

UNIVERSITY OF SÃO PAULO
School of Economics, Business and Accounting of Ribeirão Preto
Department of Economics
Graduate Program in Economics - Field: Applied Economics

CAIO AUGUSTO VIGO PEREIRA

Portfolio efficiency tests with conditioning information using
empirical likelihood estimation

Advisor: Prof. Dr. Márcio Poletti Laurini

Ribeirão Preto
2016

Prof. Dr. Marco Antonio Zago
Rector of University of São Paulo

Prof. Dr. Dante Pinheiro Martinelli
Dean of School of Economics, Business and Accounting of Ribeirão Preto

Prof. Dr. Renato Leite Marcondes
Head of Department of Economics

Prof. Dr. Cláudio Ribeiro de Lucinda
Coordinator of the Graduate Program in Economics - Field: Applied Economics

CAIO AUGUSTO VIGO PEREIRA

Portfolio efficiency tests with conditioning information using
empirical likelihood estimation

Dissertation presented to the Graduate Program in Economics - Field: Applied Economics from the School of Economics, Business and Accounting of Ribeirão Preto at University of São Paulo, in order to obtain the title of Master of Science. Corrected Version. The original is available at FEA-RP/USP.

Advisor: Prof. Dr. Márcio Poletti Laurini

Ribeirão Preto

2016

I authorize the reproduction and total or partial disclosure of this work, by any conventional or electronic means, for the purpose of study and research, provided the source is acknowledged.

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Pereira, Caio Augusto Vigo

Portfolio efficiency tests with conditioning information using empirical likelihood estimation. Ribeirão Preto, 2016

98 p. : il.; 30 cm.

Dissertação de Mestrado, apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo. Área de concentração: Economia Aplicada.

Orientador: Laurini, Márcio Poletti

1. Portfólio Eficiente. 2. Informação Condicional 3. Testes de Eficiência. 4. GEL.

APPROVAL SHEET

Name: PEREIRA, Caio Augusto Vigo

Title: Portfolio efficiency tests with conditioning information using empirical likelihood estimation

Dissertation presented to the Graduate Program in Economics - Field: Applied Economics from the School of Economics, Business and Accounting of Ribeirão Preto at University of São Paulo, in order to obtain the title of Master of Science. Corrected Version. The original is available at FEA-RP/USP.

Approved date: _____

COMMITTEE

Prof. Dr. **Márcio Poletti Laurini (Orientador)** Signature: _____

FEA-RP/USP

Prof. Dr. **Fábio Augusto Reis Gomes** Signature: _____

FEA-RP/USP

Prof. Dr. **André Alves Portela Santos** Signature: _____

Universidade Federal de Santa Catarina

Prof. Dr. **Bruno Cara Giovannetti** Signature: _____

FEA/USP

Acknowledgements

First of all, I thank all my family for their support in all decisions I made in my life. In particular I would like to thank my parents Claudiné and Maria Tereza, my brother Cezar Augusto, and my sister Ana Luiza.

I want to thank the professor, researcher and advisor Márcio Poletti Laurini. I have a debt of gratitude to him for his advices, incentives and willingness to teach. His interest in always encourage me to seek the contact with new knowledge was extremely important.

I also would like to thank professors Cláudio Ribeiro Lucinda and Fábio Augusto Reis Gomes for the relevant comments and valuable observations both in qualifying and in the pre-defense stage.

To all the professors and staff from the School of Economics, Business and Accounting of Ribeirão Preto. Throughout my academic trajectory I had the opportunity to work on some projects with several professors of the department. I want to thank prof. Sérgio Kannebley Júnior for the lessons and words of motivation, and also to prof. Renato Leite Marcondes for his support and willingness to help. I also would like to express my gratitude to the professors Luiz Guilherme Scorzafave and Daniel Domingues dos Santos.

I also thank all my friends and classmates who shared with me the journey of study and learning throughout the Masters.

Finally, I thank CAPES (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*) for the financial support that enabled the development of this work.

Abstract

PEREIRA, C.A.V. **Portfolio efficiency tests with conditioning information using empirical likelihood estimation.** 2016. 98 p. Dissertation (Master Degree). School of Economics, Business and Accounting of Ribeirão Preto, University of São Paulo, Ribeirão Preto, 2016.

We evaluate the use of Generalized Empirical Likelihood (GEL) estimators in portfolios efficiency tests for asset pricing models in the presence of conditional information. Estimators from GEL family presents some optimal statistical properties, such as robustness to misspecification and better properties in finite samples. Unlike GMM, the bias for GEL estimators do not increase as more moment conditions are included, which is expected in conditional efficiency analysis. We found some evidences that estimators from GEL class really performs differently in small samples, where efficiency tests using GEL generate lower estimates compared to tests using the standard approach with GMM. With Monte Carlo experiments we see that GEL has better performance when distortions are present in data, especially under heavy tails and Gaussian shocks.

Keywords: Portfolio Efficiency. Conditional Information. Efficiency Tests. GEL.

JEL Classification: C12, C13, C58, G11, G12

Resumo

PEREIRA, C.A.V. **Testes de eficiência com o uso de informação condicional em portfólios com estimação por verossimilhança empírica.** 2016. 98 f. Dissertação (Mestrado). Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2016.

Neste estudo avaliamos o uso de estimadores Generalized Empirical Likelihood (GEL) em testes de eficiência de portfólios para modelos apreçamento de ativos na presença de informação condicional. Estimadores da família GEL apresentam algumas propriedades estatísticas ótimas, tais como robustez à má especificação e melhores propriedades em amostras finitas. Diferentemente do GMM, o viés dos estimadores GEL não aumenta conforme se incluem mais condições de momentos, o que é esperado na análise de eficiência condicional. Encontramos algumas evidências de que os estimadores da classe GEL realmente performam diferentemente em amostras finitas, em que testes de eficiência com o uso do GEL geram estimativas menores comparadas aos testes com o uso da abordagem padrão com GMM. Através dos experimentos de Monte Carlo vemos que o GEL possui melhor performance quando distorções estão presentes nos dados, especialmente sob heavy tails e choques Gaussianos.

Palavras-chave: Portfólio Eficiente. Informação Condicional. Testes de Eficiência. GEL.

Classificação JEL: C12, C13, C58, G11, G12

List of Figures

Figure 1 – Historical series of the instruments for 720 months (from jan-1955 to dec-2014)	44
Figure 2 – Historical series of the factors for 720 months (from jan-1955 to dec-2014)	45
Figure 3 – Descriptive statistics of the monthly returns for the portfolio with 25 assets for a period of 720 months (60 years) from jan-1955 to dec-2014	48
Figure 4 – Descriptive statistics of the monthly returns for the portfolio with 49 assets for a period of 720 months (60 years) from jan-1955 to dec-2014	49
Figure 5 – Descriptive statistics of the monthly returns for the portfolio with 100 assets for a period of 720 months (60 years) from jan-1955 to dec-2014	50
Figure 6 – Boxplots for comparison of estimations by GMM and GEL using 6 portfolios formed on size and book-to-market under <i>scaled returns</i> for selected periods of time	55
Figure 7 – CAPM Model - Comparison of GMM and GEL estimated betas under <i>scaled returns</i> against the sample mean of monthly excess returns for the portfolio with 6 assets	57
Figure 8 – Boxplots for comparison of estimations by GMM and GEL using 25 portfolios formed on size and book-to-market for selected periods of time	60
Figure 9 – Boxplots for comparison of estimations by GMM and GEL for CAPM model using 49 industry portfolios for selected periods of time	65
Figure 10 – Simulated scenario 1 with Gaussian innovations ($\hat{\varepsilon}_{i,t}^{\text{Sim}^*} = \hat{\xi}_{i,t}^{\text{Sim}1}$) in Wald and GRS tests (Model=Fama-French, N=6, T=120, 500 simulations) .	79
Figure 11 – Simulated scenario 2 with shocks from a t distribution ($\hat{\varepsilon}_{i,t}^{\text{Sim}^*} = \hat{\nu}_{i,t}^{\text{Sim}2}$) in Wald and GRS tests (Model=Fama-French, N=6, T=120, 500 simulations)	81
Figure 12 – Simulated scenario 3 with shocks at $t=T/2$ defined by $\hat{\varepsilon}_{i,t}^{\text{Sim}^*} = \mathbb{1}_{t=T/2}(\hat{\kappa}_{i,t}^{\text{Sim}3})$ in Wald and GRS tests (Model=Fama-French, N=6, T=120, 500 simulations)	83

Figure 13 – Simulated scenario 4 with shocks defined by $\hat{\varepsilon}_{i,t}^{\text{Sim}^*} = \hat{\xi}_{i,t}^{\text{Sim}4} - \mathbb{1}_{\hat{p}_{i,t} < 0.05}(\hat{\kappa}_{i,t}^{\text{Sim}4})$
in Wald and GRS tests (Model=Fama-French, N=6, T=120, 500 simu-
lations) 84

Figure 14 – Boxplots for comparison of estimations by GMM and GEL for Fama-
French model using 6 portfolios formed on size and book-to-market
under *scaled returns* for selected periods of time 92

Figure 15 – Boxplots for comparison of estimations by GMM and GEL for Fama-
French model using 25 portfolios formed on size and book-to-market
for selected periods of time 92

Figure 16 – CAPM Model - Comparison of GMM and GEL estimated betas against
the sample mean of monthly excess returns for the portfolio with 25 assets 93

Figure 17 – CAPM Model - Comparison of GMM and GEL estimated betas against
the sample mean of monthly excess returns for the portfolio with 49 assets 94

List of Tables

Table 1 – Descriptive statistics of the lagged variables and factors for a period of 720 months (60 years) from jan-1955 to dec-2014	46
Table 2 – Descriptive statistics of the monthly returns for the portfolio with 6 assets for a period of 720 months (60 years) from jan-1955 to dec-2014	47
Table 3 – Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market for selected periods of time	53
Table 4 – Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market under <i>scaled returns</i> for selected periods of time	54
Table 5 – Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market for selected periods of time	59
Table 6 – Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market under <i>scaled returns</i> for selected periods of time	61
Table 7 – Tests of portfolio efficiency using 49 industry portfolios for selected periods of time	63
Table 8 – Portfolio efficiency tests in terms of maximum Sharpe ratios for all 4 portfolios to selected periods of time	66
Table 9 – Tests of portfolio efficiency (VK method) using 6 portfolios formed on size and book-to-market for selected periods of time	69
Table 10 – Tests of portfolio efficiency (VK method) using 6 portfolios formed on size and book-to-market under <i>scaled returns</i> for selected periods of time	70
Table 11 – Tests of portfolio efficiency (VK method) using 25 portfolios formed on size and book-to-market for selected periods of time	71
Table 12 – Tests of portfolio efficiency (VK method) using 25 portfolios formed on size and book-to-market under <i>scaled returns</i> for selected periods of time	72
Table 13 – Tests of portfolio efficiency (VK method) using 49 industry portfolios for selected periods of time	73

Table 14 – Tests of portfolio efficiency (VK method) using 49 industry portfolios under <i>scaled returns</i> for selected periods of time	74
Table 15 – Tests of portfolio efficiency (VK method - two-step) using 6 portfolios formed on size and book-to-market under <i>scaled returns</i> for selected periods of time	95
Table 16 – Tests of portfolio efficiency (VK method - two-step) using 25 portfolios formed on size and book-to-market for selected periods of time	96
Table 17 – Tests of portfolio efficiency (VK method - two-step) using 25 portfolios formed on size and book-to-market under <i>scaled returns</i> for selected periods of time	96
Table 18 – Tests of portfolio efficiency (VK method - two-step) using 49 industry portfolios for selected periods of time	97
Table 19 – Rejections proportions of J_{Wald} and J_{GRS} tests for nominal sizes of 1%, 2.5%, 5% e 10% from the Monte Carlo experiments for each of the 4 scenarios (Model=Fama-French, N=6, T=120, 500 simulations)	98

Contents

1	Introduction	19
2	Literature Review	22
3	Methodology	24
3.1	Incorporating Conditional Information	24
3.2	Estimation Methodology	26
3.2.1	Empirical Likelihood Estimation	28
3.2.2	HAC Estimation	31
3.2.3	Fixed-b Estimators	33
3.3	Tests of Efficiency	35
3.4	Empirical Strategy	41
4	Data	43
4.1	Instruments and Factors	43
4.2	Portfolios	46
5	Results	51
5.1	Conventional HAC Estimation	51
5.1.1	6 Portfolios Formed on Size and Book-to-Market	51
5.1.1.1	No Conditional Information	52
5.1.1.2	<i>Managed Portfolios</i>	53
5.1.2	25 Portfolios Formed on Size and Book-to-Market	56
5.1.2.1	No Conditional Information	56
5.1.2.2	<i>Managed Portfolios</i>	60
5.1.3	49 Industry Portfolios	62
5.1.3.1	No Conditional Information	62
5.1.4	100 Portfolios Formed on Size and Book-to-Market	65
5.2	Tests of Efficiency in Squared Sharpe ratio Form	65

5.3	Fixed-b HAC Estimators Methodology	67
5.3.1	6 Portfolios Formed on Size and Book-to-Market	68
5.3.1.1	No Conditional Information	68
5.3.1.2	<i>Managed Portfolios</i>	69
5.3.2	25 Portfolios Formed on Size and Book-to-Market	70
5.3.2.1	No Conditional Information	70
5.3.2.2	<i>Managed Portfolios</i>	71
5.3.3	49 Industry Portfolios	72
5.3.3.1	No Conditional Information	72
5.3.3.2	<i>Managed Portfolios</i>	73
5.4	Evaluating Robustness with Monte Carlo Simulations	75
5.4.1	Sampling Distributions of the Test Statistics	77
6	Conclusions	85
	Bibliography	87
	Appendix	91
	APPENDIX A – Additional Graphics	92
	APPENDIX B – Additional Tables	95

1 Introduction

In financial economics, the pivotal works of Markowitz (1952) and Sharpe (1964) on portfolio allocation are fundamental pieces that influenced a whole range of subsequent studies. With these works, we had the emergence of so-called Modern Portfolio Theory (MPT), in which the mean-variance structure plays a central role. The mean-variance structure had a significant impact in a number of areas, having influence not only in the finance field, in which it has been used in portfolio analysis, asset pricing and corporate finance; but also on analysis of economic policy under uncertainty, labor markets, monetary policy, as well as in hedging and even in inventory problems.

The assessment of investments funds performance is of major importance in this context. Not only for investors, but also for fund managers, in view that his remuneration may be related with such performance. The efficiency of financial allocations plays a key role in empirical asset pricing framework, with theoretical and practical importance in financial markets.

More recently, approaches to examine efficiency tests under the conditional point of view has been quickly developing (e.g. Ferson and Siegel (2009)), in contrast to tests that are performed in the unconditional form. Given that the true values are not observable (e.g. expected return and volatility), then these parameters must be estimated in some way and the results analysis should be based on statistical inference.

This study uses an alternative econometric method based on Empirical Likelihood estimators for comparing the portfolios efficiency. This class of estimators has some special characteristics that confer it to have better statistical properties, such as robustness to outliers and heavy tails distributions, and better finite and asymptotic properties compared to the usual methods based on least squares and Generalized Method of Moments.

According to Chaussé (2010), financial data, in particular the return of stocks, commonly presents heavy-tailed and asymmetric distributions. As the GMM estimators

does not impose any restrictions on the data distribution, only being based on assumptions about the moments, this method is widely used in finance. Cochrane (2009) also says that GMM structure fits naturally for the stochastic discount factors formulation of asset pricing theories, due to the easiness on the use of sample moments in the place of population moments. However, GMM estimators can be suboptimal in finite samples. In special, in two-stage and iterated GMM estimators are affected by the presence of a finite sample bias component proportional to the number of moment conditions, and the use of higher order moment conditions makes these estimators sensitive to the presence of outliers in distributions with heavy tails (ANATOLYEV; GOSPODINOV, 2011; FERSON; FOERSTER, 1994; NEWEY; SMITH, 2004).

This study analyses the use of Generalized Empirical Likelihood (GEL) to circumvent the deficiencies existing on the use of usual estimators in testing portfolio efficiency in the presence of conditional information.

Primarily, we evaluate how efficiency tests based on GEL and GMM estimations can lead to different decisions. We analyze this issue by comparing the test results for different sample sizes and portfolios types, for two asset pricing models, as well as estimations with and without conditional information. The results indicated that, in general, efficiency tests using GEL generate lower estimates when compared to tests using the standard approach based on GMM. Moreover, for the smallest sample in use, which is the one that most resembles features of a finite sample size used in finance, we see that the efficiency tests results are conflicting among GEL and GMM methodologies.

In the second part of the work we discuss the influence of the covariance matrix estimation in the results of efficiency tests, comparing the use of the standard long-run covariance HAC matrix estimators with the fixed bandwidth estimators, to overcome the problems caused by serial correlation, heteroskedasticity and observation or measurement errors that are common in financial data.

We also study tests robustness with the use of GMM and GEL estimators in a finite sample context. With Monte Carlo experiments we assess the effects that data contaminations, as outliers and the presence of heavy tails in the innovation structure, may

cause on the results of efficiency tests. In general, we see that GEL has better performance when heavy tails are present. While, regarding the presence of outliers, both GMM and GEL may have better robustness depending on the DGP we choose to use. We also see evidences that, under the null hypothesis, the tests using either GEL or GMM estimators have a tendency to over-reject the hypothesis of efficiency in finite samples.

All these results gathered may be an evidence that efficiency tests based on estimators from GEL class perform differently when compared to GMM, especially under small samples. These results extend the findings from Almeida and Garcia (2008) who obtain nonparametric estimates of bounds on the stochastic discount factors accounting for higher moments in the distribution of returns. The same authors developed an econometric method based on GEL that provides consistent estimators of information bounds and specification-error bounds that are based on the minimum discrepancy measures (ALMEIDA; GARCIA, 2012).

The structure of this work is as follows. Next section briefly review the literature, contextualizing the topic under evaluation. Section 3 introduces the methodology, presenting the asset pricing theory and models, as well as the econometric models for portfolios efficiency tests for both estimation methods assessed: GMM and GEL. Section 4 presents the data used: portfolios, factors and instruments. Section 5 presents the results of the efficiency tests performed for all portfolios, asset pricing models and estimations methodologies. It also provides the simulations experiments to evaluate tests robustness under both methods of estimation. Finally, section 6 concludes. Additional tables and figures are presented in Appendix.

