \mid \mathcal{F}_{t-1}) = 1 - \Pr(R_t \leq Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) \mid \mathcal{F}_{t-1}) = 1 - \tau^* = \tau^{**}$ based on the definition of the conditional quantile function. Therefore, $E(H_t - \tau^{**} \mid \mathcal{F}_{t-1}) = 0$.

(iib) $S1 \Rightarrow S3$ From the previous item, it follows that $S1 \Rightarrow S2$. Following Berkowitz et al. (2006), the martingale difference hypothesis ($S2$) naturally implies that the demeaned violation sequence is uncorrelated at all leads and lags. More specifically, the violation sequence has a first-order autocorrelation of zero, which is exploited by the Markov test of Christoffersen (1998). In other words, $S2 \Rightarrow S3$ and, therefore, $S1 \Rightarrow S3$. In addition, note that $E(H_t^2 \mid \mathcal{F}_{t-1}) = \Pr(R_t > V_t \mid \mathcal{F}_{t-1}) = \tau^{**}$ and $Var(H_t \mid \mathcal{F}_{t-1}) = E(H_t^2 \mid \mathcal{F}_{t-1}) - [E(H_t \mid \mathcal{F}_{t-1})]^2 = \tau^{**} - (\tau^{**})^2 = \tau^{**}(1-\tau^{**})$. Therefore, the random variable H_t follows a Bernoulli (τ^{**}) distribution.

(iic) $S1 \Rightarrow S4$ From item (iia), it follows that $S1 \Rightarrow S2$. Applying the law of iterated expectations on $S2$, it follows that $E(H_t) = \tau^{**}$.

(iiia) $S2 \not\Rightarrow S1$ Consider the following VaR model $V_t = Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) + \eta_t$, where $\eta_t \sim iid\ N(0,1)$, inspired by Engle & Manganelli (2004), which argue that: *"any noise introduced into the quantile estimate will change the conditional probability of a hit given the estimate itself"*. Firstly, note that $E(H_t \mid \mathcal{F}_{t-1}) = 1{*}\Pr(R_t > V_t \mid \mathcal{F}_{t-1}) + 0{*}\Pr(R_t \leq V_t \mid \mathcal{F}_{t-1}) = \Pr(R_t > V_t \mid \mathcal{F}_{t-1}) = \Pr(R_t > (Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) + \eta_t) \mid \mathcal{F}_{t-1})$. Now, apply Lemma 1 by defining $X = R_t - Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$ and $Y = \eta_t$. Thus, $\Pr(R_t > V_t \mid \mathcal{F}_{t-1}) = \Pr(R_t > Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1}) \mid \mathcal{F}_{t-1}) = 1 - \tau^* = \tau^{**}$, based on the definition of the conditional quantile function. This way, $E(H_t \mid \mathcal{F}_{t-1}) = \tau^{**}$ and $Var(H_t \mid \mathcal{F}_{t-1}) = \tau^{**}(1 - \tau^{**})$. Therefore, the considered VaR model V_t satisfies $S2$. On the other hand, by definition, V_t clearly does not satisfy $S1$, since $V_t \neq Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$.

(iiib) $S3 \not\Rightarrow S1$ Based on the same example of item (iiia), it follows that $E(H_t - \tau^{**} \mid \mathcal{F}_{t-1}) = 0$ and, thus, $E(H_t - \tau^{**})(H_{t-1} - \tau^{**}) = 0$, i.e., $S3$ holds, whereas, $S1$ does not hold by construction.

(iiic) $S4 \not\Rightarrow S1$ From $S4$, we have that $E(H_t) = \tau^{**}$, which is not sufficient to guarantee that $E(H_t \mid \mathcal{F}_{t-1}) = \tau^{**}$ neither $\Pr(R_t > V_t \mid \mathcal{F}_{t-1}) = \tau^{**}$, i.e. $\Pr(R_t \leq V_t \mid \mathcal{F}_{t-1}) = \tau^*$, or $V_t = Q_{R_t}(\tau^* \mid \mathcal{F}_{t-1})$. ■

