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: Carlos Enrique Carrasco Gutierrez
Professor Orientador: João Victor Issler

Rio de Janeiro
2008

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: Carlos Enrique Carrasco Gutierrez

Banca Examinadora:
João Victor Issler (Orientador, EPGE-FGV)
Caio Ibsen Rodrigues de Almeida (EPGE-FGV)
Carlos Eugenio Ellery Lustosa da Costa (EPGE-FGV)
Marco Antonio César Bonomo (EPGE-FGV)
Cristiano Augusto Fernandes Coelho (PUC-Rio)
Wagner Piazza Gaglianone (Banco Central do Brasil)

Rio de Janeiro
2008

Contents

Acknowledgements iv

List of Tables v

List of Figures vi

Introduction 1

Chapter 1 - *Selection of Optimal Lag Length in Cointegrated VAR Models with Weak Form of Common Cyclical Features* 2

Chapter 2 - *Evidence on Common Features and Business Cycle Synchronization in Mercosur* 22

Chapter 3 - *Evaluating Asset Pricing Models in a Fama-French Framework* 47

Chapter 4 - *Testing Consumption-based Asset Pricing Models with Factor Model Analysis* 67

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, esta tese não seria possível.
Aos membros da Banca Examinadora, pela paciência na leitura dos artigos, e pelas valiosas contribuições a este trabalho.
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, Daniel Ferreira, Renato Flores, Marcelo Fernandes, Paolo Klinger, Carlos Eugenio, Luis Braido, Marcelo Moreira.
A meus colegas e parceiros em alguns paper, Fábio Gomes, Osmani Texeira, Wagner Piazza.
A meus professores da PUC-Rio, Carlos Kubrusly, Reinaldo Castro, Mônica Barros, Cristiano Fernandes.
Aos funcionários da EPGE e da Biblioteca Mario Henrique Simonsen, pelo apoio técnico, indispensável.
A Christiam Gonzales pela sua companhia e amizade.
Aos colegas da EPGE, pela amizade e companheirismo durante estes anos de convivência.
Aos meus pais Lida e Mauro, meus irmãos, Lucho e Gabriela, que sempre me apoiaram, com muito amor e carinho, em todos os meus projetos e desafios, ao longo de toda a minha vida.
A meus filhos, Juan Carlitos e Camilita, por fazer-me sentir amado.
A Maria Luiza, pelo seu amor.
A todos meus amigos, dos quais sempre recebi grande incentivo.
À FGV e ao Banco BBM pela bolsa de estudos que me deu a possibilidade de me dedicar exclusivamente a os meus estudos e artigos de tese.
Por último, e mais importante, a Deus porque me acompanha, me protege e me sustenta sempre.

List of Tables

Chapter 1

Table 1 – Performance of information criterion, $IC(p)$, in selecting the lag order p 15

Table 2 – Performance of information criterion, $IC(p, s)$, in selecting p and s simultaneously 16

Chapter 2

Table 1 – Descriptive Statistics 31

Table 2 – Statistics of Unit Root Tests 32

Table 3 – Identification of the VAR order 32

Table 4 – Johansen's cointegration test 34

Table 5 – Common cycle test 35

Table 6 – Linear correlations in business cycles and in common cycle 38

Chapter 3

Table 1 – Monte Carlo Simulation Results 61

Table 2 – Descriptive statistics of the SDF proxies 63

Chapter 4

Table 1 – Power Utility Function Estimation Results 78

Table 2 – Power Utility Function Estimation Results 79

Table 3 – External Habit Utility Function Estimation Results 82

Table 4 – External Habit Utility Function Estimation Results 83

Table 5 – Kreps-Porteous Utility Function Estimation Results 84

Table 6 – Kreps-Porteous Utility Function Estimation Results 85

Table 7 – Descriptive Statistics 86

Table 8 – Cumulative Explained Variation on S&P500 87

List of Figures

Chapter 1

Figure 1 – One realization of a VAR(3) model with $n = 3$, $r = 1$ and $s = 2$ 12

Chapter 2

Figure 1 – Real GDP (in log) per capita series of Mercosur countries (1951-2003) 30
Figure 2 – The growth rates of the real GDP per capita series of Mercosur countries (1951-2003) 30
Figure 3 – Common Cycles. 36
Figure 4 – Cyclical components for $p = 1$, $s = 3$ and $r = 2$. 36
Figure 5 – Cyclical components in each country for $p = 1$, $p = 2$ and $p = 3$. 37
Figure 6 – Coherence and Phase 40
Figure 7 – Coherence and Phase 41
Figure 8 – Coherence and Phase 42
Figure 9 – Coherence and Phase 43

Chapter 3

Figure 1 – Three factors, with a sample size $T = 200$ 59

Chapter 4

Figure 1 – A Scree plot for a 106 stock on S&P500 88

Introdução

A presente tese é composta de quatro ensaios sobre macroeconometria e finanças. Em cada ensaio, que corresponde a um capítulo, o objetivo é investigar e analisar as técnicas econometrias avançadas, aplicadas às questões macroeconômicas e financeiras relevantes.

O primeiro capítulo estuda a seleção da ordem de defasagem em modelos VAR com características cíclicas comuns como introduzidas nos trabalhos seminais de Engle e Kozicki (1993). Em particular, o artigo estende resultados precedentes de Vahid e Issler (2001) para o caso dos modelos VAR com restrições nas características comuns no sentido fraco (Weak Form - WF) tal como proposto por Hecq et al. (2006).

O segundo capítulo apresenta uma aplicação empírica da teoria econometrica de características comuns. O objetivo deste trabalho é analisar os ciclos de negócios dos países membros do Mercosur a fim investigar seu grau de sincronização. O modelo usa a decomposição multivariada tendência-ciclo de Beveridge-Nelson-Estoque-Watson, tomando em consideração a presença de características comuns tais como a tendência comum e o ciclo comum.

No terceiro capítulo é proposto uma metodologia para comparar modelos de precificação de ativos através dos fatores estocásticos de desconto (SDF) baseados na informação relevante do mercado. O ponto de partida é o trabalho de Fama e French, que evidenciou que os retornos dos ativos da economia dos EUA poderiam ser explicados pelos fatores relativos às características das empresas. Uma aplicação empírica é fornecida.

Finalmente, o último capítulo apresenta o estudo de estimação e teste de hipóteses para diferentes modelos consumo no arcabouço de CCAPM. Foram usados um modelo de fatores para construir um número pequeno das carteiras e o método dos momentos generalizado (MMG) para identificar e estimar os parâmetros da função utilidade. O resultado principal é obtido pela função utilidade CRRA.

Introduction

This thesis is composed of four 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 and financial questions.

The first chapter studied the lag length selection in VAR models with common cyclical features as introduced in the early work of Engle and Kozicki (1993). In particular, the paper extends previous results by Vahid and Issler (2001) to the case of VAR models having the weak form of serial correlation common feature (WF henceforth) proposed by Hecq et al. (2006).

The second chapter presents an empirical application of common features. The aim of this work is to analyze the business cycles of the Mercosur's member countries in order to investigate their degree of synchronization. The model estimation uses the Beveridge-Nelson-Stock-Watson multivariate trend-cycle decomposition, taking into account the presence of common features such as common trend and common cycle.

In the third chapter is proposed a methodology to compare different stochastic discount factor (SDF) proxies based on relevant market information. The starting point is the work of Fama and French, which evidenced that the asset returns of the U.S. economy could be explained by relative factors linked to characteristics of the firms. An empirical application of our setup is also provided.

Finally, the last chapter presents the study of estimation and testing of different representative classes of consumption-based asset pricing models (CCAPM). We use the factor model to construct a small set of portfolios and the generalized method of moments (GMM) to identify and estimate the parameters of the utility function. The main result is obtained by the CRRA preference. We show that when the factor portfolio is used, the implication of this model is not rejected.

.

Selection of Optimal Lag Length in Cointegrated VAR Models with Weak Form of Common Cyclical Features[1]

Abstract

An important aspect of empirical research based on the vector autoregressive (VAR) model is the choice of the lag order, since all inference in the VAR model depends on the correct model specification. Literature has shown important studies of how to select the lag order of a nonstationary VAR model subject to cointegration restrictions. In this work, we consider an additional *weak form* (WF) restriction of common cyclical features in the model in order to analyze the appropriate way to select the correct lag order. Two methodologies have been used: the traditional information criteria (AIC, HQ and SC) and an alternative criterion ($IC(p, s)$) which select simultaneously the lag order p and the rank structure s due to the WF restriction. A Monte-Carlo simulation is used in the analysis. The results indicate that the cost of ignoring additional WF restrictions in vector autoregressive modelling can be high, especially when SC criterion is used.

Keywords: Cointegration; Common Cyclical Features; Reduced Rank Model; Estimation; Information Criteria.

JEL Codes: C32, C53.

[1] This article was jointly made with Osmani Guillén and Reinaldo Souza. We are grateful to comments and suggestions given by João Victor Issler, Wagner Gaglianone, Ricardo Cavalcanti, Luiz Renato Lima and participants of the Brazilian Econometric Meeting (December 2006) and European Econometric Society Meeting (August 2007). Special thanks are due to Alain Hecq for solving doubts and comments. The authors are responsible for any remaining errors in this paper. Carlos Enrique acknowledges the support of CAPES-Brazil. suggestions.

1 Introduction

In the modelling of economic and financial time series, the vectorial autoregressive (VAR) model became a standard linear model used in empirical works. An important aspect of empirical research in the specification of the VAR models is the determination of the lag order of the autoregressive lag polynomial, since all inference in the VAR model depends on the correct model specification. In several contributions, the effect of lag length selection has been demonstrated: Lütkepohl (1003) indicates that selecting a higher order lag length than the true lag length causes an increase in the mean square forecast errors of the VAR and that underfitting the lag length often generates autocorrelated errors. Braun and Mittnik (1993) show that impulse response functions and variance decompositions are inconsistently derived from the estimated VAR when the lag length differs from the true lag length. When cointegration restrictions are considered in the model, the effect of lag length selection on the cointegration tests has been demonstrated. For example, Johansen (1991) and Gonzalo (1994) point out that VAR order selection may affect proper inference on cointegrating vectors and rank.

Recently empirical works have considered another kind of restrictions on the VAR model (e.g., Engle and Issler (1995), Caporale (1997) and Mamingi and Sunday(2003)). Engle and Kozicki (1993) showed that VAR models can have another type of restrictions, called common cyclical features, which are restrictions on the short-run dynamics. These restrictions are defined in the same way as cointegration restrictions, while cointegration refers to relations among variables in the long-run, the common cyclical restrictions refer to relations in the short-run. Vahid and Engle (1993) proposed the *Serial Correlation Common Feature* (SCCF) as a measure of common cyclical feature. SCCF restrictions might be imposed in a covariance stationary VAR model or in a cointegrated VAR model. When short-run restrictions are imposed in cointegrated VAR models it is possible to define a weak version of SCCF restrictions. Hecq, Palm and Urbain (2006) defined a weak version of SCCF restrictions which they denominated it as *weak-form* (WF) common cyclical restrictions. A fundamental difference between SCCF and WF restrictions is in the form which each one imposes restrictions on the Vector Error Correction Model (VECM) representation[2]. When SCCF are imposed, all matrices of a VECM have rank less than the number of variables

[2]When a VAR model has cointegration restriction it can be represented as a VECM. This representation is also known as Granger Representation Theorem (Engle and Granger 1987).

analyzed. On the other hand with WF restrictions all matrices, except the long-run matrix, have rank less than a number of variables in analysis. Hence, WF restrictions impose less restriction on VECM parameters. Some advantages emerge when WF restrictions are considered. First, due to the fact that WF restrictions does not impose restrictions on the cointegration space; the rank of common cyclical features is not limited by the choice of cointegrating rank.

The literature has shown how to select an adequate lag order of a covariance stationary VAR model and an adequate lag order of a VAR model subject to cointegration restrictions. Among the classical procedures, there are the information criteria such as Akaike (AIC), Schwarz (SC) and Hannan-Quinn (HQ) (Lütkepohl, 1993). Kilian (2001) study the performance of traditional AIC, SC and HQ criterion of a covariance stationary VAR model. Vahid and Issler (2002) analyzed the standard information criterion in a covariance stationary VAR model subject to SCCF restriction and more recently Guillén, Issler and Athanasopoulos (2005) studied the standard information criterion in VAR models with cointegration and SCCF restrictions. However, when cointegrated VAR models contain additional weak form of common cyclical feature, there are no reported work on how to appropriately determine the VAR model order.

The objective of this paper is to investigate the performance of information criterion in selecting the lag order of a VAR model when the data are generated from a true VAR with cointegration and WF restrictions that is referred as the correct model. It will be carried out following two procedures: *a*) the use of standard criteria as proposed by Vahid and Engle (1993), referred here as IC (p), and *b*) the use of an alternative procedure of model selection criterion (see, Vahid and Issler, 2002 and Hecq, Palm and Urbain 2006) consisting in selecting simultaneously the lag order p and the number of weak form of common cyclical feature, s, which is referred to as IC(p, s)[3]. The most relevant results can be summarized as follows. The information criterion that selects simultaneously the pair (p, s) has better performance than the model chosen by conventional criteria. The cost of ignoring additional WF restrictions in vector autoregressive modelling can be high specially when SC criterion is used.

The remaining of this work is organized as follows. Section 2 shows the econometric model. In section 3 the information criteria are mentioned. Monte Carlo simulation is

[3]This is quite recent in the literature (see, Hecq *et al. 2006*).

shown in section 4 and the results in section 5. Finally, the conclusions are shown in section 6.

2 The Econometric Model

We show the VAR model with short-run and long-run restrictions. First, we consider a Gaussian vector autoregression of finite order p, so-called VAR(p), such that:

$$y_t = \sum_{i=1}^{p} A_i y_{t-i} + \varepsilon_t \tag{1}$$

where, y_t is a vector of n first order integrated series, $I(1)$, A_i, $i = 1, \ldots, p$ are matrices of dimension $n \times n$, $\varepsilon_t \sim Normal\,(0, \Omega)$ and $\{\Omega, \text{ if } t = \tau \text{ and } 0_{n\times n}, \text{ if } t \neq \tau$, where Ω is non singular$\}$. The model (1) could be written equivalently as; $\Pi\,(L)\; y_t = \varepsilon_t$ where L represents the lag operator and $\Pi\,(L) = I_n - \sum_{i=1}^{p} A_i L^i$ that when $L = 1$, $\Pi\,(1) = I_n - \sum_{i=1}^{p} A_i$. If cointegration is considered in (1) the $(n \times n)$ matrix $\Pi\,(\cdot)$ satisfies two conditions: a) Rank $(\Pi\,(1)) = r$, $0 < r < n$, such that $\Pi\,(1)$ can be expressed as $\Pi\,(1) = -\alpha\beta'$, where α and β are $(n \times r)$ matrices with full column rank, r. b) The characteristic equation $|\Pi\,(L)| = 0$ has $n - r$ roots equal to 1 and all other are outside the unit circle. These assumptions imply that y_t is cointegrated of order $(1, 1)$. The elements of α are the adjustment coefficients and the columns of β span the space of cointegration vectors. We can represent a VAR model as VECM. Decomposing the polynomial matrix $\Pi\,(L) = \Pi\,(1)\, L + \Pi^*\,(L)\,\Delta$, where $\Delta \equiv (1 - L)$ is the difference operator, a Vector Error Correction Model (VECM) is obtained:

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{2}$$

where: $\alpha\beta' = -\Pi\,(1)$, $\Gamma_j = -\sum_{k=j+1}^{p} A_k$ for $j = 1, ..., p-1$ and $\Gamma_0 = I_n$. The VAR(p) model can include additional short-horizon restrictions as shown by Vahid and Engle (1993). We consider an interesting WF restriction (as defined by Hecq, Palm and Urbain (2006)) that does not impose restrictions over long-run relations.

Definition 1 *Weak Form-WF holds in (2) if, in addition to assumption 1 (cointegration), there exists a $(n \times s)$ matrix $\tilde{\beta}$ of rank s, whose columns span the cofeature space, such that $\tilde{\beta}'(\Delta y_t - \alpha\beta' y_{t-1}) = \tilde{\beta}'\varepsilon_t$, where $\tilde{\beta}'\varepsilon_t$ is a s-dimensional vector that constitutes an innovation process with respect to information prior to period t, given by $\{y_{t-1},\ y_{t-2}, ..., y_1\}$.*

Consequently we considerate WF restrictions in the VECM if there exists a cofeature matrix $\tilde{\beta}$ that satisfies the following assumption:

Assumption 1 : $\tilde{\beta}'\,\Gamma_j = 0_{s\times n}$ *for* $j = 1, ..., p-1$.

Imposing WF restrictions is convenient because it allows the study of both cointegration and common cyclical feature without the constraint $r + s \leq n$. We can rewrite the VECM with WF restrictions as a model of reduced-rank structure. In (2) let $X_{t-1} = [\Delta y'_{t-1}, ...\Delta y'_{t-p+1}]'$ and $\Phi = [\Gamma_1, ..., \Gamma_{p-1}]$, therefore we get:

$$\Delta y_t = \alpha\beta' y_{t-1} + \Phi X_{t-1} + \varepsilon_t \tag{3}$$

If assumption (1) holds matrices $\Gamma_i, i = 1, ..., p$ are all of rank $(n-s)$ then we can write $\Phi = \tilde{\beta}_\perp \Psi = \tilde{\beta}_\perp [\Psi_1, ..., \Psi_{p-1}]$, where, $\tilde{\beta}_\perp$ is $n \times (n-s)$ full column rank matrix, Ψ is of dimension $(n-s) \times n(p-1)$, the matrices $\Psi_i, i = 1, ..., p-1$ all of rank $(n-s) \times n$. Hence, given assumption (1), there exists $\tilde{\beta}$ of $n \times s$ such that $\tilde{\beta}'\tilde{\beta}_\perp = 0$. That is, $\tilde{\beta}_\perp$ $n \times (n-s)$ is a full column rank orthogonal to the complement of $\tilde{\beta}$ with $rank(\tilde{\beta}, \tilde{\beta}_\perp) = n$. Rewriting model (3) we have:

$$\Delta y_t = \alpha\beta' y_{t-1} + \tilde{\beta}_\perp \left(\Psi_1, \Psi_2, ..., \Psi_{p-1}\right) X_{t-1} + \varepsilon_t \tag{4}$$

$$= \alpha\beta' y_{t-1} + \tilde{\beta}_\perp \Psi X_{t-1} + \varepsilon_t \tag{5}$$

Estimation of (5) is carried out via the switching algorithms (see, Centoni et. al (2007) and Hecq (2006)) that use the procedure in estimating reduced-rank regression models suggested by Anderson (1951). There is a formal connection between a reduced-rank regression and the canonical analysis as noted by Izenman (1975), Box and Tiao (1977), Tso (1981) and Velu Reinsel and Wichern (1986). When the multivariate regression has all of its matrix coefficients of full rank, it may be estimated by usual Least Square or Maximum-Likelihood procedures. But when the matrix coefficients are of reduced-rank they have to be estimated using the reduced-rank regression models of Anderson (1951). The use of canonical analysis may be regarded as a special case of reduced-rank regression. More specifically, the maximum-likelihood estimation of the parameters of the reduced-rank regression model may result in solving a problem of canonical analysis[4]. Therefore, we can use the expression $CanCorr\{X_t, Z_t|X_{t-1}\}$ that denotes the partial canonical correlations between X_{t-1} and Z_t:

[4]This estimation is referred as *Full Information Maximum Likelihood* - FIML

both sets concentrate out the effect of X_t that allows us to obtain canonical correlation, represented by the eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3 ..... > \hat{\lambda}_n$. The Johansen test statistic is based on canonical correlation. In model (2) we can use the expression $CanCorr\{\Delta y_t, y_{t-1}|X_{t-1}\}$ where $X_{t-1} = [\Delta y'_{t-1}, ...\Delta y'_{t-p+1}]'$ that summarizes the reduced-rank regression procedure used in the Johansen approach. It means that one extracts the canonical correlations between Δy_t and y_{t-1}: both sets concentrated out the effect of lags of X_{t-1}. In order to test for the significance of the r largest eigenvalues, one can rely on Johansen's trace statistic (6):

$$\xi_r = -T \sum_{i=r+1}^{n} Ln\,(1 - \hat{\lambda}_i^2) \qquad\qquad i = 1, ..., n \tag{6}$$

where the eigenvalues $0 < \hat{\lambda}_n < ... < \hat{\lambda}_1$ are the solution of : $|\lambda m_{11} - m_{10}^{-1} m_{00} m_{01}| = 0$, where $m_{ij},\ i,\ j = 0.1$, are the second moment matrices: $m_{00} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{0t}\tilde{u}'_{0t}$, $m_{10} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{1t}\tilde{u}'_{0t}$, $m_{01} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{0t}\tilde{u}'_{1t}$, $m_{11} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{1t}\tilde{u}'_{1t}$ of the residuals $\tilde{u}_{0t}$ and $\tilde{u}_{1t}$ obtained in the multivariate least squares regressions $\Delta y_t = (\Delta y_{t-1}, ...\Delta y_{t-p+1}) + u_{0t}$ and $y_{t-1} = (\Delta y_{t-1}, ...\Delta y_{t-p+1}) + u_{1t}$ respectively (see, Hecq *et al.(2006)*; Johansen (1995)). The result of Johansen test is a superconsistent estimated β. Moreover, we could also use a canonical correlation approach to determine the rank of the common features space due to WF restrictions. It is a test for the existence of cofeatures in the form of linear combinations of the variables in the first differences, corrected for long-run effects which are white noise (i.e., $\tilde{\beta}'(\Delta y_t - \alpha\beta' y_{t-1}) = \tilde{\beta}'\varepsilon_t$ where $\tilde{\beta}'\varepsilon_t$ is a white noise). Canonical analysis is adopted in the present work in estimating, testing and selecting lag-rank of VAR models as shown in next sections.

3 Model Selection Criteria

In model selection we use two procedures to identify the VAR model order. The standard selection criteria, IC(p) and the modified informational criteria, IC(p, s), novelty in the literature, which consists on identifying p and s simultaneously.

The model estimation following the standard selection criteria, IC(p), used by Vahid and Engle (1993) entails the following steps:

1. Estimate p using standard informational criteria: Akaike (AIC), Schwarz (SC) and Hanna-Quinn (HQ). We choose the lag length of the VAR in levels that minimize the information criteria.

2. Using the lag length chosen in the previous step, find the number of cointegration vector, r using Johansen cointegration test[5].

3. Conditional on the results of cointegration analysis, a final VECM is estimated and then the multi-step ahead forecast is calculated.

The above procedure is followed when there is evidence of cointegration restrictions. We check the performance of IC(p) when WF restrictions contain the true model. Additionally we check the performance of alternative selection criteria IC(p, s). Vahid and Issler (2002) analyzed a covariance-stationary VAR model with SCCF restrictions. They showed that the use of IC(p, s) has better performance than IC(p) in VAR model lag order selection. In the present work we analyze cointegrated VAR model with WF restrictions in order to analyze the performance of IC(p) and IC(p, s) for model selection. The question investigated is: is the performance of IC(p, s) superior to that of IC(p)? This is an important question we aim to answer in this work.

The procedure of selecting the lag order and the rank of the structure of short-run is carried out by minimizing the following modified information criteria (see Hecq (2006)).

$$AIC\,(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{2}{T} \times N \tag{7}$$

$$HQ(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{2\ln(\ln T)}{T} \times N \tag{8}$$

$$SC(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{\ln T}{T} \times N \tag{9}$$

$$N = [n \times (n \times (p-1)) + n \times r] - [s \times (n \times (p-1) + (n-s))]$$

The number of parameters N is obtained by subtracting the total number of mean parameters in the VECM (i.e., $n^2 \times (p-1) + nr$), for given r and p, from the number of restrictions the common dynamics imposes from $s \times (n \times (p-1)) - s \times (n-s)$. The eigenvalues λ_i

[5]Cointegration rank and vectors are estimated using the FIML as shown in Johansen (1991).

are calculated for each p. To calculate the pair (p, s) we assume that no restriction of cointegration exists, that is, $r = n$ (see Hecq (2006)). We fix p in model (3) and then find λ_i $i = 1, 2...n$ using the program $cancorr(\Delta y_t, X_{t-1} \mid y_{t-1})$. This procedure is followed for every p and in the end we choose the p and s that minimizes the IC(p, s). After selecting the pair (p, s) we can test the cointegration relation using the procedure of Johansen. Finally we estimate the model using the switching algorithms as shown in the next chapter. Notice that in this simultaneous selection, testing the cointegration relation is the last procedure to follow, so we are inverting the hierarquical procedure followed by Vahid and Engle (1993) where the first step is the selection of the number of cointegration relations. It may be an advantage specially when r is over-estimated. Few works have been dedicated to analyze the order of the VAR models considering modified IC(p, s). As mentioned, Vahid and Issler (2002) suggested the use of IC(p, s) to simultaneously choose the order p and a number of reduced rank structure s on covariance stationary VAR model subject to SCCF restrictions. However, no work has analyzed the order of the VAR model with cointegration and WF restrictions using a modified criterion, which is exactly the contribution of this paper.

To estimate the VAR model considering cointegration and WF restrictions we use the switching algorithms model as considered by Hecq (2006). Consider the VECM given by:

$$\Delta y_t = \alpha\beta' y_{t-1} + \tilde{\beta}_{\perp}\Psi X_{t-1} + \varepsilon_t \tag{10}$$

A full description of switching algorithms is presented below in four steps:

*Step*1 : Estimation of the cointegration vectors β.

Using the optimal pair $(\bar{p}, \bar{s})$ chosen by information criteria (14), (15) or (16), we estimate β (and so its rank, $r = \bar{r}$) using Johansen cointegration test.

*Step*2 : Estimation of $\tilde{\beta}_{\perp}$ and Ψ.

Taking $\hat{\beta}$ estimated in step one, we proceed to estimate $\tilde{\beta}_{\perp}$ and Ψ. Hence, we run a regression of Δy_t and of X_{t-1}on $\hat{\beta}' y_{t-1}$. We labeled the residuals as u_0 and u_1, respectively. Therefore, we obtain a reduced rank regression:

$$u_0 = \tilde{\beta}_{\perp}\Psi u_1 + \varepsilon_t \tag{11}$$

where Ψ can be written as $\Psi = \left(C_1, ..., C_{(\bar{p}-1)}\right)$ of $(n - \bar{s}) \times n(\bar{p} - 1)$ and $\tilde{\beta}_{\perp}$ of $n \times (n - \bar{s})$. We estimate (11) by FIML. Thus, we can obtain $\tilde{\beta}_{\perp}$ and $\hat{\Psi}$.

*Step*3 : Estimate of the Maximum Likelihood (ML) function.