2 Literature Review

In the investment decisions and portfolios evaluation literature, we have in the theory introduced by Markowitz (1952), and complemented by Sharpe (1964), a grounding milestone in the asset allocation field, with impacts and consequences still present in the most recent contributions in the economic finance area. Both works, given the influence and diffusion that had, gave rise to what today is known as Modern Portfolio Theory (MPT). In his study, Markowitz (1952) shows that investors face a trade-off between gains and risk, which respectively are the first two moments of the returns distribution: mean and variance. Based on this, the investor could choose the best portfolios, i.e., those with low risk for a given return, obtaining what would be the efficient frontier.

Prior to the MPT, the construction of portfolios were mostly based on the returns, taking the risk a secondary figure. With Markowitz theory, the focus on returns remained, but the risk was elevated to a similar importance given to the returns. With the MPT, we have a clear demonstration that with assets diversification it was possible improving the reduction of variance, thereby building portfolios that as a whole have better risk-return ratio, instead of building portfolios taking each asset and their characteristics individually. In short, with Markowitz's theory it was possible to understand that no positive marginal expected return can be achieved without an impact in the portfolio's risk characteristics. It is worth mentioning that, in general, under this approach it is imposed the assumption that investors have quadratic utility function. Thus, the expected utility should be a function of the first two moments of the distribution, which implies that the mean-variance framework would be sufficient to describe investors behavior.

With the broad concept of assets diversification derived from Markowitz, the question of optimizing portfolios has taken shape, and in recent decades different methods have emerged influenced by the MPT. The mean-variance framework for risk-return evaluation guided the main measures and indexes that have come to light with the purpose of assisting investors in asset performance assessment and portfolio construction. Sharpe

(1966), Treynor (1965), Jensen (1968), Sortino and Price (1994) and many others developed measurements for portfolios evaluation.

From Sharpe (1966), we have the most traditional and widely used measure for portfolios evaluation. Known as Sharpe ratio, it is given by: $\frac{E(R_t) - R_f}{\sigma}$. The numerator of this ratio can be understood as the portfolio risk premium, measured by the difference between expected return ($E(R_t)$) and the risk-free asset (R_f) representing the investor's opportunity cost, i.e., a benchmark set by himself. So, the Sharpe ratio evaluates the net expected return of the portfolio in relation to its risk (σ). Note that this measure implies that there is a uniform weighting by the investor for returns above and below the sample mean.

When we focus in the problem of testing portfolios efficiency, we step into a form of optimization problem. The optimization solution of the mean-variance structure can be made either conditionally on a set of information, as unconditionally too. Ferson and Siegel (2001) define that the conditional information is present when the optimal solution may be a function of the information received from the future returns probability distribution.

Among several studies, Chen and Knez (1996) and Ferson and Schadt (1996) show that conditioning in the state of the economy is statistically and economically significant for investment performance measurement. Christopherson, Ferson and Glassman (1998) found that the overall distribution of conditional alphas is similar to that of the unconditional alphas.

Hansen and Richard (1987) is an important study which analyzes and compares the conditional and unconditional approaches. The authors propose to study the efficiency of the mean-variance structure with respect to conditional information, and they conclude that unconditional efficiency portfolios should be conditionally efficient, but the reciprocal is not necessarily true. Ferson and Siegel (2001) derive the solution of the mean-variance structure optimization in the presence of conditional information, in what might be framed as *unconditional mean variance efficiency*. Continuing this analysis, Ferson and Siegel (2009) derive a methodology to test portfolios efficiency under the structure that they derived in 2001.

3 Methodology

This study aims to evaluate an alternative econometric method for comparing portfolios efficiency in the presence of conditional information. Ferson and Siegel (2009) presented results in this direction, and this project intends to expand the results and applications of this research, testing an alternative methodology for this analysis.

The methodology to be tested is based in the Empirical Likelihood estimation method. Thus, it is expected that estimators have better statistical properties, such as robustness and better asymptotic bias in finite samples. The study proposed here, has at this point the main difference when compared to the Ferson and Siegel (2009). How, essentially, we must work with finite sample data, we seek in this research to analyze the behavior of an estimator which has superior properties under this characteristic of the sample.

Note that, when testing portfolios efficiency with the use of conditional information, one should seek to maximize the unconditional mean relative to the unconditional variance, where portfolio composition strategies are functions of the information matrix. This is the line followed by the *unconditional mean variance efficiency* with respect to the information. It is valid to compare this framework with the conditional efficiency, where efficiency of the mean-variance structure is evaluated under conditionals means and variance. Note that in the first case, being the approach pursued by this work, the conditional information is used in the construction of the portfolio, and then the efficiency is assessed unconditionally.

3.1 Incorporating Conditional Information

Any asset pricing model may be defined following the basic pricing equation:

$$p_t = E_t(m_{t+1}x_{t+1}) \tag{3.1}$$

where p_t is the asset price, x_{t+1} the asset payoff, and m_{t+1} is the stochastic discount factor

(SDF). Depending the purpose of the research, the SDF may also be known as *change of measure*, *pricing kernel*, or even as *state-price density*.

Notice that the models need to portray the prices taking into account conditional moments. This may be observed in equation (3.1) when it represents with the use of a conditional expectation the available set of information to the investor in period t of time. Defining Z_t as the set of available information at t , the equation (3.1) may also be written as $p_t = E(m_{t+1}x_{t+1}|Z_t)$.

According to Cochrane (2009), all asset pricing models may be reduced to distinct ways to connect the SDF to the data. Restricting only to assets in stock class, succinctly, the payoff may be represented as $x_{t+1} = p_{t+1} + d_{t+1}$, where d_{t+1} is the dividend from the asset evaluated. For practical reasons, it is preferable to work with the gross return, i.e., $R_{t+1} \equiv \frac{x_{t+1}}{p_t}$. Thus, follows that the pricing models can be represented in accordance with the fundamental valuation equation:

$$E(m_{t+1}R_{t+1}|Z_t) = 1 \quad (3.2)$$

Assuming that exist a subset of variables \tilde{Z}_t such that $\tilde{Z}_t \subset Z_t$ and multiplying both sides by the elements of \tilde{Z}_t it is possible to get:

$$E_t(m_{t+1}R_{t+1} \otimes \tilde{Z}_t) = 1 \otimes \tilde{Z}_t \quad (3.3)$$

where \otimes represents a Kronecker product. If we take the unconditional expectation in (3.3) we get:

$$E(m_{t+1}R_{t+1} \otimes \tilde{Z}_t) = E(1 \otimes \tilde{Z}_t) \quad (3.4)$$

For the equation (3.1) it is also possible to incorporate instruments and work with unconditional moments as in $E(m_{t+1}x_{t+1} \otimes \tilde{Z}_t) = E(p_t \otimes \tilde{Z}_t)$. This approach is known as *managed portfolios*, being the product $R_{t+1} \otimes \tilde{Z}_t$ denominated *scaled returns*, and the product $x_{t+1} \otimes \tilde{Z}_t$ as *scaled payoffs*. As the instrument $z_t \in \tilde{Z}_t$ is inserted in the pricing

equation as a product, this approach may also be denominated as *multiplicative approach*. Intuitively, following Ferson and Siegel (2009), the equation (3.4) asks the SDF to price the dynamic strategy payoffs on average, which may also be understood in an unconditionally form.

Notice that with *managed portfolios* it is possible to incorporate conditional information and still work with unconditional moments. The main advantage of this structure is that there is no need (i) to explicit model conditional distributions, and besides (ii) it avoids the range problem of the conditional information under assumption. If it was necessary to incorporate conditional information with the use of conditional moments, from (i) would be necessary to formulate parametric models taking the risk of incorrectly define it; while from (ii) would be necessary to assume that all investors use the same set \tilde{Z}_t of instruments that was included in the conditional model, what clearly incorporates a high degree of uncertainty.

3.2 Estimation Methodology

The use of generalized method of moments (GMM) is fairly common in the estimation of asset pricing models. This happens primarily because with GMM there is no need to impose any distribution regarding the data, requiring only assumptions about the population moment conditions. In addition, for the multiplicative approach, its structure entails that the amount of instruments must exceed the moment conditions, justifying the use of the GMM. Then, from equation (3.1) we can take unconditional expectations to get:

$$\begin{aligned}
 p_t &= E_t(m_{t+1}x_{t+1}) \\
 \Rightarrow p_t &= E(m_{t+1}x_{t+1}|Z_t) \\
 \Rightarrow E(p_t) &= E(m_{t+1}x_{t+1})
 \end{aligned} \tag{3.5}$$

the asset pricing under unconditional moments must be a specific case of pricing under conditional moments. To do so, we must use the Law of Iterated Expectations.

When we use the asset pricing equation under unconditional moments, the moments

conditions necessary for the estimation by GMM become evident. Isolating the terms from equation (3.5) we can define the errors u_t , so that $u_t = 0$, i.e., $u_t = m_{t+1}R_{t+1} - 1$. Thus, the conditions under unconditional moments can be written as:

$$E(m_{t+1}x_{t+1} - p_t) = 0 \quad (3.6)$$

replicating the same procedure in the equation with the gross returns :

$$E(m_{t+1}R_{t+1} - 1) = 0 \quad (3.7)$$

For *managed portfolios*, the unconditional moments conditions are easily derived from equation (3.7):

$$E[(m_{t+1}R_{t+1} - 1) \otimes \tilde{Z}_t] = 0 \Rightarrow E[m_{t+1}(R_{t+1} \otimes \tilde{Z}_t) - (1 \otimes \tilde{Z}_t)] = 0 \quad (3.8)$$

Thus, the sample means of u_t are defined as:

$$g_T = \frac{1}{T} \sum_{t=1}^T u_t = \frac{1}{T} \sum_{t=1}^T [m_{t+1}R_{t+1} - 1] \quad (3.9)$$

while for the *managed portfolios* approach the sample means of u_t are defined as:

$$g_T = \frac{1}{T} \sum_{t=1}^T u_t = \frac{1}{T} \sum_{t=1}^T [m_{t+1}(R_{t+1} \otimes \tilde{Z}_t) - (1 \otimes \tilde{Z}_t)] \quad (3.10)$$

As the moment conditions are nothing more than the difference between observed and expected returns; then, in a graph that relates both returns, the alpha from Jensen (1968) must be the vertical distance between the points and a straight 45° line. It is worth to pay attention for the fact that, to make use of GMM all variables that comprise the moment conditions must be jointly stationary and ergodic, besides having finite fourth moments. This case highlights the need to use the equation given in (3.7), instead of equation (3.6). In the latter one, even that the payoffs are explicit in the definition of the

moments condition, as prices and dividends are expected to rise over time, this fact would cause failure in the stationarity hypothesis.

Finally, denoting by θ as the vector of parameters to be estimated, the GMM estimator can be defined as:

$$\hat{\theta}_T(\hat{W}) \equiv \arg \min_{\hat{\theta}} g_T(\hat{\theta})' \hat{W}_T g_T(\hat{\theta}) \quad (3.11)$$

where, \hat{W} is the conventional positive weighting matrix $q \times q$, for q moment conditions from GMM estimation.

3.2.1 Empirical Likelihood Estimation

Smith (1997) and Owen (2001) recently introduced a new family of estimators denominated *Generalized Empirical Likelihood* (GEL), that just as GMM, it is only based on moment conditions. According to Anatolyev and Gospodinov (2011), this is a non-parametric method that has the important attractive of asymptotic properties and finite samples with a variety of applications and can be used as a convenient structure analysis, estimation and inference. According to them, the class of estimators GEL leads to a better understanding regarding the properties of the estimators based on moments and allow more powerful tests, more efficient estimation of the density and distribution functions, and better bootstrap methods.

Following Anatolyev and Gospodinov (2011), suppose a system of restrictions on unconditional moments, such as:

$$E[g(w, \theta_0)] = 0 \quad (3.12)$$

where $\theta \in \Theta$ is a $k \times 1$ vector of the true parameters, w is a vector of observables, $\{w_i\}_{i=1}^n$ is a random sample, and $g(w, \theta)$ is a vector $q \times 1$ of the moments conditions. Let $\mathbf{p} = (p_1, p_2, \dots, p_n)$ be a collection of probability weights assigned to each sample

observation. Thus, we have the following *empirical likelihood* problem:

$$\begin{aligned} \max_{\mathbf{p}, \theta} \quad & \frac{1}{n} \sum_{i=1}^n \log(p_i) \\ \text{subject to} \quad & \sum_{i=1}^n p_i g(w_i, \theta) = 0 \\ & \sum_{i=1}^n p_i = 1 \end{aligned}$$

From this constraint maximization we obtain the *saddlepoint problem*, given by:

$$\max_{\theta \in \Theta} \min_{\lambda} \frac{1}{n} \sum_{i=1}^n -\log(1 + \lambda' g(w_i, \theta)) \quad (3.13)$$

from the solution of this problem it is possible to obtain the Empirical Likelihood estimator $\hat{\theta}$ (as well as the *GEL multipliers* $\hat{\lambda}$). If the substitution is made in the *saddlepoint problem* in (3.13) by an arbitrary criterion that is subject to certain shape conditions, it can be obtained the *GEL estimator*. To do so, let $\rho(v)$ be a strictly concave smooth function which satisfies $\rho(0) = 0$, $\partial\rho(0)/\partial v = \partial^2\rho(0)/\partial v^2 = -1$. This brings up the GEL estimator, given by $\hat{\theta}$, and the GEL multipliers, given by $\hat{\lambda}$, which are the solution of the *saddlepoint problem* below:

$$\min_{\theta \in \Theta} \sup_{\lambda \in \Lambda_n} \sum_{i=1}^n \rho(\lambda' g(w_i, \theta)) \quad (3.14)$$

where, $\Lambda_n = \{\lambda : \lambda' m(w_i, \theta) \in \Upsilon, i = 1, \dots, n\}$ and Υ is some open set containing zero (NEWHEY; SMITH, 2004).

Anatolyev and Gospodinov (2011) demonstrate how both estimators class, GEL and GMM, have equivalent asymptotic properties. The authors point out the asymptotic normality of GEL estimators. To do so, they invoke the Central Limit Theorem and show that for $\sqrt{n} \sum_{i=1}^m g(w, \theta_0)$ implies that the GEL estimator θ :

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, \Omega_\theta) \quad (3.15)$$

Thus, under overidentification conditions, the elements of the multiplier estimator $\hat{\lambda}$ are linearly dependent. The authors also show that the GEL estimators of θ_0 e λ_0 are asymptotically independent.

However, even if GEL and GMM estimators have identical asymptotic properties, in finite samples they exhibit different behaviors. According to Newey and Smith (2004), a good way to evaluate the bias of these estimators would be through the analysis of how many terms compose the analytical expressions of the second order asymptotic bias, and also by the precise manner in which their magnitudes are related to the number of moments constraints from the model.

Therefore, the examination should focus on the analysis of the higher-orders asymptotic bias expressions. Newey and Smith (2004) derive this higher-order asymptotic bias for the i.i.d. case using a random sample for both GEL and GMM estimators. They conclude that GEL estimation is preferable to GMM because GEL has one less term in its second order asymptotic bias expression. Moreover, the authors also demonstrate a practical implication when there are a considerable quantity of instruments: under this situation would not be recommended to select many instruments on a GMM estimation to avoid inflating the bias. Anatolyev (2005) get similar conclusions when comparing the second order asymptotic bias for GEL and GMM estimators in time series models. In summary, estimations based via GEL implies that, in opposition to GMM, the bias should not increases as the number of moment conditions grows.

Thus, using the criterion of bias assessment from Newey and Smith (2004), GEL estimator shows to be the best compared to GMM estimator. According to Anatolyev and Gospodinov (2011), this higher order asymptotic superiority from GEL estimator is usually attributed to its one-step nature, as opposed to the multi-step of GMM.

Succinctly, one can say that the estimation by GEL method seeks to minimize the distance between the vector of probabilities p and the empirical density $1/n$ in equation (3.13). Each of the estimators which are within the GEL class use different metrics to measure the distance. Owen (2001) defines the *Empirical Likelihood* (EL), where $\rho(v) = \ln(1 - v)$. Kitamura and Stutzer (1997) developed the estimator *exponential*

tilting (ET), where $\rho(v) = -\exp(v)$. Finally, one can still highlight the *continuous updated estimator* (CUE), where $\rho(v)$ is a quadratic function. The CUE was developed by Hansen, Heaton and Yaron (1996), but were Newey and Smith (2004) who have shown that this estimator can also be classified in GEL family.

3.2.2 HAC Estimation

The equation (3.11) that defines the GMM estimator may have an alternative representation. Let θ_0 be the true value, and assume θ_0 is an interior point of Θ , so that $\theta \in \Theta \subset R^p$. Let w_t be a vector of observed data, the estimator $\hat{\theta}(\hat{W})$ may be defined as the solution to the p first-order conditions associated with:

$$G_T(\hat{\theta}_T)' W_T g_T(\hat{\theta}_T) = \mathbf{0} \quad (3.16)$$

where $G_T(\hat{\theta}_T) = T^{-1} \sum_{j=1}^t \partial g(w_j, \theta) / \partial \theta'$, $g(\cdot)$ is a $q \times 1$ vector with q moment conditions, and $q \geq p$.

As the returns of the assets may not be an i.i.d. process, this fact creates the need to work under serial correlation or even with dependence on returns. Consider also that other violations of the i.i.d. processes may still be present, such as fat tails, or even non-linearity of returns.