Appendix B. Regulatory Framework

The Basle Accord, also known as the 1988 Bank of International Settlements (BIS) Accord, established international guidelines that linked bank's capital requirements to their credit exposures. The "1996 Amendment" extended the initial Accord to include risk-based capital requirements for the market risks that banks incur in their trading accounts, officially consecrating the use of internal models based on Value-at-Risk methodologies to assess market risk exposure. The fact that banks were required to hold capital to face market risk associated with their trading positions intends to create incentives for them to develop their own internal VaR models. The advantage for the banks using an internal model should be a substantial reduction in regulatory capital. The current regulatory framework uses a so-called "traffic-light" approach for the daily market required capital (MRC_t), which is calculated in the following way:

$$MRC_t = \max(V_t; \frac{k}{60}\sum_{i=0}^{59} V_{t-i}) + SRC_t, \tag{16}$$

where V_t is the daily global VaR calculated for the 99% one sided significance level, over a 10-day forecast horizon, SRC_t is a specific risk charge (for the portfolio's idiosyncratic risk), and k represents a multiplicative factor applied to the average VaR and depends on the backtesting results, as it follows:

Table 5 - Multiplier (k) based on the number of exceptions (N)

"Traffic-light"	N	k
Green Zone	4 or fewer	3.00
Yellow Zone	5	3.40
Yellow Zone	6	3.50
Yellow Zone	7	3.65
Yellow Zone	8	3.75
Yellow Zone	9	3.85
Red Zone	10 or more	4.00

where N is the number of violations of V_t in the previous one year of historical data (250 trading days).[24] The k factor can be set by individual supervisory authorities on the basis of their assessment of the quality of the bank's risk management system, directly related to the ex-post performance of the model, thereby introducing a built-in positive incentive to maintain the predictive quality of the model.

According to the Basle Committee (2006), it is with the statistical limitations of "backtesting" in mind that the Basle Committee introduced a framework for the supervisory interpretation of backtesting results that encompasses a range of possible responses, depending on the strength of the signal generated from the backtest. These responses are classified into three zones, distinguished by colours into a hierarchy of responses. The green zone corresponds to backtesting results that do not themselves suggest a problem with the quality or accuracy of a bank's model. The yellow zone encompasses results that do raise questions in this regard, but where such a conclusion is not definitive. In this case, the penalty is up to the supervisor, depending to the reason for the violation. The red zone indicates a backtesting result that almost certainly indicates a problem with a bank's risk model. Mr. Tommaso Padoa-Schioppa, former chairman of the Basle Committee (apud Jorion, 2007), argues that this system is *"designed to reward truthful internal monitoring, as well as developing sound risk management systems."*

[24] According to Crouhy et al. (2001), when being employed in relation to regulatory requirements, backtests must compare daily VaR forecasts against two measures of the profit & loss (P&L) results: (i) The actual net trading P&L for the next day; (ii) The theoretical P&L, also called "static P&L", that would have occurred if the position at the close of the previous day had been carried forward to the next day, i.e., the revenue that would have been realized had the bank's positions remained the same throughout the next day. The main reason is that VaR measures should not be compared against actual trading outcomes, since the actual outcomes would inevitably be "contaminated" by changes in portfolio composition during the holding period. In addition, the inclusion of fee income together with trading gains and losses resulting from changes in the composition of the portfolio should also not be included in the definition of the trading outcome because they do not relate to the risk inherent in the static portfolio that was assumed in constructing the value-at-risk measure. Since this fee income is not typically included in the calculation of the risk measure, problems with the risk measurement model could be masked by including fee income in the definition of the trading outcome used for backtesting purposes. For these reasons, Supervisors will have national discretion to require banks to perform backtesting on either hypothetical (i.e. using changes in portfolio value that would occur were end-of-day positions to remain unchanged), or actual trading (i.e. excluding fees, commissions, and net interest income) outcomes, or both.