Given the parameters estimated in steps 1 and 2 we use a recursive algorithm to estimate the Maximum Likelihood (ML) function. We calculate the eigenvalues associated with $\hat{\Psi}$, $\hat{\lambda}_i^2$ $i = 1, ..., \bar{s}$ and the matrix of residuals $\sum_{\bar{r},\ s=\bar{s}}^{\max}$. Hence, we compute the ML function:

$$L^0_{\max,\ \bar{r}<n,\ s=\bar{s}} = -\frac{T}{2}\left[\ln\left|\sum_{\bar{r}<n,\ s=\bar{s}}^{\max}\right| - \sum_{i=1}^{\bar{s}} \ln\left(1-\hat{\lambda}_i^2\right)\right] \tag{12}$$

If $\bar{r} = n$, we use instead of (12) the derived log-likelihood: $L_{\max,\ r=n,\ s=\bar{s}} = -\frac{T}{2}\ln\left|\sum_{\bar{r}=n,\ s=\bar{s}}^{\max}\right|$. The determinant of the covariance matrix for $\bar{r} = n$ cointegration vector is calculated by

$$\ln\left|\sum_{\bar{r}=n,\ s=\bar{s}}^{\max}\right| = \ln\left|m_{00} - m_{01}m_{11}^{-1}m_{10}\right| - \sum_{i=1}^{\bar{s}} \ln\left(1-\hat{\lambda}_i^2\right) \tag{13}$$

where m_{ij} refers to cross moment matrices obtained in multivariate least square regressions from Δy_t and X_{t-1} on y_{t-1}. In this case, estimation does not imply an iterative algorithm yet because the cointegrating space spans R^n.

*Step*4 : Reestimation of β.

We reestimate β to obtain a more appropriated value for the parameters. In order to reestimate β we use the program $CanCorr\left[\Delta y_t, y_{t-1} \mid \hat{\Psi} X_{t-1}\right]$ and thus using the new $\hat{\beta}$ we can repeat step 2 to reestimate $\tilde{\beta}_\perp$ and Ψ. Then, we can calculate the new value of the ML function in the step 3. Henceforth, we obtain $L^1_{\max,\ \ r=\bar{r},\ s=\bar{s}}$ for calculating $\Delta L = \left(L^1_{\max,\ r=\bar{r},\ s=\bar{s}} - L^0_{\max,\ r=\bar{r},\ s=\bar{s}}\right)$.

We repeat steps 1 to 4 to choose $\tilde{\beta}_\perp$and Ψ until convergence is reached (i.e., $\Delta L < 10^{-7}$). In the end, optimal parameters $\bar{p}$, $\bar{r}$ and $\bar{s}$ are obtained and it can be used for estimation and forecasting of a VECM with WF restrictions.

4 Monte-Carlo Design

The simple real business cycle models and also the simplest closed economy monetary dynamic stochastic general equilibrium models are three-dimensional. The consumption, saving

and output and the prices, output and money are notable examples. Motivated by these applications and according the previous work of Vahid and Issler (2001) we procedure to construct the Monte-Carlo experiment in a three-dimensional environment. Therefore, the data generating processes considering a VAR model with three variables, one cointegration vector, and two cofeatures vectors (i.e., $n = 3$, $r = 1$ and $s = 2$, respectively). β and $\tilde{\beta}$ satisfy:

$$\beta = \begin{bmatrix} 1.0 \\ 0.2 \\ -1.0 \end{bmatrix}, \tilde{\beta} = \begin{bmatrix} 1.0 & 0.1 \\ 0.0 & 1.0 \\ 0.5 & -0.5 \end{bmatrix}$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.0 & 0.6 & 0.6 \\ 0.6 & 1.0 & 0.6 \\ 0.6 & 0.6 & 1.0 \end{bmatrix} \right)$$

Consider the VAR(3) model: $y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + \varepsilon_t$. The VECM respresentation as a function of the VAR level parameters can be written as:

$$\Delta y_t = (A_1 + A_2 + A_3 - I_3) y_{t-1} - (A_2 + A_3) \Delta y_{t-1} - A_3 \Delta y_{t-2} + \varepsilon_t \tag{14}$$

The VAR coefficients must simultaneously obey the restrictions: a) The cointegration restrictions: $\alpha\beta' = (A_1 + A_2 + A_3 - I_3)$; b) WF restrictions: $\tilde{\beta}' A_3 = 0$ (iii) $\tilde{\beta}'(A_2 + A_3) = 0$ and c) covariance-stationary condition. Considering the cointegration restrictions we can rewrite (14) as the following VAR(1):

$$\xi_t = F \, \xi_{t-1} + v_t \tag{15}$$

$$\xi_t = \begin{bmatrix} \triangle y_t \\ \triangle y_{t-1} \\ \beta' y_t \end{bmatrix}, \; F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix} \text{ and } v_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \beta' \varepsilon_t \end{bmatrix}$$

Thus, the equation (15) will be covariance-stationary if all eigenvalues of matrix F lie inside the unit circle. An initial idea to design the Monte-Carlo experiment may consist of constructing the companion matrix (F) and verify whether the eigenvalues of the companion matrix all lie inside the unit circle. This may be carried out by selecting their values from a uniform distribution, 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. Hence, we follow an alternative procedure. We propose an analytical solution to generate a covariance-stationary VAR, based on the choice of the eigenvalues, and then on the generation of the respective companion matrix. In the appendix we present a detailed discussion of the final choice of these free parameters, including analytical solutions. In our simulation, we constructed 100 data generating processes and for each of these we generate 1000 samples containing 1000 observations. In order to reduce the impact of initial values, we consider only the last 100 and 200 observations. All the experiments were conducted in the MatLab environment.

5 Results

Figure 1 shows one realization of the three-dimensional VAR model with cointegration and WF restrictions.

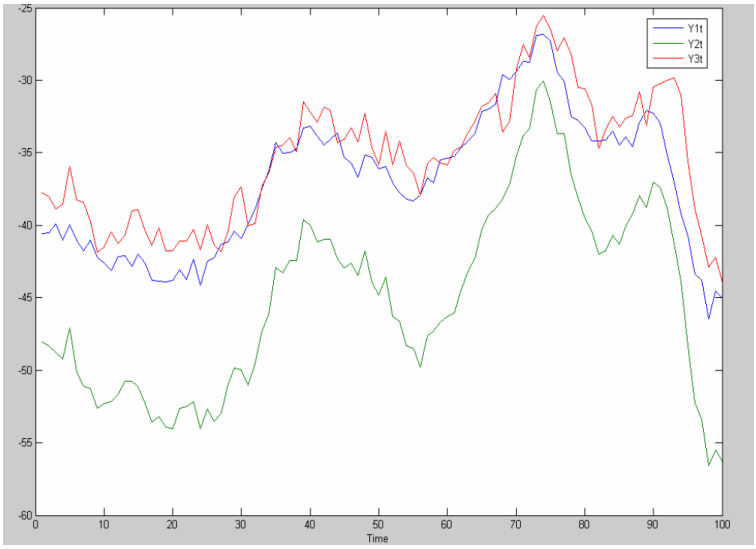


Figure 1. One realization of a VAR(3) model with $n = 3$, $r = 1$ and $s = 2$

Values in Table 1 represent the percentage of time that the model selection criterion, IC(p), chooses that cell corresponding to the lag and number of cointegration vectors in 100,000 realizations. The true lag-cointegrating vectors are identified by bold numbers and the selected lag-cointegration vectors chosen more times by the criterion are underlined. The

results show that, in general, the AIC criterion choose more frequently the correct lag length for 100 and 200 observations. For example, for 100 observations, the AIC, HQ and SC criteria chose the true lag, p, 54.08%, 35.62% and 17.49% of the times respectively. Note that all three criteria chose more frequently the correct rank of cointegration ($r = 1$). When 200 observations are considered, the correct lag length was chosen 74.72%, 57.75% and 35.28% of the time for AIC, HQ and SC respectively. Again all three criteria selected the true cointegrated rank $r = 1$. Tables 2 contains the percentage of time that the simultaneous model selection criterion, IC(p, s), chooses that cell, corresponding to the lag-rank and number of cointegrating vectors in 100,000 realizations. The true lag-rank-cointegration vectors are identified by bold numbers and the best lag-rank combination chosen more times by each criterion are underlined. The results show that, in general, the AIC criterion chooses more frequently lag-rank for 100 and 200 observations. For instance, for 100 observations, the AIC, HQ and SC criteria choose more frequently the true pair $(p, s) = (3, 1)$, 56.34%, 40.85% and 25.20% of the times respectively. For 200 observations, AIC, HQ and SC criteria choose more frequently the true pair $(p, s) = (3, 1)$, 77.07%, 62.58% and 45.03% of the times respectively. Note that all three criteria choose more frequently the correct rank of cointegration ($r = 1$) in both samples.

The most relevant results can be summarized as follows:

– All criteria (AIC, HQ and SC) choose the correct parameters more often when using IC(p, s).

– The AIC criterion has better performance in selecting the true model more frequently for both the IC(p, s) and the IC(p) criteria.

– For T=100 the SC(p, s) select 25.2% the true $p = 3$ while the SC(p) only select 17.4%. It represents gains more than 44%. For T=200 the gains are more than 27%.

– For T=100 the HQ(p, s) select 40.8% the true $p = 3$ while the HQ(p) only select 35.6%. It represents gains more than 14%. For T=200, the gains are more than 8%.

It is known that the literature suggests the use of the traditional SC and HQ criteria in VAR model selection. The results of this work indicate that if additional WF restrictions are

ignored, the standard SC and HQ criteria select few times the true value of p. That is, there is a cost of ignoring additional WF restrictions in the model specially when SC criterion is used. In general, the standard Schwarz or Hannan-Quinn selection criteria should not be used for this purpose in small samples due to the tendency of identifying an underparameterized model. In general, the use of these alternative criteria of selection, IC(p, s) has better performance than the usual criteria, IC(p), when the cointegrated VAR model has additional WF restriction.

6 Conclusions

In this work, we considered an additional *weak form* restriction of common cyclical features in a cointegrated VAR model in order to analyze the appropriate way for selecting the correct lag order. These additional WF restrictions are defined in the same way as cointegration restrictions, while cointegration refers to relations among variables in the long-run, the common cyclical restrictions refer to relations in the short-run. Two methodologies have been used for selecting lag length; the traditional information criterion, IC(p), and an alternative criterion (IC(p, s)) that selects simultaneously the lag order p and the rank structure s due to the WF restriction.

The results indicate that information criterion that selects the lag length and the rank order simultaneously has better performance than the model chosen by conventional criteria. When the WF restrictions are ignored there is a non trivial cost in selecting the true model with standard information criteria. In general, the standard Schwarz or Hannan-Quinn criteria selection criteria should not be used for this purpose in small samples due to the tendency of identifying an under-parameterized model.

In applied work, when the VAR model contains WF and cointegration restrictions, we suggest the use of AIC(p, s) criteria for simultaneously choosing the lag-rank, since it provides considerable gains in selecting the correct VAR model. Since no work in the literature has been dedicated to analyze a VAR model with WF common cyclical restrictions, the results of this work provide new insights and incentives to proceed with this kind of empirical work.

Appendix A. Results of the Monte Carlo simulation

Table I Performance of information criterion, IC(p), in selecting the lag order p

Frequency of lag(p) and cointegrating vectors (r) choice by different criteria for trivariate VAR model in levels when the true model have parameters: p = 3 and r = 1.

		Number of observations = 100 Selected cointegrated vectors				Number of observations = 200 Selected cointegrated vectors			
		0	**1**	2	3	0	1	2	3
	Selected Lag								
	1	0.000	0.996	0.359	0.031	0.000	0.095	0.016	0.003
	2	0.002	32.146	1.136	0.048	0.000	17.073	0.686	0.033
	3	2.792	**54.082**	0.902	0.041	0.012	**74.721**	1.488	0.108
AIC(p)	4	0.737	4.068	0.091	0.003	0.005	4.177	0.081	0.006
	5	0.392	0.987	0.031	0.000	0.013	0.828	0.020	0.000
	6	0.219	0.333	0.014	0.000	0.023	0.257	0.005	0.000
	7	0.166	0.173	0.006	0.000	0.039	0.133	0.002	0.000
	8	0.133	0.107	0.005	0.000	0.060	0.115	0.001	0.000
	1	0.000	3.884	1.915	0.165	0.000	1.098	0.243	0.021
	2	0.002	52.593	1.907	0.080	0.000	37.390	1.614	0.098
	3	2.600	**35.617**	0.612	0.027	0.012	**57.749**	1.146	0.082
HQ(p)	4	0.065	0.189	0.007	0.000	0.001	0.158	0.004	0.000
	5	0.059	0.037	0.000	0.000	0.009	0.082	0.001	0.000
	6	0.073	0.025	0.000	0.000	0.016	0.076	0.000	0.000
	7	0.059	0.019	0.001	0.000	0.030	0.070	0.000	0.000
	8	0.053	0.011	0.000	0.000	0.044	0.055	0.001	0.000
	1	0.000	8.344	6.609	0.511	0.000	3.964	1.385	0.093
	2	0.003	61.966	2.279	0.105	0.000	55.156	2.776	0.169
	3	2.042	**17.485**	0.313	0.015	0.012	**35.283**	0.728	0.044
SC(p)	4	0.049	0.045	0.000	0.000	0.001	0.083	0.002	0.000
	5	0.071	0.025	0.000	0.000	0.007	0.076	0.001	0.000
	6	0.057	0.016	0.000	0.000	0.013	0.063	0.000	0.000
	7	0.036	0.009	0.000	0.000	0.025	0.056	0.000	0.000
	8	0.017	0.003	0.000	0.000	0.027	0.035	0.001	0.000

Numbers represent the percentage times that the model selection criterion choice that cell corresponding to the lag and number of cointegration vectors in 100,000 realizations. The true lag-cointegrating vectors are indentified by bold numbers.

Table II. Performance of information criterion, IC(p,s), in selecting p and s simultaneously.

Johansen Tested coint. Vectors (r)		0			**1**			2			3		
	Selected rank (s)	1	2	3	**1**	2	3	1	2	3	1	2	3
	Selected lag (p)												
Sample size = 100													
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0.002	0.000	0.000	39.049	0.001	0.000	1.218	0.000	0.000	0.056	0.000	0.000
	3	0.301	0.000	0.000	**56.341**	0.003	0.000	1.559	0.000	0.000	0.053	0.000	0.000
AIC(p,s)	4	0.004	0.000	0.000	1.186	0.001	0.000	0.070	0.000	0.000	0.001	0.000	0.000
	5	0.000	0.000	0.000	0.114	0.001	0.000	0.012	0.000	0.000	0.000	0.000	0.000
	6	0.000	0.000	0.000	0.020	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
	7	0.000	0.000	0.000	0.006	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
	8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0.002	0.000	0.000	55.563	0.000	0.000	1.888	0.000	0.000	0.081	0.000	0.000
	3	0.267	0.000	0.000	**40.855**	0.000	0.000	1.207	0.000	0.000	0.043	0.000	0.000
HQ(p,s)	4	0.000	0.000	0.000	0.088	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.000
	5	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0.004	0.000	0.000	70.971	0.000	0.000	2.574	0.000	0.000	0.113	0.000	0.000
	3	0.221	0.000	0.000	**25.204**	0.000	0.000	0.887	0.000	0.000	0.025	0.000	0.000
SC(p,s)	4	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sample size = 200													
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0.000	0.000	0.000	18.797	0.000	0.000	0.681	0.000	0.000	0.038	0.000	0.000
	3	0.000	0.000	0.000	**77.065**	0.002	0.000	2.260	0.000	0.000	0.145	0.000	0.000
AIC(p,s)	4	0.000	0.000	0.000	0.908	0.000	0.000	0.035	0.000	0.000	0.001	0.000	0.000
	5	0.000	0.000	0.000	0.063	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000
	6	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0.000	0.000	0.000	33.952	0.000	0.000	1.370	0.000	0.000	0.086	0.000	0.000
	3	0.000	0.000	0.000	**62.576**	0.000	0.000	1.877	0.000	0.000	0.111	0.000	0.000
HQ(p,s)	4	0.000	0.000	0.000	0.027	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
	5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0.000	0.000	0.000	50.983	0.000	0.000	2.351	0.000	0.000	0.146	0.000	0.000
	3	0.000	0.000	0.000	**45.028**	0.000	0.000	1.416	0.000	0.000	0.076	0.000	0.000
SC(p,s)	4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	7	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Numbers represent the percentage times that the simultaneous model selection criterion IC(p,s) choice that cell, corresponding to the lag-rank and number of cointegrating vectors in 100,000 realizations. The true lag-rank-cointegration vectors are identified by bold numbers and the best lag-rank-cointegration vectors chosen by criterias are identified by underline lines.

Appendix B. VAR Restrictions for the DGPs

Let's consider the VAR(3) model :

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + \varepsilon_t \tag{16}$$

with parameters: $A_1 = \begin{bmatrix} a_{11}^1 & a_{12}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 & a_{22}^1 \\ a_{31}^1 & a_{32}^1 & a_{32}^1 \end{bmatrix}$, $A_2 = \begin{bmatrix} a_{11}^2 & a_{12}^2 & a_{12}^2 \\ a_{21}^2 & a_{22}^2 & a_{22}^2 \\ a_{31}^2 & a_{32}^2 & a_{32}^2 \end{bmatrix}$ and $A_3 = \begin{bmatrix} a_{11}^3 & a_{12}^3 & a_{12}^3 \\ a_{21}^3 & a_{22}^3 & a_{22}^3 \\ a_{31}^3 & a_{32}^3 & a_{32}^3 \end{bmatrix}$

We consider the cointegration vectors $\beta = \begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \beta_{31} \end{bmatrix}$, the cofeatures vectors $\tilde{\beta} = \begin{bmatrix} \tilde{\beta}_{11} & \tilde{\beta}_{12} \\ \tilde{\beta}_{21} & \tilde{\beta}_{22} \\ \tilde{\beta}_{31} & \tilde{\beta}_{32} \end{bmatrix}$ and the adjustament matrix $\alpha = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix}$. The long-run relation is defined by $\alpha\beta' = (A_1 + A_2 + A_3 - I_3)$. The VECM respresentation is:

$$\Delta y_t = \alpha\beta' y_{t-1} - (A_2 + A_3)\Delta y_{t-1} - A_3 \Delta y_{t-2} + \varepsilon_t \tag{17}$$

Considering the cointegration restrictions we can rewrite (17) as the following VAR(1)

$$\xi_t = F\ \xi_{t-1} + v_t \tag{18}$$

where $\xi_t = \begin{bmatrix} \triangle y_t \\ \triangle y_{t-1} \\ \beta' y_t \end{bmatrix}$, $F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix}$ and $v_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \beta'\varepsilon_t \end{bmatrix}$

1) Short-run restrictions (WF)

Let us, $G = -[R_{21}K + R_{31}]$, $K = [(R_{32} - R_{31})/(R_{21} - R_{22})]$, $R_{j1} = \tilde{\beta}_{j1}/\tilde{\beta}_{11}$, $R_{j2} = \tilde{\beta}_{j2}/\tilde{\beta}_{12}$ $(j = 2, 3)$ and $S = \beta_{11}G + \beta_{21}K + \beta_{31}$

(i) $\tilde{\beta}' A_3 = 0 ==>\ A_3 = \begin{bmatrix} -Ga_{31}^3 & -Ga_{32}^3 & -Ga_{33}^3 \\ -Ka_{31}^3 & -Ka_{32}^3 & -Ka_{33}^3 \\ -a_{31}^3 & -a_{32}^3 & -a_{33}^3 \end{bmatrix}$

(ii) $\tilde{\beta}'(A_2 + A_3) = 0 ==> \tilde{\beta}' A_2 = 0 ==>\ A_2 = \begin{bmatrix} -Ga_{31}^2 & -Ga_{32}^2 & -Ga_{33}^2 \\ -Ka_{31}^2 & -Ka_{32}^2 & -Ka_{33}^2 \\ -a_{31}^2 & -a_{32}^2 & -a_{33}^2 \end{bmatrix}$

2) Long-run restrictions (cointegration)

(iv) $\beta'(A_2 + A_3) = [-(a_{31}^2 + a_{31}^3)S \quad -(a_{32}^2 + a_{32}^3)S \quad -(a_{33}^2 + a_{33}^3)S]$ and $\beta' A_3 = [-a_{31}^3 S \quad -a_{32}^3 S \quad -a_{33}^3 S]$

$$(v)\ \beta'\alpha + 1 = \beta = \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \end{bmatrix} \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix} + 1 = \beta_{11}\alpha_{11} + \beta_{21}\alpha_{21} + \beta_{31}\alpha_{31} + 1$$

Therefore, considering short- and long-run restrictions, the companion matrix F is represented as:

$$F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix}$$

$$= \begin{bmatrix} -G(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -G(a_{33}^2 + a_{33}^3) & -Ga_{31}^3 & -Ga_{32}^3 & -Ga_{33}^3 & \alpha_{11} \\ -K(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -G(a_{33}^2 + a_{33}^3) & -Ka_{31}^3 & -Ka_{32}^3 & -Ka_{33}^3 & \alpha_{21} \\ -(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -(a_{33}^2 + a_{33}^3) & -a_{31}^3 & -a_{32}^3 & -a_{33}^3 & \alpha_{31} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -(a_{31}^2 + a_{31}^3)S & -(a_{32}^2 + a_{32}^3)S & -(a_{33}^2 + a_{33}^3)S & -a_{31}^3 S & -a_{32}^3 S & -a_{33}^3 S & b \end{bmatrix}$$

with $b = \beta'\alpha + 1 = \beta_{11}\alpha_{11} + \beta_{21}\alpha_{21} + \beta_{31}\alpha_{31} + 1$

3) Restrictions of covariance-stationary in equation (18)

The equation (18) will be covariance-stationary, all eigenvalues of matrix F lie inside the unit circle. Therefore, the eigenvalues of the matrix F is a number λ such that:

$$|F - \lambda I_7| = 0 \tag{19}$$

The solution of (19) is:

$$\lambda^7 + \Omega\lambda^6 + \Theta\lambda^5 + \Psi\lambda^4 = 0 \tag{20}$$

where the parameters Ω, Θ, and Ψ are: $\Omega = G(a_{31}^2 + a_{31}^3) + K(a_{32}^2 + a_{32}^3) + a_{33}^2 + a_{33}^3 - b$, $\Theta = Ga_{31}^3 + Ka_{32}^3 - (a_{33}^2 + a_{33}^3)b - Gb(a_{31}^2 + a_{31}^3) - Kb(a_{32}^2 + a_{32}^3) + \alpha_{31}S(a_{33}^2 + a_{33}^3) + S\alpha_{21}(a_{32}^2 + a_{32}^3) + S\alpha_{11}(a_{31}^2 + a_{31}^3) + a_{33}^3$ and $\Psi = -a_{33}^3 b - Ga_{31}^3 b - Ka_{32}^3 b + \alpha_{31}a_{33}^3 S + a_{32}^3 S\alpha_{21} + a_{31}^3 S\alpha_{11}$, and the first four roots are $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$.We calculated the parameters of matrices A_1, A_2 and A_3 as function of roots (λ_5, λ_6 and λ_7) and free parameters. Hence we have three roots satisfying equation (20)

$$\lambda^3 + \Omega\lambda^2 + \Theta\lambda + \Psi = 0 \tag{21}$$

for λ_5, we have: $\lambda_5^3 + \Omega\lambda_5^2 + \Theta\lambda_5 + \Psi = 0$ .....$Eq1$
for λ_6, we have: $\lambda_6^3 + \Omega\lambda_6^2 + \Theta\lambda_6 + \Psi = 0$ .....$Eq2$
for λ_7, we have: $\lambda_7^3 + \Omega\lambda_7^2 + \Theta\lambda_7 + \Psi = 0$ .....$Eq3$

Solving $Eq1, Eq2$ and $Eq3$ we have: $\Omega = -\lambda_7 - \lambda_6 - \lambda_5$, $\Theta = \lambda_6\lambda_7 + \lambda_6\lambda_5 + \lambda_5\lambda_7$ and $\Psi = -\lambda_5\lambda_6\lambda_7$. Equaling these parameters with relations above we have:

$a_{31}^2 = -(-Ka_{32}^2 - Ka_{32}^2 b + \alpha_{31}Sa_{33}^2 - \lambda^6\lambda^7 - \lambda^6 - \lambda^7 - a_{33}^2 b - \lambda^5\lambda^6\lambda^7 + b - \lambda^5\lambda^7 - \lambda^5\lambda^6 - a_{33}^2 + Sa_{32}^2\alpha_{21} - \lambda^5)/(S\alpha_{11} - G - Gb)$

$a_{32}^3 = (-S^2\lambda^7\alpha_{11}\alpha_{31} - b^2\lambda^7 G - \lambda^6 Gb^2 + b\lambda^7 S\alpha_{11} + \lambda^6 S\alpha_{11}b - a_{31}^3 S\alpha_{11}G + a_{31}^3 S^2\alpha_{11}^2 - Ga_{31}^3 bS\alpha_{11} - \lambda^5 Gb^2 + \lambda^5 S\alpha_{11}b - \lambda^7\lambda^6\alpha_{31}SG - \lambda^7\lambda^5\alpha_{31}SG - S^2\alpha_{11}\lambda^5\alpha_{31} - S^2\alpha_{11}\lambda^6\alpha_{31} + S\lambda^5 Gb\alpha_{31} + S\alpha_{31}\lambda^6 Gb - \lambda^5\lambda^7\lambda^6 G + \lambda^6\lambda^7 Gb + \lambda^5\lambda^7 Gb + \lambda^5\lambda^6 Gb - SGb^2\alpha_{31} + S^2\alpha_{11}b\alpha_{31} - S^2\alpha_{11}\alpha_{31}a_{33}^2 + S^2\alpha_{31}^2 a_{33}^2 G + SG^2 a_{31}^3\alpha_{31} + S\alpha_{11}a_{33}^2 b + Gb^3 - S\alpha_{11}b^2 - S^2\alpha_{11}Ka_{32}^2\alpha_{31} - S^2\alpha_{11}\alpha_{31}Ga_{31}^3 + S^2 a_{32}^2\alpha_{21}G\alpha_{31} - Sa_{32}^2\alpha_{21}Gb + S\alpha_{31}G^2 a_{31}^3 b - S\alpha_{31}a_{33}^2 Gb + S\alpha_{11}Ka_{32}^2 b + S\lambda^7 Gb\alpha_{31} - \lambda^5\lambda^6\alpha_{31}SG - \lambda^5\lambda^7\lambda^6\alpha_{31}SG + \lambda^5\lambda^7\lambda^6 S\alpha_{11})/(S\alpha_{11}K\alpha_{31} - KG\alpha_{31} + bG\alpha_{21} - K\alpha_{31}Gb - S\alpha_{11}\alpha_{21} + G\alpha_{21})/S$

$a_{33}^3 = -(Kb^3 G - \lambda^5 Gb^2 K + S\alpha_{11}\lambda^6 K\lambda^7\lambda^5 + Kb\lambda^7 S\alpha_{11} - Kb^2\lambda^7 G - S^2\alpha_{21}\lambda^7\alpha_{11} + \lambda^6 GbS\alpha_{21} + S\alpha_{21}\lambda^7 Gb - \lambda^6 Gb^2 K + \lambda^6 S\alpha_{11}Kb - \lambda^6 S^2\alpha_{11}\alpha_{21} + \lambda^5 GbS\alpha_{21} + \lambda^5 S\alpha_{11}Kb - \lambda^5 S^2\alpha_{11}\alpha_{21} - \lambda^7\lambda^6 S\alpha_{21}G + Kb\lambda^7\lambda^6 G + Kb\lambda^7\lambda^5 G + Kb\lambda^5\lambda^6 G - \lambda^7\lambda^6 KG\lambda^5 - S^2\alpha_{11}\alpha_{21}Ka_{32}^2 + S^2\alpha_{11}\alpha_{21}b - S^2\alpha_{11}\alpha_{21}a_{33}^2 + S^2\alpha_{21}^2 a_{32}^2 G - S\alpha_{11}Kb^2 + S\alpha_{21}G^2 a_{31}^3 - S\alpha_{21}Gb^2 + S^2 a_{31}^3 K\alpha_{11}^2 - S^2\alpha_{11}\alpha_{21}Ga_{31}^3 + S^2\alpha_{21}a_{33}^2 G\alpha_{31} + S\alpha_{11}K\hat{}2ba_{32}^2 + S\alpha_{11}Kba_{33}^2 - S\alpha_{11}a_{31}^3 KG - S\alpha_{11}KbGa_{31}^3 - SKba_{33}^2 G\alpha_{31} + S\alpha_{21}G^2 a_{31}^3 b - S\alpha_{21}\lambda^5\lambda^6 G - S\alpha_{21}\lambda^5\lambda^7\lambda^6 G - S\alpha_{21}Ka_{32}^2 Gb - S\alpha_{21}\lambda^7\lambda^5 G)/(S\alpha_{11}K\alpha_{31} - KG\alpha_{31} + bG\alpha_{21} - K\alpha_{31}Gb - S\alpha_{11}\alpha_{21} + G\alpha_{21})/S$

We can calculate a_{31}^2, $a_{32}^3$and a_{33}^3 fixing the set $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ and sort independently from uniform distributions $(-0.9; 0.9)$ the values of a_{31}^3, a_{32}^2, a_{33}^2, λ_5, λ_6 and λ_7.