Thus, it is necessary to work with estimators that have robust qualities for these deviations. One possibility to overcome these situations is with the use of estimators based on long-run covariance matrix. To this end, under the assumption of weak stationarity and ergodicity, the long-run covariance matrix Ω can be defined as the optimal matrix W from the equation (3.11), as follows:

$$W^* = \{ \lim_{n \rightarrow \infty} \text{Var}(\sqrt{T} \bar{g}_T(\theta_0)) \equiv \Omega(\theta_0) \}^{-1} \quad (3.17)$$

where,

$$\Omega(\theta_0) = \sum_{j=-\infty}^{\infty} \gamma(j) \quad (3.18)$$

where $\gamma(j)$ the autocovariances are defined by $\gamma(j) = E[(w_t - E(w_t))(w_{t-j} - E(w_{t-j}))]$ for the j -th order. For the multivariate version, the long-run covariance matrix Ω has the widely known expression:

$$\Omega(\theta_0) = \sum_{j=-\infty}^{\infty} \Gamma_j = \Gamma_0 + \sum_{j=1}^{\infty} (\Gamma_j + \Gamma_j') = \Lambda \Lambda' \quad (3.19)$$

where Λ is a lower triangular matrix given by the Cholesky decomposition of $\Omega(\theta_0)$, and Γ_j is the j -th order autocovariance matrix defined by:

$$\Gamma_j = E(\mathbf{g}_t \mathbf{g}_{t-j}'), \quad j = 0, \pm 1, \pm 2, \dots \quad (3.20)$$

An important estimator class for $\Omega(\theta_0)$ matrix are the non-parametric estimators. As the autocovariances are unknown; then, they may be replaced by sample autocovariances:

$$\begin{aligned} \hat{\Gamma}_j &= T^{-1} \sum_{t=j+1}^T \hat{\mathbf{g}}_t \hat{\mathbf{g}}_{t-j}', & j = 0, 1, \dots, T-1 \\ \hat{\Gamma}_j &= T^{-1} \sum_{t=j+1}^T \mathbf{g}(w_t, \hat{\boldsymbol{\theta}}_T) \mathbf{g}(w_{t-j}, \hat{\boldsymbol{\theta}}_T)', & \text{para } j \geq 0, \\ \hat{\Gamma}_j &= \hat{\Gamma}_{-j}', & \text{para } j < 0 \end{aligned} \quad (3.21)$$

While it is possible to estimate the corresponding samples, the estimator (3.21) is not consistent because the number of parameters grows in proportion to the sample size. To overcome this difficulty Newey and West (1987) and Andrews (1991) formulated a now widely used class of non-parametric estimators for the optimal long-run covariance matrix consistent to heteroskedasticity and autocorrelation (HAC) defined by:

$$\hat{\Omega}_{HAC}(\theta_0) = \sum_{j=-(T-1)}^{T-1} k(j/b) \hat{\Gamma}_j \quad (3.22)$$

where $k(\cdot)$ is a kernel function $k : R \rightarrow R$ satisfying the properties $k(x) = k(-x)$, $k(0) = 1$, $|k(x)| \leq 1$, being $k(x)$ continuous at $x = 0$ and $\int_{-\infty}^{\infty} k^2(x)dx < \infty$. The necessary conditions for $\mathbf{\Omega}_{HAC}(\boldsymbol{\theta}_0)$ be consistent requires that the bandwidth b grows in a lower rate when compared to the sample size, so that $b \rightarrow \infty$ and $b/T \rightarrow 0$ when $T \rightarrow \infty$.

Some options have been proposed for the kernel function, as well as the bandwidth in (3.22). There are a variety of choices for HAC matrix Andrews (1991) and Newey and West (1987) propose some possibilities for the kernel function and procedures for the bandwidth selection. It is known that the asymptotic properties of the GMM are not affected by the choice of a kernel or a specific bandwidth.

The estimation of the optimal long-run covariance matrix consistent to heteroskedasticity and autocorrelation $\mathbf{\Omega}_{HAC}(\boldsymbol{\theta}_0)$ should be used in the estimation of the parameters via GMM methods. Therefore, the GMM estimator $\hat{\boldsymbol{\theta}}_T$ may be defined as:

$$\hat{\boldsymbol{\theta}}(\hat{\mathbf{\Omega}}_{HAC}(\boldsymbol{\theta}_0)) \equiv \arg \min_{\hat{\boldsymbol{\theta}}} g_T(\hat{\boldsymbol{\theta}})' \hat{\mathbf{\Omega}}_{HAC}^{-1}(\boldsymbol{\theta}_0) g_T(\hat{\boldsymbol{\theta}}) \quad (3.23)$$

3.2.3 Fixed-b Estimators

A recent econometric literature has developed a number of methods to circumvent the problems caused by serial correlation, heteroskedasticity and observation or measurement errors. The conventional asymptotic theory of estimators HAC has a critical point in its structure when assuming that the estimator variance depends on a fraction of the sampling autocovariances, wherein the number of these tend to infinity, while the fraction tends to zero when the sample size increases. On that basis, assuming that the variance is known it is possible to derive the asymptotic distribution of the coefficients estimations.

Despite the important outcome of consistency with the autocorrelation and heteroskedasticity (HAC) of the variance estimators based on a long-run covariance matrix, this methodology has some weaknesses. Müller (2007) shows that this asymptotic estimator class does not performs well in case of finite samples. This characteristic is more evident especially in processes with strong dependence and temporal heterogeneity, as well as

in the presence of contamination and measurement errors. Kiefer and Vogelsang (2005) discuss that HAC robust tests have a tendency to over-reject in finite samples under the null hypothesis. According to the authors, there are two major sources of distortion in finite samples: one related to low accuracy through the Central Limit Theorem, and another related to the sampling variability of the asymptotic variance using the HAC methodology. It is focused on this second source of distortion that Kiefer and Vogelsang (2002b) propose an alternative method to build more robust tests. One of the usual problems is that conventional HAC estimators have a bias in the variance. Although this bias converges to zero when $T \rightarrow \infty$, in finite samples it is relevant.

To circumvent this deficiency Kiefer and Vogelsang (2002b) propose a new HAC class of robust tests that have bandwidth equal to sample size. Under this approach, the tests have pivotal asymptotic distributions, i.e., free from any parameter to be estimated. This estimator class is known as *fixed-b estimators*. According to the authors, although it is common to use a consistent estimator for the variance, this feature is not necessary to provide a pivotal asymptotic test. As a result, a weaker condition for valid tests is that the variance estimator should be asymptotically proportional to the unknown variance.

For this, suppose we use a bandwidth equal the sample size, i.e., $M = T$ in $\hat{\Omega}_{HAC}(\theta_0)$ from equation (3.23). Denoting by $\hat{\Omega}_{M=T}(\theta_0)$ this situation, it is clear that this estimator is inconsistent, because M violates the rule $M/T \rightarrow 0$ when $M, T \rightarrow \infty$. Defining $W_p(r)$ as p -vector of independent standard Wiener processes, where $r \in [0, 1]$, and $B_p(r) = W_p(r) - rW_p(1)$ as a $p \times 1$ vector of standard Brownian bridges. Then, under a pair of sufficient assumptions about the DGP (for details, see Kiefer and Vogelsang (2002b)) and assuming that $k''(x)$ exist for $x \in [-1, 1]$ and it is continuous. So, when $T \rightarrow \infty$,

$$\hat{\Omega}_{M=T}(\theta_0) \Rightarrow^* \Lambda \int_0^1 \int_0^1 -k''(r-s)B_p(r)B_p(s)'drds\Lambda' \quad (3.24)$$

where \Rightarrow^* denotes weak convergence. Thus, pivotal asymptotically tests for the null hypothesis may be derived, because $\hat{\Omega}_{M=T}(\theta_0)$ is asymptotically proportional to $\hat{\Omega}_{HAC}(\theta_0)$

through $\Lambda\Lambda'$. When using the Bartlett kernel and a bandwidth equal to the sample size, Kiefer and Vogelsang (2002a) show that:

$$\hat{\Omega}_{M=T}(\boldsymbol{\theta}_0) = \hat{\Gamma}_0 + \sum_{j=1}^{T-1} k(j/T)(\hat{\Gamma}_j + \hat{\Gamma}'_j) \Rightarrow^* \Lambda \int_0^1 \int_0^1 -k''(r-s)B_p(r)B_p(s)'drds\Lambda' \quad (3.25)$$

which can be written as:

$$\hat{\Omega}_{M=T}(\boldsymbol{\theta}_0) = T^{-1} \sum_{i=1}^T \sum_{j=1}^T \hat{v}_i \left(1 - \frac{|1-j|}{T}\right) \hat{v}'_j \quad (3.26)$$

where $\hat{v}_t = x_t u_t$, being \hat{u}_t the sample counterpart of the error term possibly exhibiting autocorrelation and conditional heteroskedasticity. In spite of $\hat{\Omega}_{M=T}(\boldsymbol{\theta}_0)$ be inconsistent, Kiefer and Vogelsang (2002b) conclude that (i) the asymptotic approaches for the size tests derived from this approach often are better in finite sample than those based on tests via conventional HAC estimations, and (ii) the power is also better compared to conventional statistics for the Bartlett kernel.

3.3 Tests of Efficiency

To test portfolios efficiency within the mean-variance framework, we need to predefine the asset pricing model in use. This study chose to use two linear pricing factors models which are the most commonly used: the CAPM from Sharpe (1964) and Lintner (1965), in addition to the Fama-French model with three factors from Fama and French (1993). At this point, a note is valid. This work has no interest to assess whether any asset pricing model correctly price assets; but, as different estimation methodologies may impact on inference regarding efficiency for each model.

It should also be noted that the market efficiency hypothesis is what motivates the construction of the null hypothesis to be tested. Fama (1970) defines the model of efficient markets such as the hypothesis on which the security prices “completely reflect” all or a particular subset of the information available at any point of time.

In his article, Fama (1970) also classifies in three groups the coverage of this model: (1) the *weak form* in which the information subset of interest only represents historical past prices (or returns); (2) the *semi-strong form* concerns to the speed of price adjustments related to other publicly available information; and finally (3) the *strong form* in which the concern is whether any investor or groups (e.g., mutual funds or institutional investors) have monopolistic access to any relevant information for pricing. Therefore, within this classification, the interest for this research is to use the information matrix in a consistent way to the *semi-strong form* of market efficiency.

For the CAPM model, the SDF can be defined as:

$$m_{t+1} = a + bR_{t+1}^W \quad (3.27)$$

where a and b are constants. The R^W is defined as the *wealth portfolio*, where generally is used a proxy that reflects the market behavior for empirical works. For the Fama-French three-factors model the SDF is defined as:

$$m_{t+1} = a + b_1Mkt_{t+1} + b_2SMB_{t+1} + b_3HML_{t+1} \quad (3.28)$$

where a and bs are constants, Mkt defines the return of a market proxy, SMB *small minus big* factor, and HML *high minus low* factor. Both models, CAPM and Fama-French, may also be derived in beta approach. To see this fact, before it is necessary to demonstrate that there is a connection between the stochastic discount factor representation in (3.1) and the beta representation. More precisely, one can say that both formats are equivalent and carry the same information, so that it is possible to move from one representation to another with no loss. At this point, we want to show that $p_t = E_t(m_{t+1}x_{t+1}) \Rightarrow E(R_{i,t+1}) = \alpha + \beta_{i,m}\lambda_m$. For simplicity, take equation (3.2) without incorporating conditional information and apply the decomposition of covariance,

$$\begin{aligned} 1 &= E(m_{t+1}R_{i,t+1}) \\ \Rightarrow 1 &= E(m_{t+1})E(R_{i,t+1}) + Cov(m_{t+1}, R_{i,t+1}) \end{aligned} \quad (3.29)$$

Thus,

$$\begin{aligned}
E(R_{i,t+1}) &= \frac{1}{E(m_{t+1})} - \frac{\text{Cov}(m_{t+1}, R_{i,t+1})}{E(m_{t+1})} \\
\Rightarrow E(R_{i,t+1}) &= \frac{1}{E(m_{t+1})} + \left(\frac{\text{Cov}(m_{t+1}, R_{i,t+1})}{\text{Var}(m_{t+1})} \right) \left(-\frac{\text{Var}(m_{t+1})}{E(m_{t+1})} \right) \quad (3.30) \\
\Rightarrow E(R_{i,t+1}) &= R^f + \beta_{i,m} \lambda_m
\end{aligned}$$

so that we get the beta representation, where $R^f = \frac{1}{E(m_{t+1})}$ denotes a risk-free rate (or zero-beta rate) if present, $\beta_{i,m} = \frac{\text{Cov}(m_{t+1}, R_{i,t+1})}{\text{Var}(m_{t+1})}$, i.e., the regression coefficient of the return $R_{i,t+1}$ on m (SDF), and $\lambda_m = -\frac{\text{Var}(m_{t+1})}{E(m_{t+1})}$. The pricing model in beta format seeks to explain the variation in average returns across assets to express that expected return should be proportional to the regression coefficient $\beta_{i,m}$. Note that, in this format betas $\beta_{i,m}$ are explanatory variables varying for each asset i , while R^f and λ_m represent, respectively, the intercept and the common slop for all assets i in a cross-section regression. In this approach, assets with higher betas should get higher average returns. Thus, $\beta_{i,\pi}$ is interpreted as the amount of risk that the asset i is exposed to the risk factor π , and the term λ_π is interpreted as the price of such risk exposure.

For simplicity, define \mathbf{f}_t as a vector with dimension $K \times 1$ for the factors that compose each of the models, and assume that now on we are working with excess returns ($R_{i,t+1}^e = R_{i,t+1} - R_{t+1}^f$), where it will not be used anymore the superscript notation e . For Fama-French model, the vector \mathbf{f}_t has dimension 3×1 , while for the CAPM it is a scalar. Then, one can write:

$$R_{i,t} = \alpha + \beta_i \mathbf{f}_t + \varepsilon_t, \quad t = 1, \dots, T \quad ; \text{ and } \quad i = 1, \dots, N \quad (3.31)$$

where, for practicality $R_{i,t}$ and \mathbf{f}_t already represent, respectively, excess returns for the N securities and for the K factors in both models in a sample size T . For a system with N

assets we have the following statistical structure for these models:

$$\begin{aligned} \mathbf{R}_t &= \boldsymbol{\alpha} + \boldsymbol{\beta} \mathbf{f}_t + \boldsymbol{\varepsilon}_t \\ E[\boldsymbol{\varepsilon}_t] &= \mathbf{0} \\ E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] &= \boldsymbol{\Sigma} \\ Cov[\mathbf{f}_t, \boldsymbol{\varepsilon}_t'] &= \mathbf{0} \end{aligned} \tag{3.32}$$

where \mathbf{R}_t , $\boldsymbol{\alpha}$ e $\boldsymbol{\varepsilon}_t$ have $N \times 1$ dimension; while \mathbf{f}_t has $K \times 1$ dimension, and $\boldsymbol{\beta}$ is a $N \times K$ matrix. Defining $\boldsymbol{\Sigma}$ as the Variance-Covariance of the disturbances $\boldsymbol{\varepsilon}_t$. The theoretical framework for these asset pricing models imply that the vector $\boldsymbol{\alpha} = \mathbf{0}$. In this case, under the correct pricing assumption, these models can be written as:

$$E(\mathbf{R}) = \boldsymbol{\beta} E(\mathbf{f}) \tag{3.33}$$

Therefore, the portfolio defined by K factors derived from a linear pricing model is said to be efficient only when the N estimated intercepts are not jointly statistically significant. One can even say that the N intercepts of regressions must be equal to u_t , i.e., the pricing errors.

The test of efficiency to assess whether all pricing errors u_t are jointly equal to zero can be done through a Wald test. The null and alternative hypotheses are given by:

$$\begin{aligned} H_0 : \boldsymbol{\alpha} &= \mathbf{0} \\ H_A : \boldsymbol{\alpha} &\neq \mathbf{0} \end{aligned} \tag{3.34}$$

whereas the test statistic is given by:

$$J_{Wald} = \hat{\boldsymbol{\alpha}}' [Cov(\hat{\boldsymbol{\alpha}})]^{-1} \hat{\boldsymbol{\alpha}} \tag{3.35}$$

so that under the null hypothesis J_{Wald} must have distribution χ^2 with N degrees of freedom. However, one should remember of the limitation from the Wald test that underlies in the large samples distribution theory. According to Cochrane (2009), the test remains valid

asymptotically even if the factor is stochastic and the Var-Cov matrix of the disturbances Σ is estimated. If, on the one hand there is no need to assume that the errors are normally distributed; on the other, this test ignores the sources of variation in finite samples. Bolstered on the Central Limit Theorem, the test is based primarily on the fact that $\hat{\alpha}$ has normal distribution.

Gibbons, Ross and Shanken (1989) derive the finite sample distribution of the null hypothesis in which the alphas are jointly equal to zero. In contrast to the J_{Wald} test, this test recognizes sample variation in the estimated Var-Cov matrix of the disturbances $\hat{\Sigma}$. However, the test requires that the errors are normally distributed, homoskedastic and uncorrelated. This test is defined by:

$$J_{GRS} = \frac{T - N - K}{N} \left(1 + E_T(\mathbf{f})' \hat{\Omega}^{-1} E_T(\mathbf{f}) \right)^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} \quad (3.36)$$

where it is used the same notation from Hansen and Singleton (1982) in which $E_T(\cdot)$ represents the sample mean, and

$$\hat{\Omega} = \frac{1}{T} \sum_{t=1}^T [\mathbf{f}_t - E_T(\mathbf{f})][\mathbf{f}_t - E_T(\mathbf{f})]'$$

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t'$$

Therefore, under i.i.d. and normally distributed errors, the statistic test J_{GRS} has a non-conditional distribution as a F with N degrees of freedom in the numerator, and $T - N - K$ degrees of freedom in the denominator. Note that, assuming $\varepsilon_t \sim N.I.D.$ effortlessly one can shown that $\hat{\alpha}$ has normal distribution, and $\hat{\Sigma}$ has Wishart distribution, more precisely:

$$\hat{\alpha} \sim N \left(\boldsymbol{\alpha}, \frac{1}{T} \left[1 + E_T(\mathbf{f})' \hat{\Omega}^{-1} E_T(\mathbf{f}) \right] \Sigma \right) \quad (3.37)$$

$$T \hat{\Sigma} \sim W_N(T - 2, \Sigma)$$

which, being the Wishart distribution a multivariate χ^2 , implies that $\hat{\boldsymbol{\alpha}}'[\mathbf{Cov}(\hat{\boldsymbol{\alpha}})]^{-1}\hat{\boldsymbol{\alpha}}$ should result in a F distribution.