Furthermore, regulators accept that it is the nature of the modern banking world that institutions will use different assumptions and modeling techniques (see Crouhy et al. (2001)). The regulators take account of this for their own purposes by requiring institutions to scale up the VaR number derived from the internal model by a k factor, which can be viewed as "insurance" against model misspecification or can also be regarded as a safety factor against "non-normal market moves". In the same line, Jorion (2007) argues that k also accounts for additional risks not modeled by the usual applications of VaR. According to the author, studies of portfolios based on historical data, reporting the performance of the MRC_t during the turbulence of 1998, have shown that while 99% VaR is often exceeded, a multiplier of 3 provides adequate protection against extreme losses. Jorion (2007) also provides a very interesting (possible) rationale for the multiplicative k factor, due to Stahl (1997), based on the Chebyshev's inequality.[25]

However, there are several critiques to the k multiplier in the literature. For instance, Danielsson et al. (1998) argue that current VaR regulation may, perversely, provide incentives for banks to underestimate VaR as much as possible. The ISDA/LIBA 1996 Joint Models task force (apud Crouhy et al., 2001) considers that a multiplier of any size is an unfair penalty on banks that are already sophisticated in the design of their own risk management system. In addition, ISDA also argues that an arbitrarily high scaling factor may even provide perverse incentives to abandon initiatives to implement prudent modifications of the internal model.

In this sense, Berkowitz and O'Brien (2002) report too few violations of actual VaRs in the U.S., indicating overly conservative models for six large commercial banks. These results are quite surprising because they imply that the market risk charges are too high. Recall that a poor VaR specification might lead to a higher capital requirement, which provides an incentive for the banks to improve their internal risk models. However, the capital requirement might not be a binding condition, since the capital that U.S. banks currently hold is above the regulatory capital. Another potential explanation is the existence of incentives for no violations: banks could prefer to report higher VaR numbers to avoid the possibility of regulatory intrusion.

Another important issue regarding the regulatory framework is the "square-root-of-time rule". The current regulatory framework requires that financial institutions use their own internal risk models to calculate and report their 99% VaR over a 10-day horizon.[26] In practice, however, banks are allowed (during an initial phase of the implementation of the internal model) to compute their 10-day ahead VaR by scaling up their 1-day VaR by $\sqrt{10}$, i.e., banks may use VaR numbers calculated according to shorter holding periods scaled up to ten days by the square root of time.

[25] The main idea is to generate a robust upper limit to the VaR when the model is misspecified. Let x be a random variable with expected value μ and finite variance σ^2. Then for any real number $r > 0$ it follows that $\Pr(|x-\mu| > r\sigma) \leq \frac{1}{r^2}$. By assuming a symmetric distribution, we have that $\Pr((x-\mu) < -r\sigma) \leq \frac{1}{2r^2}$. Now, set the desired confidence level $\tau^* = 1\%$ on the right side of the previous expression in order to obtain the respective value of r, i.e., provided that $1/2r^2 = 0.01 \therefore r = 7.071$. Thus, last expression becomes $\Pr((x-\mu) < -7.071\sigma) \leq 0.01$, where the maximum VaR measure is $V_t^{\max} = 7.071\sigma$. Say that the bank report its VaR model using a normal distribution, we have that $V_t = 2.326\sigma$. If the true distribution is misspecified, the correction factor is then $k = \frac{V_t^{\max}}{V_t} = \frac{7.071\sigma}{2.326\sigma} = 3.03$, which is an attempt to justify the correction factor adopted by the Basel Committee.

[26] The 10-day holding period means that regulators are asking banks to consider that they might not be able to liquidate their positions for a 2-week period.

If we assume that returns are $iid \sim N(\mu, \sigma^2)$, then, the 10-day return is also normally distributed with mean 10μ and variance $10\sigma^2$. Thus, it follows that $V_t^{(10-day)} = \sqrt{10}V_t^{(1-day)}$. It is well known that the self-additivity of normal distributions implies the $\sqrt{T}$ scaling factor for multiperiod VaR. However, for heavy-tailed distributions this factor can be different for the largest risks. Danielsson and de Vries (2000) argue that the appropriate method for scaling up a single day VaR to a multiperiod VaR is the "alpha-root rule", where alpha is the number of finite bounded moments, also known as the tail index. According to the authors, heavy tailed distributions are self–additive in the tails, implying a scaling factor $T^{1/\alpha}$. Danielsson and Zigrand (2005) argue that the "square-root-of-time rule" could lead to a systematic underestimation of risk. See also Taylor (1999), which proposed a procedure to estimate a conditional quantile model over the next n periods, and Chen (2001) for a forecasting multiperiod VaR based on a quantile regressions.