Hencefore, each parameter of the matrices A_1, A_2 and A_3 are defined and so we can generate the DGPs of VAR(3) model with cointegration and WF restrictions.

References

[1] Lütkepohl H. Introduction to Multiple Time Series Analysis. Springer, New York, 1993.

[2] Braun PA, Mittnik S. Misspecifications in Vector Autoregressions and Their Effects on Impulse Responses and Variance Decompositions. *Journal of Econometrics*, 1993, 59, 319-41.

[3] Johansen, S. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, 1991, 59, 1551-1580.

[4] Gonzalo, J. Five Alternative Methods of Estimating Long-Run Equilibrium Relationships. *Journal of Econometrics*, 1994, 60 issue 1-2, 203-233

[5] Centoni M, Cubadda G, Hecq. Common Shocks, Commo Dynamics, and the International Business Cycle. *Economic-Modelling*, 24, 149-166.

[6] Engle R, Issler JV. Estimating Common Sectoral Cycles. *Journal of Monetary Economics*, 1995, 35**:**83-113.

[7] Caporale GM. Common Features and Output Fluctuations in the United Kingdom. *Economic-Modelling*, 1997, 14: 1-9.

[8] Mamingi N, Iyare SO. Convergence and Common Features in International Output: A Case Study of the Economic Community of West African States, 1975-1997. *Asian-African-Journal-of-Economics-and-Econometrics*, 2003, **3**: 1-16.

[9] Engle R, Kozicki S. Testing for common features. *Journal of Business and Economic Statistics*, 1993, 11**:** 369-395.

[10] Vahid F, Engle R. Common trends and common cycles. *Journal of Applied Econometrics*, 1993, 8: 341-360.

[11] Hecq A, Palm FC, Urbain JP. Testing for Common Cyclical Features in VAR Models with Cointegration, *Journal of Econometrics*, 2006, 132: 117-141.

[12] Engle R, Granger C. Cointegration and error correction: representation, estimation and testing. *Econometrica*, 1987, 55, 251-76.

[13] Kilian L. Impulse response analysis in vector autoregressions with unknown lag order. *Journal of Forecasting*, 2001, 20: 161-179.

[14] Vahid F, Issler JV. The Importance of Common Cyclical Features in VAR Analysis: A Monte Carlo Study. *Journal of Econometrics*, 2001, 109: 341-363.

[15] Guillén, Issler, Athanasopoulos. Forecasting Accuracy and Estimation Uncertainty using VAR Models with Short- and Long-Term Economic Restrictions: A Monte-Carlo Study. *Ensaios Econômicos EPGE*, 2005, 589.

[16] Hecq A. Cointegration and Common Cyclical Features in VAR Models: Comparing Small Sample Performances of the 2-Step and Iterative Approaches. Mimeo, 2006.

[17] Anderson TW. Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Ann. Math. Statist.* 1951, 22: 327-351. [Correction (1980) *Ann. Statist* 8 1400.]

[18] Izenman AJ. Reduced rank regression for the multivariate linear model, *Journal of. multivariate Analysis*, 1975, 5: 248-264.

[19] Box GE, Tiao GC. A canonical analysis of multiple time series. *Biometrika*, 1977, 64:355-365.

[20] Tso MK 1981. Reduced rank regression and Canonical analysis. *Journal of the Royal Statistical Society*, 1981, 43: 183-189.

[21] Velu RP, Reinsel GC, Wichern DW. Reduced rank models for multiple times series. *Biometrika,* 1986, 73: 105-118.

[22] Johansen S. Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford Univ. Press. 1995. Z .

Chapter 2

Evidence on Common Features and Business Cycle Synchronization in Mercosur[1]

Abstract

The aim of this work is to analyze the business cycles of the Mercosur's member countries in order to investigate their degree of synchronization. The econometric model uses the Beveridge-Nelson-Stock-Watson multivariate trend-cycle decomposition, taking into account the presence of common features such as common trend and common cycle. Once the business cycles are estimated, their degree of synchronization is analyzed by means of correlation in time domain and coherence and phase in frequency domain. Despite the evidence of common features, the results suggest that the business cycles are not synchronized. This may generate an enormous difficulty to intensify the agreements into Mercosur.

Key-words: Mercosur, business cycles, trend-cycle decomposition, common features, spectral analysis.

Jel Codes: C32, E32, F02, F23.

[1]This article was jointly made with Fábio Gomes. We would like to thank Wagner Piazza, João Issler, Luiz Renato, Cristiano Fernandes and Christiam Gonzales for helpful discussions. We are also grateful to Editor Sergio Firpo and two anonymous referees. In addition, many seminar participants at LAMES-LACEA – México (November, 2006), Encontro Brasileiro de Econometria XXVIII and XXXIV Encontro Nacional de Economia (ANPEC) Meetings for helpful discussions and coments.

1 Introduction

The design of economic blocks, such as the European Union and the Mercosur, has the purpose to amplify society welfare through the unification of economic policies and commercial agreements. According to Backus and Kehoe (1992) and Chistodoulakis and Dimelis (1995), the success of these policies depends on the similarities of the business cycles of the member states. A business cycle is a periodic but irregular up-and-down movement in economic activity, measured by fluctuations in real GDP and other macroeconomic variables. However, in compliance with Lucas (1977), many authors focus the analysis on GDP, defining business cycles as the difference between the actual GDP and its long-run trend.

The aim of this paper is to analyze the business cycles of the Mercosur member countries. The Mercosur or southern common market is a regional trade agreement created in 1991 by the treaty of Asunción. Its members are: Argentina, Brazil, Paraguay, Uruguay and Venezuela, welcomed as the fifth member in 2006. These countries differ in their institutions, economic policies and industrial structures, creating an enormous internal asymmetry in Mercosur (Flores, 2005). Although the block was created in 1991, we will analyze a broader period, from 1951 to 2003. Therefore, if we find evidence in favor of similarity we can safely assume that it cannot be attributed only to Mercosur[2]. In fact, an inverse causality is investigated: if the similarities among the countries lead to commercial integration.

In the empirical literature, there is no consensus about how to estimate the trend-cycle components of economic time series and how to analyze the so-called co-movements[3] in their business cycles. In the past decades a rich debate on the abilities of different statistical methods to decompose time series in long-run and short-run fluctuations has taken place (Baxter and King, 1995; Guay and St-Amant, 1996). The Hodrick-Prescott (HP) filter and the linear detrending are the usual univariate methodologies applied. However, these methodologies do not take in account the existence of common features among the economic series. In addition to that, as shown by Harvey and Jaeger (1993), the HP filter can induce spurious cyclicality when applied to integrated data. Therefore, in order to obtain a measure of the business cycles, we employ the Beveridge-Nelson-Stock-Watson (BNSW) multivariate trend-cycle decomposition (Beveridge and Nelson, 1981), considering the occurrence of cointegration and serial correlation common feature among the variables.

[2] Besides, there is not a consensus that Mercosur led to an increase in the flow of commerce among its integrated parts.

[3] Two countries present comovements when their real GDP expansions and downturns are simultaneous.

In order to investigate the degree of synchronization or co-movement of their business cycles an extra effort is necessary. Many authors have used the linear correlations between cycles; however, this analysis gives a static measure of the co-movements since it is not a simultaneous analysis of the persistence of co-movement (Engle and Kozick, 1993). To avoid this critique, the measures of coherence and phase in frequency domain are employed in order to investigate how synchronized the business cycles are (Wang, 2003). These frequency domain techniques constitute a straightforward way to represent economic cycles, because they provide information for all frequencies.

Finally, the results indicate the existence of common trends and common cycles among the economies studied. Thus, we confirm the need to use a multivariate approach, which is our first contribution. Time domain analysis found synchronization in two sub-groups: Paraguay-Uruguay and Argentina-Brazil. However, frequency domain findings did not corroborate these results. Thus, in general, the countries of the Mercosur are not synchronized.

Besides this introduction, the paper is organized as following. Section 2 presents the econometric methodology. Section 3 reports the econometric results while the section 4 analyzes the degree of synchronization of the business cycles. Finally, the conclusions are summarized in the last section.

2 Econometric Model

Common features may be seen as restrictions over the dynamics of the countries and, consequently, over the dynamics of their business cycles. While cointegration refers to long-run relationships, common cyclical restrictions refer to short-run dynamics. Engle and Kozicki (1993) and Vahid and Engle (1993) proposed the serial correlation common feature (SCCF) as a measure of common cyclical feature in the short-run, which is applied in many empirical works. For example, Gouriéroux and Peaucelle (1993) analyzed some issues on purchase power parity; Campbell and Mankiw (1990) found a common cycle between consumption and income for most G-7 countries; Engle and Kozicki (1993) found common international cycles in GNP data for OECD countries; Engle and Issler (2001) found common cycles among sectoral output for US; Candelon and Hecq (2000) tested the Okun's law.

To implement the BNSW decomposition, taking into account the common features restrictions, a VAR model is estimated and the existence of long-run and short-run common

dynamics is tested. Consider a Gaussian Vector Autoregression of finite order p, VAR(p):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \tag{1}$$

where y_t is a vector of n first order integrated series, $I(1)$, and ϕ_i, $i = 1, \ldots, p$ are matrices of dimension $n \times n$ and $\varepsilon_t \sim Normal(0, \Omega)$, $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_\tau) = \{\Omega, \ \ se\ t = \tau$ and $0_{n \times n}, \ \ se\ t \neq \tau$; where Ω is no singular$\}$. The model (1) can be written equivalently as:

$$\Pi(L)\ y_t = \varepsilon_t \tag{2}$$

where $\Pi(L) = I_n - \sum_{i=1}^{p} \phi_i L^i$ and L represents the lag operator. Besides, $\Pi(1) = I_n - \sum_{i=1}^{p} \phi_i$ when $L = 1$.

2.1 Long run restrictions (Cointegration)

The following hypotheses are assumed:

Assumption 1 : The $(n \times n)$ matrix $\Pi(\cdot)$ satisfies:

1. Rank $(\Pi(1)) = r$, $0 < r < n$, *such that* $\Pi(1)$ *can be expressed as* $\Pi(1) = -\alpha\beta'$, where α *and* β *are* $(n \times r)$ *matrices with full column rank* r.

2. The characteristic equation $|\Pi(L)| = 0$ *has* $n - r$ *roots equal to* 1 *and all other are outside the unit circle.*

Assumption 1 implies that y_t is cointegrated of order $(1, 1)$. The elements of α are the adjustment coefficients and the columns of β span the cointegration space. Decompounding the polynomial matrix $\Pi(L) = \Pi(1) L + \Pi^*(L) \Delta$, where $\Delta \equiv (1 - L)$ is the difference operator, a Vector Error Correction (VEC) model is obtained:

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t \tag{3}$$

where $\alpha\beta' = -\Pi(1), \Gamma_j = -\sum_{k=j+1}^{p} \phi_k \ \ (j = 1, ..., p-1)$ and $\Gamma_0 = I_n$.

2.2 Common cycles restrictions

The VAR(p) model can have short-run restrictions as shown by Vahid and Engel (1993).

Definition 2 *Serial Correlation Common Feature holds in (3) if there is a* $(n \times s)$ *matrix* $\tilde{\beta}$ *of rank* s*, whose columns span the cofeature space, such as* $\tilde{\beta}'\Delta y_t = \tilde{\beta}'\varepsilon_t$*, where* $\tilde{\beta}'\varepsilon_t$ *is a* s*-dimensional vector that constitute an innovation process with respect to all information prior to period* t*.*

Consequently, the SCCF restrictions occur if there is a cofeature matrix $\tilde{\beta}$ that satisfies the following assumption:

Assumption 2 : $\tilde{\beta}'\,\Gamma_j = 0_{s\times n} \qquad j = 1, ..., p-1$

Assumption 3 : $\tilde{\beta}'\,\alpha\beta' = 0_{s\times n}$

2.3 Trend-Cycle decomposition

The BNSW trend-cycle decomposition can be introduced by means of the Wold representation of the stationary vector Δy_t given by:

$$\Delta y_t = C(L)\varepsilon_t \tag{4}$$

where $C(L) = \sum_{i=0}^{\infty} C_i L^i$ is polynomial matrix in the lag operator, $C_0 = I_n$ and $\sum_{i=1}^{\infty} j\ |C_j| < \infty$. Using the following polynomial factorization $C(L) = C(1) + \Delta C^*(L)$, it is possible to decompose Δy_t such that:

$$\Delta y_t = C(1)\,\varepsilon_t + \Delta C^*(L)\,\varepsilon_t \tag{5}$$

where $C_i^* = \sum_{j>i}^{\infty}(-C_j)$, $i \geq 0$, and $C_0^* = I_n - C(1)$. Ignoring the initial value y_0 and integrating both sides of (5), we obtain:

$$y_t = C(1)\sum_{j=1}^{T}\varepsilon_t + C^*(L)\varepsilon_t \ = \tau_t + c_t \tag{6}$$

Equation (6) represents the BNSW decomposition where y_t is decomposed in "n" random walk process named "stochastic trend" and "n" stationary process named "cycles". Thus,

$\tau_t = C(1)\sum_{j=1}^{T}\varepsilon_t$ and $c_t = C^*(L)\varepsilon_t$ represent trend and cycle components, respectively. Assuming that long-run restrictions exist, then r cointegration vectors exist ($r < n$). These vectors eliminate the trend component which implies that $\beta' C(1) = 0$. Thus, $C(1)$ has dimension $n - r$, which means that there are $n - r$ common trends. Analogously, assuming short-run restrictions, there are s cofeature vectors that eliminate the cycles, $\tilde{\beta}' C^*(L) = 0$, which implies that $C^*(L)$ has dimension $n - s$, which is the number of common cycles. It is worth noting that $r + s \leq n$ and the cointegration and cofeatures vectors are linearly independent (Vahid and Engle, 1993). In order to obtain the common trends, it is necessary (and sufficient) to multiply equation (6) by $\tilde{\beta}'$, such that

$$\tilde{\beta}' y_t = \tilde{\beta}' C(1) \sum_{j=1}^{T} \varepsilon_t \ = \tilde{\beta}' \tau_c$$

This linear combination does not contain cycles because the cofeatures vectors eliminate them. In the same way, to get the common cycles it is necessary to multiply equation (6) by β', and so

$$\beta' y_t = \beta' C^*(L)\varepsilon_t \ = \beta' c_t$$

This linear combination does not contain the stochastic trend because the cointegration vectors eliminate the trend component. A special case emerges when $r + s = n$. In this case, it is extremely easy to estimate the trend and cycle components of y_t. As $\tilde{\beta}'$ and β' are linearly independent matrices, it is possible to build a matrix A, such as $A_{n\times n} = (\tilde{\beta}', \beta')'$ has full rank and, therefore, is invertible. Notice that, the inverse matrix can be partitioned as $A^{-1} = (\tilde{\beta}^- \ \ \beta^-)$ and the trend and cycle components can be obtained as follows:

$$y_t = A^{-1} A y_t = \tilde{\beta}^- \left(\tilde{\beta}' y_t\right) + \beta^- (\beta' y_t) = \tau_t + c_t \tag{7}$$

This implies that $\tau_t = \tilde{\beta}^- \tilde{\beta}' y_t$ and $c_t = \beta^- \beta' y_t$. Therefore, trend and cycle are linear combinations of y_t. Note that τ_t is generated by a linear combination of y_t using the cofeature vectors, containing the long-run component (because $\tilde{\beta}' y_t$ is a random walk component). On the other hand, c_t is generated by a linear combination of y_t using the cointegration vectors, containing the short-run component (because $\beta' y_t$ is $I(0)$ and serially correlated).

2.4 Estimation and testing

Considering the SCCF and the cointegration restrictions, we can rewrite the vector error correction as a model of reduced-rank structure. In (3) we define a vector $X_{t-1} =$

$[y_{t-1}\beta', \Delta y'_{t-1}, ...\Delta y'_{t-p+1}]'$ of dimension $(n(p-1)+r)\times 1$ and a $n\times(n(p-1)+r)$ matrix $\Phi = [\alpha,\ \Gamma_1, ..., \Gamma_{p-1}]$. Therefore (3) is written as:

$$\Delta y_t = \Phi X_{t-1} + \varepsilon_t \tag{8}$$

If assumptions (1), (2) and (3) hold, then the matrices $\Gamma_i, i = 1, ..., p-1$ are all of reduced rank $(n-s)$ and they can be written as $\Phi = A[\Psi_0, \Psi_1, ..., \Psi_{p-1}] = A\Psi$, where A is $n\times(n-s)$ full column rank matrix and Ψ has dimension $(n-s)\times(n\,(p-1)+r)$ and $\tilde{\beta}'A\Psi = 0$, that is, $\tilde{\beta}\in sp(A_\perp)$ where $A_\perp$ is the orthogonal complement of A. Therefore, let $A = \tilde{\beta}_\perp$.[4] Hence the model (8) can be expressed as a dynamic factor model with $n-s$ factor, given by ΨX_{t-1}, which are linear combinations of the right hand side variables in (3).

$$\Delta y_t = \tilde{\beta}_\perp\,(\Psi_0, \Psi_1, ..., \Psi_{p-1})\,X_{t-1} + \varepsilon_t \tag{9}$$

$$= \tilde{\beta}_\perp \Psi X_{t-1} + \varepsilon_t \tag{10}$$

To estimate the coefficient matrices $\tilde{\beta}_\perp$ and Ψ in the reduced rank model (10) we use the Anderson's (1951) procedure (see additionally Anderson, 1988, Johansen, 1995). This procedure is based on canonical analysis, which is a special case of a reduced-rank regression. More specifically, the maximum-likelihood estimation of the parameters of the reduced-rank regression model may result a problem of canonical analysis[5]. Therefore, we can use the expression $CanCorr\{X_t, Z_t|W_t\}$ that denotes the partial canonical correlations between X_t and Z_t: both sets concentrate out the effect of W_t that allows us to obtain canonical correlation, represented by the eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3..... > \hat{\lambda}_n$.

By the way, the Johansen cointegration test statistic is also based on canonical correlation. In model (3) we can use the expression $CanCorr\{\Delta y_t, y_{t-1}|W_t\}$ where $W_t = [\Delta y_{t-1}, \Delta y_{t-2}, ..., \Delta y_{t+p-1}]$ that summarizes the reduced-rank regression procedure used in the Johansen approach. It means that one extracts the canonical correlations between Δy_t and y_{t-1}: both sets concentrated out the effect of lags of W_t.

Moreover, we could also use a canonical correlation approach to determine the rank of the common features space due to SCCF restrictions. It is a test for the existence of cofeatures in the form of linear combinations of the variables in first differences, which are white noise

[4] The orthogonal complement of the $n\times s$ matrix B, $n > s$ and $\text{rank}(B) = s$, is the $n\times(n-s)$ matrix $B_\perp$ such that $B'_\perp B = 0$ and $\text{rank}(B : B_\perp) = n$. Hence, $B_\perp$ spans the null space of B and B' spans the left null space of $B_\perp$. The space is denoted by sp.

[5] This estimation is referred as Full Information Maximum Likelihood - FIML.

(i.e., $\tilde{\beta}'\Delta y_t = \tilde{\beta}'\varepsilon_t$ where $\tilde{\beta}'\varepsilon_t$ is a white noise). Based on Tiao and Tsay (1985), Vahid and Engle (1993) proposed a sequential test for SCCF, assuming that the rank of β is known. The sequence of hypotheses to be tested are: $H_0 : rank\left(\tilde{\beta}\right) \geq s$ against $H_a : rank\left(\tilde{\beta}\right) < s$, (see Lütkepohl, 1993; Velu et al, 1986) starting with $s = 1$ against the alternative model with $s = 0$ (there is no common cycle). If the null hypotheses is not rejected, we implement the test for $s = 2$, and so on.

In the VEC model the significance of the s smallest eigenvalues is determined through the following statistic:

$$\xi_s = -T\sum_{i=1}^{s} Ln\left(1-\lambda_i^2\right) \sim \chi^2_{(v)}, \qquad s = 1, ..., n-r \tag{11}$$

$\lambda_1 < \lambda_2 ....., < \lambda_{n-r} < 1$, with $v = s\left[n\left(p-1\right)+r\right] - s(n-s)$ degrees of freedom, where n is the dimension of the system and p the lag order of the VAR model.[6] Suppose that the statistical test (11) has found s independent linear combinations of the elements of Δy_t unpredictable. This implies that there is an $n \times s$ matrix $\tilde{\beta}$ of full rank s with s eigenvectors associated with the s smallest eigenvalues. Reinsel and Ahn (1992) propose a correction in statistic (11) in small samples $\xi_s^{corr} = \frac{T-n(p-1)-r}{T}\xi_s$, where T is the real number of observations after the deduction of initial points in regressions containing lags.

3 Empirical results

3.1 Database

The database used was extracted from *Penn World Table*, corresponding to Real GDP *per capita* series of Mercosur countries.[7] The frequency is annual, ranging from 1951 to 2003.[8] We consider the model $Y_t = T_t\ C_t$, where C_t is the cycle and T_t the trend of the series.

[6] For $p = 1$ the degrees of freedom is $(r+s)^2$. Notice in the model $\Delta y_t = \alpha\beta' y_{t-1} + \varepsilon_t$, the $rank(\alpha\beta') = \tilde{r} = n-s-r$, hence $\nu = (n-\tilde{r}) \times (np-\tilde{r}) = (n-(n-s-r))^2 = (r+s)^2$.

[7] Heston, Robert Summers and Bettina Aten, Penn World Table Version 6.2, Center for International Comparisons at the University of Pennsylvania (CICUP). Real GDP per capita (Constant Prices: Chain series) *http : //pwt.econ.upenn.edu/php_site/pwt_index.php*

[8] Penn World Table Version 6.2, Center for International Comparisons at the University of Pennsylvania (CICUP). Version 6.2 contains data from 1950 to 2004. However, some countries present missing data, like Brazil in 2004 and Paraguay in 1950 and 2004. http://pwt.econ.upenn.edu/php_site/pwt62/pwt62_form.php

Define $y_t \equiv \log Y_t$, $\tau_t \equiv \log T_t$ and $c_t \equiv \log C_t$. Then, $y_t = \tau_t + c_t$. The Figure 1 reports the GDP expressed in log terms. After 1975, in general, the series become closer - a behavior that may be generated by a common trend. Figure 2 the growth rates of real gross domestic product, i.e., $\ln(Y_t/Y_{t-1})$. It is possible to see the recession in Argentina, in 1989-1990.

Figure 1. Real GDP (in log) *per capita* series of Mercosur countries (1951-2003)

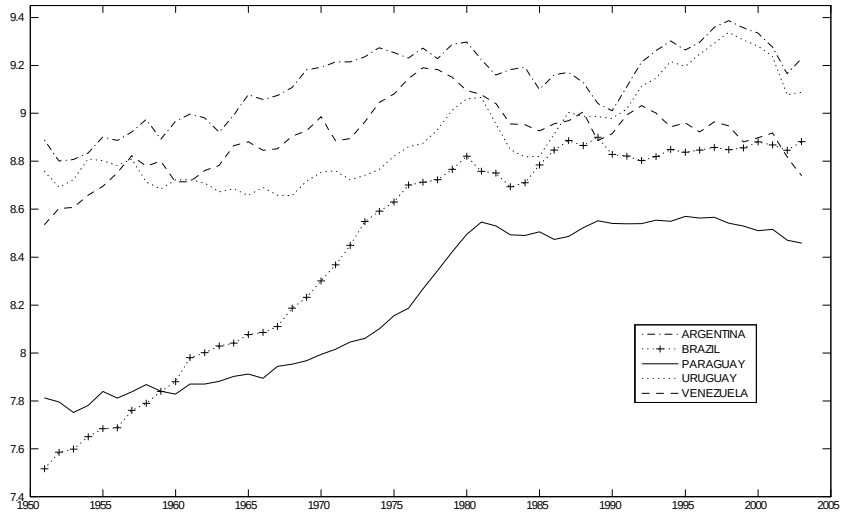


Figure 2. The growth rates of the real GDP *per capita* series of Mercosur countries (1951-2003)

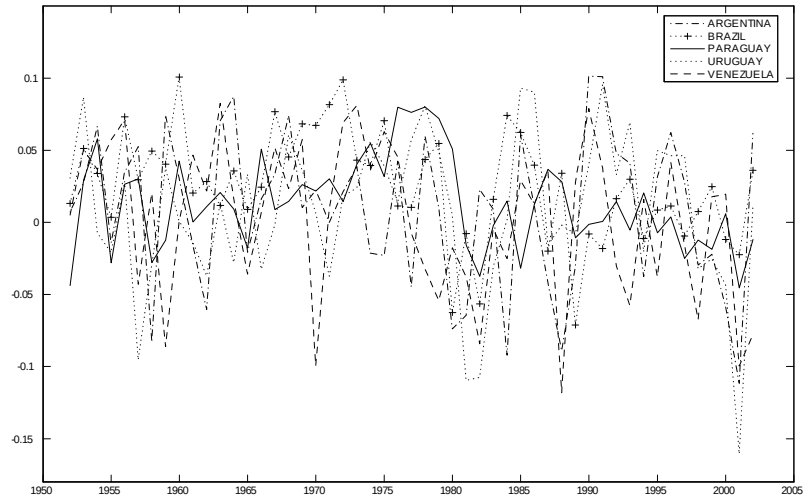


Table 1 displays the descriptive statistics of the real GDP growth rates. In general, the average growth rate is very lower; the exception is Brazil (2.63%). Indeed, the Figure 1 showed that Brazil had the lowest income level in 1951 and becomes an intermediated country in 2003. In Figure 1 also shows a kind of convergence of Paraguay to rich countries. Indeed, Paraguay has the second largest growth rate (1.24%). The other countries are below the 1% rate. While Brazil and Paraguay show an up-award trend, the other countries oscillated around a similar level and the standard deviation reflects these behaviors. Argentina, Uruguay and Venezuela are more volatile than Brazil and Paraguay. All countries experienced years of high growth, some of them above 10%, like Argentina and Brazil; but, episodes of sharp decreases are also present.