The authors also demonstrated a test variation that can be expressed as the distance between the returns of factors \mathbf{f} and the ex-post mean-variance frontier:

$$J_{GRS-SR} = \frac{T - N - K}{N} \frac{(\mu_q/\sigma_q)^2 - E_T(\mathbf{f})'\hat{\boldsymbol{\Omega}}^{-1}E_T(\mathbf{f})}{1 + E_T(\mathbf{f})'\hat{\boldsymbol{\Omega}}^{-1}E_T(\mathbf{f})} \quad (3.38)$$

where $(\mu_q/\sigma_q)^2$ is the squared Sharpe ratio of the ex-post tangency portfolio, that consequently, must be the maximum ex-post Sharpe ratio. Notice that the term $E_T(\mathbf{f})'\hat{\boldsymbol{\Omega}}^{-1}E_T(\mathbf{f})$ found both in the numerator and in the denominator, can be seen to the CAPM model as the squared Sharpe ratio of the factor. For the CAPM, we can write this test in the following format:

$$J_{GRS-SR} = \frac{T - N - K}{N} \frac{(\mu_q/\sigma_q)^2 - (E_T(\mathbf{f})/\hat{\sigma}(\mathbf{f}))^2}{1 + (E_T(\mathbf{f})/\hat{\sigma}(\mathbf{f}))^2} \quad (3.39)$$

note that in the numerator of the second term are the difference of two Sharpe ratios:

$$\left(\frac{\mu_q}{\sigma_q}\right)^2 - \left(\frac{E_T(\mathbf{f})}{\hat{\sigma}(\mathbf{f})}\right)^2 = sr_q^2 - sr_m^2 \quad (3.40)$$

that is, the difference between the Sharpe ratio of the tangency portfolio q , i.e., sr_q^2 ; and the portfolio's market factor m , i.e., sr_m^2 . It should be mentioned that the tangent portfolio must be computed considering all N assets, including the market factor. From (3.36) one can easily note that the term $sr_q^2 - sr_m^2$ must be equal to $\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$.

Jobson and Korkie (1982) also studied portfolios efficiency tests using Sharpe ratios. Following the derivation from Ferson and Siegel (2009) the "JK" test can be written as:

$$J_{JK} = T \frac{(\mu_q/\sigma_q)^2 - E_T(\mathbf{f})'\hat{\boldsymbol{\Omega}}^{-1}E_T(\mathbf{f})}{1 + E_T(\mathbf{f})'\hat{\boldsymbol{\Omega}}^{-1}E_T(\mathbf{f})} \quad (3.41)$$

in this case should be distributed according to a χ^2 with N degrees of freedom. Note that the J_{JK} test assume only that errors have no autocorrelation or heteroskedasticity.

Moreover, unlike the finite sample counterpart J_{GRS-SR} test, this test is based on the asymptotic distribution theory of the parameters.

3.4 Empirical Strategy

Following the *unconditional mean variance efficiency* with respect to information that this work intends to study, we seek to compare the approach without conditional information with the case which makes use of the information matrix. To evaluate the GEL estimators performance in efficiency tests of portfolios, this research aims to evaluate this family under each of the previously discussed approaches.

For the situation without the conditional information, it is conducted the estimation of the pricing equation given in (3.7). Now, in the case of including information about the state of the economy, it is used the *managed portfolios* structure presented in section 3.1, in which the moment conditions are derived from equation (3.4). Notice that in this situation the instruments used, which are lagged variables, are only used in a multiplicative form in the sample moments.

For each of these approaches, estimations are made by GMM and GEL, for the same sample and period of time. It is assessed portfolios with different amounts of assets. The same applies to sample size T of the time series, in which it is considered various time intervals. In addition, the analysis is done for both asset pricing models already mentioned: the CAPM and the Fama-French three-factors. Having calculated the estimations for each scenario, it is conducted efficiency tests according to section 3.3.

Firstly, the performance is evaluated when a HAC estimated matrix via kernel is used according to section 3.2.2. After that, the tests performance are assessed under both the GMM and GEL when fixed-b estimators are used, as presented in section 3.2.3.

After this analysis, it is assessed the statistical properties of the efficiency tests using GMM and GEL estimations in a finite sample context. To this end, four Monte Carlo experiments are performed in order to examine the tests robustness to disturbances in the data when both estimators are used. In particular, we analyze the presence of heavy

tails and outliers.

4 Data

Since the goal of this study is to evaluate an alternative econometric methodology to test portfolios efficiency with the use of conditional information, there is no interest to examine any particularity or features of a particular market. Thus, it was decided to use the US financial market data for obvious reasons.

4.1 Instruments and Factors

Following previous studies, we selected a limited number of instruments from those commonly used to measure the state of the economy. One can say that the lagged variables chosen are part of a standard set of instruments for this purpose.

The set of lagged variables consists of 5 instruments. The first is the lagged value of a 3-month Treasury-bill yield (see Ferson and Qian (2004)). The second is the spread between corporate bond yields with different ratings. This spread is derived from the difference between the Moody's Baa and Aaa corporate bond yields (see Keim and Stambaugh (1986); Ferson and Siegel (2009)). Another instrument is the spread between the 10-year and 1-year Treasury-bill yield with constant maturity (see Fama and French (1989); Ferson and Siegel (2009)). Following Ferson and Qian (2004), it was also included the percentage change in the U.S. inflation, measured by the *Consumer Price Index* (CPI). Finally, the monthly growth rate of the seasonally adjusted industrial production is also used, measured by the *Industrial Production Index* (see Ferson and Qian (2004)). All data were extracted from the historical time series provided by the Federal Reserve.

Regarding the factors that compose the asset pricing models, in view of that will be used the CAPM and Fama-French three-factors model, it was extracted the factors from both approaches. The market portfolio used, also known as the *wealth portfolio* from CAPM and Fama-French, consists of the weighted return of the value of all companies

Figure 1 – Historical series of the instruments for 720 months (from jan-1955 to dec-2014)

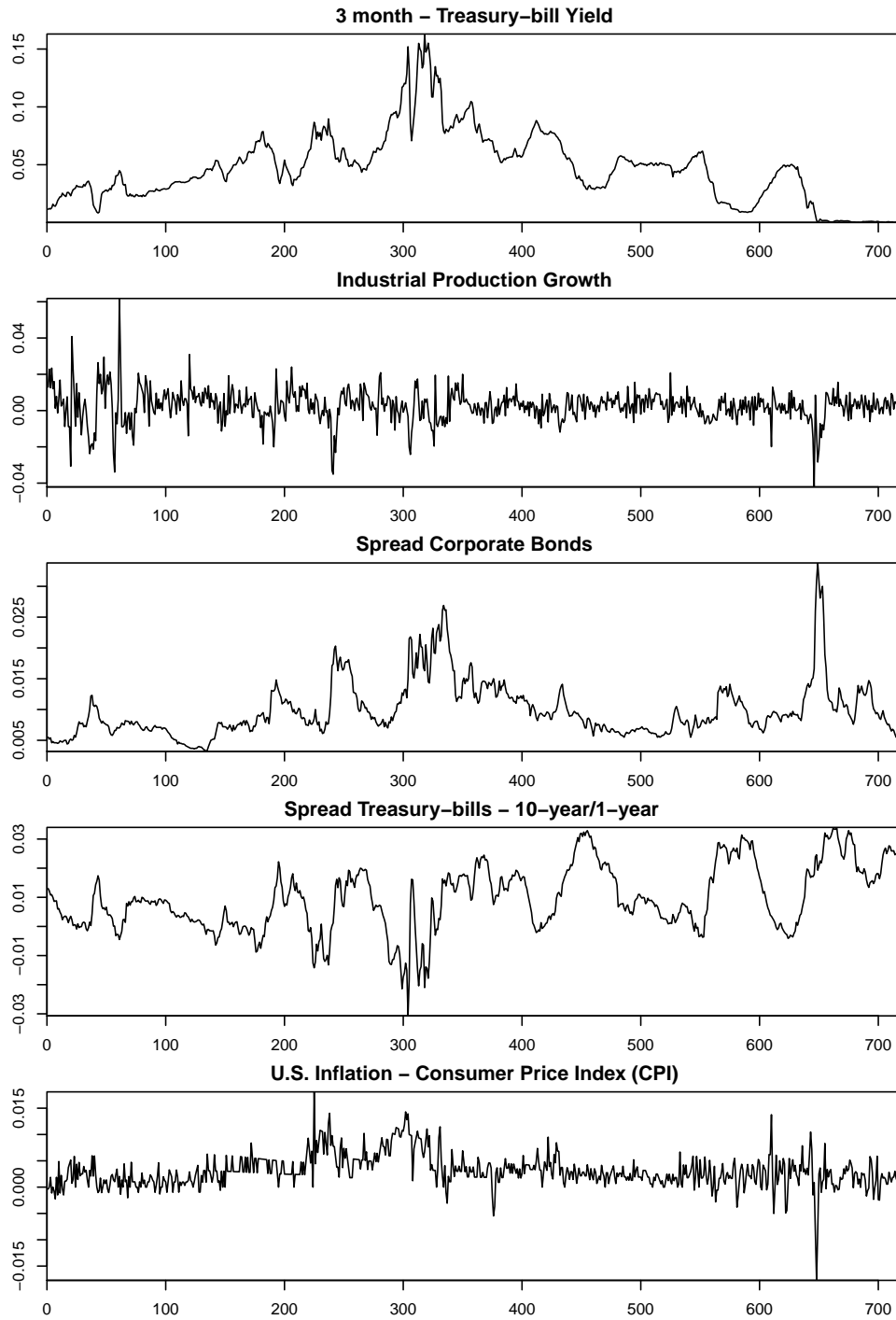
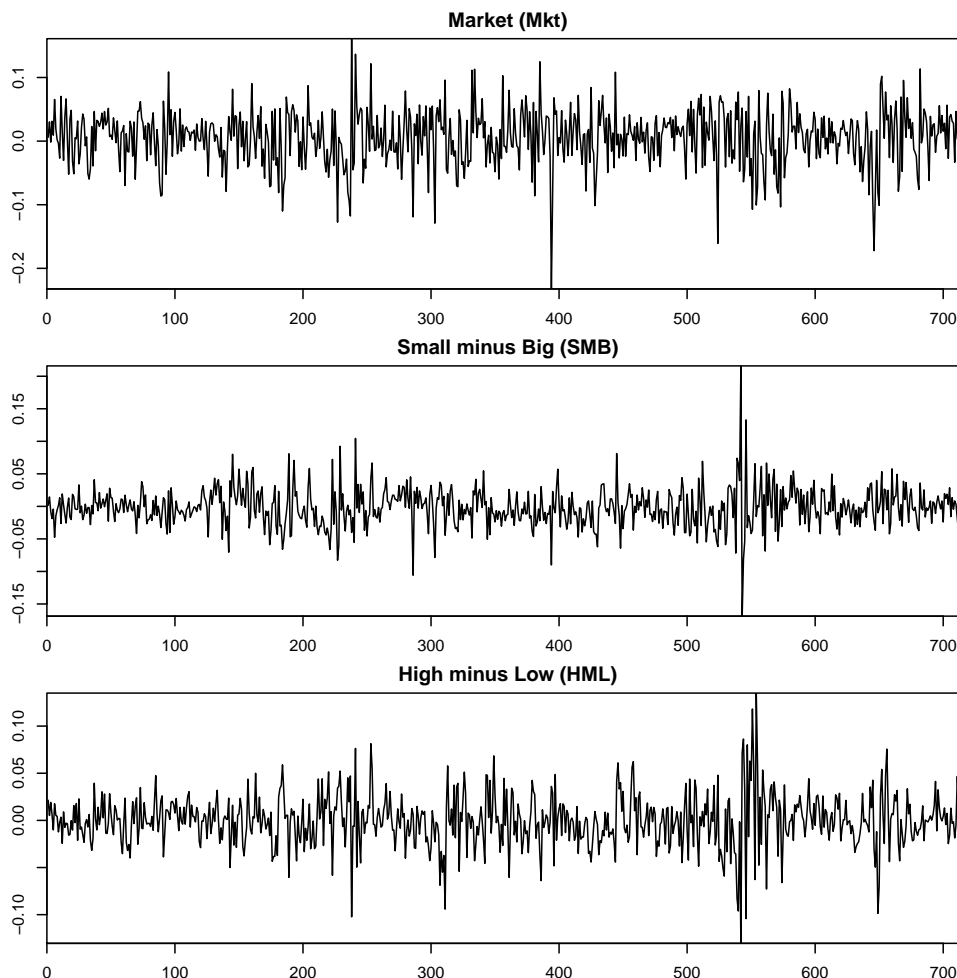


Figure 2 – Historical series of the factors for 720 months (from jan-1955 to dec-2014)



listed on the NYSE, AMEX and NASDAQ¹. The SMB and HML factors are computed in accordance with Fama and French (1993), i.e., the first is the average return of 3 smaller portfolios subtracted by the average return of the three largest portfolios; while the second is the average return of the two portfolios with high book-to-market subtracted from the average return of the two portfolios with low book-to-market. Data from all the factors were extracted from the database provided by Kenneth R. French².

Figures 1 and 2 respectively present the complete historical series of lagged state variables and factors used. The common maximum time span to all instruments is 720 months (60 years) prior to December 2014.

¹ Specifically, the market portfolio consists of the value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11 at the beginning of month t , good shares and price data at the beginning of t , and good return data for t .

² <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/datalibrary.html>

Table 1 – Descriptive statistics of the lagged variables and factors for a period of 720 months (60 years) from jan-1955 to dec-2014

	Mean	Std. Dev.	Min	Max	ρ_1
<i>Lagged Variables</i>					
3 month - Treasury-bill Yield	0.047	0.030	0.000	0.163	0.99
Industrial Production Growth	0.002	0.009	-0.042	0.062	0.37
Spread Corporate Bonds	0.010	0.004	0.003	0.034	0.97
Spread Treasury-bills - 10-year/1-year	0.010	0.011	-0.031	0.034	0.97
U.S. Inflation - Consumer Price Index (CPI)	0.003	0.003	-0.018	0.018	0.61
<i>Factors</i>					
Market (Mkt)	0.005	0.044	-0.232	0.161	0.08
Small minus Big (SMB)	-0.002	0.030	-0.169	0.216	0.06
High minus Low (HML)	-0.000	0.027	-0.130	0.135	0.16

Monthly returns of the 5 lagged variables and the 3 factors from asset pricing models. First column presents the sample mean, the second shows the sample standard deviation, the third and fourth column present minimum and maximum returns, and the last column presents the first-order autocorrelation. The sample period is January 1955 through December 2014 (720 observations).

In the graphs from the five instruments, important events that shook the economy in this 60-year range are easily noted through peaks and valleys. The oil crisis and the 2008 financial crisis are examples that impacted the state lagged variables of the economy.

In the table 1 it is presented the most common descriptive statistics of instruments and factors for the maximum period of 720 months. Observing the first order autocorrelation, we can see that instruments are highly persistent series, while this characteristic is not true for factors. Note that most of the five instruments, the first order autocorrelation is 97% or higher. The only instrument that can not be considered persistent is the *Industrial Production Index*, which has the first order autocorrelation of 37%. The three factors have first order autocorrelation lower than 20%.

4.2 Portfolios

In order to examine the estimations behavior for various sample sizes, we selected four portfolios with increasing number of assets. It sought not to be limited solely on a single portfolio composition methodology. Among the selected portfolios, three of them have their composition based on size and book-to-market, while one is composed with

Table 2 – Descriptive statistics of the monthly returns for the portfolio with 6 assets for a period of 720 months (60 years) from jan-1955 to dec-2014

	Mean	Std. Dev.	Min	Max	ρ_1	R^2
Small - 1 (Low)	0.005	0.072	-0.329	0.434	0.19	0.02
Small - 2	0.010	0.056	-0.286	0.316	0.22	0.02
Small - 3 (High)	0.012	0.057	-0.273	0.359	0.26	0.02
Big - 1 (Low)	0.006	0.053	-0.265	0.205	0.12	0.02
Big - 2	0.007	0.047	-0.242	0.218	0.13	0.02
Big - 3 (High)	0.009	0.049	-0.209	0.264	0.12	0.02

Monthly returns of the 6 portfolios formed on size and book-to market (2x3). The portfolios are based on equal-weighted returns extracted from the database provided by Kenneth R. French. First column presents the sample mean, the second shows the sample standard deviation, the third and fourth column present minimum and maximum returns, ρ_1 is the first-order autocorrelation, and the R^2 is the adjusted coefficient of determination from the regression of the returns on the lagged instruments. The sample period is January 1955 through December 2014 (720 observations).

categories derived from industries classification according to the business segment. All observations of the 4 portfolios used was taken from the historical data series provided by French with data extracted from the *Center for Research in Security Prices* (CRSP), and compiled in various types of portfolios composition.

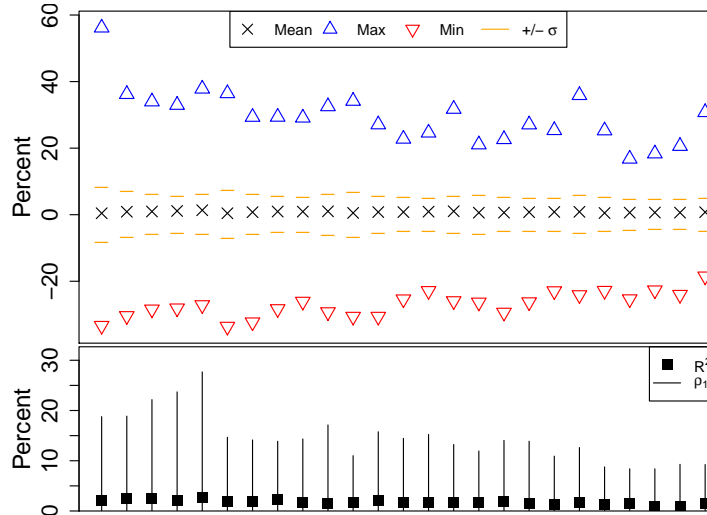
As for the amount of assets, it was used increasing time intervals for all portfolios. The evaluated periods begin with 120 months prior December 2014. Each of the remaining periods extend another 120 months (10 years) up to the limit of the dataset, which is 1020 months (85 years).

One of the portfolio has six assets selected with equal weights by size book-to-market³. Following the methodology from Fama and French (1993), this portfolio is constructed from the intersection of two portfolios composed by market cap with portfolios composed by the ratio of book equity with market cap. Table 2 reports the descriptive statistics of the 6 monthly returns. The table presents the 720 months period prior to December 2014, because this is the longest period of data for all instruments. The lagged variables are used to compute the R^2 statistic. Note that the mean ranges from 0.5% to 1.2% and the standard deviation from 4.7% to 7.2%. In the same table also can be seen the first order autocorrelation, which is generally low and between 12% and 26%; as well as the R^2 from the regression of the 5 instruments on the returns. Note that the adjustment

³ 6 Portfolios Formed on Size and Book-to-Market (2 x 3)

coefficient is very low for all 6 assets, being of the order of 2%.