A final remark is about the backtest implicitly incorporated into the BIS formulation. Campbell (2005) notes that the k multiplier is solely determined by the number of hits in the past 250 trading days in the same manner as the Kupiec (1995) test. This way, the market capital requirement can be interpreted as an unconditional coverage test that mandates a larger market risk capital set-aside as the evidence that the VaR model under consideration is misspecified. Jorion (2007) argues that regulators operate under different constraints from financial institutions and, since they do not have access to every component of the models, the approach is at a broader level. However, a serious caveat of the Kupiec (1995) test is the difficulty to detect VaR models that systematically under report risk (low power) in sample sizes consistent with the regulatory framework (i.e., $T = 250$). According to Jorion (2007), the lack of power of this framework is due to the choice of the high VaR confidence level (99%) that generates too few exceptions for a reliable test.[27]

A second drawback is that unconditional coverage tests may fail to detect VaR models with adequate unconditional coverage, but with dependent VaR violations, in which case the independence test is recommended. According to Campbell (2005), the unconditional coverage and independence (no clustering) properties are separate and distinct, and must (both) be satisfied by an accurate VaR model. In this paper, we also showed that these two conditions are necessary but not sufficient conditions for a desirable VaR measure (Proposition 2), due to the limited information contained in the hit sequence that ignores the respective magnitude of violations. This issued is also discussed by Campbell (2005), which states that the reported quantile provides a quantitative and continuous measure of the magnitude of realized profits and losses while the hit indicator only signals whether a particular threshold was exceeded. In this sense, the author suggests that quantile tests can provide additional power to detect an inaccurate risk model, which is exactly the idea we discussed throughout this paper.

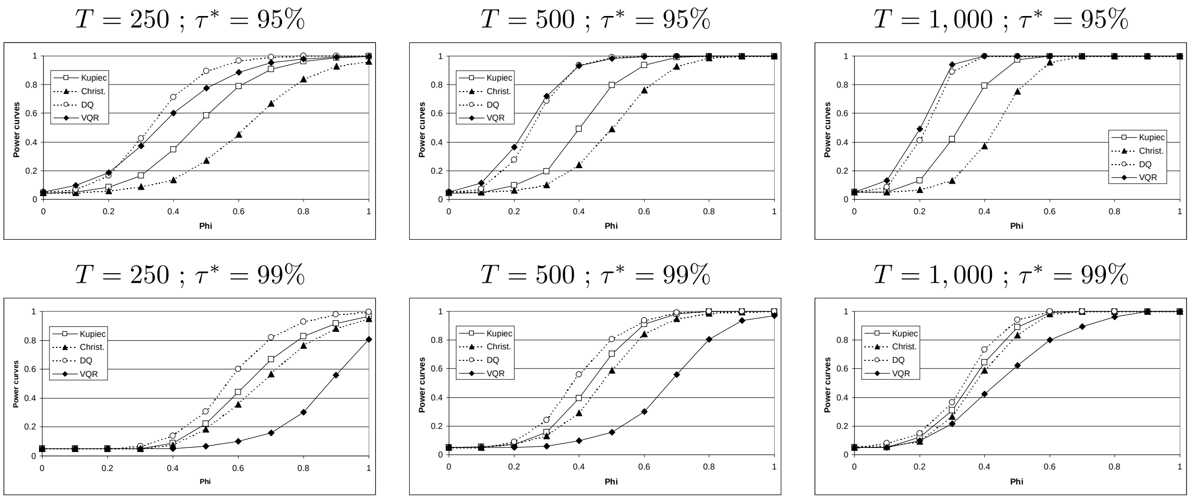
Jorion (2007) says that capital requirements will evolve automatically at the same speed as risk measurement techniques. According to the Basle Committee (2006), the essence of all backtesting efforts is the comparison of actual trading results with model-generated risk measures. If this comparison is close enough, the backtest raises no issues regarding the quality of the risk measurement model. In some cases, however,

[27] Note that there are many combinations of confidence level, the horizon and the multiplicative factor that would yield the same capital charge MRC_t. This way, some suggestions to increase the power of the backtest, already pointed out by Jorion (2007, p.150-151), are to increase the number of observations T from 250 to 1,000 or decrease the confidence level from 99% to 95%.

the comparison uncovers sufficient differences that problems almost certainly must exist, either with the model or with the assumptions of the backtest. In between these two cases there is a gray area where the test results are, on their own, inconclusive. Based on a quantile regression framework, we try to contribute to the debate inside the "gray area".