Table 1 - Descriptive Statistics

	Argentina	Brazil	Paraguay	Uruguay	Venezuela
Mean	0.65%	2.63%	1.24%	0.64%	0.39%
Standart Dev.	5.40%	3.79%	3.16%	5.38%	5.23%
Maximum	10.15%	10.08%	8.00%	9.56%	8.25%
Minimum	-11.16%	-7.12%	-4.55%	-16.05%	-11.82%

3.2 Common Features results

To implement the methodology previously stated, a hyerarquical procedure is followed to estimate the parameter of the model (see, Vahid and Engle,1993). First, the VAR order, p, is estimated via information criteria: Akaike (AIC), Schwarz (SC) and Hannan-Quinn (HQ) (Lütkepohl, 1993). After that, we identify the number of long-run restrictions, r, through Johansen cointegration test. Then the number of short-run restrictions due to SCCF, s, is estimated using χ^2 test. Finally, the matricial parameters are estimated in model (3) using the FIML procedure (Vahid and Issler (1993)).

Since BNSW decomposition assumes that the series are $I(1)$, we begin the analysis using the augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and DF-GLS unit root tests. In addition, we apply the KPSS procedure, which differently from previous tests, has a

stationary null hipothesis. The results for all countries are reported in Table 2.[9] The ADF, PP and DF-GLS tests do not reject the unit root null hypothesis, at 5% level of significance, for all countries. At 5% level, the KPSS do not reject the stationarity null hypothesis only for Uruguay. Even in this case, at 10% level, the null hypothesis is rejected. After all, the results suggests that series are I(1).

Table 2. Statistics of Unit Root Tests

Country	ADF	PP	DF-GLS	KPSS
Argentina	-1.8691	-1.9276	-1.9447	0.1543**
Brazil	-0.2404	-0.4308	-0.5998	0.2411***
Paraguay	-0.5757	-0.7392	-1.0805	0.1475**
Uruguay	-2.6644	-2.0443	-2.5328	0.1433*
Venezuela	-1.0972	-1.0780	-0.8094	0.2318***

Note: *, **, *** means rejection at 10%, 5% and 1% level of significance.

To estimate the order of the VAR, the AIC, HQ and SC information criteria are used. Table 3 shows the results for $p \in \{1, 2, 3, 4, 5\}$. As the data are annual we consider that an upper bound of 5 lags is sufficient. We observe that the three criteria suggest $p = 1$, indicating a VAR(1) model. Although the p selected by the criteria was one, to check the robustness of the results, we additionally test the model for $p = 2$ and $p = 3$.

Table 3. Identification of the VAR order

	Lag				
Criteria	1	2	3	4	5
AIC	-18.4272*	-18.3237	-18.2791	-18.3222	-18.2962
SC	-17.2577*	-16.1797	-15.1605	-14.2289	-13.2284
HQ	-17.9852*	-17.5135	-17.1006	-16.7753	-16.3811

Note: * indicates the lag suggested by information criteria

[9] In the case of ADF and DF-GLS tests, the choice of lags of the dependent variable in the right side of the test equation is based on the Schwarz criterion. In the PP and KPSS tests we use the nucleus of Bartlett and the window of Newey-West. All test equations have a constant and a linear trend. In any case, the results are robust to exclude of the linear trend.

Considering $p = 1, 2, 3$ the usual diagnostic tests are applied in order to verify if these specifications are suitable. For $p = 1$ and $p = 2$ the LM test does not indicate the presence of serial autocorrelation in the residuals, at 5% level of significance[10]. On the other hand, for $p = 3$ the opposite result is obtained. The White heteroskedasticity test (without cross terms) does not find evidence of heteroskedasticity, at 5% level of significance, for $p = 1, 2, 3$. The Jarque-Bera normality test does not reject the null hypothesis of normal distribution of residuals only for $p = 1$, at 5% level of significance[11]. Consequently, the best specification is obtained when $p = 1$.

To test if the series are cointegrated, the Johansen's (1988) procedure is used. We introduced a constant in the cointegration equation. In Table 4 the results for the cointegration test are shown. The trace and the maximum eigenvalue test indicate $r = 2$ for $p = 1, 2$ while for $r = 1$ for $p = 3$. Even though, we use $r = 2$ for $p = 3$ to check robustness of the subsequent analysis.

Table 5 shows the SCCF test for $p = 1, 2, 3$ using the correction given by Reinsel and Ahn (1992). For $p = 1$ the test indicates that $s = 4$, at 5% level of significance, but as the p-value is close to 5% we may assume $s = 3$ without trouble (see Table 5 (a)). For $p = 2, 3$ the test indicates $s = 3$ (see Table 5 (b) e (c)). Therefore, in all cases $s + r = n$. These results confirm the necessity to use a multivariate approach to identify the business cycles. In the next section we analyze the economic cycles obtained from the BNSW decomposition, considering the common cycles and the common trend restrictions. Once $s + r = n$, it is possible to find the trend and cycle components as shown above. Figure III shows the common cycles for each value of p. We observe that for $p = 1, 2$ common cycles are very similar.

[10] The null hypothesis of the LM test is the absence of serial correlation until the lag h. We consider h from 1 to 5.

[11] The normality test uses the orthogonalization of Cholesky.

Table 4. Johansen's cointegration test

a) Johansen cointegration test for $p = 1$

		Trace Test		Maximum Eigenvalue Test		
Null hypothesis	Statistic	Critical value	p-value	Statistic	Critical value	p-value
$r = 0$	103.2097*	69.81889	0.0000	52.74712*	33.87687	0.0001
$r \leq 1$	50.46258*	47.85613	0.0279	33.30722*	27.58434	0.0082
$r \leq 2$	17.15536	29.79707	0.6286	11.76805	21.13162	0.5707
$r \leq 3$	5.387311	15.49471	0.7664	4.090286	14.26460	0.8497
$r \leq 4$	1.297025	3.841466	0.2548	1.297025	3.841466	0.2548

Note: *indicates rejection of null hypothesis, at 5% level of significance

b) Johansen cointegration test for $p = 2$

		Trace Test		Maximum Eigenvalue Test		
Null hypothesis	Statistic	Critical value	p-value	Statistic	Critical value	p-value
$r = 0$	95.68994*	69.81889	0.0001	38.13180*	33.87687	0.0146
$r \leq 1$	57.55814*	47.85613	0.0047	36.27881*	27.58434	0.0030
$r \leq 2$	21.27933	29.79707	0.3404	11.42115	21.13162	0.6053
$r \leq 3$	9.858185	15.49471	0.2918	7.026434	14.26460	0.4860
$r \leq 4$	2.831751	3.841466	0.0924	2.831751	3.841466	0.0924

Note: *indicates rejection of null hypothesis, at 5% level of significance

c) Johansen cointegration test for $p = 3$

		Trace Test		Maximum Eigenvalue Test		
Null hypothesis	Statistic	Critical value	p-value	Statistic	Critical value	p-value
$r = 0$	96.33886*	69.81889	0.0001	49.69316*	33.87687	0.0003
$r \leq 1$	46.64570	47.85613	0.0647	24.17728	27.58434	0.1287
$r \leq 2$	22.46842	29.79707	0.2732	13.51631	21.13162	0.4059
$r \leq 3$	8.952108	15.49471	0.3698	7.064437	14.26460	0.4816
$r \leq 4$	1.887672	3.841466	0.1695	1.887672	3.841466	0.1695

Note: *indicates rejection of null hypothesis, at 5% level of significance

Table 5. Common cycle test

a) $r = 2$, $n = 5$, $p = 1$ (constant)

Null hypothesis	λ^2	$\xi_{(p,s)}$	$[r+s]^2$	p-value
$s > 0$	0.0246	1.2971	9	0.9984
$s > 1$	0.0756	5.3875	16	0.9935
$s > 2$	0.2025	17.1553	25	0.8761
$s > 3$	0.4730	50.4638	36	0.0554
$s > 4$*	0.6373	103.2064	49	0.0000

Note: *indicates rejection of null hypothesis, at 5% level of significance

b) $r = 2$, $n = 5$, $p = 2$ (constant)

Null hypothesis	λ^2	$\xi^{corr}_{(p,s)}$	$s\,[n(p-1)+r]+s^2-sn$	p-value
$s > 0$	0.0059	0.2606	3	0.9673
$s > 1$	0.1215	5.9605	8	0.6517
$s > 2$	0.1856	14.9927	15	0.4519
$s > 3$*	0.5996	55.2643	24	0.0003
$s > 4$*	0.6781	105.1426	35	0.0000

Note: *indicates rejection of null hypothesis, at 5% level of significance

c) $r = 2$, $n = 5$, $p = 3$ (constant)

Null hypothesis	λ^2	$\xi^{corr}_{(p,s)}$	$s\,[n(p-1)+r]+s^2-sn$	p-value
$s > 0$	0.0644	2.5316	8	0.9602
$s > 1$	0.2441	13.1665	18	0.7816
$s > 2$	0.4422	35.3461	30	0.2303
$s > 3$*	0.6317	73.3036	44	0.0036
$s > 4$*	0.7392	124.3791	60	0.0000

Note: *indicates rejection of null hypothesis, at 5% level of significance

Figure 3- Common Cycles.

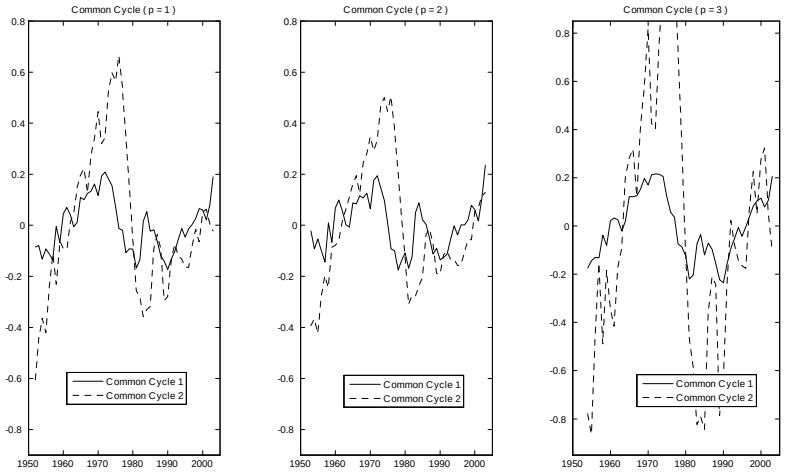


Figure 4 Cyclical components for $p = 1$, $s = 3$ and $r = 2$.

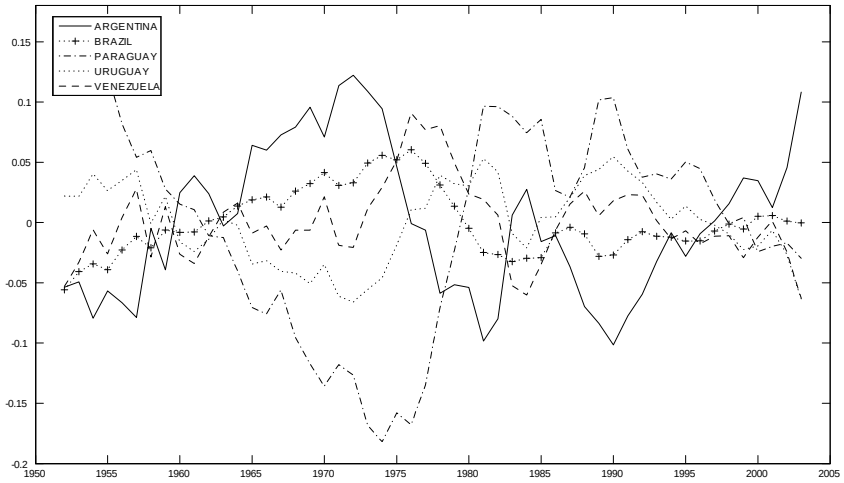
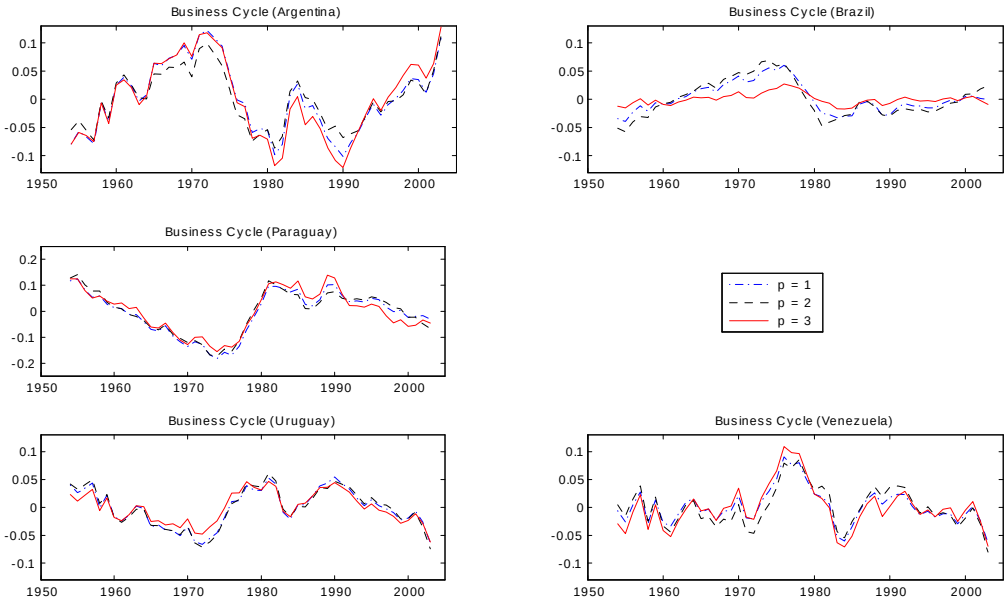


Figure 4 shows the business cycle components for our best specification: $p = 1$, $s = 3$ and $r = 2$. We notice an enormous contraction in Argentina in 1990's, as expected. As for Brazil, the period of the economic miracle is apparent. To analyze the robustness of the results we

estimate business cycles for each country for $p = 1, 2, 3$. Figure V shows the business cycle for each country. It is possible to see that the business cycles obtained from different p are similar.

Figure 5. Cyclical components in each country for $p = 1$, $p = 2$ and $p = 3$.



4 Business cycle analysis

The degree of association among the contemporaneous movements may be measured through the pairwise linear correlation as reported in Table 6 for $p = 1, 2, 3$. We can observe for $p = 1$ that Paraguay and Uruguay have high positive correlation. The same occurs for Brazil and Argentina. So far, based on correlation analysis there are two pairs of countries with similar patterns. The correlations of each country with the cycles 1 and 2 explain these results. Paraguay and Uruguay have a negative correlation with both cycles, while Brazil and Argentina are negatively related to both cycles. Not surprisingly, Venezuela has

a different behavior, being negatively correlated with the first cycle and positively with the second cycle.

Despite the fact that $p = 1$ is our best specification, the association between Paraguay and Uruguay remains high for $p = 2, 3$. However, the association between Brazil and Argentina plunged for $p = 3$. The correlations with the common cycles are robust in the following sense: Paraguay and Uruguay has a negative correlation with both cycles for $p = 1, 2, 3$ while Brazil and Argentina are negatively related to both cycles for $p = 1, 2, 3$. Venezuela results are more sensitive to p. Indeed, for $p = 3$, its correlation with the first common cycle becomes positive, although close to zero.

Table 6. Linear correlations in business cycles and in common cycle

Countries	Argentina	Brazil	Paraguay	Uruguay	Venezuela	C. Cycle 1	C. Cycle 2
VAR(1)							
Argentina	1.0000					0.9994	0.6165
Brazil	0.6383	1.0000				0.6110	0.9996
Paraguay	-0.7476	-0.9885	1.0000			-0.7239	-0.9838
Uruguay	-0.9944	-0.5533	0.6731	1.0000		-0.9975	-0.5297
Venezuela	-0.2806	0.5597	-0.4277	0.3806	1.0000	-0.3140	0.5827
VAR(2)							
Argentina	1.0000					0.9910	0.5637
Brazil	0.6306	1.0000				0.5212	0.9965
Paraguay	-0.6536	-0.9995	1.0000			-0.5466	-0.9936
Uruguay	-0.9926	-0.7201	0.7406	1.0000		-0.9675	-0.6597
Venezuela	-0.7706	0.0088	0.0213	0.6876	1.0000	-0.8488	0.0921
VAR(3)							
Argentina	1.0000					0.9824	0.5526
Brazil	0.2604	1.0000				0.4360	0.9486
Paraguay	-0.7832	-0.8042	1.0000			-0.8855	-0.9510
Uruguay	-0.9547	0.0387	0.5627	1.0000		-0.8824	-0.2796
Venezuela	-0.1772	0.9041	-0.4731	0.4620	1.0000	0.0096	0.7223

Once the analysis through linear correlation gives a static measure of the co-movements, as noted by Engle and Kozick (1993), we complement this analysis using techniques based on the frequency domain. Two measures are employed in frequency domain: coherence and phase.[12]

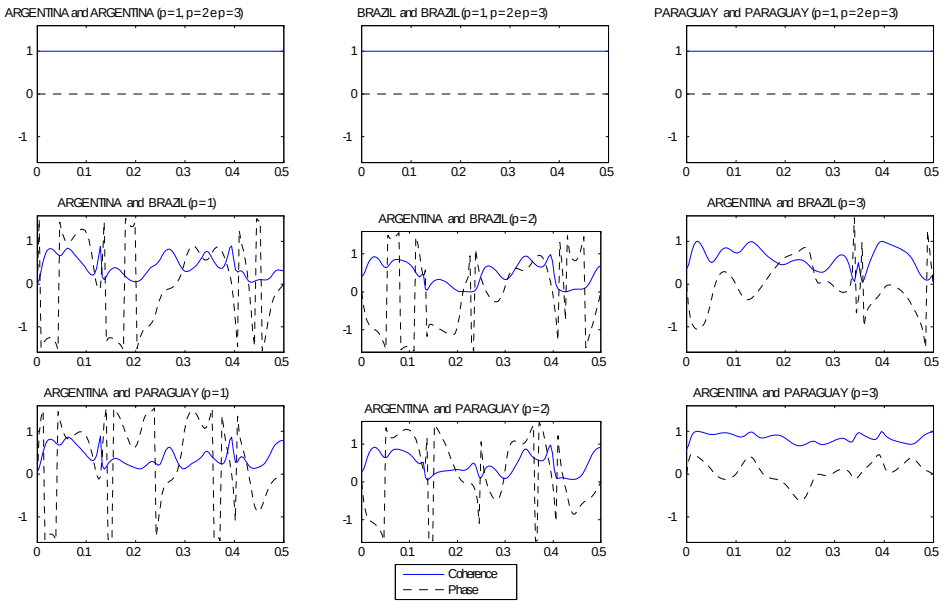
The coherence between two time series is a measure of the degree to which the series are jointly influenced by cycles of frequency w. Coherence belongs to the interval $[0, 1]$. If two time series have perfect linear correlation (positive or negative) the coherence is equal to one. It happens because the same cycles of frequency w are present in both time series. The phase spectrum measures phase difference between two cycles at frequency w. Two oscillators that have the same frequency and different phases have a phase difference, and the oscillators are said to be out of phase with each other.

In summary, we can define co-movements using information in time and frequency domains. Therefore, two time series are synchronized when they have a positive linear correlation, their coherence is close to one and their phase difference is close to zero at each frequency w. An example where one of this condition fails is shown in Figure 7, Argentina and Uruguay have coherence close to one and phase close to zero, but the first condition is not satisfied which means that they have a high negative linear correlation (see Table 7).

Given that, we focus our analysis on the sub-groups identified by the time domain approach. Two groups have high positive linear correlation; i) Brazil and Argentina (0.6383 for $p = 1$) and ii) Paraguay and Uruguay (0.6731 for $p = 1$). Results for $p = 2, 3$ are also reported.

[12] See Appendix B.

Figure 6. Coherence and Phase



Figures 6 to 9 show the coherence and phase between pairs of the business cycles of the Mercosur members[13]. These pictures show values of coherence varying between zero and one (vertical axis) for each value of frequency (horizontal axis). Values of phase (vertical axis) are calculated for each value of frequency (horizontal axis). At the final point of the horizontal axis, the frequency 0.5 corresponds to period of two years, the point 0.25 corresponds to four years, and frequency 0.1 corresponds to ten years, and so on.

The first row of Figure 6 shows the ideal values of coherence and phase are shown, that is, coherence one and phase zero in all frequencies. For example, this picture shows results for synchronization of business cycle of Argentina with himself at each value p, and, after, the same is made for Brazil and Paraguay.

[13]The coherence is estimated using the the *mscohere* function of Matlab 7.0 which considers smoothed with Hamming window of 30 with 50% overlap. The function *cpsd* is used to estimate the Cross Power Spectral Density (CPSD) via Welch's method.

Figure 7. Coherence and Phase

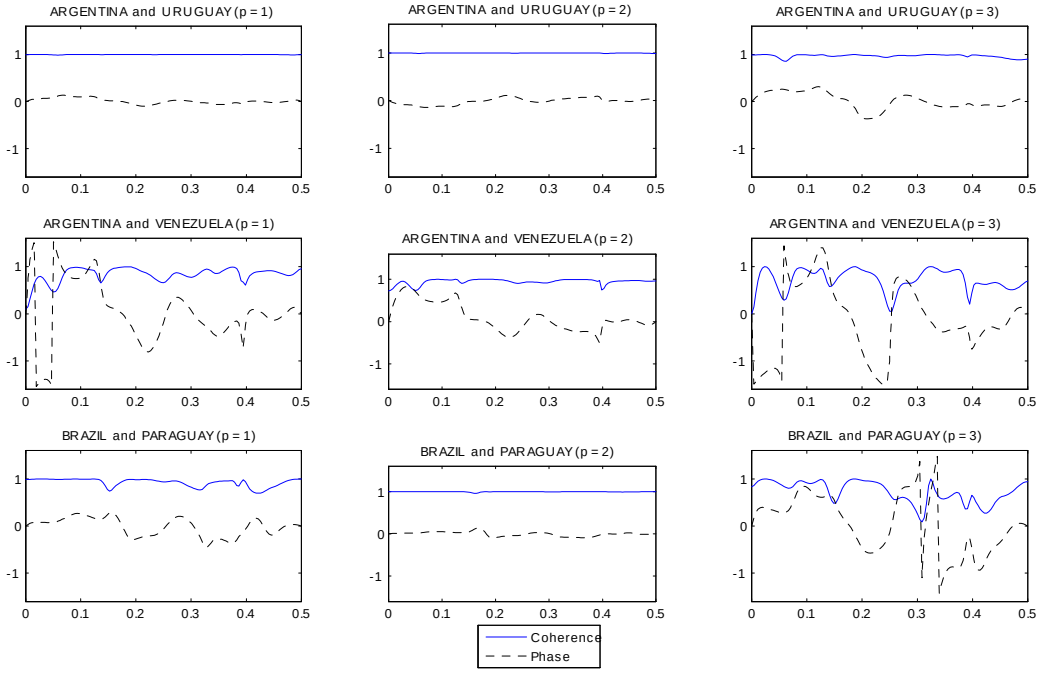
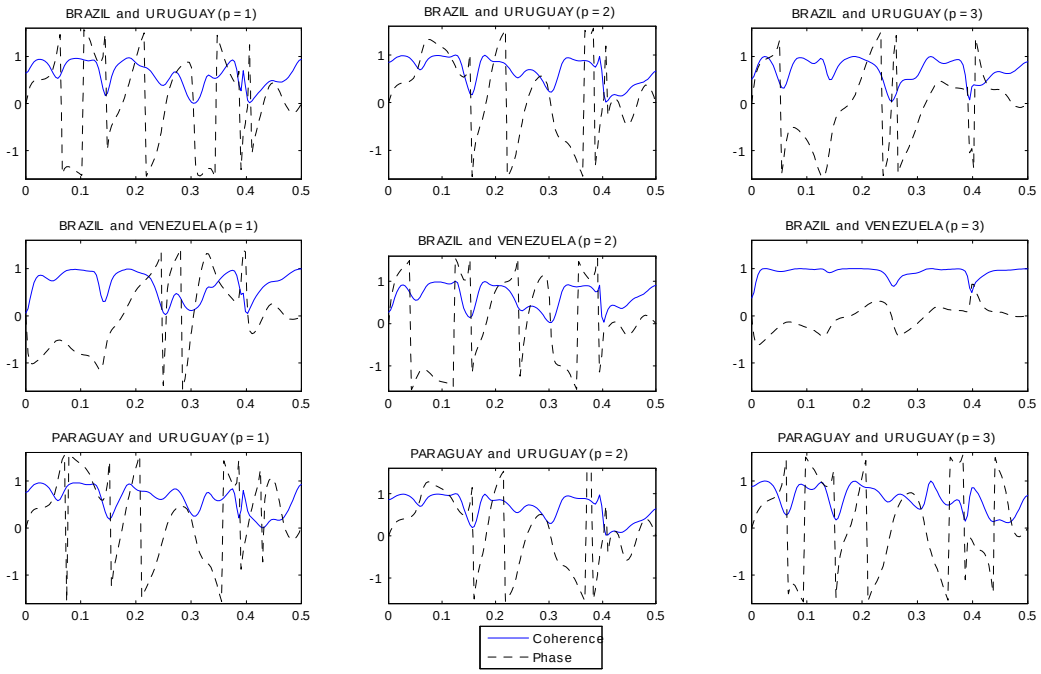


Figure 6 also shows results of coherence and phase for first group: Argentina and Brazil. Focusing on $p = 1$, the coherence are close to one for some frequencies; however, the phase are not close to zero in most frequencies. Thus, when we scrutinized the time domain results, using frequency domain tools, the degree of association between Argentina and Brazil is severely reduced. The results are similar for $p = 2, 3$.