Figure 3 – Descriptive statistics of the monthly returns for the portfolio with 25 assets for a period of 720 months (60 years) from jan-1955 to dec-2014



The top panel shows the sample mean statistics (represented by “X”), Max (represented by a blue triangle), Min (represented by an upside down red triangle), and the distance between the two horizontal bars represent the range of $\pm\sigma$ for a 720 months period. The bottom panel shows the first-order autocorrelation ρ_1 (bar) and R^2 (square point), which is the adjusted coefficient of determination in percent from the regression of the returns on the 5 instruments. In both panels, the x-axis represents the 25 assets, and the y-axis is expressed in percentage.

A second portfolio composed with 25 assets selected with equal weights by size and book-to-market ⁴ was chosen. Given the amount of observations, we decided to present in the figure 3 the main common descriptive statistics. For this portfolio, the statistics are similar to the portfolio with six assets, except that it can be observed portfolios with first order autocorrelation less than 10%.

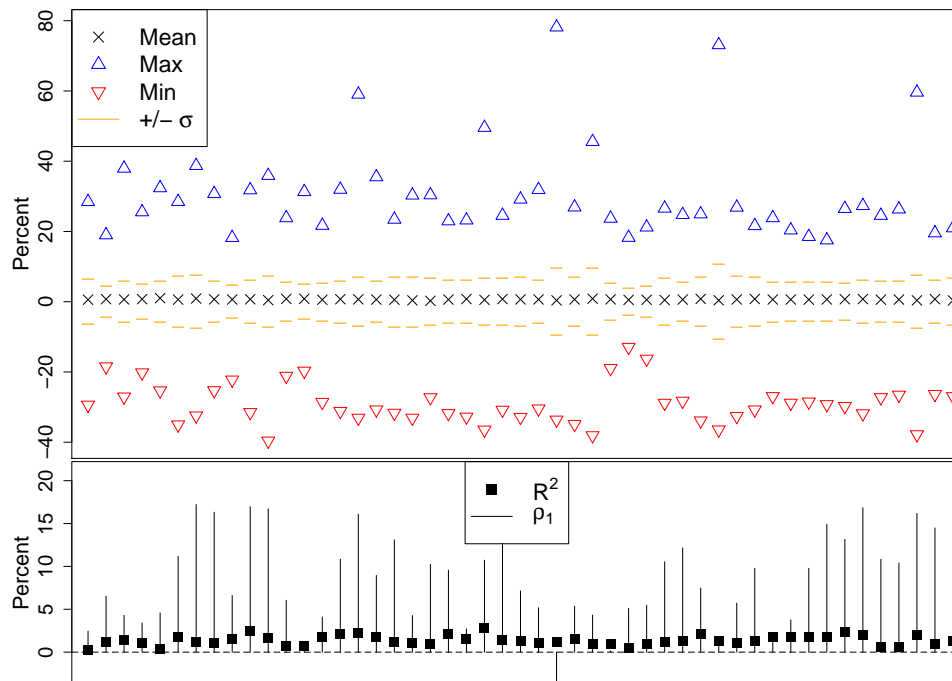
Another chosen portfolio consists of 49 assets representing industrial portfolios selected with equal weights ⁵. Finally, the last portfolio of 100 assets consists of 100 portfolios selected with equal weights by size and book-to-market ⁶. Again, it was decided to present the descriptive statistics for portfolios with charts. Figures 4 and 5 present for portfolios with 49 and 100 assets respectively, the descriptive statistics of the 720 monthly returns. Notice that both mean and standard deviation are similar to previous portfolios, while the maximum and minimum returns are magnified. Note that most of the first order

⁴ 25 Portfolios Formed on Size and Book-to-Market (5 x 5)

⁵ 49 Industry Portfolios

⁶ 100 Portfolios Formed on Size and Book-to-Market (10 x 10)

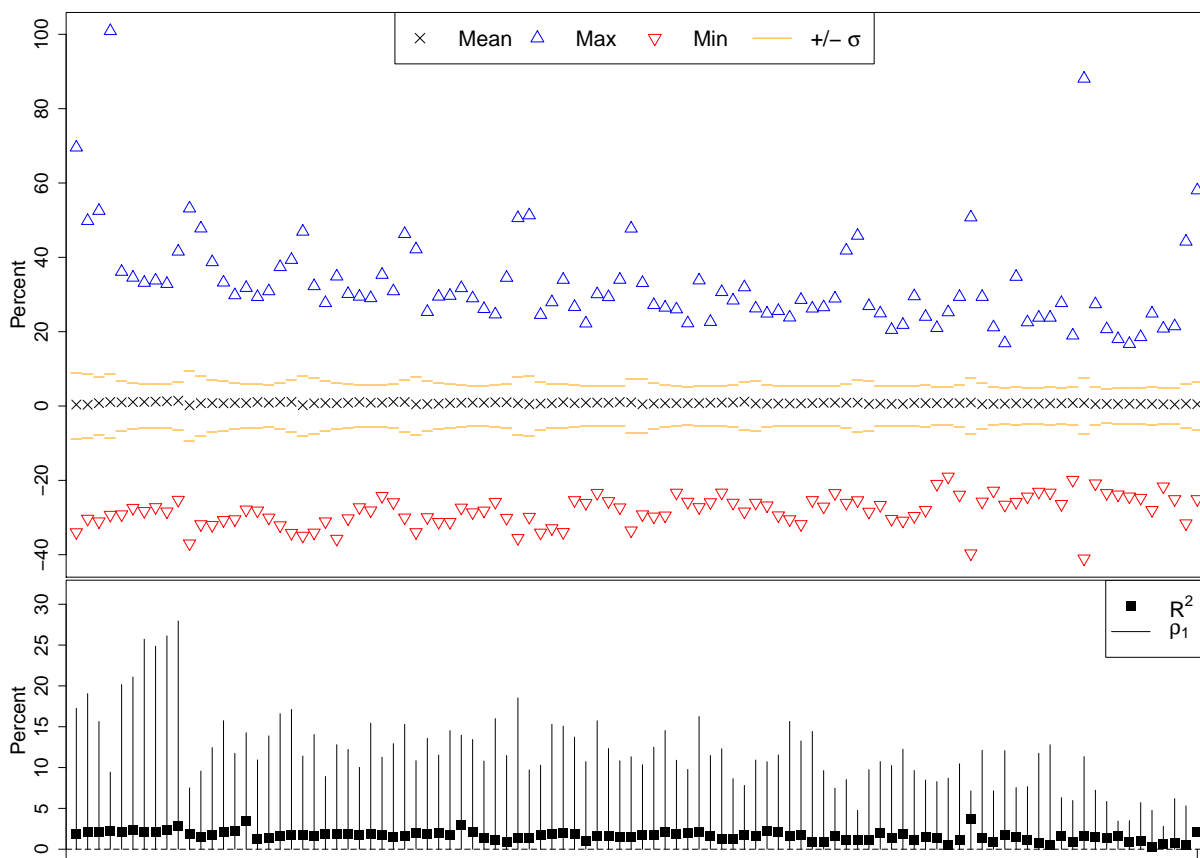
Figure 4 – Descriptive statistics of the monthly returns for the portfolio with 49 assets for a period of 720 months (60 years) from jan-1955 to dec-2014



The top panel shows the sample mean statistics (represented by “X”), Max (represented by a blue triangle), Min (represented by an upside down red triangle), and the distance between the two horizontal bars represent the range of $\pm\sigma$ for a 720 months period. The bottom panel shows the first-order autocorrelation ρ_1 (bar) and R^2 (square point), which is the adjusted coefficient of determination in percent from the regression of the returns on the 5 instruments. In both panels, the x-axis represents the 49 assets, and the y-axis is expressed in percentage.

autocorrelation is lower than 20% and there is only one asset with negative value for both portfolios. Only for the portfolio with 100 assets we can find five of them with first order autocorrelation higher than 20%. The R^2 keeps low values, with adjustments coefficients no higher than 5%.

Figure 5 – Descriptive statistics of the monthly returns for the portfolio with 100 assets for a period of 720 months (60 years) from jan-1955 to dec-2014



The top panel shows the sample mean statistics (represented by “X”), Max (represented by a blue triangle), Min (represented by an upside down red triangle), and the distance between the two horizontal bars represent the range of $\pm\sigma$ for a 720 months period. The bottom panel shows the first-order autocorrelation ρ_1 (bar) and R^2 (square point), which is the adjusted coefficient of determination in percent from the regression of the returns on the 5 instruments. In both panels, the x-axis represents the 100 assets, and the y-axis is expressed in percentage.

5 Results

In short, we are interested to know how efficiency tests based on GEL and GMM estimations can lead to different decisions. Under the *unconditional mean variance efficiency* with respect to information, in this section we compare both methods when (i) no conditional information is used, and when (ii) the *managed portfolios* structure is used.

In order to better fulfill this analysis, the assessment is made by comparing the test results for different sample sizes and portfolios types, as well as for two asset pricing models (CAPM and the Fama-French three-factors). The first part of this section presents the results when the standard long-run covariance HAC matrix is used during the parameters estimations.

In the second part we assess the efficiency tests when fixed-b estimators are used. At this point, as the recent literature indicates, we want to circumvent the problems caused by serial correlation, heteroskedasticity and observation or measurement errors. In order to prevent problems of over-rejection, we used the Vogelsang-Kiefer method in which one would expect more robust tests for finite samples.

Finally, the last part of this section performs Monte Carlo experiments to evaluate the tests robustness when GEL and GMM estimators are used in a finite sample context. Here, we want to study the effects that data contaminations, as outliers and the presence of heavy tails in the innovation structure, may cause on the results of efficiency tests.

5.1 Conventional HAC Estimation

5.1.1 6 Portfolios Formed on Size and Book-to-Market

First we start analyzing the portfolio consisting of the six portfolios formed on size and book-to-market. For these assets, we seek to test whether the factors from each of the asset pricing models explain the portfolios average returns. For CAPM, we are interested to

evaluate the efficiency of the market portfolio discussed in the previous section. Remember that this portfolio is composed of a value-weight return of all companies listed on the NYSE, AMEX and NASDAQ.

5.1.1.1 No Conditional Information

Table 3 presents the estimation results by GMM and GEL for an increasing sequence of months, starting with the last 120 months and adding 120 months up to the limit of sample, i.e., 1020 months (85 years). Each sample begins in January of a given year, and ends in December 2014. In the same table is also presented for each time interval, the estimation of both asset pricing models of interest, CAPM and Fama-French.

Initially examining test results when using the estimation by GMM, we note that both CAPM and Fama-French models show strong evidence to reject the hypothesis of efficiency for each model, especially for periods of 240 months or higher. This happens due to the p-values of Wald and GRS tests are practically zero.

For the period of 10 years and the CAPM model, the p-value of the Wald statistic is 0.02, while the p-value of the F distribution under the assumption of normality which is given by GRS test is 0.03. For the same period and Fama-French model, the p-values are very similar: 0.02 and 0.04 for the Wald test and GRS respectively. So, just for the period of January 2005 to December 2014 we have not negligible p-values.

Analyzing results of the tests using GEL, we have very similar results to the estimations by GMM for almost all periods. However, for the shorter time period, i.e., 120 months, the test using the GEL estimations lead to divergent conclusions. Note that the p-values for Wald and GRS tests are 0.30 and 0.35, respectively for the CAPM model. So, with these results we do not have evidence to reject the market portfolio efficiency. For the Fama-French model, the absence of such evidence in rejecting is more tenuous. The results are very close to the significance limit, the p-value for the Wald test is 0.05 and for the GRS is 0.08.

If we examine the complete table for both estimation methods, and take into account all periods of time, we may have some evidence for the fact that the DGP may

Table 3 – Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	15.0	0.020	2.4	0.035	14.6	0.023	2.3	0.043
	240	28.4	0.000	4.6	0.000	64.4	0.000	10.3	0.000
	360	60.6	0.000	9.9	0.000	137.7	0.000	22.4	0.000
	480	73.0	0.000	12.0	0.000	219.3	0.000	35.9	0.000
	600	74.2	0.000	12.2	0.000	274.8	0.000	45.1	0.000
	720	78.5	0.000	13.0	0.000	285.8	0.000	47.0	0.000
	840	81.6	0.000	13.5	0.000	277.6	0.000	45.8	0.000
	960	83.2	0.000	13.8	0.000	268.5	0.000	44.3	0.000
	1020	70.2	0.000	11.6	0.000	242.7	0.000	40.1	0.000
GEL	120	7.2	0.304	1.1	0.351	12.5	0.052	1.9	0.083
	240	26.0	0.000	4.2	0.000	51.6	0.000	8.3	0.000
	360	54.0	0.000	8.8	0.000	109.6	0.000	17.8	0.000
	480	61.0	0.000	10.0	0.000	185.6	0.000	30.4	0.000
	600	51.5	0.000	8.5	0.000	221.6	0.000	36.4	0.000
	720	56.8	0.000	9.4	0.000	236.6	0.000	38.9	0.000
	840	56.6	0.000	9.4	0.000	227.4	0.000	37.5	0.000
	960	72.7	0.000	12.0	0.000	176.6	0.000	29.2	0.000
	1020	47.6	0.000	7.9	0.000	171.9	0.000	28.4	0.000

Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market (2x3) for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case.

have changed over time. This explains why for the shortest and most recent period of time, both methodologies point to the same direction, suggesting no evidence to reject efficiency. However, as we expand the period of time, aggregating more information from older data, we start to have stronger evidences to reject efficiency. One possible explanation for these results is that the DGP changed over time.

5.1.1.2 Managed Portfolios

Table 4 presents the results of the market portfolio efficiency tests for the *scaled returns* approach. Here, we use the *managed portfolios* perspective that can be understood when investing according to a sign. To this end, the five instruments previously described are used.

Table 4 – Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
<i>GMM</i>	120	128.0	0.000	20.1	0.000	15.7	0.015	2.4	0.031
	240	32.1	0.000	5.2	0.000	97.5	0.000	15.6	0.000
	360	79.0	0.000	12.9	0.000	209.7	0.000	34.1	0.000
	480	111.5	0.000	18.3	0.000	310.4	0.000	50.8	0.000
	600	111.1	0.000	18.3	0.000	398.1	0.000	65.3	0.000
	720	102.4	0.000	16.9	0.000	380.0	0.000	62.5	0.000
<i>GEL</i>	120	5.6	0.466	0.9	0.509	16.3	0.012	2.5	0.025
	240	31.6	0.000	5.1	0.000	82.3	0.000	13.2	0.000
	360	63.3	0.000	10.3	0.000	174.9	0.000	28.4	0.000
	480	78.9	0.000	13.0	0.000	275.2	0.000	45.0	0.000
	600	72.2	0.000	11.9	0.000	241.8	0.000	39.7	0.000
	720	76.0	0.000	12.5	0.000	332.5	0.000	54.7	0.000

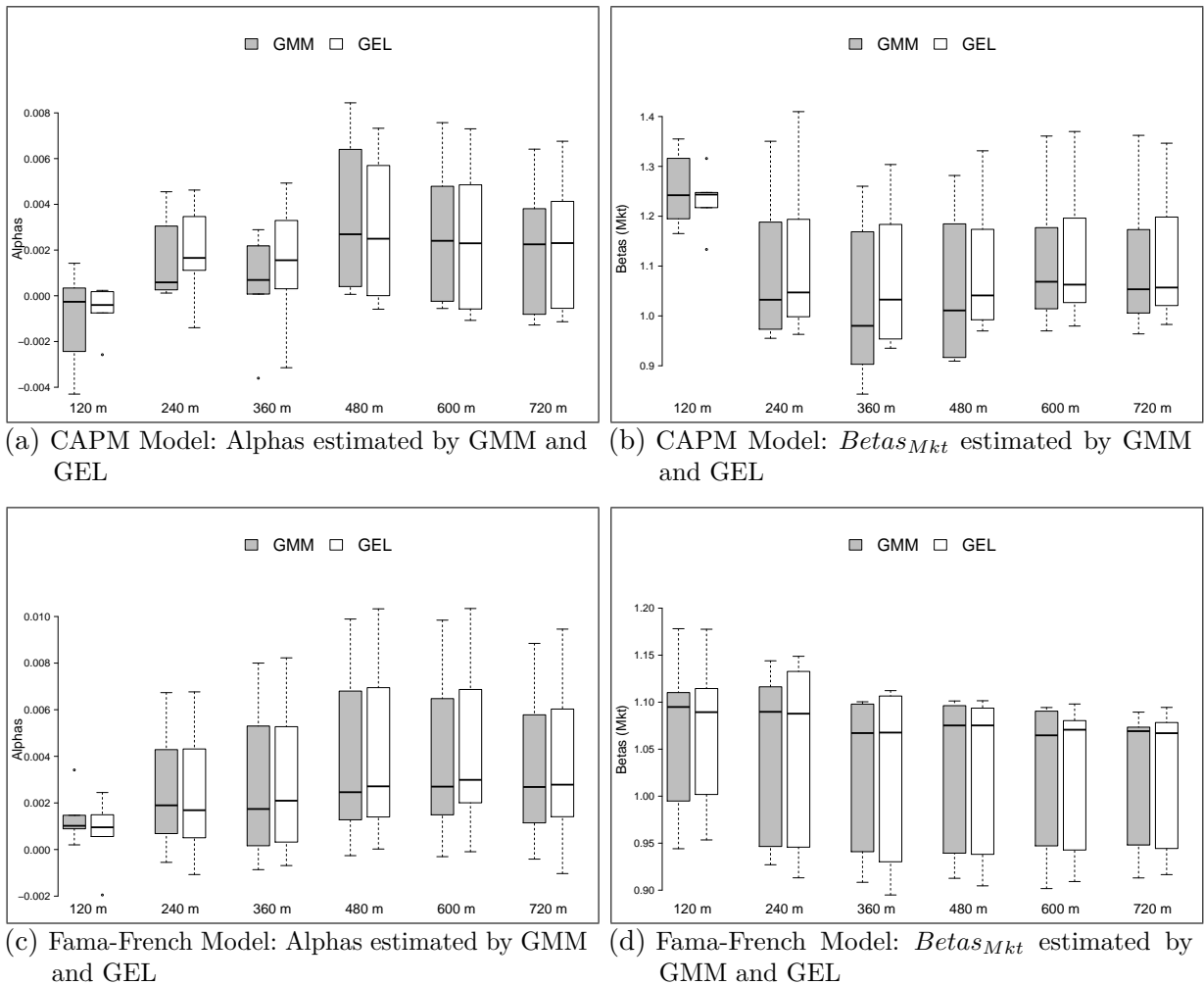
Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market (2x3) for 6 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years) and T=720 (60 years). The tests are evaluated using conditioning information when instruments are incorporated to the pricing equation. The lagged variables consisting the conditioning information are: (i) 3 month Treasury-bill yield, (ii) industrial production growth, (iii) yield spreads of low-grade over high-grade corporate bonds, (iv) yield spreads of long-term over short-term Treasury-bills (10-year/1-year) and (v) U.S. inflation (CPI). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case.