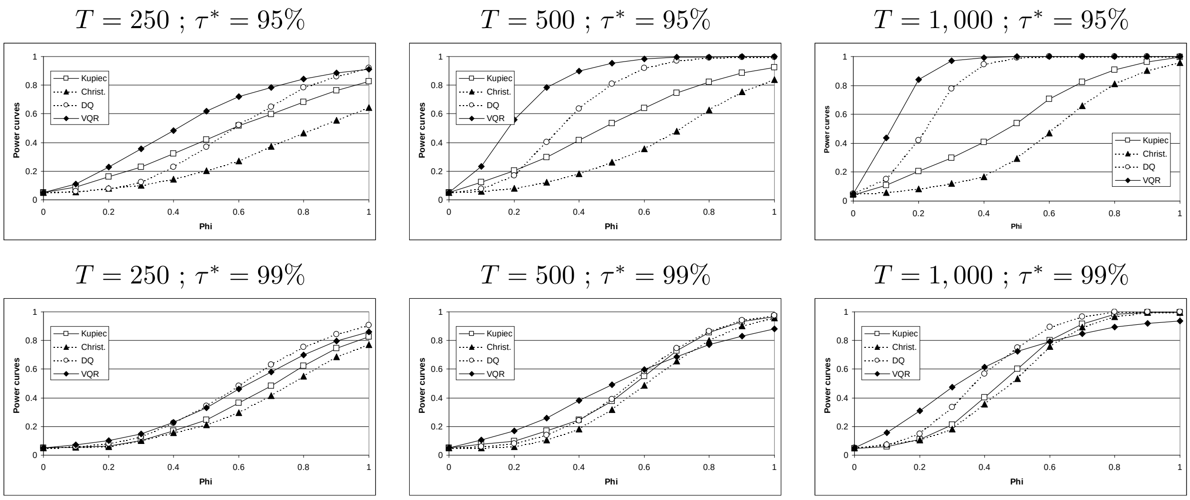
Appendix C. Monte Carlo simulation

Figure 7 - Size-corrected Power Curves - Method 2



Note: Nominal level of significance is $\alpha = 5\%$.

Figure 8 - Size-corrected Power Curves - Method 3



Note: Nominal level of significance is $\alpha = 5\%$.

References

[1] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., 1999. Coherent Measures of Risk. Mathematical Finance 9, 203-228.

[2] Basle Committee on Banking Supervision, 1996. Amendment to the Capital Accord to Incorporate Market Risks. BIS - Bank of International Settlements.

[3] ______, 2006. International Convergence of Capital Measurement and Capital Standards. BIS - Bank of International Settlements.

[4] Berkowitz, J., O'Brien, J., 2002. How Accurate are the Value-at-Risk Models at Commercial Banks? Journal of Finance 57, 1093-1111.

[5] Berkowitz, J., Christoffersen, P., Pelletier, D., 2006. Evaluating Value-at-Risk Models with Desk-Level Data, mimeo.

[6] Campbell, S.D., 2005. A Review of Backtesting and Backtesting Procedures. Finance and Economics Discussion Series, Working Paper 21.

[7] Chaudhuri, P., 1996. On a geometric notion of quantiles for multivariate data. Journal of the American Statistical Association 91, 862-872.

[8] Chen, J.E., 2001. Investigations on Quantile Regression: Theories and Applications for Time Series Models. National Chung Cheng University, mimeo.

[9] Chernozhukov, V., Umantsev, L., 2001. Conditional Value-at-Risk: Aspects of Modelling and Estimation. Empirical Economics 26 (1), 271-292.

[10] Christoffersen, P.F., 1998. Evaluating interval forecasts, International Economic Review 39, 841-862.

[11] ______, 2003. Elements of Financial Risk Management, San Diego: Academic Press.

[12] ______, 2006. Value-at-Risk Models. Desautels Faculty of Management, McGill University, mimeo.

[13] Christoffersen, P.F., Hahn, J., Inoue, A., 2001. Testing and Comparing Value-at-Risk Measures. Journal of Empirical Finance 8, 325-342.