Figure 8 reports the results of coherence and phase for the second group; Uruguay and Paraguay. For $p = 1$, Uruguay and Paraguay has coherence close to one for some frequencies; however, their phase is, in general, far from zero. Thus, this deeper analysis in frequency domain casts doubts on the Uruguay-Paraguay association. Qualitatively, the results are the same for $p = 2, 3$. Hence, the findings suggested no association between this pair.

Therefore, the lack of synchronization among the business cycles confirms that the presence of common cycles does not imply synchronization and corroborates the importance to conduct this analysis in frequency domain.

Figure 8. Coherence and Phase



5 Conclusion

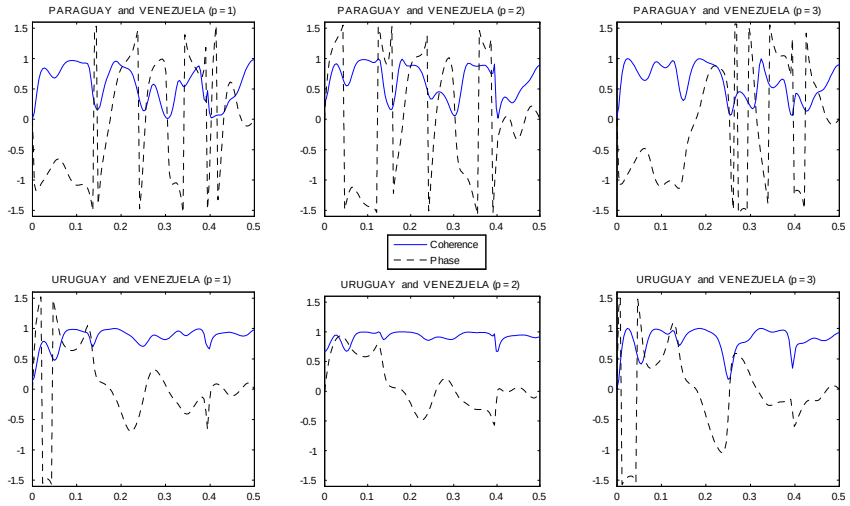
The design of economic blocks is based on the harmonization of economic and commercial policies. However, as argued by Backus and Kehoe (1992) and Chistodoulakis and Dimelis (1995), this harmonization is well succeeded when the member states are sufficiently similar. If this is true, it is of utmost importance to analyze the dynamics of the members and investigate the degree of synchronization of their business cycles. Regarding the Mercosur, it is common to see in the media discussions on the intensification of this economic block. However, it is not usual to argue which the necessary conditions for this intensification are and if they are valid. Considering the members of Mercosur (Argentina, Brazil, Paraguay, Uruguay and Venezuela), this paper analyzes if there are any common dynamic in their economies and if their business cycles are synchronized.

To implement the analysis we estimate a VAR model and test the presence of common trends and common cycles. Using the BNSW trend-cycle decomposition, the business cycles were estimated, taking in account the cointegration and serial correlation common feature restrictions. Then, beyond the usual correlation analysis, measures of coherence and phase, in the frequency domain, are used to examine the degree of co-movements in business cycles.

The results suggest that there are three common trends and two common cycles among the countries. These results confirm the necessity to use a multivariate approach to obtain the business cycles, the first contribution of this work. Time domain results identified evidence of co-movements in two sub-groups: Paraguay-Uruguay and Argentina-Brazil. However, frequency domain tools casts doubts on the synchronization of these pairs. Hence, the lack of synchronism or symmetry in the business cycle of Mercosur makes difficult a greater integration into this economic block.

Appendix A. Coherence and phase results

Figure 9. Coherence and Phase



Appendix B. Coherence and phase

Consider a vector of two stationary variables $y_t = (X_t, Y_t)$. Let $S_{YY}(w)$ represent the *population spectrum* of Y and $S_{YX}(w)$ the *population cross spectrum* between X, Y. The *population cross spectrum* can be written in term of its real and imaginary components as $S_{YX}(w) = C_{YX}(w) + i\ Q_{YX}(w)$, where $C_{YX}(w)$ and $Q_{YX}(w)$ are labeled the *population cospectrum* and *population quadrature spectrum* between X, Y respectively. The *population coherence* between X and Y is a measure of the degree to which X and Y are jointly influenced by cycles of frequency w.

$$h_{YX}(w) = \frac{[C_{YX}(w)]^2 + [Q_{YX}(w)]^2}{S_{YY}(w)\ S_{XX}(w)}$$

Coherence takes values in $0 \leq h_{YX}(w) \leq 1$. A value of one for coherence at a particular point means the two series are altogether in common at that frequency or cycle; if coherence is one over the whole spectrum then the two series are common at all frequencies or cycles. The *cross spectrum* is in general complex, and may express in its polar form as:

$$S_{YX}(w) = C_{YX}(w) + i\ Q_{YX}(w) = R(w)\exp(i\ \theta(w))$$

where $R(w) = \left\{[C_{YX}(w)]^2 + [Q_{YX}(w)]^2\right\}^{\frac{1}{2}}$ and $\theta(w)$ represent the gain and the angle in radians at the frequency w. The angle satisfies $tang(\theta(w)) = \frac{Q_{YX}(w)}{C_{YX}(w)}$. More details in Hamilton (1994).

References

[1] Anderson, T. W, (1951) Estimating linear restrictions on regression coefficients for multivariate normal distributions. Ann. Math. Statist. 22 327-351. [Correction (1980) Ann. Statist. 8 1400].

[2] Ahn, S. K.; Reinsel, G. C. (1988) Nested reduced-rank autoregressive models for multiple time series. J. Amer. Statist. Assoc. 83 849 856. Z.

[3] Backus, David K & Kehoe, Patrick J, (1992). International Evidence of the Historical Properties of Business Cycles. *American Economic Review*, vol. 82(4), pages 864-88, September.

[4] Baxter, M & King, R. G. (1995). Measuring Business Cycles Approximate Band-Pass Filters for Economic Time Series. NBER Working Papers 5022, National Bureau of Economic Research, Inc.

[5] Beveridge, S. & Nelson, C.R. (1981). A new approach to decomposition of economic time series into a permanent and transitory components with particular attention to measurement of the "business cycle". *Journal of Monetary Economics*, 7:151-74.

[6] Candelon, B. & A. Hecq (2000). Stability of Okun's Law in a Codependent System. *Applied Economic Letters*, 7, 687-693.

[7] Christodoulakis, Nicos; Dimelis, Sophia P. & Kollintzas, Tryphon (1995). Comparisons of Business Cycles in the EC: Idiosyncrasies and Regularities. *Economica*, 62 (245): 1-27.

[8] Cubadda G. (1999). Common serial correlation and common business cycles: A cautious note. *Empirical Economics*, vol. 24, issue 3, pages 529-535.

[9] Engle & Kozicki (1993). Testing for common features. *Journal of Business and Economic Statistics*, 11, 369-395.

[10] Guay, A & St-Amant, P (1996). Do Mechanical Filters Provide a Good Approximation of Business Cycles?. Technical Reports 78, Bank of Canada.

[11] Gouriéroux, C & I. Peaucelle (1993), Séries codépendantes : application à l'hypothèse de parité du pouvoir d'achat, in: *Macroéconomie, Développements Récents*, (Economica: Paris).

[12] Harvey, A C & Jaeger, A, (1993). Detrending, Stylized Facts and the Business Cycle. *Journal of Applied Econometrics*, vol. 8(3), pages 231-47.

[13] Issler, J. V. & F. Vahid (2001). Common cycles and the importance of transitory shocks to macroeconomic aggregates. *Journal of Monetary Economics*, vol. 47, no. 3, pp. 449-475.

[14] Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, 231-54.

[15] Johansen, S. (1995). Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford Univ. Press. Z.

[16] Lucas, R. E., Jr., (1977). Understanding business cycles. *Stabilization of the domestic and international economy*, Carnegie-Rochester Conference Series on Public Policy 5 (North-Holland, Amsterdam).

[17] Lütkepohl H. (1993). Introduction to Multiple Time Series Analysis. Springer, New York.

[18] Renato Galvão Flôres Junior, (2005). Investimento Direto Estrangeiro no Mercosur: Uma Visão Geral. *Ensaios Econômicos* EPGE no. 598.

[19] Reinsel, G. C. & S. K Ahn (1992). Vector Autoregressive Models with Unit Roots and Reduced Rank Structure : Estimation, Likelihood Ratio Tests, and Forecasting. *Journal of Time Series Analysis*, 13, 353-375.

[20] Vahid & Engle (1993). Common trends and common cycles. *Journal of Applied Econometrics*, 8, 341-360.

[21] Velu, R. P., Reinsel, G. C. & Wichern, D. W. (1986) . Reduced rank models for multiple times series. *Biometrika* 73, 105 118.

[22] Wang (2003). A frequency domain analysis of common cycles in property and related. sectors. *Journal of Real Estate Research*, 25(3), 325-346 .

[23] Tiao, G. C. & Tsay, R. S. (1985). A canonical correlation approach to modeling multivariate time series. I*n Proceedings of the Business and Economic Statistics Section, American statistical Association*, pp. 112-120.

Chapter 3

Evaluating Asset Pricing Models in a Fama-French Framework[1]

Abstract

In this work we propose a methodology to compare different stochastic discount factor (SDF) proxies based on relevant market information. The starting point is the work of Fama and French, which evidenced that the asset returns of the U.S. economy could be explained by relative factors linked to characteristics of the firms. In this sense, we construct a Monte Carlo simulation to generate a set of returns perfectly compatible with the Fama and French factors and, then, investigate the performance of different SDF proxies. Some goodness-of-fit statistics and the Hansen and Jagannathan distance are used to compare asset pricing models. An empirical application of our setup is also provided.

Keywords: Asset Pricing, Fama and French model, Stochastic Discount Factor, Hansen-Jagannathan distance.

JEL Codes: G12, C15, C22.

[1]This article was jointly made with Wagner Piazza. We are indebted to João Victor Issler, Caio Almeida, Marco Bonomo, Carlos Eugênio, Luis Braido, Christiam Gonzales. In addition, to seminar participants at The 8th Brazilian Finance Society Meeting (Rio de Janeiro, Brazil) and seminar participants at EPGE-FGV (August, 2008) for valuable comments.

1 Introduction

In this work, we propose a new methodology to compare different *stochastic discount factor* or *pricing kernel* proxies.[2] In asset pricing theory, one of the major interests for empirical researchers is oriented by testing whether a particular asset pricing model is (indeed) supported by the data. In addition, a formal procedure to compare the performance of competing asset pricing models is also of great importance in empirical applications.

In both cases, it is of utmost relevance to establish an objective measure of model misspecification. The most useful measure is the well-known Hansen and Jagannathan (1997) distance (or simply HJ-*distance*), which has been used both as a model diagnostic tool and as a formal criterion to compare asset pricing models. This type of comparison has been employed in many recent papers.[3]

As argued by Hansen and Richard (1987), observable implications of candidate models of asset markets are conveniently summarized in terms of their implied stochastic discount factors. As a result, some recent studies of the asset pricing literature have been focused on proposing an estimator for the SDF and also on comparing competing pricing models in terms of the SDF model. For instance, see Lettau and Ludvigson (2001b), Chen and Ludvigson (2008), Araujo, Issler and Fernandes (2006).

A different route to investigate and compare asset pricing models has also been suggested in the literature. The main idea is to assume a data generation process (DGP) for a set of asset returns, based on some assumptions about the asset prices and, then, create a *controlled framework*, which is used to evaluate and compare the asset pricing models.

In this sense, Fernandes and Vieira (2006) study through Monte Carlo simulations the performance of different SDF estimatives at different environments. For instance, the authors

[2] We use the term "stochastic discount factor" as a label for a state-contingent discount factor.

[3] For instance, by using the HJ-distance, Campbell and Cochrane (2000) explain why the CAPM and its extensions better approximate asset pricing models than the standard consumption based model; Jagannathan and Wang (2002) compare the SDF method with Beta method in estimating a risk premium; Dittmar (2002) uses the HJ-distance to estimate the nonlinear pricing kernels in which the risk factor is endogenously determined and preferences restrict the definition of the pricing kernel. Other examples in the literature include Jagannathan, Kubota and Takehara (1998), Farnsworth, Ferson, Jackson, and Todd (2002), Lettau and Ludvigson (2001a) and Chen and Ludvigson (2008).

consider that all asset prices follow a geometric Brownian motion. In this case, one should expect that a SDF proxy based on a geometric Brownian motion assumption would have a better performance, in comparison to an asset pricing model that does not assume this hypothesis. The authors also study competing asset pricing models in a stationary Ornstein-Uhlenbeck process as done in Vasicek (1977).

However, a critical issue of this procedure is that the best asset pricing model inside these particular environments (i.e., when the asset prices are supposed to follow a geometric Brownian motion or a stationary Ornstein-Uhlenbeck process), might not be a good model in the real world. In other words, the best estimator for each *controlled framework* might not necessarily exhibit the same performance for observed stock market prices of a real economy.

In this paper, we use the controlled approach of Fernandes and Vieira (2006), but instead of generating the asset returns from an *ad-hoc* assumption about the DGP of returns, we use related market information from the real economy. Our starting point is the work of Fama and French, which evidenced that asset returns of the U.S. economy could be explained by few relative factors linked to characteristics of the firms[4].

Based on the Fama and French factors, we firstly construct a Monte Carlo simulation to generate a set of returns that is perfectly compatible with these factors. The next step is to create a framework to compare the competing asset pricing models. To do so, we consider two sets of returns: The first sample is used to estimate the different SDF proxies, whereas the remaining sample is used to analyze the out-of-sample performance of each asset pricing model. Although we do not directly use market returns data in this paper, we are able to compare different SDFs by using important market information provided by the Fama-French factors.[5]

Finally, because our approach enables us to construct a data generation process of the SDF provided by the Fama and French specification, it is possible to compare competing proxies through some goodness-of-fit statistics. In addition, it is relevant to test if a set of SDF candidates satisfy the law of one price, such that $1 = E_t\left(m_{t+1}R_{i,t+1}\right)$, where m_{t+1} is

[4] Fama and French (1993, 1995) argue that a three-factor model is successful because it proxies for unobserved common risk in portfolio returns.

[5] Notice that this procedure could also be adopted to compare models by using real data, but with some limitations since the DGP would be unknown.

referred to the investigated stochastic discount factor. Thus, we say that a SDF correctly *prices* the assets if this equation is (in fact) satisfied. In this sense, we test the previous restriction by evaluating, out-of-sample, the HJ-distance of each SDF candidate model.

As shown by Hansen and Jagannathan, the HJ-distance $\delta = \min_{m \in \mathcal{M}} \|y - m\|$, defined in the L^2 space, is the distance of the SDF model y to a family of SDFs, $m \in \mathcal{M}$, that correctly price the assets. In other interpretation, Hansen and Jagannathan show that the HJ-distance is the pricing error for the portfolio that is most mispriced by the underlying model. In this sense, even though the investigated SDF models are misspecified, in practical terms, we are interested in those models with the lowest HJ-distance.

The main objective here is not to propose a DGP process of actual market returns, but to provide a controlled environment that allows one to properly compare and evaluate different SDF proxies. This work follows the idea of Farnsworth et al. (2002), which study different SDFs by constructing artificial mutual funds using real stock returns from the CRSP data.

To illustrate our methodology, we present an empirical application, in which three SDF models are compared: a) The novel nonparametric estimator of Araujo, Issler and Fernandes (2006); b) The Brownian motion pricing model studied in Brandt, Cochrane and Saint-Clara (2006); and c) The (traditional) unconditional linear CAPM.

This work is organized as follows: Section 2 presents the Fama and French model and describes the Monte Carlo simulation strategy; Section 3 presents the results of the empirical application; and Section 4 shows the main conclusions.

2 The stochastic discount factor and the Fama and French model

A general framework to asset pricing is well described in Harrison and Kreps (1979), Hansen and Richard (1987) and Hansen and Jagannathan (1991), associated to the stochastic discount factor (SDF), which relies on the pricing equation:

$$p_t = E_t \left(m_{t+1} x_{i,t+1} \right), \tag{1}$$

where $E_t(\cdot)$ denotes the conditional expectation given the information available at time t, p_t is the asset price, m_{t+1} the stochastic discount factor, $x_{i,t+1}$ the asset payoff of the i-th asset in $t+1$. This pricing equation means that the market value today of an uncertain payoff tomorrow is represented by the payoff multiplied by the discount factor, also taking into account different states of nature by using the underlying probabilities.

The stochastic discount factor model provides a general framework for pricing assets. As documented by Cochrane (2001), asset pricing can basically be summarized by two equations:

$$p_t = E_t \left[m_{t+1} x_{t+1} \right], \tag{2}$$

$$m_{t+1} = f \left(\text{data, parameters} \right). \tag{3}$$

where the model is represented by the function $f(\cdot)$, and the (2) can lead to different predictions stated in terms of returns. For instance, in the Consumption-based Capital Asset Pricing Model (CCAPM) context, the first-order conditions of the consumption-based model, summarized by the well-known Euler equation: $p_t = E_t \left[\beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1} \right]$. The specification of m_{t+1} corresponds to the intertemporal marginal rate of substitution. Hence, $m_{t+1} = f(c, \beta) = \beta \frac{u'(c_{t+1})}{u'(c_t)}$, where β is the discount factor for the future, c_t is consumption and $u(\cdot)$ is a given utility function. The pricing equation (2) mainly illustrates the fact that consumers (optimally) equate marginal rates of substitution to prices.

2.1 Fama and French framework

Fama and French (1992) show that, besides the market risk, there are other important factors that help explain the average return in the stock market. This evidence has been demonstrated in several works for different stock markets (see Gaunt (2004) and Griffin (2005) for a good review). Although there is not a clear link between these factors and the economic theory (e.g., CAPM model), these evidences show that some additional factors might (quite well) help to understand the dynamics of the average return.

These factors are known as the *size* and the *book-to-market equity* and represent special features about firms. Fama and French (1992) argue that size and book-to-market equity are indeed related to economic fundamentals. Although they appear to be "*ad hoc* variables" in an average stock returns regression, these authors justify them as expected and natural proxies for common risk factors in stock returns.

The factors

(i) The SMB (Small Minus Big) factor is constructed to measure the *size premium*. In fact, it is designed to track the additional return that investors have historically received by investing in stocks of companies with relatively small market capitalization. A positive SMB in a given month indicates that small cap stocks have outperformed the large cap stocks in that month. On the other hand, a negative SMB suggests that large caps have outperformed.

(ii) The HML (High Minus Low) factor is constructed to measure the *premium-value* provided to investors for investing in companies with high book-to-market values. A positive HML in a given month suggests that "value stocks" have outperformed the "growth stocks" in that month, whereas a negative HML indicates that growth stocks have outperformed.[6]

(iii) The Market factor is the market excess return in comparison to the risk-free rate. For instance, we proxy the excess return on the market $(R_M - R_f)$, in the U.S. economy, by the value-weighted portfolio of all stocks listed on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and NASDAQ stocks (from CRSP) minus the one-month Treasury Bill rate.

The Model

Fama and French (1993, 1996) propose a three-factor model for expected returns (see also Fama and French (2004) for a good survey).

$$E(R_{it}) - R_{ft} = \beta_{im}\left[E(R_{Mt}) - R_{ft}\right] + \beta_{is}E(SMB_t) + \beta_{ih}E(HML_t), \quad i \in \{1, ..., N\}, \quad (4)$$

where the betas β_{im}, β_{is} and β_{ih} are slopes in the multiple regression (4). Hence, one implication of the expected return equation of the three-factor model is that the intercept in the time-series regression (5) is zero for all assets i:

$$R_{it} - R_{ft} = \beta_{im}\left(R_{Mt} - R_{ft}\right) + \beta_{is}SMB_t + \beta_{ih}HML_t + \varepsilon_{it}. \quad (5)$$

[6] Notice that, in respect to SMB, small companies logically are expected to be more sensitive to many risk factors, as a result of their relatively undiversified nature, and also their reduced ability to absorb negative financial events. On the other hand, the HML factor suggests higher risk exposure for typical value stocks in comparison to growth stocks.

Using this criterion, Fama and French (1993, 1996) find that the model captures much of the variation in the average return for portfolios formed on size, book-to-market equity and other price ratios.

Expected return - beta representation

The Fama and French approach is (in fact) a multifactor model that can be seen as an expected-beta[7] representation of linear factor pricing models of the form:

$$E(R_i) = \gamma + \beta_{im}\lambda_m + \beta_{is}\lambda_s + \beta_{ih}\lambda_h + \alpha_i, \qquad i \in \{1, ..., N\}. \tag{6}$$

If we run this cross sectional regression of average returns on betas, one can estimate the parameters (γ, λ_m, λ_s, λ_h). Notice that γ is the intercept and λ_m, λ_s and λ_h the slope in this cross-sectional relation. In addition, the β_{im}, β_{is} and β_{ih} are the unconditional sensitivities of the i-th asset to the factors[8]. Moreover, β_{ij}, for some $j \in \{m, s, h\}$, can be interpreted as the amount of risk exposure of asset i to factor j, and λ_j as the price of such risk exposure. Hence, the betas are defined as the coefficients in a multiple regression of returns on factors:

$$R_{it} - R_{ft} = \beta_{im}R^{ex}_{Mt} + \beta_{is}SMB_t + \beta_{ih}HML_t + \varepsilon_{it}, \qquad t \in \{1, ..., T\}, \tag{7}$$

where $R^{ex}_{Mt} = (R_{Mt} - R_{ft})$. Following the equivalence between this beta-pricing model and the linear model for the discount factor M, in an unconditional setting (see Cochrane, 2001), we can estimate M as:

$$M = a + b'f, \tag{8}$$

where $f = [R^{ex}_M,\ SMB,\ HML]'$, and the relations between λ e γ, and a and b, are given by:

$$a = \frac{1}{\gamma} \quad \text{and} \quad b = -\gamma \left[cov\left(ff'\right)\right]^{-1}\lambda. \tag{9}$$

[7] The main objective of the beta model is to explain the variation in terms of average returns across assets.

[8] An unconditional time-series approach is used here. The conditional approaches to test for international pricing models include those by Ferson & Harvey (1994, 1999) and Chan, Karolyi and Stulz (1992).

2.2 Evaluating the performance of competing models

In the asset pricing literature, some measures are suggested to compare competing asset pricing models. The most famous measure is the Hansen and Jagannathan distance. However, as long as the data generation process (DGP) is known in each specification of the Fama and French model, it is also possible to use some simple sample statistics. In addition, we use the Hansen and Jagannathan distance to test for model misspecification and to compare the performance of different asset pricing models.

The Hansen-Jagannathan (1997) distance measure is a summary of the mean pricing errors across a group of assets. It may also be interpreted as the distance between the SDF candidate and one that would correctly price the primitive assets. The pricing error can be written by $\alpha_t = E_t\left(m_{t+1}R_{i,t+1}\right) - 1$. Notice, in particular, that α_t depends on the considered SDF, and the SDF is not unique (unless markets are complete). Thus, different SDF proxies can produce similar HJ measures. In this sense, even though the investigated SDF models are misspecified, in practical terms, we are interested in those models with the lowest HJ-distance.

Goodness-of-fit statistics

We use two goodness-of-fit statistics to compare different SDF proxies. The $\widehat{MSE}_s$ is merely a standardized version of the mean squared error of the SDF proxies, whereas the $\widehat{\gamma}_s$ compares the sample correlation between the actual and estimated stochastic discount factors. Let M_t be the stochastic discount factor generated by the Fama and French specification (DGP), and $\widehat{M}_t^s$ the SDF proxy provided by model s in a family S of asset pricing models. The standardized mean squared error is computed as:

$$\widehat{MSE}_s = \frac{\sum_{t=1}^{T}\left(\widehat{M}_t^s - M_t\right)^2}{\sum_{t=1}^{T} M_t^2}, \quad for\ s \in S. \tag{10}$$

and the sample correlation between the actual and estimated SDF is given by:

$$\widehat{\gamma}_s = corr(\widehat{M}_t^s, M_t), \quad for\ s \in S. \tag{11}$$

2.3 Constructing the Fama and French environment

Based on the assumption that R_{Mt}, SMB_t and HML_t are known variables, we can reproduce a Fama and French environment following the three factors of the Fama and French model:

$$R_{i,t} - R_{ft} = \beta_{im}\left(R_{Mt} - R_{ft}\right) + \beta_{is} SMB_t + \beta_{ih} HML_t + \varepsilon_{it}. \tag{7}$$

The simulated asset returns are generated using equation (7). This way, we propose the following steps of a Monte Carlo simulation:

1) Firstly, calibrate each parameter β_{ij}^k, for $j \in \{m, s, h\}$ and $i \in \{1, ...N\}$ according to previous estimations of Fama and French (1992,1993). Therefore, we will generate for each j a N-dimensional vector of asset returns.

2) By considering β_{ij}^k created in step 1 for some $i \in \{1, ...N\}$ and using the known factors R_{Mt}, SMB_t and HML_t, we generate a vector of returns along the time dimension, through equation (7). The *iid* shock ε_{it} is assumed to be a white noise with zero mean and constant variance.

3) Repeating step 2 for each $i \in \{1, ...N\}$, we create the matrix $\mathbf{R}^k$ of asset returns, in which rows are formed by different returns and columns represent the time dimension.

4) Evaluate the mean of $\mathbf{R}^k$ across each row to generate a cross-section vector. Now, it is possible to estimate the parameters γ^k and λ^k through equation (6).

5) Estimate parameters a^k and b^k from the equivalence relation shown in equation (9). Finally, the stochastic discount factor can be estimated by using equation (8).

6) Repeat steps 1 to 5 for an amount of K replications in order to construct the Monte Carlo simulation.