Again we have compelling evidence to reject efficiency for all intervals above 120 months. This result is true for the estimates using either GMM or GEL; and for both models under analysis, CAPM and Fama-French. Note that the longest period in which we can perform the tests is 720 months due to possess data from all five instruments up to this limit.

While for longer periods p-values were virtually zero, to 120 months estimation by GEL again shows no indication to reject efficiency for the CAPM. When we use fixed weights in the table 3, the estimations by GEL were also pointing to this direction. However, with the use of instruments, the p-values increase to 0.47 (Wald) and 0.51 (GRS) for CAPM; and decrease for Fama-French model 0.01 (Wald) and 0.03 (GRS). Again, if we analyze the whole results from table 4, we also have evidence that the DGP may have changed over time.

As it was possible to perform the estimation under both methodologies, for both

Figure 6 – Boxplots for comparison of estimations by GMM and GEL using 6 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time



In all 4 panels, the x-axis represents the time intervals, starting with 120 months before December 2014 to 720 months (60 years) prior to this date. The y-axis are the estimated coefficients values by GMM and GEL. Estimation by GMM is represented by gray boxplots, while GEL estimation is represented by white boxplots.

models, and all periods of time; figure 6 summarizes the estimated coefficients of α and β_{Mkt} with the use of instruments for CAPM and Fama-French models. As for every T and model there are 6 estimated alphas and betas (Mkt), we decided to present the estimated coefficients for each methodology (GMM and GEL) using boxplots. The graphs of $\hat{\beta}_{SMB}$ and $\hat{\beta}_{HML}$ for Fama-French model can be seen in appendix A.

Notice that as the boxplots distributions are not identical, GMM and GEL produce different coefficients estimates, i.e., not only the Var-Cov matrix from these estimates. For CAPM, which is located in the top two panels, note that for shorter periods differences are evident between both methodologies. Regarding the $\hat{\alpha}$, note that the biggest differences are for T equal to 120, 240 and 360. For $\hat{\beta}_{Mkt}$, this is more clear for $T = 120$. For Fama-French model, differences between estimates by GMM and GEL are more subtle. However, the $\hat{\alpha}$ for very short samples, as for 120 months, differences in estimates are not negligible.

From asset pricing theory it is known that the returns of any asset should be higher if this asset have higher betas. Figure 7 graphically displays for CAPM the estimated betas ($\hat{\beta}_{Mkt}$) against the sample mean of monthly excess returns ($\hat{E}(R_i)$) of each of the six assets. The model states that the average returns should be proportional to betas. Each panel represents one of the 6 time periods, and the GMM estimates are located on the left, while GEL estimates are on the right.

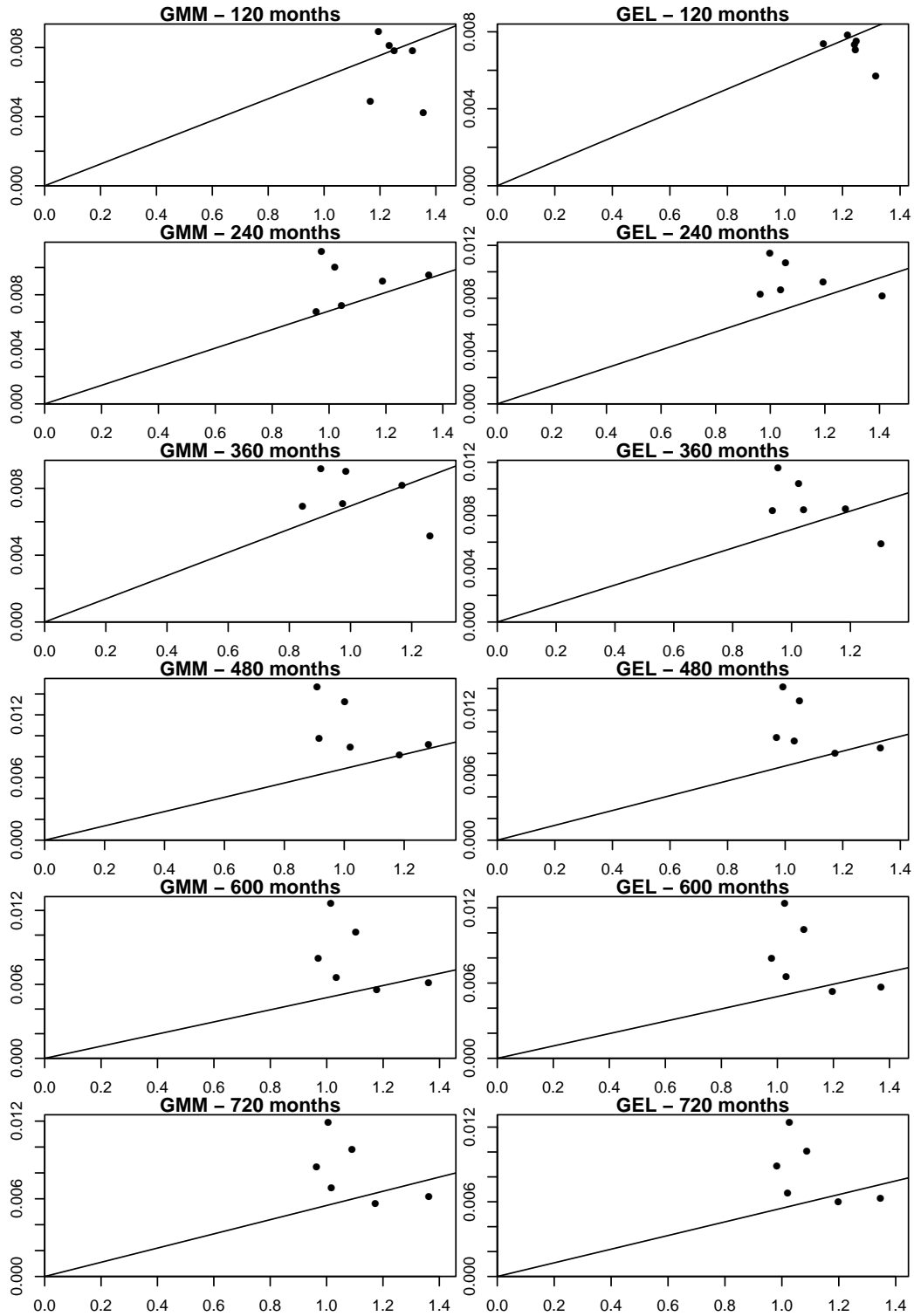
The distance between the points and the straight line must represent the pricing errors, i.e., estimated alphas. With this figure it is clear the difference in estimation between these methodologies, being the most symptomatic divergence for shorter T . Note that in the top two panels, while the estimates by GMM are more dispersed, those with GEL are more grouped and closer with slope line equal $\hat{E}(R_{Mkt})$.

5.1.2 25 Portfolios Formed on Size and Book-to-Market

5.1.2.1 No Conditional Information

In this section we move to a portfolio with higher amount of assets. Here, we use the 25 portfolios formed on size and book-to-market. Table 5 shows the results of estimations

Figure 7 – CAPM Model - Comparison of GMM and GEL estimated betas under *scaled returns* against the sample mean of monthly excess returns for the portfolio with 6 assets



In all panels, the estimated betas ($\hat{\beta}_{Mkt}$) are in the x-axis, and the sample mean of the monthly excess returns for each of the $n = 6$ assets in the portfolio are in the y-axis. Estimations via GMM are on the left, while via GEL are on the right. Each panel represents one of the time intervals, starting with 120 months before December 2014 to 720 months (60 years) prior to this date.

by GMM and GEL for the 9 months periods previously defined.

When we compare the estimated tests by GMM and GEL for CAPM, there is no doubt that both provide us with compelling evidence to reject the market portfolio efficiency hypothesis under consideration, remembering that this market portfolio is composed with companies listed on NYSE, AMEX and NASDAQ, as it was done in the previous section. When analyzing the Wald or even GRS statistics we see that estimates by GEL consistently have lower values compared with GMM. The p-values for both approaches are practically zero for every T , except for 120 months with GEL estimation. One could conclude, in general, for the efficiency rejection of the market proxy in use. If we compare these results with other studies regarding the efficiency of a market index, many studies come to similar conclusions (FERSON; SIEGEL, 2009; FAMA; FRENCH, 1992)¹.

From the formulas in (3.35) and (3.36), inverting the Var-Cov matrix of the alphas can become an impediment to the estimation of the tests, given the fact that singular matrices can arise. The cases presented by “NA” in table 5 stem primarily this problem. However, it may also occur that even estimations could not be performed for a certain methodology. This situation is also represented by “NA” in the table. The first point to note is that almost all the tests were able to be calculated.

Analyzing the tests for the Fama-French model, the conclusions regarding the efficiency do not change: all periods in which the test calculation was possible lead us to reject the null hypothesis. Here, we had the only case (120 months by GMM) where it was not possible to invert the Var-Cov matrix of the alphas. Similarly, for the CAPM, both statistics, in general have lower values when the estimation is calculated using GEL, when compared to those obtained by GMM (except for 240 months in Fama-French model). This difference seems to become more relevant as the analysis period increases, i.e., we have an evidence of increasing difference relationship when the sample expands.

In Figure 8 can be seen the boxplots of the estimated coefficients of α and β_{Mkt} for CAPM and Fama-French models. Each boxplot summarizes the 25 estimated values using

¹ One should just be cautious because the portfolio from these works and the exact time intervals have a strong resemblance, but are not exactly the same.

Table 5 – Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market for selected periods of time

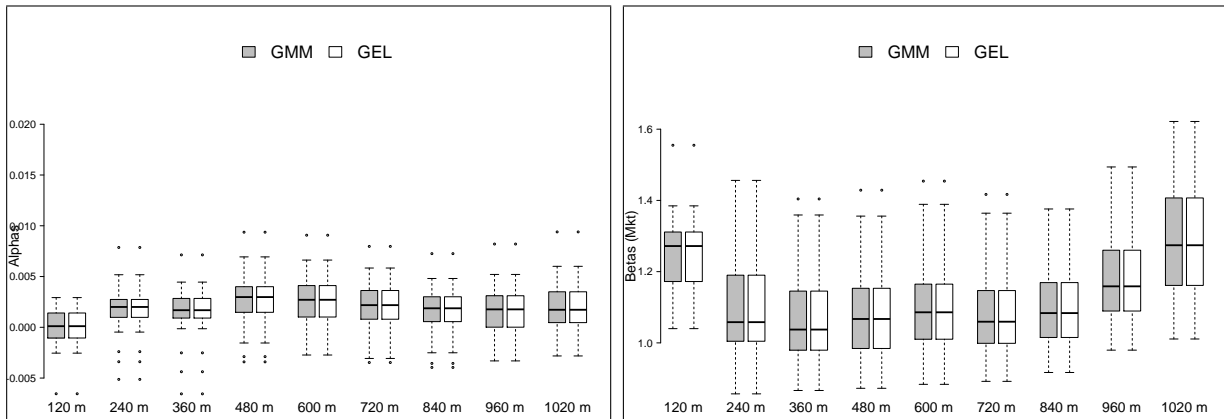
	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
<i>GMM</i>	120	98.5	0.000	3.1	0.000	NA	NA	NA	NA
	240	93.0	0.000	3.3	0.000	114.9	0.000	4.1	0.000
	360	168.7	0.000	6.3	0.000	257.5	0.000	9.5	0.000
	480	123.4	0.000	4.7	0.000	335.7	0.000	12.6	0.000
	600	133.4	0.000	5.1	0.000	465.2	0.000	17.7	0.000
	720	125.5	0.000	4.8	0.000	459.7	0.000	17.7	0.000
	840	122.5	0.000	4.8	0.000	413.2	0.000	16.0	0.000
	960	131.9	0.000	5.1	0.000	398.2	0.000	15.5	0.000
	1020	129.2	0.000	5.0	0.000	403.8	0.000	15.7	0.000
<i>GEL</i>	120	53.6	0.001	1.7	0.039	92.1	0.000	2.8	0.000
	240	85.4	0.000	3.0	0.000	138.9	0.000	4.9	0.000
	360	135.0	0.000	5.0	0.000	253.8	0.000	9.4	0.000
	480	90.6	0.000	3.4	0.000	329.0	0.000	12.4	0.000
	600	85.1	0.000	3.3	0.000	362.4	0.000	13.8	0.000
	720	83.2	0.000	3.2	0.000	374.1	0.000	14.4	0.000
	840	80.6	0.000	3.1	0.000	300.4	0.000	11.6	0.000
	960	90.3	0.000	3.5	0.000	300.5	0.000	11.7	0.000
	1020	84.1	0.000	3.3	0.000	310.5	0.000	12.1	0.000

Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market (5x5) for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

GMM or GEL for each time interval. Note that for every T estimation was possible to be performed. Even if the test for 120 months for the Fama-French could not be computed, the estimation of the coefficients was feasible. For the graphs of β_{SMB} and β_{HML} with Fama-French model, see appendix A.

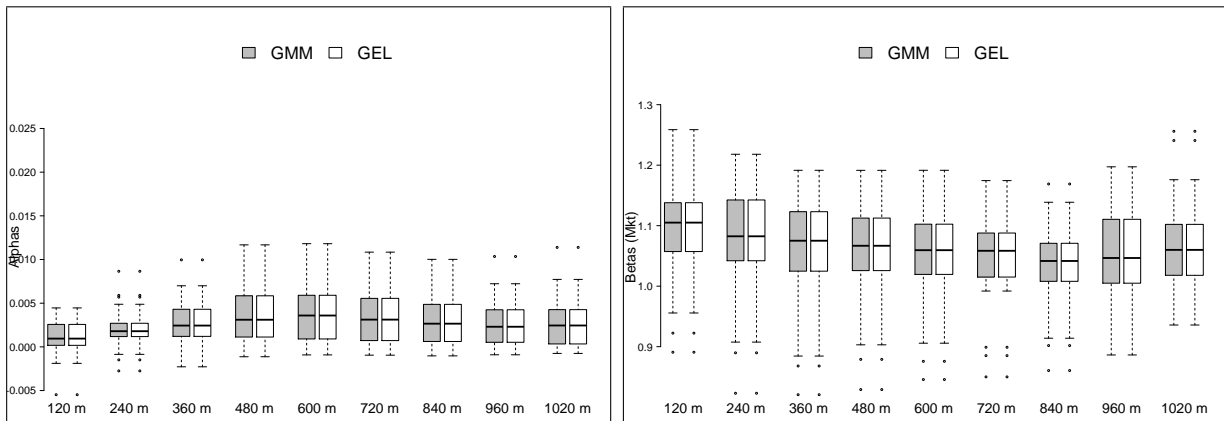
Regarding the alphas and betas, from the boxplots distribution it can be inferred that GMM and GEL showed identical values. The GRS and Wald tests have different values between GMM and GEL because the Var-Cov matrix of the estimates diverge. Anyway, from the panels it is possible to infer that for both models, the CAPM and Fama-French, the estimated alphas are low for the shortest period ($T = 120$) and increasing their values as t grows; and from 720 months the alphas return to decrease their estimates. The betas have similar characteristics; however, in opposite direction: starting with high values, then

Figure 8 – Boxplots for comparison of estimations by GMM and GEL using 25 portfolios formed on size and book-to-market for selected periods of time



(a) CAPM Model: Alphas estimated by GMM and GEL

(b) CAPM Model: $Betas_{Mkt}$ estimated by GMM and GEL



(c) Fama-French Model: Alphas estimated by GMM and GEL

(d) Fama-French Model: $Betas_{Mkt}$ estimated by GMM and GEL

In all 4 panels, the x-axis represents the time intervals, starting with 120 months before December 2014 to 1020 months prior to this date. The y-axis are the estimated coefficients values by GMM and GEL. Estimation by GMM is represented by gray boxplots, while GEL estimation is represented by white boxplots.

they decrease, and finally, increase again.

In Appendix A is the figure 16 plotting for CAPM the sample mean of the monthly excess returns ($\hat{E}(R_i)$) against the estimated betas ($\hat{\beta}_{Mkt}$) for the 25 assets.

5.1.2.2 Managed Portfolios

The results of the efficiency tests for the market proxy when using the multiplicative approach with 5 instruments can be seen in table 6. A quick inspection of the table show that in many cases it was not possible to compute the tests. In this case, all “NA” occur due to the impossibility of the methods to estimate even the parameters. This fact occurs

Table 6 – Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	NA	NA	NA	NA	NA	NA	NA	NA
	240	NA	NA	NA	NA	NA	NA	NA	NA
	360	NA	NA	NA	NA	NA	NA	NA	NA
	480	802.7	0.000	30.4	0.000	849.3	0.000	32.0	0.000
	600	480.7	0.000	18.4	0.000	967.6	0.000	36.9	0.000
	720	320.5	0.000	12.4	0.000	851.1	0.000	32.7	0.000
GEL	120	NA	NA	NA	NA	NA	NA	NA	NA
	240	NA	NA	NA	NA	NA	NA	NA	NA
	360	NA	NA	NA	NA	NA	NA	NA	NA
	480	NA	NA	NA	NA	NA	NA	NA	NA
	600	NA	NA	NA	NA	NA	NA	NA	NA
	720	NA	NA	NA	NA	NA	NA	NA	NA

Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market (5x5) for 6 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years) and T=720 (60 years). The tests are evaluated using conditioning information when instruments are incorporated to the pricing equation. The lagged variables consisting the conditioning information are: (i) 3 month Treasury-bill yield, (ii) industrial production growth, (iii) yield spreads of low-grade over high-grade corporate bonds, (iv) yield spreads of long-term over short-term Treasury-bills (10-year/1-year) and (v) U.S. inflation (CPI). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

due to singularity problems, and it is a fairly common issue, especially with portfolios with high amount of assets and under the multiplicative approach with instruments. Ferson and Siegel (2009) also had to deal with this issue (see e.g. Peñaranda and Sentana (2012) for advanced treatment). Since we have for this case 25 portfolios formed on size and book-to-market, and five instruments, it is possible to notice that the optimal long-run covariance matrix $\hat{\Omega}(\theta_0)$ from the estimator may turn out to be singular. In these cases, it is not possible to construct any of the efficiency tests.