[14] Christoffersen, P.F., Pelletier, D., 2004. Backtesting Value-at-Risk: A Duration-Based Approach. Journal of Financial Econometrics 2 (1), 84-108.

[15] Crnkovic, C., Drachman, J., 1997. Quality Control in VaR: Understanding and Applying Value-at-Risk. London: Risk Publications.

[16] Crouhy, M., Galai, D., Mark, R., 2001. Risk Management. McGraw-Hill.

[17] Danielsson, J., de Vries, C., 1997. Tail index and quantile estimation with very high frequency data. Journal of Empirical Finance 4, 241-257.

[18] ______, 2000. Value–at–Risk and Extreme Returns. London School of Economics, available at www.riskresearch.org.

[19] Danielsson, J., Hartmann, P., de Vries, C., 1998. The Cost of Conservatism: Extreme Returns, Value-at-Risk, and the Basle 'Multiplication Factor'. London School of Economics, available at www.riskresearch.org.

[20] Danielsson, J., Zigrand, J.-P., 2005. On time-scaling of risk and the square-root-of-time rule. London School of Economics, available at www.riskresearch.org.

[21] Ding, Z., Granger, C.W.J, Engle, R.F. 1993. A Long Memory Property of Stock Market Returns and a New Model. Journal of Empirical Finance 1, 83-106.

[22] Dowd, K., 2005. Measuring Market Risk. John Wiley and Sons Ltd.

[23] Embrechts, P., Kluppelberg, C., Mikosch, T., 1997. Modelling Extremal Events. Springer-Verlag.

[24] Embrechts, P., Resnick, Samorodnitsky, G., 1999. Extreme value theory as a risk management tool. North American Actuarial Journal 3, 30-41.

[25] Engle, R.F., 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. Econometrica 50, 987-1008.

[26] Engle, R.F., Manganelli, S., 2004. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. Journal of Business and Economic Statistics 22 (4), 367-381.

[27] Engle, R.F., Patton, A.J., 2001. What good is a volatility model? Quantitative Finance, Institute of Physics Publishing 1, 237–245.

[28] Giacomini, R., Komunjer, I., 2005. Evaluation and Combination of Conditional Quantile Forecasts. Journal of Business and Economic Statistics 23 (4), 416-431.

[29] Hafner, C.M., Linton, O., 2006. Comment on "Quantile Autoregression". Journal of the American Statistical Association 101 (475), 998-1001.

[30] Hendricks, W., Koenker, R., 1992. Hierarchical Spline Models for Conditional Quantiles and the Demand for Electricity. Journal of the American Statistical Association 87 (417), 58–68.

[31] Huber, P.J., 1967. The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. University of California Press, Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability 4, 221-233.

[32] Huisman, R., Koedijk, K.G., Kool, C.J.M., Palm, F. 2001. Tail-Index Estimates in Small Samples. Journal of Business & Economic Statistics 19(2), 208-216.

[33] Hull, J.C., 2005. Options, Futures, and Other Derivatives. Prentice-Hall, 6th edition.

[34] Johnson, N.L., Kotz, S., 1970. Distributions in statistics: Continuous univariate distributions. Wiley Interscience.

[35] Jorion, P., 2007. Value-at-risk: The new benchmark for managing financial risk. McGraw Hill, 3rd edition.

[36] J.P. Morgan, 1996. RiskMetrics, Technical Document. New York, 4th edition.

[37] Kim, T.H., White, H., 2003. Estimation, Inference, and Specification Testing for Possibly Misspecified Quantile Regression. Advances in Econometrics 17, 107–132.

[38] Koenker, 2004. The Quantreg Package v3.35 (quantile regression and related methods). R Foundation for Statistical Computing.

[39] ______, 2005. Quantile Regression. Cambridge University Press.

[40] Koenker, R., Bassett, G., 1978. Regression Quantiles. Econometrica 46, 33-50.

[41] ______, 1982a. Robust Tests for Heteroscedasticity Based on Regression Quantiles. Econometrica 50 (1), 43–62.

[42] ______, 1982b. Tests of Linear Hypotheses and L1 Estimation. Econometrica 50 (6), 1577–1584.

[43] Koenker, R., Machado, J.A.F., 1999. Goodness of Fit and Related Inference Processes for Quantile Regression. Journal of the American Statistical Association 94 (448), 1296–1310.