7) Since our approach enables us to construct a data generation process of the SDF provided by the Fama and French specification (computed with N assets), it is possible to compare the competing SDF proxies, obtained in steps 1 to 6, through the goodness-of-fit statistics described in the previous section, as it follows:

7.a) Split the set of N assets into two groups (with the same number of time series observations in each group). Firstly, consider an amount of $\tilde{N} < N$ assets to estimate the SDF candidates (henceforth, this first group of assets will be denominated *in-sample*). Based on the estimated SDF proxies ($\widehat{M_t^s}$) we compute the *in-sample* goodness-of-fit statistics $\widehat{MSE}_s$ and $\widehat{\gamma}_s$, in order to compare every SDF proxy with the correct SDF provided by the Fama and French setup. Secondly, the remaining $(N - \tilde{N})$ assets are used to generate the *out-of-sample* to compute the Hansen and Jagannathan distance. That is, we want to know how well the proxies are carried on when new information is considered.

3 Empirical Application

In this section, we present a simple empirical exercise of our proposed framework for the U.S economy. Three asset pricing models discussed in the literature are compared:

A. The Brownian motion pricing model (studied in Brandt et al., 2006)

Brandt, Cochrane and Santa-Clara (2006) consider that the asset prices follow a geometric Brownian motion (GBM). Such hypothesis is defined by the following partial differential equation:

$$\frac{dP}{P} = \left(R^f + \mu\right) dt + \Sigma^{\frac{1}{2}} dB, \tag{12}$$

where, $\frac{dP}{P} = \left(\frac{dP_1}{P_1} +, ..., \frac{dP_N}{P_N}\right)'$, $\mu = (\mu_1, ..., \mu_n)'$, Σ is a $N \times N$ positive definite matrix, P_i is the price of the asset i, μ the risk premium vector, R^f the risk free rate, and B a standard GBM of dimension N. Using Itô theorem, it is possible to show that:

$$R^i_{t+\Delta t} = \frac{P^i_{t+\Delta t}}{P^i_t} = e^{\left(R^f + \mu_i - \frac{1}{2}\Sigma_{i,i}\right)\Delta t + \sqrt{\Delta t}\left(\Sigma_i^{\frac{1}{2}}\right)' Z_t}, \tag{13}$$

where Z_t is a vector of N independent variables with Gaussian distribution. Therefore, the SDF proposed by these authors is calculated as

$$M_{t+\Delta t} = e^{-\left(R^f + \frac{1}{2}\mu'\Sigma^{-1}\mu\right)\Delta t - \sqrt{\Delta t}\mu\left(\Sigma^{-\frac{1}{2}}\right)' Z_t}. \tag{14}$$

Thus, Brandt, Cochrane and Santa-Clara (2006) suggest the following SDF estimator:

$$\widehat{M_t} = e^{-\left(R^f + \frac{1}{2}\widehat{\mu}'\widehat{\Sigma}^{-1}\widehat{\mu}\right)\Delta t - \widehat{\mu}'\widehat{\Sigma}^{-1}\left(R_t - \bar{R}\right)}, \tag{15}$$

where, $\widehat{\mu}, \overline{R}$ and $\widehat{\Sigma}$ are estimated by:

$$\widehat{\mu} = \frac{\bar{R} - R^f}{\Delta t}, \tag{16}$$

$$\widehat{\Sigma} = \frac{1}{\Delta t}\frac{1}{T}\sum_{t=1}^{T}\left(R_t - \bar{R}\right)\left(R_t - \bar{R}\right)', \tag{17}$$

such that, $R_t = \left(R_t^1, ..., R_t^N\right)'$ and $\bar{R} = \frac{1}{T}\sum_{t=1}^{T} R_t$.

B. Araujo, Issler and Fernandes (2006)

A novel estimator for the stochastic discount factor (within a panel data context) is proposed by Araujo, Issler and Fernandes (2006). This setting is slightly more general than the GBM setup put forth by Brandt, Cochrane and Santa-Clara (2006). In fact, this estimator assumes that, for every asset $i \in \{1, ..., N\}$, $M_{t+1}R_{t+1}^i$ is conditionally homoskedastic and has a lognormal distribution. In addition, under asset pricing equation (1) and some mild additional conditions, they show that a consistent estimator for M_t is given by:

$$\widehat{M_t} = \left(\frac{\bar{R}_t^G}{\frac{1}{T}\sum_{t=1}^{T}\bar{R}_t^A\bar{R}_t^G}\right), \tag{18}$$

where $\bar{R}_t^A = \frac{1}{N}\sum_{i=1}^{N} R_{i,t}$ and $\bar{R}_t^G = \Pi_{i=1}^{N} R_{i,t}^{-\frac{1}{N}}$ are respectively the cross-sectional arithmetic and geometric average of all gross returns. Therefore, this nonparametric estimator depends exclusively on appropriate averages of asset returns that can easily be implemented.

C. *Capital Asset Pricing Model - CAPM*

Assuming the unconditional CAPM, the SDF is a linear function of market returns calculated as: $m_{t+1} = a + bR_{w,t+1}$, where $R_{w,t+1}$ is the gross return on the market portfolio of all assets. For instance, in the U.S. economy, in order to implement the static CAPM, for practical purposes, it is commonly assumed that the return on the value-weighted portfolio of all stocks listed on NYSE, AMEX, and NASDAQ is a reasonable proxy for the return on the market portfolio of all assets of the U.S. economy.

3.1 Monte Carlo design

In order to compare these three SDF proxies we construct the Monte Carlo experiment following the procedure showed in section 2.3. For the U.S. economy, the factors $(R_{Mt} - R_{ft})$, SMB_t and HML_t are extracted from the Kenneth R. French website[9]. Next, we calibrate the parameters β_{im}, β_{is} and β_{ih} according to previous estimations of Fama and French (1992,1993) and estimate the parameters $(\gamma,\ \lambda_m,\ \lambda_s,\ \lambda_h)$ from the cross-sectional regression (6), observing their significance through the F-statistic or the t-statistic for individual parameters.

We set $N = 36$ as our set of primitive assets, which are divided into two groups: The first one contains $\tilde{N} = 18$ assets that are used for the *in-sample* estimation. The second group has $(N - \tilde{N}) = 18$ assets, which are thus used for the *out-of-sample* analysis. We also consider, for each generated asset i, three sample sizes $T = \{200; 300; 400\}$.

In this way, we estimate the stochastic discount factors for the three-factor model of Fama and French, and repeat the mentioned procedure for an amount of $K = 1,000$ replications. Some descriptive statistics of the generated SDFs are presented in the appendix. Finally, the evaluation of the SDF proxies is conducted and the Monte Carlo results are summarized by two goodness-of-fit statistics (besides the HJ-distance), which are averaged across all replications.

We denote the SDF proxies, estimated in each replication, as $\widehat{M}_t^a$, $\widehat{M}_t^b$ and $\widehat{M}_t^c$ to Araujo, Issler and Fernandes (2006), Brandt, Cochrane and Santa-Clara (2006) and the unconditional CAPM respectively. In addition, the stochastic discount factor implied by the Fama and French setup (DGP) is denoted by M_t.

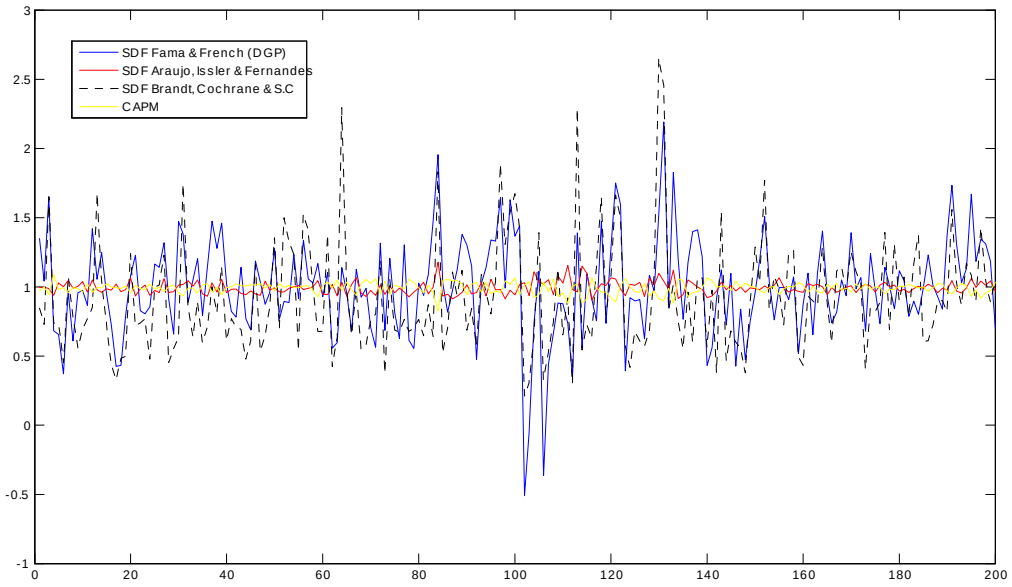
3.2 Results

In Figure 1, the estimates of the SDF proxies are shown for one replication of the Monte Carlo simulation, with a sample size $T = 200$. A simple graphical investigation reveals that

[9] More information about data can be found in: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html For other economies, the factors can be constructed as showed in Fama and French (1992, 1993).

the Brandt, Cochrane and Santa-Clara, $\widehat{M_t^b}$, and the CAPM proxy, $\widehat{M_t^c}$, are respectively the most and less volatile, which is a result confirmed by the descriptive statistics of Table 2 (in appendix). In addition, $\widehat{M_t^b}$ appears to be the SDF proxy that best tracks the DGP M_t.

Figure 1 - Three factors, with a sample size $T = 200$



Notes: a) Figure 1 shows one replication out of the total amount of 1,000 replications.

b) We adopt $\tilde{N} = 18$ assets and T=200 observations.

Regarding the performance of the SDF proxies, Table 1 reports the evaluation statistics obtained from Monte Carlo simulations. Notice that results are robust to sample size. In all cases, the mean square error of Brandt, Cochrane and Santa-Clara (2006) SDF proxy ($\widehat{MSE_b}$) shows quite a good performance, whereas the CAPM proxy seems to exhibit the worst one. Nonetheless, the magnitude of the standard deviation might suggest that all these values are quite close to each other.

In respect to the correlation of the true SDF with the considered SDF proxies, we have obtained the following ranking order for all sample sizes: $\widehat{M_t^b} \succ \widehat{M_t^a} \succ \widehat{M_t^c}$. This implies

that the Brandt, Cochrane and Santa-Clara (2006) proxy (in general) best tracks the dynamic path of the true SDF. On the other hand, the CAPM model exhibits again the worst performance (with a negative correlation in some cases!)

Finally, in respect to the *out-of-sample* analysis, the HJ distance results[10] (which should be as close as possible to zero in a correctly-specified model) indicate that for $T = 200$ and $T = 300$: $\widehat{HJ}_b < \widehat{HJ}_a < \widehat{HJ}_c$, revealing that the Brandt, Cochrane and Santa-Clara (2006) is the best proxy for forecasting purposes, followed by the Araujo et al. (2006) SDF estimator. For $T = 400$ we obtained similar results, except that in this case the CAPM model has a lower HJ-distance in comparison to the Araujo et al. (2006) proxy.[11]

Putting all together, the numerical results show that (in general) the Brandt, Cochrane and Santa-Clara (2006) has the best *out-of-sample* performance. Notice that Figure 1 already showed this tendency, since the referred SDF best tracked the respective Fama-French DGP.

Finally, the CAPM model shows a negative correlation with the true SDF, revealing its weakness in tracking the real dynamic of the true SDF. This result is because the linear CAPM only uses one single factor, out of the three factors correct-specification in the Fama-French setup. This way, our methodology allows one to rank the competing SDF models (according to different evaluation criteria), based on simulated data generated from U.S. market information.

[10] We compute the HJ distance based on the MatLab codes of Mike Cliff, available at: http://mcliff.cob.vt.edu/

[11] The standard error of the HJ-distance is estimated by a Newey & West (1987) HAC procedure, in which the optimal bandwidth (number of lags=5) is given by $m(T) = int(T^{1/3})$, where $int(.)$ represents the integer part of the argument, and T is the sample size. The adopted kernel used to smooth the sample autocovariance function is given by a standard modified Bartlett kernel: $w(j, m(T)) = 1 - [j/\{m(T)+1\}]$. See Newey & West (1994) for an extensive discussion about lag selection in covariance matrix estimation, and also Kan & Robotti (2008).

Table 1 - Monte Carlo Simulation Results

sample size: 200 (Over the time period from 09/1999 to 12/2007)					
$\widehat{MSE}_a$	$\widehat{MSE}_b$	$\widehat{MSE}_c$	$\widehat{\gamma}_a$	$\widehat{\gamma}_b$	$\widehat{\gamma}_c$
0.0962	0.1070	0.1056	0.2645	0.6429	−0.0113
(0.0228)	(0.0374)	(0.0298)	(0.1106)	(0.0720)	(0.4387)
$\widehat{HJ}_a$-distance		$\widehat{HJ}_b$-distance		$\widehat{HJ}_c$-distance	
0.4114		0.3227		0.4207	
(0.0806)		(0.0760)		(0.0792)	

sample size: 300 (Over the time period from 05/1991 to 12/2007)					
$\widehat{MSE}_a$	$\widehat{MSE}_b$	$\widehat{MSE}_c$	$\widehat{\gamma}_a$	$\widehat{\gamma}_b$	$\widehat{\gamma}_c$
0.0796	0.0722	0.0923	0.3301	0.6989	−0.1041
(0.0182)	(0.0221)	(0.0242)	(0.0895)	(0.0626)	(0.4399)
$\widehat{HJ}_a$-distance		$\widehat{HJ}_b$-distance		$\widehat{HJ}_c$-distance	
0.3489		0.2588		0.3631	
(0.0660)		(0.0606)		(0.0643)	

sample size: 400 (Over the time period from 09/1974 to 12/2007)					
$\widehat{MSE}_a$	$\widehat{MSE}_b$	$\widehat{MSE}_c$	$\widehat{\gamma}_a$	$\widehat{\gamma}_b$	$\widehat{\gamma}_c$
0.0779	0.0608	0.0702	0.3423	0.7182	0.4319
(0.0153)	(0.0161)	(0.0160)	(0.0933)	(0.0551)	(0.2351)
$\widehat{HJ}_a$-distance		$\widehat{HJ}_b$-distance		$\widehat{HJ}_c$-distance	
0.3305		0.2275		0.3227	
(0.0553)		(0.0520)		(0.0556)	

Notes: a) We simulate a panel with 25 asset returns from a Fama and French model of the form: $R_{i,t} - R_{ft} = \beta_{im}\left(R_{Mt} - R_{ft}\right) + \beta_{is}SMB_t + \beta_{ih}HML_t + \varepsilon_{it}$.

b) All results are averaged across the 1,000 replications. The MSE and γ are computed "in-sample", i.e., N=18, whereas the HJ-distance is calculated from the "out-of-sample" set of (N-$\tilde{\text{N}}$)=18 assets. The standard deviation is presented in parentheses.

c) The calibrated parameters varies from $\beta_{im} \in [0.1, 0.9]$; $\beta_{is} \in [-1.4, 1.6]$; $\beta_{ih} \in [-0.73, 8.7]$ in each replication of the Monte Carlo simulation.

4 Conclusions

In the present work, we propose a methodology to compare different stochastic discount factor models based on relevant market information. Based on the Fama and French factors, which are linked to characteristics of the firms in a particular economy, a Monte Carlo simulation strategy is proposed in order to generate a set of artificial returns that is perfectly compatible with those factors.

This way, we construct a *Fama-French world* through numerical simulations, in which SDF proxies are compared through some goodness-of-fit statistics and the Hansen and Jagannathan distance. An empirical application is provided to illustrate our methodology, in which returns time series are produced from factors such as the market portfolio return, size and book-to-market equity of the U.S. economy. The results reveal that the Brandt, Cochrane and Saint-Clara (2006) proxy dominates the other considered SDF estimators.

Therefore, the main contribution of this paper consists in a methodology to compare SDF models in a setup where the Fama and French factors are supposed to summarize the economic environment. This controlled framework allows one to use simple sample statistics to compare SDF candidates with the *true SDF* implied by the Fama and French DGP and, then, rank competing asset pricing models. In this case, the hypothesis of geometric Brownian motion, usually adopted in several empirical studies, seems to be quite reasonable for the simulated set of returns.

A further extension of the proposed method could be to compare asset pricing models in more than one framework. For instance, the asset returns could be generated following the Epstein and Zin's (1989) preferences or the consumption capital asset pricing model (CCAPM) as shown in Campbell and Cochrane (1999).

Appendix. Descriptive Statistic

Table 2 - Descriptive statistics of the SDF proxies

sample size = 200				
	Araujo	**Saint Clara**	**CAPM**	**Fama & French DGP**
Mean	0.9945	0.9185	0.9921	0.9967
Median	0.9900	0.8380	0.9927	1.0002
Maximum	1.1918	2.9764	1.1627	2.1010
Minimum	0.8860	0.1867	0.8121	-0.5184
Std. Dev	0.0482	0.4194	0.0531	0.3346
Skewness	0.7922	1.5141	-0.0567	-0.5456
Kurtosis	4.6444	7.4416	4.1835	6.0446
Freq. Jarque-Bera	0.0150	0.0000	0.0000	0.0000
sample size = 300				
Mean	0.9933	0.9196	0.9902	0.9959
Median	0.9889	0.8564	0.9917	0.9878
Maximum	1.2849	2.9480	1.1451	2.1842
Minimum	0.8728	0.2381	0.7905	-0.2985
Std. Dev	0.0506	0.3647	0.0426	0.3058
Skewness	1.1507	1.5303	-0.2606	-0.2345
Kurtosis	7.4321	8.2266	6.2824	5.3050
Freq. Jarque-Bera	0.0000	0.0000	0.0000	0.0000
sample size = 400				
Mean	0.9925	0.9181	0.9942	0.9952
Median	0.9887	0.8672	0.9875	1.0042
Maximum	1.2838	3.0148	1.5317	2.1668
Minimum	0.8661	0.1674	0.6924	-0.6743
Std. Dev	0.0504	0.3355	0.0998	0.3049
Skewness	0.9412	1.6279	0.5455	-0.9058
Kurtosis	6.4386	9.6505	5.4933	9.2686
Freq. Jarque-Bera	0.0000	0.0000	0.0000	0.0000

Notes: These statistics are computed in-sample. DGP (FF) means Data-Generating Process of the Fama & French model. The number of assets in-sample and out-of-sample is N=18. The descriptive statistics are averaged across the K=1,000 replications based on the sample sizes T={200,300,400}. For instance, for T=200 the Jarque-Bera statistic indicates the frequency of rejection of the normality hypothesis across the 1,000 replications (based on a 5% significance level). In this case, T=200, for the Araujo et al. (2006) proxy, the statistic Freq. Jarque-Bera is equal to 0.015, which means that in 1.5% of the replications the normality hypothesis is rejected at a 5% significance level.

References

[1] Araujo, F., Issler, J. V., Fernandes, M., 2006. Estimating the stochastic discount factor without a utility function. Princeton University, Getulio Vargas Foundation, and Queen Mary, University of London.

[2] Brandt, M. W., Cochrane, J. H., Santa-Clara, P., 2006. International risk sharing is better than you think, or exchange rates are too smooth, forthcoming in Journal of Monetary Economics.

[3] Campbell, John Y., and John H. Cochrane, 1999, By force of habit: a consumption-based explanation of aggregate stock market behavior, Journal of Political Economy 107, 205-251.

[4] _____, 2000. Explaining the poor performance of consumption-based "asset pricing models". Journal of Finance 55, 2863-2878.

[5] Chan, K.C., Karolyi, G.A., Stulz, R.M., 1992. Global Financial Markets and the Risk Premium on U.S. Equity. NBER Working Paper n.4074.

[6] Chen and Ludvigson, 2008. Land of Addicts? An Empirical Investigation of Habit-Based Asset Pricing Models. Working paper, New York University.

[7] Cochrane, J. H., 2001. Asset Pricing, Princeton University Press, Princeton.

[8] Dittmar, 2002. Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns. Journal of Financial Economics 33, 3-56.

[9] Epstein, Lawrence, and Stanley Zin, 1989, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, Econometrica 57, 937-969.

[10] Fama, E. F. and K.R. French, 1992. The Cross Section of Expected Stock Returns, Journal of Finance 47, 427-465.

[11] _____, 1993. Common risk factors in the returns on stocks and bonds, Journal of Financial Economics 33, 3-56.

[12] _____, 1995. Size and Book-to-Market Factors in Earnings and Returns, Journal of Finance 50, 131-155.

[13] _____, 1996. Multifactor Explanations of Asset Pricing Anomalies, Journal of Finance 51(1), 55-84.

[14] _____, 1998. Value versus growth: the international evidence, Journal of Finance 53, 1975-1979.

[15] _____, 2004. The Capital Asset Pricing Model: Theory and Evidence. The Journal of Economic Perspectives 18(3), 25-46.

[16] Farnsworth, H., Ferson, W.E., Jackson, D., Todd, S., 2002. Performance evaluation with stochastic discount factors, Journal of Business 75, 473–503.

[17] Fernandes, M., Vieira, G., 2006. Revisiting the efficiency of risk sharing between UK and US: Robust estimation and calibration under market incompleteness. Mimeo.

[18] Ferson, W.E., Harvey, C.R., 1994. Sources of Risk and Expected Returns in Global Equity Markets. Journal of Banking and Finance 18, 775-803.

[19] _____, 1999. Conditioning Variables and the Cross Section of Stock Returns. Journal of Finance 54(4), 1325-1360.

[20] Gaunt, C., 2004. Size and book to market effects and the Fama French three factor asset pricing model: evidence from the Australian stock market. Accounting and Finance 44, 27-44.

[21] Griffin, J.M., 2002. Are the Fama and French Factors Global or Country-Specific? The Review of Financial Studies 15(3), 783-803.

[22] Hansen, L.P., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies, Journal of Political Economy 99, 225-262.

[23] _____, 1997. Assessing Specification Errors in Stochastic Discount Factor Models. Journal of Finance 52(2), 557-590.

[24] Jagannathan, R., Kubota, K., Takehara, H., 1998. Relationship between Labor-Income Risk and Average Return: Empirical Evidence from the Japanese Stock Market. Jounal of Business 71, 319-348.

[25] Jagannthan, R., Wang, Z., 2002. Empirical evaluation of asset-pricing models: A comparison of the SDF and beta methods" Journal of Finance 57, 2337-2367.

[26] Johnson, R.A., Wichern, D.W., 1992. Applied multivariate statistical analysis. Third edition. New Jersey: Prentice-Hall.

[27] Kan, R., Robotti, C., 2008. Model Comparison Using the Hansen-Jagannathan Distance. Working paper, University of Toronto.

[28] Lettau, M., Ludvigson, S., 2001a. Consumption, Aggregate Wealth, and Expected Stock Returns. Journal of Finance 56(3), 815-849.

[29] _____, 2001b. Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia are Time-Varying. Journal of Political Economy 109(6), 1238-1287.

[30] Lintner, J., 1965. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. Review of Economics and Statistics 47(1), 13-37.

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

[32] _____, 1994. Automatic lag selection in covariance matrix estimation. The Review of Economic Studies 61(4), 631-653.

[33] Sharpe,W.F., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. Journal of Finance 19(3), 423-442.

[34] Vasicek, O., 1977. An equilibrium characterization of the term structure, Journal of Financial Economics 5, 177-188.

Chapter 4

Testing Consumption-based Asset Pricing Models with Factor Model Analysis[1]

Abstract

In this paper we test different representative classes of consumption-based asset pricing models (CCAPM). A critical point of testing these models, which is implied by the Euler conditions of the agent's optimization problem, is the choice of a measure for the asset returns. We use the factor model to construct a small set of portfolios that summarize the common information available in the stock market. This has an economic interpretation, since consumption should only respond to systematic changes in returns, which is exactly the information contained in these portfolios. In addition, from an econometric point of view, this methodology allows reducing the number of variables, making feasible the model estimation. We use the generalized method of moments (GMM) to identify and estimate the parameters of the utility function. The main result is obtained by the CRRA preference. We show that when the factor portfolio is used, the implication of this model is not rejected. Indeed, the relative risk-aversion coefficient is significant and ranges from 0.026 to 0.338 while the discount factor is significant and ranges from 0.998 to 0.999.

Keywords: Factor Model, CCAPM, GMM.
JEL Codes: G12, C13, C23, E44.

[1] This article was jointly made with João Victor Issler. We are indebted to seminar participants at EPGE-FGV (August, 2008), in special to Bonomo and Carlos Eugenio.

1 Introduction

This paper provides an empirical investigation of the testable restrictions implied by different representative classes of consumption-based asset pricing models (CCAPM).

Since the study of Hall (1978), the literature on consumption has focused on testing diverse implications on the behavior of individuals using aggregate data on consumption. Flavin (1981), Campbell (1987) and Campbell & Deaton (1989) are notable examples. Generally, these works concluded that the theory of consumer behavior can be rejected. In particular, Hansen and Singleton (1982, 1983) and Epstein & Zin (1991) rejected these implications in the CCAPM environment. In an effort to find evidence to support the theory of consumption, some authors have tested these models using disaggregated data, showing that the implications of the model of consumption were less often rejected (see Runkle (1991), Atanasio & Weber (1995) and Lusardi (1996)).

Hansen & Singleton (1982) and Epstein & Zin (1991) used the orthogonality restrictions implied by the Euler equations of the agent's optimization problem to identify and estimate the parameters of the utility function using the generalized method of moments (GMM). A critical point of testing CCAPM is the choice of a measure for the asset returns[2]. These authors assumed as a standard the use of a riskless return and some index as a measure of an optimal portfolio. For instance, Hansen and Jaganathan (1983) used the S&P 500 index and Treasury Bill as a risk-free rate of return. Epstein and Zin (1991) used five individual stock return indices that give value-weighted returns for broad groups of industrial stocks and Treasury Bill as a risk-free rate of return.

In this paper we use aggregated data on consumption to test the models in the CCAPM, as in Hansen and Singleton (1982, 1983) and Epstein & Zin (1991). Contrary to these studies, however, we consider a larger set of asset returns by incorporating the statistical factor model in the analysis. Using the factor model we can summarize the common information available in the stock market in a small set of portfolios. This methodology allows us to reduce the number of variables, making feasible the model estimation.

Estimation theory of intertemporal substitution in consumption requires the use of all assets in the economy. Since it may be unfeasible to use all returns in the economy, we propose an aggregation which takes into account the common information of returns. This

[2]Mulligan (2002) consider the interest rate of the U.S. national accounts data. Mulligan (2002) proposed the use of the return to aggregate capital (or the capital rental rate) to R_t

procedure has an interesting economic interpretation. Because the common effects refer to the systematic risk of the economy, it is in line with the theory of intertemporal consumption smoothing, which implies that consumption should only respond to systematic changes in returns. This is exactly the information contained in these portfolios.