Note that only for $T = 480, 600$ and 720 it was possible to perform the tests using GMM. The GEL estimators class were not able to estimate coefficients when $n = 25$ for any T . Whichever is the model in use, with instruments the efficiency of the models are rejected. Note that the tests generated considerably higher estimates, mainly for the Wald test, when compared with the fixed weights approach, leading to very strong evidences to

reject the null hypothesis.

5.1.3 49 Industry Portfolios

5.1.3.1 No Conditional Information

In this section we move to a portfolio with 49 assets. In this case, we use the portfolio composed with the 49 industrial categories previously described. We keep the structure of the tests, i.e., we seek to test the efficiency of the models. In table 7 are presented the estimation results by GMM and GEL for the 9 periods of time and both models.

Examining the results by GMM, it is clear that for the CAPM model, from a sample with 600 months or higher, it is not possible to invert the matrix. The same is true for 120 months. For three periods of time it was possible to estimate both the Wald test, as well as the finite sample counterpart GRS test, in which they point to the same direction: both have high p-values, showing that it is not possible to reject the efficiency hypothesis of the market proxy. Note that for these cases, the p-values decrease as the amount of months analyzed increases .

Regarding the Fama-French model, the results were similar to the CAPM, but with higher p-values. Only for 360 and 480 months scenarios we obtained results: the p-values of the Wald test is 0.98 and 0.69 for 360 and 480 months, respectively. But, the GRS test for finite sample, which assumes distribution F , obtained 0.99 and 0.84 p-values for the same periods.

From these results we note some evidences that the Wald test, based on large sample distribution, tends to reject the null hypothesis more often compared to tests that rely on finite sample distributions, like the GRS test. This characteristic from these tests is in line with the analysis of the size and power of efficiency tests in Campbell et al. (1997). However, for the GRS test, one should remember its limitation by requiring that errors are normally distributed, in addition to assume homoskedasticity and absence of autocorrelation.

Table 7 – Tests of portfolio efficiency using 49 industry portfolios for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	NA	NA	NA	NA	NA	NA	NA	NA
	240	41.9	0.753	0.7	0.946	NA	NA	NA	NA
	360	50.6	0.411	0.9	0.685	31.3	0.977	0.5	0.994
	480	60.7	0.123	1.1	0.292	43.7	0.689	0.8	0.838
	600	NA	NA	NA	NA	NA	NA	NA	NA
	720	NA	NA	NA	NA	NA	NA	NA	NA
	840	NA	NA	NA	NA	NA	NA	NA	NA
	960	NA	NA	NA	NA	NA	NA	NA	NA
	1020	NA	NA	NA	NA	NA	NA	NA	NA
GEL	120	108.7	0.000	1.3	0.160	NA	NA	NA	NA
	240	48.0	0.514	0.8	0.852	NA	NA	NA	NA
	360	44.1	0.672	0.8	0.860	NA	NA	NA	NA
	480	52.4	0.344	1.0	0.557	NA	NA	NA	NA
	600	81.9	0.002	1.5	0.014	NA	NA	NA	NA
	720	142.8	0.000	2.7	0.000	NA	NA	NA	NA
	840	206.4	0.000	4.0	0.000	NA	NA	NA	NA
	960	212.8	0.000	4.1	0.000	NA	NA	NA	NA
	1020	231.9	0.000	4.5	0.000	NA	NA	NA	NA

Tests of portfolio efficiency using 49 industry portfolios for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

When we analyze the results with the use of GEL class estimators, the first obvious observation is that the estimates for all periods with Fama-French model were not able to be computed. This fact occurred because with GEL methodology was not feasible to estimate the coefficients when $n = 49$, for any time period in use.

Focusing on the CAPM, there is strong evidence to reject the null hypothesis of market portfolio efficiency for long periods starting with 600 months, which includes intervals of 50, 60, 70, 80 and 85 years. This finding holds for both tests, either Wald, or GRS that do not have to support in asymptotic characteristics. Note that these five scenarios were not able to be calculated when using the GMM estimation.

Examining specifically 360 and 480 months periods, and comparing to the GMM estimation, we see that the results are in line: both provide evidence to reject the null

hypothesis. However, one can note that the p-values from GEL are higher for both tests. For instance, the GRS test using GEL obtained p-values of 0.86 and 0.56, for 360 and 480 months, respectively; while, with GMM the p-values were relatively small, i.e., 0.69 and 0.29 for the same periods. For 240 months, the estimation by GEL also conforms to the one held by GMM. However, in this scenario, the p-values were lower by GEL.

Finally, for 120 months was possible to calculate both statistics, something that was not possible when using the GMM. For this case, we can see a significant difference between both tests. Here, while the Wald test strongly rejects the efficiency, the finite sample counterpart GRS test showed no evidence of that with a p-value of 0.16.

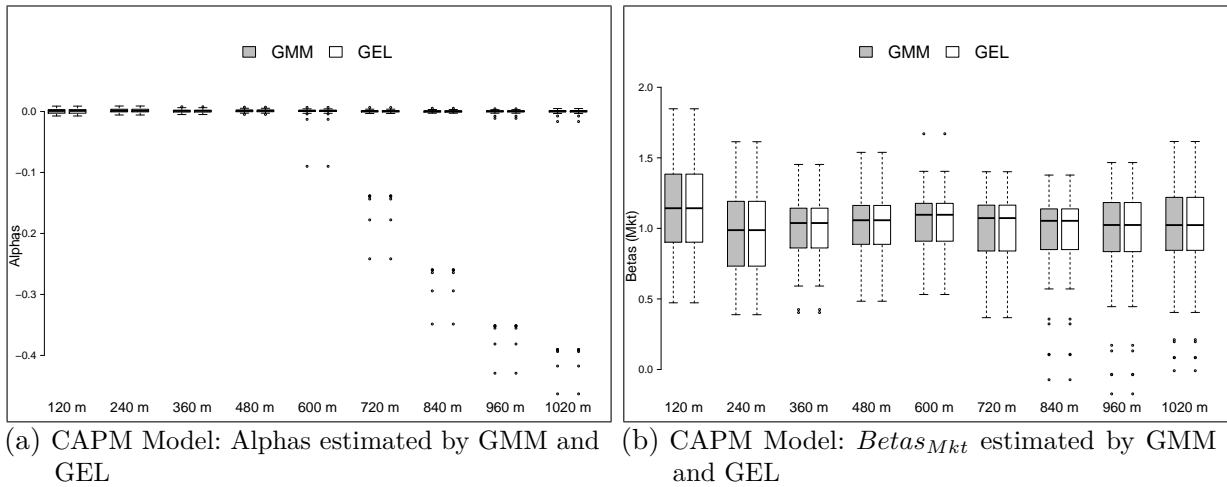
As only the estimation of CAPM model was possible using both GMM and GEL, in figure 9 is presented the boxplots graphics of the alphas and betas estimates from this model. For Fama-French model, the GEL estimation was not possible to be performed in any of the nine time scenarios. Note that the coefficients estimates showed exactly the same statistics. More precisely, the estimates of $\hat{\alpha}$ and $\hat{\beta}_{mkt}$ were identical. What caused differences in tests were the coefficient estimates of Var-Cov matrix which presented different values.

A point that worth a note is the existence of outliers in alpha estimates, occurring mainly for longer periods, from 600 months or higher. In addition, notice also that the estimates for both, $\hat{\alpha}$, as well as for $\hat{\beta}_{mkt}$ in CAPM model change marginally as time intervals increases.

In appendix A can be found the figure 17 plotting for CAPM the sample mean of the monthly excess returns ($\hat{E}(R_i)$) against the estimated betas ($\hat{\beta}_{Mkt}$) for the 49 assets.

Regarding the estimates of the multiplicative approach or *managed portfolios*, by neither method it was possible to even compute the parameters estimates. Thus, it was not feasible to evaluate the tests of efficiency using the five instruments previously described under *scaled returns* approach.

Figure 9 – Boxplots for comparison of estimations by GMM and GEL for CAPM model using 49 industry portfolios for selected periods of time



In both panels, the x-axis represents the time intervals, starting with 120 months before December 2014 to 1020 months prior to this date. The y-axis are the estimated coefficients values by GMM and GEL. Estimation by GMM is represented by gray boxplots, while GEL estimation is represented by white boxplots.

5.1.4 100 Portfolios Formed on Size and Book-to-Market

We attempted to test the efficiency of the market proxy using the 100 portfolios formed on size and book-to-market, which was presented in section 4.2. However, for any period and model, tests using GMM or GEL were not possible to be computed, either with or without the instruments approach. When the analysis is done with fixed weights without instruments, the estimation via GMM was possible to be performed. However, singularities problems occur impeding the inversion of the Var-Cov matrix. GEL estimations were not possible to be performed. When using *managed portfolios*, GMM and GEL were not able to estimate given the high n .

5.2 Tests of Efficiency in Squared Sharpe ratio Form

Jobson and Korkie (1982) and Gibbons, Ross and Shanken (1989) derived efficiency tests using squared Sharpe ratio. In this section are conducted efficiency tests for the models using equations (3.39) and (3.41) that make use of the tangent portfolios.

For the portfolio with 6 assets, from table 8 we note very similar results to those

Table 8 – Portfolio efficiency tests in terms of maximum Sharpe ratios for all 4 portfolios to selected periods of time

Portfolio	Months	JK Test		GRS Test		JK Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
<i>Portfolio N = 6</i>	120	3.3	0.767	0.5	0.790	9.3	0.159	1.4	0.210
	240	23.9	0.001	3.9	0.001	33.9	0.000	5.4	0.000
	360	29.2	0.000	4.8	0.000	55.4	0.000	9.0	0.000
	480	55.7	0.000	9.1	0.000	70.8	0.000	11.6	0.000
	600	52.3	0.000	8.6	0.000	64.7	0.000	10.6	0.000
	720	59.8	0.000	9.9	0.000	70.7	0.000	11.6	0.000
	840	66.3	0.000	11.0	0.000	71.7	0.000	11.8	0.000
	960	69.7	0.000	11.5	0.000	64.1	0.000	10.6	0.000
	1020	50.9	0.000	8.4	0.000	51.3	0.000	8.5	0.000
<i>Portfolio N = 25</i>	120	16.4	0.903	0.5	0.970	71.8	0.000	2.2	0.004
	240	50.7	0.002	1.8	0.013	88.6	0.000	3.1	0.000
	360	60.2	0.000	2.2	0.001	127.2	0.000	4.7	0.000
	480	97.7	0.000	3.7	0.000	117.8	0.000	4.4	0.000
	600	87.9	0.000	3.4	0.000	107.5	0.000	4.1	0.000
	720	93.2	0.000	3.6	0.000	103.0	0.000	4.0	0.000
	840	97.1	0.000	3.8	0.000	97.6	0.000	3.8	0.000
	960	115.6	0.000	4.5	0.000	108.8	0.000	4.2	0.000
	1020	93.6	0.000	3.7	0.000	94.2	0.000	3.7	0.000
<i>Portfolio N = 49</i>	120	64.2	0.072	0.8	0.840	82.0	0.002	0.9	0.574
	240	47.4	0.540	0.8	0.865	57.3	0.194	0.9	0.631
	360	40.5	0.802	0.7	0.926	49.9	0.439	0.9	0.717
	480	47.0	0.554	0.9	0.739	50.9	0.398	0.9	0.616
	600	49.7	0.446	0.9	0.612	97.3	0.000	1.8	0.001
	720	63.2	0.083	1.2	0.169	280.5	0.000	5.3	0.000
	840	63.7	0.077	1.2	0.146	494.8	0.000	9.5	0.000
	960	80.6	0.003	1.6	0.009	766.2	0.000	14.8	0.000
	1020	83.5	0.002	1.6	0.005	944.7	0.000	18.3	0.000
<i>Portfolio N = 100</i>	120	302.5	0.000	0.5	0.990	531.1	0.000	0.8	0.809
	240	118.3	0.103	0.7	0.977	338.2	0.000	1.9	0.000
	360	122.8	0.060	0.9	0.761	430.7	0.000	3.1	0.000
	480	169.2	0.000	1.3	0.029	376.4	0.000	3.0	0.000
	600	144.6	0.002	1.2	0.106	415.1	0.000	3.4	0.000
	720	140.3	0.005	1.2	0.098	492.1	0.000	4.2	0.000
	840	158.2	0.000	1.4	0.010	587.8	0.000	5.2	0.000
	960	219.2	0.000	2.0	0.000	639.3	0.000	5.7	0.000
	1020	228.6	0.000	2.1	0.000	709.7	0.000	6.4	0.000

Tests of portfolio efficiency using all 4 portfolios samples ($N = 6$, $N = 25$, $N = 49$ and $N = 100$) for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). The tests are performed using squared Sharpe ratio. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case.

obtained using Wald and GRS tests via GMM and GEL. For $T = 240$ or higher, it is strongly rejected the hypothesis of efficiency for CAPM and Fama-French models. Just for the CAPM and $T = 120$, when GMM is rejected (under fixed weights and *managed portfolios*), and GEL shows no evidence in this direction (under fixed weights and *managed portfolios*), through tests with squared Sharpe ratio we can conclude that, in line with GEL estimations, there is no evidence for rejection.

Now, for the portfolio with 25 assets, succinctly, when it was possible to conduct the tests, the efficiency hypothesis was rejected. In the table 8 results are in conformity with the tests via GMM and GEL for $T = 240$ or higher. For $T = 120$, the efficiency is rejected under the Fama-French model, but not for the CAPM.

For the portfolio with 49 assets from the industrial categories, the table results are partially in accordance. For the Fama-French model, when $T = 600$ or higher, it is possible to conclude for the rejection of efficiency. Using GMM, compared to results from previous section, only for 360 and 480 months was possible to carry out the tests, being clear that the results are in harmony. Regarding the CAPM, the tests under GEL shows rejection evidence for intervals above 600 months. Table 8 reveals that when $T = 960$ or $T = 1020$ there are no evidences to reject the market proxy.

Finally, for the 100 portfolios formed on size and book-to-market, with Fama-French model the efficiency is rejected except for 120 months (GRS). However, for the CAPM, focusing on GRS test there are no evidence in this direction for 120, 240, 360, 600 and 720 months. But the JK test practically rejects the null hypothesis for all scenarios, except when $T = 240$.

5.3 Fixed-b HAC Estimators Methodology

As presented in section 3.2.3 asymptotic estimators based on a long-run covariance matrix estimated with the conventional use of kernels presents deficiencies in finite samples. So, if we stick the statistical inference only in those estimators, there is reason to argue that the performance can be compromised due to the existence of serial correlation,

measurement errors and moving average contamination or outliers. In such cases, it is expected that the estimators based on the methodology from Kiefer and Vogelsang (2002b) have better performance.

In this section we use the estimator of $\hat{\Omega}_{M=T}$ from the equation (3.26) using the Bartlett kernel with a bandwidth equal to the sample size, according to the methodology of Vogelsang-Kiefer (henceforth VK). The empirical strategy is the same as the previous sections, wherein is evaluated the test results for CAPM and Fama-French models, for selected monthly windows, and both approaches, without conditional information, as well as for *managed portfolios*. It is expected that in finite samples these tests have better size and power properties than conventional asymptotic tests.

5.3.1 6 Portfolios Formed on Size and Book-to-Market

5.3.1.1 No Conditional Information

Table 9 shows the estimation results by GMM and GEL for the 9 periods of months previously defined. The first point to note is that the tests were not possible to be conducted with GEL. This fact was due that with the VK method it was not feasible to estimate a nonsingular variance matrix $\hat{\Omega}_{M=T}$ for any of the samples in this study. So, it is not possible any assessment about the efficiency tests using the VK method in GEL approach. Since this fact is a common result for all other portfolios cases, we decided to present only the GMM panel for all tables in this section.

If we compare the results from this table with the results using standard asymptotic HAC estimator in table 3, there are some interesting results. Note that in general the estimates for both tests (Wald and GRS) are substantially higher than the usual HAC case. This is valid for any time period analyzed. The p-values are virtually zero for all scenarios, indicating strong evidence to reject the null hypothesis of efficiency.

Note that for the shortest period, i.e., 120 months, we have slightly different results in the table 3. While for the latter we have 2% and 4% for Wald and GRS test, respectively; the p-values for tests based on VK are negligible. It is possible that, indeed, in samples

Table 9 – Tests of portfolio efficiency (VK method) using 6 portfolios formed on size and book-to-market for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	139.4	0.000	21.9	0.000	125.7	0.000	19.4	0.000
	240	262.2	0.000	42.4	0.000	256.5	0.000	41.1	0.000
	360	332.8	0.000	54.4	0.000	470.2	0.000	76.4	0.000
	480	417.6	0.000	68.6	0.000	532.5	0.000	87.1	0.000
	600	528.6	0.000	87.1	0.000	994.4	0.000	163.2	0.000
	720	233.7	0.000	38.6	0.000	390.5	0.000	64.3	0.000
	840	276.1	0.000	45.6	0.000	665.4	0.000	109.7	0.000
	960	302.6	0.000	50.1	0.000	699.0	0.000	115.4	0.000
	1020	294.6	0.000	48.8	0.000	497.3	0.000	82.2	0.000

Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market (2x3) for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). Estimations are performed following the methodology of Vogelsang-Kiefer. The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

with finite characteristics the VK method has a different estimation performance via standard HAC. If we make a correction so that VK method uses an optimal weighting matrix extracted from a previous estimation by standard GMM, we get the same tests results reported in table 9.