[44] Koenker, R., Portnoy, S., 1999. Quantile Regression. Unpublished Manuscript, University of Illinois.

[45] Koenker, R., Xiao, Z., 2002. Inference on the Quantile Regression Process. Econometrica 70 (4), 1583–1612.

[46] ______, 2006a. Quantile Autoregression. Journal of the American Statistical Association 101 (475), 980-990.

[47] ______, 2006b. Rejoinder of "Quantile Autoregression". Journal of the American Statistical Association 101 (475), 1002-1006.

[48] Koenker, R., Zhao, Q., 1996. Conditional Quantile Estimation and Inference for ARCH Models. Econometric Theory 12, 793-813.

[49] Kupiec, P., 1995. Techniques for Verifying the Accuracy of Risk Measurement Models. Journal of Derivatives 3, 73-84.

[50] Laine, B., 2001. Depth contours as multivariate quantiles: a directional approach. Master's thesis. Université Libre de Bruxelles.

[51] Laurent S., Peters, J.-P. 2006. G@RCH 4.2, Estimating and Forecasting ARCH Models. Timberlake Consultants Press. London.

[52] Lima, L.R., Neri, B.A.P., 2006. Comparing Value-at-Risk Methodologies. Ensaios Economicos EPGE, 629. Mimeo. Getulio Vargas Foundation.

[53] Linton, O., Whang, Y.J., 2007. The quantilogram: With an application to evaluating directional predictability. Journal of Econometrics, Forthcoming.

[54] Lopez, J.A., 1999. Methods for Evaluating Value-at-Risk Estimates. Federal Reserve Bank of San Francisco, Economic Review 2, 3-17.

[55] Machado, J.A.F., Mata, J., 2001. Earning functions in Portugal 1982-1994: Evidence from quantile regressions. Empirical Economics 26, 115-134.

[56] McNeil, A.J., Frey, R. 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance 7(3), 271-300.

[57] Mincer, J., Zarnowitz, V., 1969. The evaluation of economic forecasts and expections. In: Mincer, J. (ed.), Economic Forecasts and Expections. National Bureau of Economic Research, New York.

[58] Mittnik, S., Paolella, M.S. 2000. Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates. Journal of Forecasting 19, 313-333.

[59] Nankervis, J., Sajjad, R., Coakley, J., 2006. Value-at-Risk for long and short positions: A comparison of regime-switching GARCH and ARCH family models. University of Essex, mimeo.

[60] Powell, J., 1991. Estimation of monotonic regression under quantile restrictions. In: Barnett, W., et al. (eds.), Nonparametric and Semiparametric Methods in Economics and Statistics. Cambridge University Press.

[61] Pritsker, M., 2001. The Hidden Dangers of Historical Simulation. Washington: Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series 27.

[62] Robinson, P.M., 2006. Comment on "Quantile Autoregression". Journal of the American Statistical Association 101 (475), 1001-1002.

[63] Ruppert, D., Carroll, R.J., 1980. Trimmed Least Squares Estimation in the Linear Model, Journal of the American Statistical Association 75 (372), 828-838.

[64] Schmeidler, D., 1986. Integral representation without additivity. Proceedings of the American Mathematical Society 97, 255-261.

[65] Schulze, N., 2004. Applied Quantile Regression: Microeconometric, Financial, and Environmental Analyses. Dissertation zur Erlangung des Doktorgrades, Universität Tübingen.

[66] Stahl, G., 1997. Three Cheers. Risk 10, 67-69.

[67] Stahl, G., Carsten, W., Zapp, A., 2006. Backtesting beyond the trading book. Journal of Risk 8 (winter), 1-16.

[68] Stock, J.H., Watson, M.W., 2002. Forecasting Using Principal Components From a Large Number of Predictors. Journal of the American Statistical Association 97 (460), 1167-1179.

[69] Taylor, J.W., 1999. A Quantile Regression Approach to Estimating the Distribution of Multiperiod Returns. Journal of Derivatives 7 (1), 64–78.

[70] White, H., 1984. Asymptotic Theory for Econometricians. Academic Press, San Diego.

[71] Wu, G., Xiao, Z., 2002. An Analysis of Risk Measures. Journal of Risk 4, 53-75.