On the other hand, from an econometric point of view, with large number assets the model's estimation may be econometrically unfeasible. For instance, in GMM estimation the number of moment conditions,[3] q, must be satisfy $T > q(q+1)/2$, where T is the number of time observations. This means that asymptotic results will require T to grow at rate q^2. Our methodology allows constructing a small set of portfolios ($\tilde{N} << N$) which reduces the number of moment conditions.

In this work, we chose a representative class of consumption-based asset pricing models; the CRRA specification studied in Hansen & Singleton (1982), which is the most popular specification in the financial and macroeconomic literature; the parametric utility function, by allowing for relative consumption external-habit specification, as in Abel (1990); and the non-additive generalized expected utility specification of Kreps and Porteus (1978), as used in Epstein and Zin (1991). Our empirical work uses monthly U.S. data spanning the 1970 to 2007 period. The overidentifying restrictions can be tested to examine how close the sample versions of population orthogonality conditions are to zero.

The main result is obtained by the CRRA preference. We show that when the factor portfolio is used, the simple representative agent model with a CRRA utility is able to explain the time series data on consumption. Indeed, the intertemporal discount factor is significant and ranges from 0.998 to 0.999, while the relative risk-aversion coefficient is precisely estimated, ranging from 0.026 to 0.338. Moreover, there is no evidence of rejection in over-identifying-restriction tests.

The paper proceeds as follows: Section 2 shows the utility specifications and the factor model used to construct portfolios; Section 3 shows the results; and Section 4 contains the conclusions and possibilities for future research.

[3] The number of moment conditions, q, is computed as the number of asset times the number of instruments.

2 Utility Functions and Asset Returns

We consider, for purposes of simplification, an economy endowed with an infinitely lived representative agent who receives utility from the consumption of a single good. The dynamic rational expectation model is given by;

$$E_t\left[\sum_{i=0}^{\infty}\beta^i u\left(C_{t+i}\right)\right]\ ,\ 0<\beta<1 \tag{1}$$

Where C_t represents aggregate consumption at time t, $\beta \in (0,1)$ is the time discount factor and $u(\cdot)$ is a strictly concave function. In any period t, current consumption C_t, is deterministic but future consumption is uncertain. Suppose that at date t there are N different assets and that a dollar invested in asset i at date t will yield a gross return of $(1+r_{i,t+1})$ at date $t+1$. Therefore, the optimization problem of (1) with certain regularity conditions gives the N first-order necessary conditions:

$$u'\left(C_t\right)=\beta E_t\left[(1+r_{i,t+1})u'\left(C_{t+1}\right)\right] \quad \text{for} \quad i=1,2....N \tag{2}$$

or equivalently written as:

$$1=E_t\left[(1+r_{i,t+1})\beta\frac{u'\left(C_{t+1}\right)}{u'\left(C_t\right)}\right] \quad \text{for} \quad i=1,2....N \tag{3}$$

Define the intertemporal marginal rate of substitution $m_{t+1}=\beta\frac{u'(c_{t+1})}{u'(c_t)}$ and consider the gross return $R_{i,t+1}=1+r_{i,t+1}$ in (3) It follows that

$$1=E_t\left[m_{t+1}R_{i,t+1}\right] \quad \text{for} \quad i=1,2....N \tag{4}$$

Equation (4) is the pricing equation established by Harrison & Kreps (1979), Hansen & Richard (1987) and Hansen & Jagannathan (1991) where the intertemporal marginal rate of substitution m_{t+1} is the stochastic discount factor (SDF). This equation means that asset prices today are a function of their expected future payoffs discounted by the stochastic discount factor (SDF). The general framework for asset pricing, associated with the stochastic discount factor, affirms that the pricing equation is valid for all assets at all times. Therefore, ideally we should consider all assets available in the economy.

In this paper we use the factor models to construct a small set of portfolios that summarize the common information available in the stock market.

Notice that if equation (4) is valid for assets, it is also valid for portfolios. To see that, define $\tilde{R}_{j,t+1} = \sum_{i=1}^{N} w_i R_{i,t+1}$, for $j = 1, ...\tilde{N}$, portfolios, $\sum_{i=1}^{N} w_i = 1$, where $\tilde{N} << N$. Thus, pre-multiplying each equation (4) by w_i produces:

$$(w_i)\,\mathbb{E}_t\ [m_{t+1}R_{i,t+1}] = 1 \times (w_i\ )\,, \qquad i = 1, ..., N \tag{5}$$

and aggregating in N;

$$\mathbb{E}_t\ \left[m_{t+1}\sum_{i=1}^{N} wR_{i,t+1}\right] = 1$$

Therefore, the pricing equation for portfolios is given by:

$$\mathbb{E}_t\ \left[m_{t+1}\tilde{R}_{j,t+1}\right] = 1\ , \qquad j = 1, .., \tilde{N} \tag{6}$$

Given a utility function, $u(\cdot)$, we estimate (6) as suggested by Hansen and Singleton (1982) using the non linear GMM estimation. Notice that equation (6) is a conditional expectation, but for GMM estimation it should be in terms of unconditional expectations. We estimate the conditional expectation assuming x_t^* denote the information set implicit in (6), so

$$\mathbb{E}\ \left[\left(\beta\frac{u'(C_{t+1})}{u'(C_t)}\tilde{R}_{j,t+1} - 1\right) \mid x_t^*\right] = 0\ , \qquad j = 1, .., \tilde{N} \tag{7}$$

equation (7) requires $\beta u'(C_{t+1})/u'(C_t)\,\tilde{R}_{j,t+1} - 1$ to be uncorrelated with any variable in x_t^*. Thus, from the law of iterated expectations, the conditional model (7) implies the unconditional model

$$\mathbb{E}\ \left[\left(\beta\frac{u'(C_{t+1})}{u'(C_t)}\tilde{R}_{j,t+1} - 1\right) x_t\right] = 0\ , \qquad j = 1, ..\tilde{N} \tag{8}$$

where x_t is $(M \times 1)$ observable subset of x_t^*. Since we have $\tilde{N}$ portfolios and M instruments the number of moment conditions is $q = \tilde{N}M$, which implies:

$$h(\theta, w_t)_{q\times 1} = \begin{bmatrix} \left(\beta\frac{u'(C_{t+1})}{u'(C_t)}\tilde{R}_{1,t+1} - 1\right) x_t \\ \left(\beta\frac{u'(C_{t+1})}{u'(C_t)}\tilde{R}_{2,t+1} - 1\right) x_t \\ \vdots \\ \left(\beta\frac{u'(C_{t+1})}{u'(C_t)}\tilde{R}_{\tilde{N},t+1} - 1\right) x_t \end{bmatrix} = 0 \tag{9}$$

The q orthogonal conditions in (9) are used to estimate the structural parameters through the GMM method. Notice that we only have $\tilde{N}$ equations ($\tilde{N} << N$) given by equation

(8). Thus we have significantly reduced the number of asset returns used in testing the consumption models. This makes GMM estimation feasible and efficient.

It is known that there is a natural econometric restriction in GMM estimation. The number of moment conditions, q, must be satisfy $T > q\,(q+1)\,/2$, where T is the number of time observations. In practice this constraint may be impossible to satisfy. For example, if we consider twenty assets and two instruments, for $T = 400$ observations available, then the number of moment conditions results in $q = 40$. Thus, we should have at least $T > 820$ observations to estimate the model adequately. As time series data will not be sufficient to satisfy this constraint, we are forced to reduce the number of assets, which implies losing estimation efficiency.

Preferences representations in the CCAPM framework

Economic theory provides a number of preference representations, of which the most popular is the time-separable power utility defined over aggregate consumption:

$$u\left(C_{t+1}\right) = \frac{C_{t+1}^{1-\gamma} - 1}{1-\gamma} \tag{10}$$

where the parameter $\gamma > 0$ defines the curvature of the utility function and hence represents the constant relative risk aversion in the economy. The resulting Euler equation is

$$E_{t-1}\left[\beta\left(\frac{C_{1,t}}{C_{1,t-1}}\right)^{-\gamma} R_{i,t}\right] = 1,\ i = 1,..,N \tag{11}$$

The parametric utility function, by allowing for relative consumption external-habit specification, as in Abel (1990);

$$u\left(C_{t+1}\right) = \frac{\left(C_{t+1}/\bar{C}_t^{\kappa}\right)^{1-\gamma} - 1}{1-\gamma} \tag{12}$$

where habits arise from the average past consumption $\bar{C}_t$ in the economy, the parameter $\kappa > 0$ gauges the strength of the catching-up-with-the-Joneses effect and hence controls the time separability of the utility function. The resulting Euler equation is

$$E_{t-1}\left[\beta\left(\frac{C_{1,t}}{C_{1,t-1}}\right)^{-\gamma}\left(\frac{C_{1,t-1}}{C_{1,t-2}}\right)^{-\kappa(\gamma-1)} R_{i,t}\right] = 1,\ i = 1,..,N \tag{13}$$

Finally, the Kreps and Porteus (1978) specification used in Epstein and Zin (1991) is

$$U_t = \left[(1-\beta) C_t^{\rho} + \left(E_t U_{t+1}^{\alpha} \right)^{\frac{\rho}{\alpha}} \right]^{\frac{1}{\rho}} \tag{14}$$

The resulting Euler equation is

$$E_{t-1} \left[\beta^{\frac{\alpha}{\rho}} \left(\frac{C_{1,t}}{C_{1,t-1}} \right)^{-\frac{\alpha(\rho-1)}{\rho}} B_t^{\frac{\alpha}{\rho}-1} R_{i,t} \right] = 1, \ i = 1, .., N \tag{15}$$

where $\psi = 1/(1-\rho)$ is the intertemporal elasticity of substitution in consumption and $\gamma = 1 - \alpha$ determines the agent's behavior towards risk. In addition, B_t denotes the gross return on the optimum portfolio.

Factor model

We use the factor model to construct our portfolios. Consider the weak stationary vector of return $R_t = (R_{1,t}, ..., R_{p,t},)'$ of p assets at time period t with $E(R_t) = \mu$ and covariance matrix, $Cov(R_t) = \sum_R$. Therefore, the statistical factor model is in the form

$$R_t = \mu + \beta F_t + \varepsilon_t \tag{16}$$

where $F_t = (f_{1,t}, ..., f_{m,t})'$ is a vector of a reduced number of unobservable random variables ($m < p$), $f_{i,t}$ are the common factors, $\beta = [\beta_{ij}]_{p \times m}$ is the matrix of factor loadings, β_{ij} is the loading of the ith variable on the jth factor, $\varepsilon_t = (\varepsilon_{1t}, ..., \varepsilon_{pt},)'$ are p additional noises, such that, each ε_{it} is the specific error of $R_{i,t}$. Notice that unlike in the multivariate regression model, β and F_t are not observed. In this work we use an orthogonal factor model that satisfies the following assumptions:

1. $E(F_t) = 0$ and $Cov(F_t) = I_m$;
2. $E(\varepsilon_t) = 0$ and $Cov(\varepsilon_t) = D$ where, D is a $p \times p$ diagonal matrix ($D =$dig$\{\sigma_1^2, ..., \sigma_p^2\}$);
3. $Cov(F_t, \varepsilon_t) = E(F_t \varepsilon_t') = 0$, the $m \times p$ matrix of zeros

Under these hypothesis, it is possible to show that (see Johnson & Wichern, 1992):

1. $\sum_R = \beta\beta' + D$

   or

$$\begin{aligned} Var(R_{i,t}) &= \beta_{i1}^2 + ... + \beta_{im}^2 + \sigma_i^2 \\ Cov(R_{i,t}, R_{j,t}) &= \beta_{i1}\beta_{j1} + ... + \beta_{im}\beta_{jm} \end{aligned} \tag{17}$$

2. $Cov(R_t, F_t) = \beta$

   or $\qquad Cov(R_{i,t}, f_{j,t}) = \beta_{ij}$

where the quantity $\beta_{i1}^2 + ... + \beta_{im}^2$ is known as *communality* which is the portion of the variance of $R_{i,t}$ contributed by the m factors. The remaining portion σ_i^2 is called the uniqueness, or *specific variance*. Denoting the ith communality by h_i^2, produces

$$Var(r_{it}) = \ h_i^2 + \ \sigma_i^2, \qquad i = 1, 2, ..., p \tag{18}$$

In this way, we use R_t assets to estimate the factor-loadings and the m common factors among R_t.

Constructing the portfolios

We use the orthogonal factor model to estimate the portfolios used in GMM estimation. Let the factor model be:

$$(R_t)_{p\times 1} \ - (\mu)_{p\times 1} = (\beta)_{p\times m} (F_t)_{m\times 1} + (\varepsilon_t)_{p\times 1} \ , \qquad m << p \tag{19}$$

Proposition 1 *: Let the $p \times m$ matrix β of rank m and F_t, respectively, be the factor loading and the common factors in the factor model. There exists a left inverse matrix $\beta^{-1} = [b_{i,j}]$ such that $\beta_{m\times p}^{-1} \beta_{p\times m} = I_m$. Define the diagonal matrix $A = diag\{a_{11}, ..., a_{mm}\}$, where $a_{ii} = \frac{1}{\sum_{j=1}^{p} b_{i,j}}$ for all $i = 1, ...m$, where $b_{i,j}$ are the elements of the matrix β^{-1}. Therefore, the vector $\tilde{R}_t$ defined as $\left(\tilde{R}_t\right)_{m\times 1} = \left(A\beta^{-1}\right)_{m\times p} (R_t)_{p\times 1} = A\beta^{-1}\mu + AF_t + A\beta^{-1}\varepsilon_t$ is a vector of portfolios. In addition $\tilde{R}_t$ has the property $Cov(\tilde{R}_t) = I_m + \tilde{\beta}^{-1}D$, where $I_m + \tilde{\beta}^{-1}D$ is a $m-$diagonal matrix.*

Proof. : Since $p > m$, there exists a left inverse matrix $\beta^{-1} = [b_{i,j}]$ such that $\beta^{-1}_{m\times p}\,\beta_{p\times m} = I_m$. Pre-multiplying the factor model by $A\beta^{-1}$ produces

$$A\beta^{-1}R_t \;= A\beta^{-1}\mu + \tilde{F}_t \;+\; \tilde{\beta}^{-1}\varepsilon_t \tag{20}$$

where $\tilde{F}_t = AF_t = [\frac{F_{1t}}{a_{11}}, ..., \frac{F_{mt}}{a_{mm}}]'$ is the original factor scaled by constants. In addition, the sum of each row of the matrix $A\beta^{-1}$ is one. This is easy to see, since $A\beta^{-1}$with elements $[a_{ii}b_{ij}]$, thus the ith row is $\sum_{j=1}^{p} a_{ii}b_{ij} = a_{ii}\sum_{j=1}^{p} b_{i,j} = a_{ii}\times\frac{1}{a_{ii}} = 1$. Therefore, we define m portfolios as $\tilde{R}_t = A\beta^{-1}R_t$. Finally, an important property is shown. Given the hypothesis $E(F_t) = 0$ and $Cov(F_t) = I_m$, the mean and covariance of $\tilde{R}_t$ are, respectively, $E(\tilde{R}_t) = A\beta^{-1}\mu$ and $Cov(\tilde{R}_t) = E\left[(\tilde{F}_t \;+\; \tilde{\beta}^{-1}\varepsilon_t)(\tilde{F}_t \;+\; \tilde{\beta}^{-1}\varepsilon_t)'\right] = E(\tilde{F}_t\tilde{F}_t\,') \;+\; \tilde{\beta}^{-1}E(\varepsilon_t\varepsilon_t') = I_m + \tilde{\beta}^{-1}D$, where the $m\times m$ matrix $I_m + \tilde{\beta}^{-1}D$ is diagonal. ■

In the next section we show how to estimate the factor loading β, which will be used to estimate the matrix A. Thus, we use $\hat{A}$ to estimate the m portfolios as $\tilde{R}_t = \hat{A}\hat{\beta}^{-1}R_t$.

Factor Loading Estimation

Considering observations $R = (R_1', ..., R_T')'$ of p generally correlated variables. The orthogonal factor model (19) is estimated by the method of principal component analysis. We can estimate the unknown population covariance matrix $\sum_r$ by the sample covariance matrix $\widehat{\sum}_r$. The principal component method allows having eigenvalue-eigenvector pairs $(\hat{\lambda}_1, \hat{e}_1), ..., (\hat{\lambda}_k, \hat{e}_p)$ with $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq ...\hat{\lambda}_p \geq 0$. Let $m < p$ be the number of common factors. Then the matrix of factor loading is given by $\hat{\beta} = \left(\hat{\beta}_{ij}\right) = \left(\sqrt{\hat{\lambda}_1}\hat{e}_1 \mid \sqrt{\hat{\lambda}_2}\hat{e}_2 \mid ...\mid \sqrt{\hat{\lambda}_1}\hat{e}_m\right)$. The estimated specific variances are the diagonal elements of the matrix $\widehat{\sum}_r - \hat{\beta}\hat{\beta}'$. Thus, the error matrix caused by approximation is $\widehat{\sum}_r - \left[\hat{\beta}\hat{\beta}' + \hat{D}\right]$. For the case $m = p$, the specific variances are $\sigma_i^2 = 0$ for all i. The jth column of the loading matrix is given by $\sqrt{\hat{\lambda}_j}\hat{e}_j$. That is, we can write $\widehat{\sum}_r = \hat{\beta}\hat{\beta}' + 0 = \hat{\beta}\hat{\beta}'$. It is importan to note that the estimated factor loadings for a given factor do not change as the number of factors is increased.

3 Estimation and Testing

3.1 Data

We constructed a number of monthly data sets for the United States beginning in January 1970 and ending in September 2008. Monthly data have been used in previous studies, such as Epstein and Zin (1991) and Hansen and Singleton (1983). We examine two measures of per capita consumption: the expenditures on nondurable goods and nondurable and service goods together. We obtained the consumption measures from the FRED database.

Previous works used a limited number of asset returns. Hansen and Jaganathan (1983) used two assets; the S&P500 index and Treasury Bill as a risk-free rate of return. Epstein and Zin (1991) used five individual stock return indices reflecting value-weighted returns for broad groups of the standard industrial classification and the Treasury Bill return. In this work, we consider a larger set of asset returns, corresponding to the monthly S&P500 stock returns[4] to construct a small set of portfolios by using the factor model. We consider only companies for which data are available throughout the period, reducing the sample to $N = 106$ for the S&P500 dataset. We also consider U.S. Treasury Bill return as a measure of the risk-free asset and S&P500 index as a measure of the optimal portfolio. All nominal asset returns are converted to real returns using the respective price deflator for nondurable and for nondurable and services consumption goods.

3.2 Results

We test three specifications: the time-separable CRRA utility, a time non-separable utility that exhibits external habit formation (Abel, 1990) and the Kreps and Porteus (1978) specification. We use different sets of instruments that include a constant, consumption growth lagged, and portfolio lagged. The models are tested for three sets of asset returns; i) six factor portfolios constructed on S&P500 data; ii) the S&P500 index as studied by Hansen and Jaganathan (1983); and iii) the average return, by month, of 106 stocks. The average return is the simple arithmetic mean of a series of returns generated over a period of time. In all sets, the Treasury bill return is considered as a riskless asset return.

[4]The data were taken from the monthly tape of the Center for Research in Security Prices (CRSP) of the University of Chicago.

In Table 7 in the appendix we report a summary statistic of the portfolios constructed by factor model over the period 1970:1 to 2007:9. In addition, the correlation among these sets of asset returns are shown. The first portfolio with the S&P500 index and the average return are highly correlated, with respective correlation equal to 0.89 and 0.88. The others factor portfolios are orthogonal to each other by the factor model construction.

Table 8 in the appendix reports the cumulative total for the first 15 portfolios. The six factor portfolios considered account for 36.6% of the total variability of the original dataset when deflated by non durables goods and 36.2% when deflated by non durable and service goods. It is clear that whereas the six-factor portfolios explain around 37% of the variation, further factors can add only marginally to the explanatory power of the set of portfolios, as reported in Table 8. It is noteworthy that although 37% may appear to be a small part of the variation, it represents a set of portfolios that can explain on average 37% of the variation for 106 assets, which is a considerable achievement .

Tables 1 and 2 show the results of the estimation for the CRRA preference. Table 1 shows estimates for the S&P500 index and the factor portfolios. Table 2 shows the results for average return. Parameters that are not significant at 5% are identified by *. For the factor portfolios and nondurables and services consumption goods (right site of Table 1) the estimates of the relative risk aversion coefficient γ are in general significant, ranging from 0.026 to 0.338 for the factor portfolios. The estimates of the intertemporal discount rate β are significant and range from 0.998 to 0.999. For the factor portfolio and non durable goods, only the parameter β is significant and in general $\gamma > 0$. For the S&P500 index and nondurable and service goods (left side of Table 1), the parameter γ is not significant, ranging from -0.064 to 0.266. In Table 2, for the average return and nondurable and services goods, the estimates of γ range from 0.0129 to 0.3195. Again γ is not significant. In all cases the estimates of the intertemporal discount rate β exceed 0.99 but are less than unity. Indeed, the values of β and γ are significant only for nondurables and services consumption goods and for the set of factor portfolios. In summary, for the set of factor portfolios the estimates of the coefficient of relative risk aversion are $\gamma > 0$, which is according the theory. The p-value of the chi-square tests (TJ) is also displayed in Tables 1 and 2. For the set of factor portfolios, the identification test (TJ) does not reject any of the implicit over-identifying restrictions at 5%. Contrary to the S&P500 and average return, the TJ test is rejected. Therefore, these results support the evidence that using the factor model to build portfolios dominates the model when using only two assets (as estimated by Hansen and Jagannathan,

1983) in the CRRA specification.

Table 1 - Power Utility Function Estimation Results

$$E_{t-1}\left[\beta\left(\frac{C_{1,t}}{C_{1,t-1}}\right)^{-\gamma}R_{i,t}\right]=1,\ i=1,..,N$$

N = 2; $R_{1,t}$ = T-bill and $R_{2,t}$=S&P500 Bt=					N = 7; $R_{1,t}$ = T-bill and the portfolios F1t, F2t, F3t, F4t, F5t, F6t .			
Consumption *Proxy*	**Instruments** μ; ct-i/ct-1-i R2t-j	**β** *s.e.*	**ɣ** *s.e.*	**TJ** *p-value*	**Instruments** μ; ct-i/ct-1-i; F1t-j F2t-k; F3t-m	**β** *s.e.*	**ɣ** *s.e.*	**TJ** *p-value*
Non-durable and Services per capita	i=1	0.9988	0.0770*		i=1; j=1; k=1	0.9991	0.2011	
		0.0003	*0.1412*	*0.0062*		*0.0002*	*0.0835*	*0.0601*
	i=1,2	0.9987	0.0428*		i=1; j=1; k=2	0.9990	0.2196	
		0.0003	*0.0707*	*0.0142*		*0.0002*	*0.0868*	*0.0633*
	i=1; j=1	0.9988	0.1647*		i=1; j=1; m=1	0.9988	0.1129*	
		0.0003	*0.1017*	*0.0085*		*0.0002*	*0.0873*	*0.0615*
	i=1,2; j=1	0.9986	0.1093*		i=1; j=1; m=2	0.9989	0.1847	
		0.0003	*0.0620*	*0.0216*		*0.0003*	*0.0885*	*0.0533*
	i=1,2; j=1,2	0.9985	0.1133*		i=1; j=1	0.9992	0.1732	
		0.0003	*0.0619*	*0.0237*	k=1; m=1	*0.0002*	*0.0769*	*0.0703*
	i=1,2,3; j=1	0.9987	0.1096		i=1; j=1	0.9992	0.2535	
		0.0003	*0.0535*	*0.0234*	k=2; m=2	*0.0002*	*0.0849*	*0.0694*
Non-durable per capita	i=1	0.9985	-0.0066*		i=1; j=1; k=1	0.9985	0.0588*	
		0.0004	*0.0762*	*0.0018*		*0.0003*	*0.0681*	*0.0597*
	i=1,2	0.9985	-0.0192*		i=1; j=1; k=2	0.9985	0.0778*	
		0.0004	*0.0656*	*0.0061*		*0.0003*	*0.0672*	*0.0667*
	i=1; j=1	0.9984	0.0520*		i=1; j=1; m=1	0.9986	0.1214*	
		0.0004	*0.0659*	*0.0107*		*0.0003*	*0.0644*	*0.0648*
	i=1,2; j=1	0.9982	0.0268*		i=1; j=1; m=2	0.9986	0.1137*	
		0.0004	*0.0605*	*0.0181*		*0.0003*	*0.0630*	*0.0526*
	i=1,2; j=1,2	0.9981	0.0296*		i=1; j=1	0.9989	0.1476	
		0.0004	*0.0604*	*0.0213*	k=1; m=1	*0.0002*	*0.0621*	*0.0739*
	i=1,2,3; j=1	0.9982	0.0467*		i=1; j=1	0.9988	0.1430	
		0.0004	*0.0586*	*0.0210*	k=2; m=2	*0.0003*	*0.0597*	*0.0701*

Table 2 - Power Utility Function Estimation Results

N = 2; $R_{1,t}$ = T-bill and $R_{2,t}$=Rmean

Consumption	Instruments	β	γ	TJ
Proxy	μ; ct-i/ct-1-i R2t-j	*s.e.*	*s.e.*	*p-value*
	i=1	0.9989	0.1804*	
Non-durable		*0.0003*	*0.1391*	*0.0119*
	i=1,2	0.9987	0.0857*	
and		*0.0003*	*0.0728*	*0.0206*
	i=1; j=1	0.9988	0.1547*	
Services		*0.0003*	*0.1135*	*0.0125*
	i=1,2; j=1	0.9987	0.0985*	
per		*0.0003*	*0.0698*	*0.0223*
	i=1,2; j=1,2	0.9987	0.0878*	
capita		*0.0003*	*0.0678*	*0.0235*
	i=1,2,3; j=1	*0.9986*	0.1039*	
		0.0002	*0.0581*	*0.0222*
	i=1	0.9983	0.0172*	
Non-durable		*0.0004*	*0.0749*	*0.0098*
	i=1,2	0.9982	-0.0043*	
per		*0.0004*	*0.0723*	*0.0122*
	i=1; j=1	0.9982	0.0526*	
capita		*0.0004*	*0.0679*	*0.0131*
	i=1,2; j=1	0.9982	0.0438*	
		0.0004	*0.0686*	*0.0220*
	i=1,2; j=1,2	0.9982	0.0154*	
		0.0004	*0.0623*	*0.0284*
	i=1,2,3; j=1	0.9981	0.0074*	
		0.0004	*0.0607*	*0.0251*

Note: the Average return, Rmean, is the simple arithmetic mean of a series of returns generated over a period of time

Tables 3 and 4 contain results obtained by using the external habit preference. In all cases, the estimatives of γ and κ are not significant. Considering the S&P500 index and the nondurables and services consumption goods, the estimates of the time separability κ vary from -0.414 to 0.108 (upper part of Table 3). On the other hand, κ range from -0.676 to 0.161 for factor portfolios (lower part of Table 3). In Table 4, for the average return and nondurable and service goods, the estimates of κ range from -1.055 to 0.623. In all cases the parameter κ is not significant. By making $\kappa = 0$ in Abel preference we reproduce the CRRA utility. The p-value of the chi-square tests (TJ) are also displayed in Tables 3 and 4. They show that only for the case of factor portfolio is the TJ test not rejected for most

of the instruments.

Tables 5 and 6 report the results for the Kreps-Porteous preference. Using the Euler equations, we estimate the parameters β, α and ρ though GMM estimation. The coefficient of relative risk aversion γ and the intertemporal elasticity of substitution in consumption ψ are estimated indirectly by estimation of α and ρ, respectively, by using $\gamma = 1 - \alpha$ and $\psi = 1/(1-\rho)$. The delta method is used to estimate the standard error of these parameters. The parameters γ and ψ are not significant. Again, the results show that only for the case of factor portfolio is the TJ test not rejected for most of the instruments. The recursive utility model of Epstein and Zing performs the worst with US data.

4 Conclusion

This paper provides an empirical investigation of the testable restrictions implied by different representative classes of consumption-based asset pricing models (CCAPM). We used aggregated data on consumption to test the models in the CCAPM, as in Hansen & Singleton (1982, 1983) and Epstein & Zin (1991). Contrary to these studies, we considered a larger set of asset returns by incorporating the statistical factor model in the analysis. By using the factor model we can summarize the common information available in the stock market in a small set of portfolios. This has an economic interpretation, since consumption should only respond to systematic changes in returns, which is exactly the information contained in these portfolios. In addition, from an econometric point of view this method allows reducing the number of variables, making feasible the model estimation.

We tested three specifications; i) the CRRA preferences; ii) the external habit formation model of Abel (1990) and iii) the Kreps-Porteus preference representation specification (Epstein and Zin, 1991) for the utility function. The main results, which are shown in Tables 1 and 2, are obtained by the CRRA preference. When the factor portfolio is used, the implication of this model is not rejected. Indeed, the relative risk-aversion coefficient is significant and ranges from 0.026 to 0.338, while the discount factor is significant and ranges from 0.998 to 0.999.

Hansen & Singleton (1983) and Prescott & Mehra (1985) found risk aversion values of $\gamma \in [30; 60]$ and more recently Mulligan (2002) found values of $\gamma \in [0.6; 3.2]$. Our results show that the risk aversion varies in the interval $\gamma \in [0.026; 0.338]$ This result is in line with

Mulligan (2002) and with some results of real business cycle models ($\gamma < 2.5$). An extension of this work would be to study the equity premium puzzle in the framework of factor models using a non-linear GMM approach.

Appendix A. Utility Function Estimation Results

Table 3 - External Habit Utility Function Estimation Results

$$E_{t-1}\left[\beta\left(\frac{C_{1,t}}{C_{1,t-1}}\right)^{-\gamma}\left(\frac{C_{1,t-1}}{C_{1,t-2}}\right)^{-\kappa(\gamma-1)}R_{i,t}\right]=1,\ i=1,..,N$$

Consumption *Proxy*	**Instruments** μ; ct-i/ct-1-i R2t-j	**β** *s.e.*	**γ** *s.e.*	**κ**	**TJ** *p-value*
N = 2; $R_{1,t}$ = T-bill and $R_{2,t}$=S&P500					
	i=1; j=1	0.9995	0.4084*	-0.1467*	
Non-durable		*0.0008*	*0.2610*	*0.2046*	*0.0066*
	i=1; j=1,2	0.9996	0.4822	-0.1891*	
and		*0.0008*	*0.2407*	*0.2255*	*0.0087*
	i=1,2,3	0.9990	0.1506*	-0.0285*	
Services		*0.0007*	*0.2078*	*0.0781*	*0.0165*
	i=1; j=2,3	0.9992	0.3054*	-0.0746*	
per		*0.0009*	*0.3247*	*0.1828*	*0.0110*
	i=1,3; j=1	0.9990	0.1808*	-0.0034*	
capita		*0.0007*	*0.1858*	*0.0843*	*0.0124*
	i=1; j=1,2,3	0.9993	0.3349*	-0.0946*	
		0.0007	*0.1978*	*0.1229*	*0.0117*
	i=1,2	0.9990	0.2715*	-0.1294*	
Non-durable		*0.0007*	*0.3407*	*0.2099*	*0.0050*
	i=1; j=1,2	0.9994	0.5896	-0.5446*	
per		*0.0007*	*0.2655*	*0.5912*	*0.0039*
	i=1,2; j=1	0.9993	0.4995	-0.3500*	
capita		*0.0006*	*0.1840*	*0.2509*	*0.0084*
	i=1,2; j=2	0.9989	0.2775*	-0.1397*	
		0.0006	*0.2380*	*0.1509*	*0.0063*
	i=1,2; j=1,2	0.9993	0.4902	-0.3444*	
		0.0006	*0.1661*	*0.2158*	*0.0098*
	i=1,2,3; j=1	0.9993	0.4633	-0.2792*	
		0.0006	*0.1663*	*0.1923*	*0.0107*
N = 7; $R_{1,t}$ = T-bill and the factor portfolios F1t, F2t, F3t, F4t, F5t, F6t					
	μ; ct-i/ct-1-i; F1t-j, F2t-k; F3t-m				
	i=1,2; j=1	0.9992	0.3052*	-0.0893*	
Non-durable		*0.0007*	*0.3019*	*0.1497*	*0.0421*
	i=2; j=1; m=1	0.9977	-0.5828*	-0.0259*	
and		*0.0011*	*0.5345*	*0.0758*	*0.0530*
	i=1,2; j=1; k=1	0.9986	-0.0433*	0.0333*	
Services		*0.0005*	*0.1936*	*0.0468*	*0.0691*
	i=1, 2; j=1; k=2	0.9999	0.5692	-0.3634*	
per		*0.0005*	*0.2188*	*0.3129*	*0.0631*
	i=1; j=1,2; k=1	0.9976	-0.3372*	0.1314	
capita		*0.0005*	*0.1972*	*0.0299*	*0.0658*
	i=2; j=1; k=1,2	0.9999	0.3215	-0.3306*	
		0.0004	*0.1469*	*0.1696*	*0.0709*
	i=1; j=1; k=2	0.9995	0.6842	-0.8788*	
Non-durable		*0.0005*	*0.2961*	*1.1617*	*0.0619*
	i=1, 2; j=1	0.9989	0.3946	-0.2376*	
per		*0.0004*	*0.1894*	*0.1761*	*0.0379*
	i=1; k=1; m=1	0.9992	0.3583*	-0.1285*	
capita		*0.0008*	*0.4300*	*0.3379*	*0.0658*
	i=1,2; k=1	0.9985	0.0724*	-0.0376*	
		0.0005	*0.2368*	*0.0933*	*0.0470*
	i=1,2; j=1,2	0.9989	0.3678	-0.2139*	
		0.0004	*0.1731*	*0.1486*	*0.0523*
	i=1,2; j=1; k=1	0.9993	0.4199	-0.2590*	
		0.0004	*0.1808*	*0.1834*	*0.0665*

Table 4 - External Habit Utility Function Estimation Results

N = 2; $R_{1,t}$ = T-bill and $R_{2,t}$=Rmean				
Instruments	**β**	**γ**	**κ**	**TJ**
μ; ct-i/ct-1-i	*s.e.*	*s.e.*		*p-value*
R2t-j				
i=1; j=2	0.9974	-0.4246*	0.1069*	
	0.0016	*0.6306*	*0.0813*	*0.0125*
i=1,2; j=2	0.9982	-0.1353*	0.05167*	
	0.0013	*0.5105*	*0.0985*	*0.0210*
i=1,2,3	0.9990	0.1877*	-0.0329*	
	0.0007	*0.2034*	*0.0796*	*0.0197*
i=1; j=2,3	0.9982	-0.1086*	0.0613*	
	0.0011	*0.4026*	*0.0878*	*0.0132*
i=1,3; j=1	0.9986	0.0752*	0.0285*	
	0.0007	*0.2011*	*0.0732*	*0.0129*
i=1,2; j=1,2	0.9991	0.2589*	-0.0621*	
	0.0009	*0.3644*	*0.1603*	*0.0221*
i=1,2	0.9992	0.5244*	-0.3826*	
	0.0008	*0.4091*	*0.6165*	*0.0145*
i=1; j=1,3	0.9986	0.2317*	-0.1090*	
	0.0006	*0.2785*	*0.1828*	*0.0137*
i=1,2; j=1	0.9992	0.5377	-0.4113*	
	0.0006	*0.2249*	*0.3559*	*0.0169*
i=1,2; j=2	0.9986	0.2079*	-0.0959*	
	0.0006	*0.2418*	*0.1265*	*0.0169*
i=1,2; j=1,2	0.9989	0.4249	-0.2672*	
	0.0005	*0.2037*	*0.2068*	*0.0218*
i=1,2,3; j=1	0.9991	0.4038	-0.2591*	
	0.0005	*0.1744*	*0.1730*	*0.0188*

Note: the Average return, Rmean, is the simple arithmetic mean of a series of returns generated over a period of time

Table 5 - Kreps-Porteous Utility Function Estimation Results

Consumption Proxy	Instruments μ; ct-i/ct-1-l,R2t-j	β s.e.	γ s.e.	α s.e.	ρ s.e.	ψ s.e.	TJ p-value
N = 2; $R_{1,t}$ = T-bill and $R_{2,t}$=S&P500							
	i=1,2,3	0.9993	-0.2221	1.2221	1.2923	-3.4217	
Non-durable		*0.0006*	*0.1507*	*0.1507*	*0.1876*	*2.1959*	*0.0161*
	i=1,2; j=1	0.9989	-0.0804	1.0804	1.1099	-9.1002	
and		*0.0004*	*0.0744*	*0.0744*	*0.0783*	*6.4821*	*0.0160*
	i=1; j=1,2	0.9988	-0.1718	1.1718	1.1806	-5.5373	
Services		*0.0004*	*0.0978*	*0.0978*	*0.0983*	*3.0136*	*0.0106*
	i=1; j=1,3	0.9988	-0.1413	1.1413	1.1493	-6.6982	
per		0.0004	0.0902	0.0902	0.0889	*3.9878*	0.0084
	i=1,2; j=1,2	0.9988	-0.1042	1.1042	1.1246	-8.0252	
capita		*0.0003*	*0.0694*	*0.0694*	*0.0727*	*4.6851*	*0.0198*
	i=1,2,3	0.9990	-0.1093	1.1093	1.1402	-7.1315	
	j=1	*0.0004*	*0.0648*	*0.0648*	*0.0684*	*3.4788*	*0.0192*
	i=1; j=1	0.9986	0.0849	0.9151	0.9573	23.4340	
Non-durable		*0.0005*	*0.0950*	*0.0950*	*0.0877*	*48.1498*	*0.0026*
	i=1; j=1,2	0.9985	0.0614	0.9386	0.9789	47.3844	
per		*0.0005*	*0.0843*	*0.0843*	*0.0800*	*179.7010*	*0.0054*
	i=1,2; j=1	0.9986	0.1042	0.8958	0.9479	19.1883	
capita		*0.0005*	*0.0903*	*0.0903*	*0.0853*	*31.3891*	*0.0093*
	i=1,2,3	0.9994	0.2419	0.7581	1.0999	-10.0145	
		0.0011	*0.2414*	*0.2414*	*0.2750*	*27.5752*	*0.0085*
	i=1,2; j=1,2	0.9984	0.0822	0.9178	0.9657	29.1460	
		0.0005	*0.0831*	*0.0831*	*0.0792*	*67.3145*	*0.0120*
	i=1,2,3; j=1	0.9985	0.0589	0.9411	0.9912	113.4301	
		0.0005	*0.0814*	*0.0814*	*0.0767*	*987.3156*	*0.0161*
N = 7; $R_{1,t}$ = T-bill and the factor portfolios F1t, F2t, F3t, F4t, F5t, F6t							
	μ; ct-i/ct-1-i; F1t-j,F2t-k; F3t-m						
	i=1; j=1; k=1	0.9992	-0.1861	1.1861	1.2044	-4.8922	
Non-durable		*0.0003*	*0.0854*	*0.0854*	*0.0873*	*2.0885*	*0.0597*
	i=1; j=1; k=2	0.9992	-0.1566	1.1566	1.1862	-5.3714	
and		*0.0002*	*0.0946*	*0.0946*	*0.0964*	*2.7801*	*0.0632*
	i=1; j=1; m=1	0.9989	-0.0790	1.0790	1.0910	-10.9855	
Services		*0.0002*	*0.0913*	*0.0913*	*0.0912*	*11.0083*	*0.0616*
	i=1; j=1; m=2	0.9990	-0.1586	1.1586	1.1749	-5.7190	
per		*0.0003*	*0.0904*	*0.0904*	*0.0917*	*2.9991*	*0.0517*
	i=1; j=1; k=1; m=1	0.9994	-0.1071	1.1071	1.1384	-7.2244	
capita		*0.0002*	*0.0821*	*0.0821*	*0.0842*	*4.3925*	*0.0701*
	i=1; j=1; k=2; m=2	0.9994	-0.1860	1.1860	1.2283	-4.3796	
		0.0003	*0.0945*	*0.0945*	*0.0976*	*1.8714*	*0.0673*
	i=1; j=1; k=1	0.9987	0.0084	0.9916	1.0262	-38.2307	
Non-durable		*0.0003*	*0.0798*	*0.0798*	*0.0763*	*111.4678*	*0.0586*
	i=1; j=1; m=1	0.9988	-0.0698	1.0698	1.0949	-10.5383	
per		*0.0003*	*0.0728*	*0.0728*	*0.0696*	*7.7348*	*0.0650*
	i=1; j=1; m=2	0.9986	-0.0846	1.0846	1.1003	-9.9686	
capita		*0.0003*	*0.0700*	*0.0700*	*0.0661*	*6.5701*	*0.0520*
	i=1; j=2; m=1	0.9983	0.0023	0.9977	0.9973	374.2515	
		0.0005	*0.2505*	*0.2505*	*0.0933*	*13068.41*	*0.0486*
	i=1; j=1; k=1; m=1	0.9992	-0.0467	1.0467	1.0993	-10.0668	
		0.0002	*0.0685*	*0.0685*	*0.0679*	*6.8836*	*0.0739*
	i=1; j=1; k=2; m=2	0.9989	-0.0746	1.0746	1.1082	-9.2405	
		0.0003	*0.0692*	*0.0692*	*0.0657*	*5.6101*	*0.0695*

Table 6 - Kreps-Porteous Utility Function Estimation Results

N = 2; $R1_{,t}$ = T-bill and $R2_{,t}$ = Rav
Bt is the value-weighted index of stocks in the NYSE and American Stock Exchange (AMEX)

Consumption	**Instruments**	**β**	**γ**	**α**	**ρ**	**ψ**	**TJ**
Proxy	μ; ct-i/ct-1-i	*s.e.*	*s.e.*	*s.e.*	*s.e.*	*s.e.*	*p-value*
	R2t-j						
	i=1,2,3	0.9994	-0.2160	1.2160	1.2627	-3.8061	
Non-durable		0.0006	0.1559	0.1559	0.1876	*2.7180*	*0.0200*
	i=1,2; j=1	0.9988	-0.1123	1.1123	1.1228	-8.1449	
and		0.0003	0.0772	0.0772	0.0811	*5.3805*	*0.0219*
	i=1; j=1,2	0.9988	-0.1115	1.1115	1.1183	-8.4515	
Services		0.0003	0.1044	0.1044	0.1052	*7.5149*	*0.0141*
	i=1; j=1,3	0.9987	-0.1318	1.1318	1.1279	-7.8168	
per		0.0003	0.0986	0.0986	0.0976	*5.9640*	*0.0130*
	i=1,2; j=1,2	0.9989	-0.1109	1.1109	1.1239	-8.0741	
capita		0.0003	0.0771	0.0771	0.0810	*5.2777*	*0.0222*
	i=1,2,3	0.9988	-0.1092	1.1092	1.1197	-8.3566	
	j=1	0.0003	0.0624	0.0624	0.0648	*4.5225*	*0.0220*
	i=1; j=1	0.9984	0.0239	0.9761	1.0044	-225.39	
Non-durable		0.0004	0.0934	0.0934	0.0857	*4354.84*	*0.0106*
	i=1; j=1,2	0.9986	0.0781	0.9219	0.9668	30.1133	
per		0.0005	0.0967	0.0967	0.0924	*83.7784*	*0.0150*
	i=1,2; j=1	0.9985	0.0412	0.9588	0.9977	429.45	
capita		0.0005	0.0849	0.0849	0.0796	*14685.91*	*0.0207*
	i=1,2,3	0.9985	0.0511	0.9489	0.9858	70.3594	
		0.0005	0.0953	0.0953	0.0783	*387.7293*	*0.0187*
	i=1,2; j=1,2	0.9986	0.0502	0.9498	0.9943	175.80	
		0.0004	0.0858	0.0858	0.0817	*2526.11*	*0.0214*
	i=1,2,3; j=1	0.9985	0.0501	0.9499	0.9854	68.7027	
		0.0004	0.0818	0.0818	0.0775	*365.89*	*0.0212*

Note: the Average return, Rav, is the simple arithmetic mean of a series of returns generated over a period of time

Appendix B. Descriptive Statistics

Table 7 - Descriptive Statistics

Return deflated using the consumer price index for nondurable and service goods							Consumption growth	Equally-weighted stock returns	Index
	Factors Portfolio								
	F1	F2	F3	F4	F5	F6	ct/ct-1	Rmean	S&P500
Proportion variance	20.81%	27.69%	30.65%	32.71%	34.47%	36.18%			
Mean	0.9967	0.9977	0.9804	0.9958	0.9899	1.0013	1.0019	0.9967	1.0031
Median	0.9996	0.9990	0.9239	1.0491	0.9829	0.9964	1.0019	0.9993	1.0061
Maximum	1.1627	1.2826	3.5862	4.8949	3.0276	1.6863	1.0162	1.1607	1.1546
Minimum	0.7242	0.6576	-2.0789	-2.3510	-0.2289	0.1966	0.9886	0.7673	0.7802
Std. Dev	0.0482	0.0862	0.7082	0.7391	0.3497	0.2031	0.0035	0.0437	0.0439
Skewness	-0.5944	-0.0049	0.2572	0.0487	0.4490	-0.1597	0.0743	-0.4898	-0.3984
Kurtosis	5.7171	3.8501	4.1524	6.6803	6.7546	4.5782	4.2140	5.3398	4.7518
Freq. Jarque-Bera (pvalue)	0.0000	0.0012	0.0000	0.0000	0.0000	0.0000	**0.0000**	0.0000	0.0000
Return deflated using the consumer price index for nondurable goods									
	F1	F2	F3	F4	F5	F6	ct/ct-1	Rmean	S&P500
Proportion variance	21.38%	28.19%	31.15%	33.19%	34.94%	36.64%			
Mean	0.9968	0.9979	0.9741	0.9956	0.9901	1.0015	1.001216449	0.9969	1.0033
Median	0.9987	0.9990	0.8889	1.0470	0.9807	0.9986	1.000847866	0.9988	1.0049
Maximum	1.1646	1.2929	4.6883	4.6832	3.0546	1.6996	1.031381836	1.1624	1.1553
Minimum	0.7256	0.6550	-3.2773	-2.1694	-0.2761	0.1758	0.973104816	0.7678	0.7807
Std. Dev	0.0488	0.0877	0.9886	0.6959	0.3570	0.2076	0.006781101	0.0444	0.0446
Skewness	-0.5895	0.0131	0.2665	0.0382	0.4342	-0.1767	0.082956483	-0.4893	-0.3963
Kurtosis	5.5752	3.8523	4.1434	6.7419	6.6713	4.6072	4.480934104	5.2210	4.6392
Freq. Jarque-Bera (pvalue)	0.0000	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Return deflated using the consumer price index for nondurable and service goods								
	SP500	Rmean	F1	F2	F3	F4	F5	F6
SP500	1.000							
Rmean	0.883	1.000						
F1	0.893	0.992	1.000					
F2	-0.082	0.117	0.000	1.000				
F3	0.013	0.012	0.000	0.000	1.000			
F4	0.066	0.007	0.000	0.000	0.000	1.000		
F5	0.091	0.015	0.000	0.000	0.000	0.000	1.000	
F6	0.028	0.024	0.000	0.000	0.000	0.000	0.000	1.000
Return deflated using the consumer price index for nondurable goods								
	SP500	Rmean	F1	F2	F3	F4	F5	F6
SP500	1.000							
Rmean	0.887	1.000						
F1	0.896	0.993	1.000					
F2	-0.081	0.114	0.000	1.000				
F3	0.015	0.009	0.000	0.000	1.000			
F4	0.063	0.008	0.000	0.000	0.000	1.000		
F5	0.091	0.014	0.000	0.000	0.000	0.000	1.000	
F6	0.027	0.023	0.000	0.000	0.000	0.000	0.000	1.000

Note: Summary statistic of the portfolios constructed by factor model over the period 1970:1 to 2007:9. We observe that the first portfolio and the S&P500 index are highly correlated with correlation equal to 0.89. The others portfolio are orthogonal each other by factor model construction. The Average return, Rmean, is the simple arithmetic mean of a series of returns generated over a period of time.

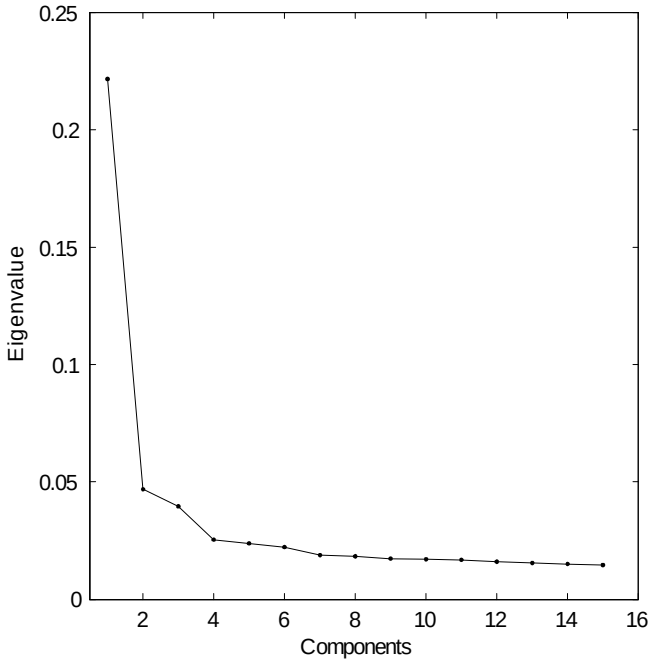
Appendix C. Analysis of Factor Model

Table 8 - Cumulative Explained Variation on S&P500

No of Portfolios	Return deflated using Nondurable and Services goods	Return deflated using Nondurable goods
1	0.208	0.214
2	0.277	0.282
3	0.306	0.311
4	0.327	0.332
5	0.345	0.349
6	0.362	0.366
7	0.378	0.383
8	0.393	0.397
9	0.408	0.412
10	0.422	0.426
11	0.436	0.440
12	0.449	0.453
13	0.461	0.465
14	0.474	0.478
15	0.486	0.490

Note: the first six factors explain around 37% of the variation, further factors can add only marginally to the explanatory power of the set of portfolios .

Figure 1 - A Scree plot for a 106 stock on S&P500



Note: The number of factor portfolios is determinated by the number of principal components. A useful visual aid to determining an appropriate number of principal components is the scree plot. The number of components is taken to be the point at which the remaining eigenvalues are relatively small and all about the same size. An elbow occurs in the plot in these figures at about 3. We consider in this study the first six components to test the consumption models. This picture corresponde to 106 stock returns which are deflated by the non durables and service good deflator.

References

[1] Attanasio,O. & Browning, M. (1995), "Consumption over the Life Cycle and over the Business Cycle". American Economic Review, 85(5):1118-1137.

[2] Campbell, J. & Deaton, A. (1989), "Why is Consumption so Smooth?". Review of Economic Studies, 56:357-374.

[3] Campbell, J. (1987), "Does Saving Anticipate Declining Labor Income? An AlternativeTest of the Permanent Income Hypothesis". Econometrica, 55(6):1249-73.

[4] Flavin, M. (1981), "The Adjustment of Consumption to Changing Expectations About Future Income". Journal of Political Economy, 89(5):974-1009.

[5] Hall, R. (1978), "Stochastic Implications of the Life Cycle Permanent Income Hypothesis: Theory and Evidence". Journal of Political Economy, 86(6):971-987.

[6] Hansen, Lars Peter & Singleton, Kenneth J, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," Econometrica, Econometric Society, vol. 50(5), pages 1269-86. 1982.

[7] Hansen, L.P., Singleton, K.J. "Stochastic Consumption, Risk-Aversion, and the Temporal Behavior of Asset Returns." Journal of Political Economy 91: 249-265 1983.

[8] Hansen, L. P., Singleton, K. J., 1984, Erratum: Generalized instrumental variables estimation of nonlinear rational expectations models, Econometrica 52, 267–268.

[9] Epstein, L., and S. Zin (1989): "Substitution Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, Econometrica, 57, 937–968.

[10] Epstein, L. and Zin, S. (1991), "Substitution, risk aversion and the temporal behavior of consumption and asset returns: an empirical investigation," Journal of Political Economy 99, 263 - 286.

[11] Forni, M.; Hallin, M.; Lippi, M. & Reichlin, L. (2000), "The Generalized Dynamic Factor Model: Identification and Estimation". The Review of Economics & Statistics, 2000, vol. 82, issue 4, pages 540-554.

[12] Johnson, R.A., Wichern, D.W., 1992. Applied multivariate statistical analysis. Third edition. New Jersey: Prentice-Hall.

[13] Lucas,R. E., JR.:"Asset Prices in an Exchange Economy," Econometrica, 46(1978), 1429-1446.

[14] Lusardi, A. (1996), "Permanent Income, Current Income and Consumption: Evidence from Two Panel Datasets". Journal of Business and Economics Statistics, 14(1):81-90

[15] Mulligan, C. (2002), "Capital, Interest, and Aggregate Intertemporal Substitution during the 20th Century". Mimeo, University of Chicago.

[16] Brock,W. A.: "An Integration of Stochastic Growth Theory and the Theory of Finance," Manuscript, University of Chicago, 1978.

[17] Prescott, E. C., & R. Mehra: "Recursive Competitive Equilibrium: The Case of Homogeneous Households," Econometrica, 48(1980), 1365- 1380.

[18] Runkle, D. (1991), "Liquidity Constraints and the Permanent Income Hypothesis: Evidence from Panel Data". Journal of Monetary Economics, 27(1):73-98.