5.3.1.2 Managed Portfolios

The table 10 presents the tests results for the approach with the use of instruments. Compared with results from the standard HAC estimation presented in table 4, the tests based on GMM estimations with VK methodology strongly indicate rejection of the null hypothesis for time periods ranging from 120 to 720 months. If we analyze, for example, the 120 months period, the rejection of efficiency in Fama-French model is questionable depending on the significance level in use for results obtained via standard HAC (with 1% it is not possible to reject the efficiency for Wald and GRS tests). On the other hand, using VK the evidences to reject are more robust. Not only for the case with 120 observations. Note that the difference in statistics for both tests between tables are considerable. For 360 months the CAPM model has a Wald test of 79 when estimation is done via conventional

HAC. For the same case, using VK the statistic is 4902. These results show evidence that there is a difference in power and size when the test makes use of the VK method.

Table 10 – Tests of portfolio efficiency (VK method) using 6 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	610.9	0.000	95.9	0.000	783.7	0.000	120.8	0.000
	240	917.8	0.000	148.5	0.000	4,184.1	0.000	671.2	0.000
	360	4,902.3	0.000	801.2	0.000	6,532.9	0.000	1,061.6	0.000
	480	3,258.6	0.000	535.2	0.000	10,250.2	0.000	1,676.3	0.000
	600	4,021.4	0.000	662.4	0.000	16,048.9	0.000	2,634.7	0.000
	720	5,396.1	0.000	890.6	0.000	10,124.1	0.000	1,666.3	0.000

Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market (2x3) for 6 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years) and T=720 (60 years). Estimations are performed following the methodology of Vogelsang-Kiefer. The tests are evaluated using conditioning information when instruments are incorporated to the pricing equation. The lagged variables consisting the conditioning information are: (i) 3 month Treasury-bill yield, (ii) industrial production growth, (iii) yield spreads of low-grade over high-grade corporate bonds, (iv) yield spreads of long-term over short-term Treasury-bills (10-year/1-year) and (v) U.S. inflation (CPI). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

The table 15 in appendix B presents the test results for GMM in the case where it is used an optimal weighting matrix extracted from a previous estimation by standard GMM. Note that the results are very similar to those reported in table 10.

5.3.2 25 Portfolios Formed on Size and Book-to-Market

5.3.2.1 No Conditional Information

Table 11 presents the estimations results from the portfolio with 25 assets. As the tests with the use of estimates were not feasible to be conducted via GEL, we focus only on GMM. Compared with the table 5 which uses the standard HAC, the results are in line with the previous portfolio with 6 assets: the test statistics are substantially larger and p-values are all virtually zero for any period. One difference is that in table 5 the GRS and Wald tests were not performed for Fama-French model. With the VK method, it was possible to get tests estimates that indicate a strong rejection of the null hypothesis.

Table 11 – Tests of portfolio efficiency (VK method) using 25 portfolios formed on size and book-to-market for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	2,120.7	0.000	66.4	0.000	787.9	0.000	24.2	0.000
	240	4,206.3	0.000	150.0	0.000	4,022.6	0.000	142.1	0.000
	360	7,141.5	0.000	265.0	0.000	7,423.2	0.000	273.8	0.000
	480	2,498.5	0.000	94.5	0.000	7,267.7	0.000	273.7	0.000
	600	3,437.0	0.000	131.5	0.000	11,034.2	0.000	420.8	0.000
	720	3,563.1	0.000	137.4	0.000	5,877.6	0.000	226.0	0.000
	840	3,453.8	0.000	133.9	0.000	5,568.0	0.000	215.3	0.000
	960	3,537.9	0.000	137.7	0.000	5,985.3	0.000	232.4	0.000
	1020	3,463.4	0.000	135.0	0.000	4,973.2	0.000	193.5	0.000

Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market (5x5) for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). Estimations are performed following the methodology of Vogelsang-Kiefer. The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

In table 16 in appendix B we have the test results with GMM where a correction is done so that VK method uses an optimal weighting matrix extracted from a previous estimation by standard GMM. This way we have a consistent estimator at this stage. Basically, we get the same results reported in table 11, and only for $T = 120$ the Fama-French model tests were not feasible.

5.3.2.2 Managed Portfolios

In table 12 we can see the test results for the approach with instruments using the same portfolio with 25 assets. For GMM the results point to the same direction. But there are two points that need further comment. First, compared to table 6, both tests for CAPM and Fama-French model show extremely high statistics, even if we take into account that this fact also occurred in tables above. Second, while it was not possible to obtain estimates for 240 and 360 months periods when we used non-parametric consistent estimators, with VK method tests were estimable for both models. For these periods, not surprisingly, there is a overwhelming evidence to reject the efficiency in both models.

Table 12 – Tests of portfolio efficiency (VK method) using 25 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	NA	NA	NA	NA	NA	NA	NA	NA
	240	36,739.5	0.000	1,310.4	0.000	7,349.8	0.000	259.7	0.000
	360	55,742.8	0.000	2,068.7	0.000	47,493.6	0.000	1,752.0	0.000
	480	45,916.7	0.000	1,737.2	0.000	65,673.5	0.000	2,473.7	0.000
	600	40,339.6	0.000	1,543.7	0.000	100,450.1	0.000	3,830.5	0.000
	720	37,845.3	0.000	1,459.1	0.000	98,475.3	0.000	3,785.8	0.000

Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market (5x5) for 6 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years) and T=720 (60 years). Estimations are performed following the methodology of Vogelsang-Kiefer. The tests are evaluated using conditioning information when instruments are incorporated to the pricing equation. The lagged variables consisting the conditioning information are: (i) 3 month Treasury-bill yield, (ii) industrial production growth, (iii) yield spreads of low-grade over high-grade corporate bonds, (iv) yield spreads of long-term over short-term Treasury-bills (10-year/1-year) and (v) U.S. inflation (CPI). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

The test results for the case in which the VK method uses an optimal weighting matrix extracted from a prior estimation by standard GMM can be found in table 17 (appendix B). Basically, the results are in the same direction of the tests in table 12, and, in general, the test statistics are slightly higher. Moreover, for T equal to 240 and 360, the tests could not be calculated.

5.3.3 49 Industry Portfolios

5.3.3.1 No Conditional Information

In table 13 we can see the test results for the portfolio formed with 49 industry categories. Here, we have the most interesting results compared with previous portfolios. Without the use of instruments we can summarize that, although GMM generally has substantially higher statistics and p-values for both tests, being all practically zero for both models (CAPM and Fama-French), what indicates the rejection of efficiency. However, note that compared to the table 7 we have different results.

First, note that the tests were not able to be calculated only for 120 months period

Table 13 – Tests of portfolio efficiency (VK method) using 49 industry portfolios for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	1,796.7	0.000	21.4	0.000	NA	NA	NA	NA
	240	2,140.4	0.000	34.6	0.000	908.2	0.000	14.5	0.000
	360	2,815.1	0.000	49.5	0.000	2,011.7	0.000	35.1	0.000
	480	4,723.6	0.000	86.4	0.000	3,340.9	0.000	60.8	0.000
	600	7,349.5	0.000	137.5	0.000	5,609.1	0.000	104.6	0.000
	720	10,214.4	0.000	194.0	0.000	7,932.5	0.000	150.2	0.000
	840	13,464.8	0.000	258.4	0.000	10,397.1	0.000	199.0	0.000
	960	13,823.3	0.000	267.4	0.000	11,094.8	0.000	214.2	0.000
	1020	14,080.1	0.000	273.3	0.000	9,840.7	0.000	190.6	0.000

Tests of portfolio efficiency using 49 industry portfolios for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). Estimations are performed following the methodology of Vogelsang-Kiefer. The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

with the Fama-French model. Without the use of VK method, only periods of 240, 360 and 480 months were estimable (where for 240 months, this was only possible with CAPM). Second, for the same three periods, VK method points to the opposite direction compared to the common HAC. Using VK there are strong evidences against the null hypothesis. On the other hand, for the conventional test based on the asymptotic theory, we found high p-values indicating the opposite.

The table 18 in appendix B shows the tests results when in VK method a correction is done in the weighting matrix. The test results are the same as those reported in table 13. But, for $T = 120$ and periods above 600 months, the tests could not be performed.

5.3.3.2 Managed Portfolios

Finally, the table 14 presents the test results for the multiplicative approach. For the case where the standard HAC was used, tests could not be realized for GMM and GEL, both models, and all periods of time. Therefore, there is no possibility to compare these results with the conventional asymptotic tests.

Solely for GMM, in table 14, the tests were not estimable for samples with finite

Table 14 – Tests of portfolio efficiency (VK method) using 49 industry portfolios under *scaled returns* for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	NA	NA	NA	NA	NA	NA	NA	NA
	240	NA	NA	NA	NA	NA	NA	NA	NA
	360	17,841.7	0.000	313.5	0.000	NA	NA	NA	NA
	480	49,600.7	0.000	906.8	0.000	10,822.6	0.000	196.9	0.000
	600	122,819.1	0.000	2,297.6	0.000	55,048.6	0.000	1,026.1	0.000
	720	262,905.2	0.000	4,992.8	0.000	154,114.2	0.000	2,918.0	0.000

Tests of portfolio efficiency using 49 industry portfolios for 6 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years) and T=720 (60 years). Estimations are performed following the methodology of Vogelsang-Kiefer. The tests are evaluated using conditioning information when instruments are incorporated to the pricing equation. The lagged variables consisting the conditioning information are: (i) 3 month Treasury-bill yield, (ii) industrial production growth, (iii) yield spreads of low-grade over high-grade corporate bonds, (iv) yield spreads of long-term over short-term Treasury-bills (10-year/1-year) and (v) U.S. inflation (CPI). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

characteristics (120 and 240 months). For other periods, the conclusion is in the same direction of the case without the use of instruments: we have extremely high statistics indicating rejection of the efficiency for each of the models in any period of time.

In summary, analyzing all the results presented in this fixed-b estimators section, for GMM all portfolios with 6, 25 and 49 assets show evidences that when the sample size tends to have finite characteristics, especially for 120 months, we note that the VK method may present distinct evidence regarding efficiency than the results from estimators based on a long-run covariance matrix conventionally estimated with the use of kernels.

As the performance of the tests for conventional HAC estimators may be compromised due to the existence of serial correlation, measurement errors, moving average contamination and outliers; thus, it is possible that these diverging results can be explained by the presence of some (or all) of these distortions. In this case, as the HAC estimators performs poorly in the presence of moving average contamination and outliers, for instance, while VK estimators have better performance in finite samples under those situations. Therefore, if we find evidence of the existence of such characteristics, we should take as

more reliable the test results using VK due to its robustness.

For GEL nothing can be concluded. As in all cases the VK method did not provide results to be analyzed, we can not extract any information if the estimators from GEL class would have a gain in performance, or even would show results in line with or without the use of fixed-b estimators. One possible explanation for this inability in GEL estimations with VK method, may be the use of smoothing that might cause difficulties in the estimation process. For the case in which is used a correction in the weighting matrix in VK method, all tests for all periods were not estimable.

5.4 Evaluating Robustness with Monte Carlo Simulations

In this section we are interested in evaluating the properties of the test statistics using GMM and GEL estimators in a finite sample context. The goal here is to analyze the size and power of the Wald and GRS tests under different situations. Therefore, we seek to analyze the robustness of the tests on each of the two different estimation methodologies.

The previous section showed evidence that characteristics such as measurement errors, moving average contamination, heavy tail and outliers may be present in the data. The exercises performed in this section seek to specifically assess the tests under (i) heavy-tailed distributions and (ii) outliers.

To generate data consistent with the null hypothesis that a portfolio is efficient, we restricted the DGP so that this portfolio must be efficient. Then, it is analyzed the tests simulating four different scenarios for the portfolio with six assets under the *managed portfolios* approach, for 120 months period ($N = 6$, $T = 120$), and the Fama-French model. To construct the dataset of artificial returns we used the following model:

$$R_{i,t} = \beta_{i,1}Mkt_t + \beta_{i,2}SMB_t + \beta_{i,3}HML_t + \varepsilon_{i,t}, \quad t = 1, \dots, T \quad ; \quad i = 1, \dots, N \quad (5.1)$$

For each of the six assets in the portfolio we obtained the 3 estimated coefficients of the parameters $\beta_{i,1}, \beta_{i,2}, \beta_{i,3}$ from OLS. It was decided to carry out 500 artificial returns

datasets simulations for each of the four scenarios, mainly because that during estimations of the parameters for the efficiency tests, especially GEL method has a high computational cost. Thus, it is possible to obtain a new artificial dataset for each of the 500 simulations, assuming the following data generating process:

$$R_{i,t}^{\text{Sim}^*} = \hat{\beta}_{i,1}^{\text{OLS}} Mkt_t + \hat{\beta}_{i,2}^{\text{OLS}} SMB_t + \hat{\beta}_{i,3}^{\text{OLS}} HML_t + \hat{\varepsilon}_{i,t}^{\text{Sim}^*}, \quad t = 1, \dots, 120 \quad ; \quad i = 1, \dots, 6 \quad (5.2)$$

All four scenarios have this generating process, where just the shock term construction $\hat{\varepsilon}_{i,t}^{\text{Sim}^*}$ is what differentiates each one of them.

Scenario 1 - Gaussian Shocks: The first scenario seeks to assess the two tests with the presence of Gaussian innovations. The shock generating process is defined by:

$$\begin{aligned} \hat{\varepsilon}_{i,t}^{\text{Sim}^*} &= \hat{\xi}_{i,t}^{\text{Sim}1}, \quad t = 1, \dots, 120 \quad ; \quad i = 1, \dots, 6 \\ \hat{\xi}_{i,t}^{\text{Sim}1} &\sim N(0, \hat{\sigma}_i^2 \text{OLS}) \end{aligned} \quad (5.3)$$

Scenario 2 - Shocks from a t distribution: In the second simulated scenario the aim is to evaluate the efficiency tests under the presence of heavy tails. The construction of the artificial dataset is similar to the previous scenario. However, as heavy tails are characterized by more extreme values in shocks; then, the most appropriate way to model this phenomenon is to use innovations drawn from a t-Student distribution. All three beta coefficients for each of the six assets in the portfolio are estimated from original data using OLS. With this estimation it is possible to construct 500 artificial dataset simulations following the shock generator process below.

$$\begin{aligned} \hat{\varepsilon}_{i,t}^{\text{Sim}^*} &= \hat{\nu}_{i,t}^{\text{Sim}2}, \quad t = 1, \dots, 120 \quad ; \quad i = 1, \dots, 6 \\ \hat{\nu}_{i,t}^{\text{Sim}2} &\sim t(4) \end{aligned} \quad (5.4)$$

Scenario 3 - Outlier on a fixed date: The third and fourth simulation scenarios seek to evaluate the Wald and GRS tests when outliers are present in data. In the third case, we model the generating process to insert a large magnitude shock on a fixed date in

our the sample. Arbitrarily, we chose to include an error in the middle of the sample, i.e., when $t = 60$. Following the structure of the previous scenarios, the beta coefficients of each asset in the portfolio are estimated by OLS, and when $t = T/2 = 60$ there is a negative shock of 5 standard deviations randomly drawn from a Normal distribution with variance estimated with the original data. So, it is possible to obtain 500 simulated returns datasets for the six assets of the portfolio with the following generating process for innovation $\hat{\varepsilon}_{i,t}$:

$$\begin{aligned} \hat{\varepsilon}_{i,t}^{\text{Sim}^*} &= \mathbb{1}_{t=T/2}(\hat{\kappa}_{i,t}^{\text{Sim}3}), \quad t = 1, \dots, 120 \quad ; \quad i = 1, \dots, 6 \\ \mathbb{1}_{t=T/2}(\hat{\kappa}_{i,t}^{\text{Sim}3}) &= \begin{cases} -\hat{\kappa}_{i,t}^{\text{Sim}3} & , \text{ if } t = T/2 \\ 0 & , \text{ if } t \neq T/2 \end{cases} \\ \hat{\kappa}_{i,t}^{\text{Sim}3} &\sim N(0, 5\hat{\sigma}_i^2 \text{OLS}) \end{aligned} \quad (5.5)$$

Scenario 4 - Outlier with 5% probability: The fourth scenario takes another direction to simulate outliers. We used a probability process of extreme events, arbitrarily assuming that there is a probability of 5% that an outlier exists. Thus, we drawn a random value from a Uniform distribution from 0 to 1 for each observation. In case of success, it is added an outlier with 5 standard deviations randomly drawn from a Normal distribution with variance estimated with the original data. In this case, the DGP innovation is given by:

$$\begin{aligned} \hat{\varepsilon}_{i,t}^{\text{Sim}^*} &= \hat{\xi}_{i,t}^{\text{Sim}4} - \mathbb{1}_{\hat{p}_{i,t} < 0.05}(\hat{\kappa}_{i,t}^{\text{Sim}4}), \quad t = 1, \dots, 120 \quad ; \quad i = 1, \dots, 6 \\ \mathbb{1}_{\hat{p}_{i,t} < 0.05}(\hat{\kappa}_{i,t}^{\text{Sim}4}) &= \begin{cases} \hat{\kappa}_{i,t}^{\text{Sim}4} & , \text{ if } \hat{p}_{i,t} < 0.05 \\ 0 & , \text{ if } \hat{p}_{i,t} \geq 0.05 \end{cases} \\ \hat{p}_{i,t}^{\text{Sim}4} &\sim \text{unif}(0, 1) \\ \hat{\xi}_{i,t}^{\text{Sim}4} &\sim N(0, \hat{\sigma}_i^2 \text{OLS}) \\ \hat{\kappa}_{i,t}^{\text{Sim}4} &\sim N(0, 5\hat{\sigma}_i^2 \text{OLS}) \end{aligned} \quad (5.6)$$

5.4.1 Sampling Distributions of the Test Statistics

To analyze the results of the Monte Carlo experiments, we chose to use the graphical method proposed by Davidson and MacKinnon (1998). The authors advocate in favor of

this method, as opposed to present the Monte Carlo simulations results in tables, primarily because of the amplitude in transmitting information that these graphics possess in an easily assimilated format.

The graphs to study the test size are based on the empirical distribution function (EDF) of p-values from Wald and GRS tests, i.e., J_{Wald} and J_{GRS} . Recalling that EDF is defined by:

$$\hat{F}(x_i) \equiv \frac{1}{N} \sum_{j=1}^N \mathbb{1}_{p_j \leq x_i}$$

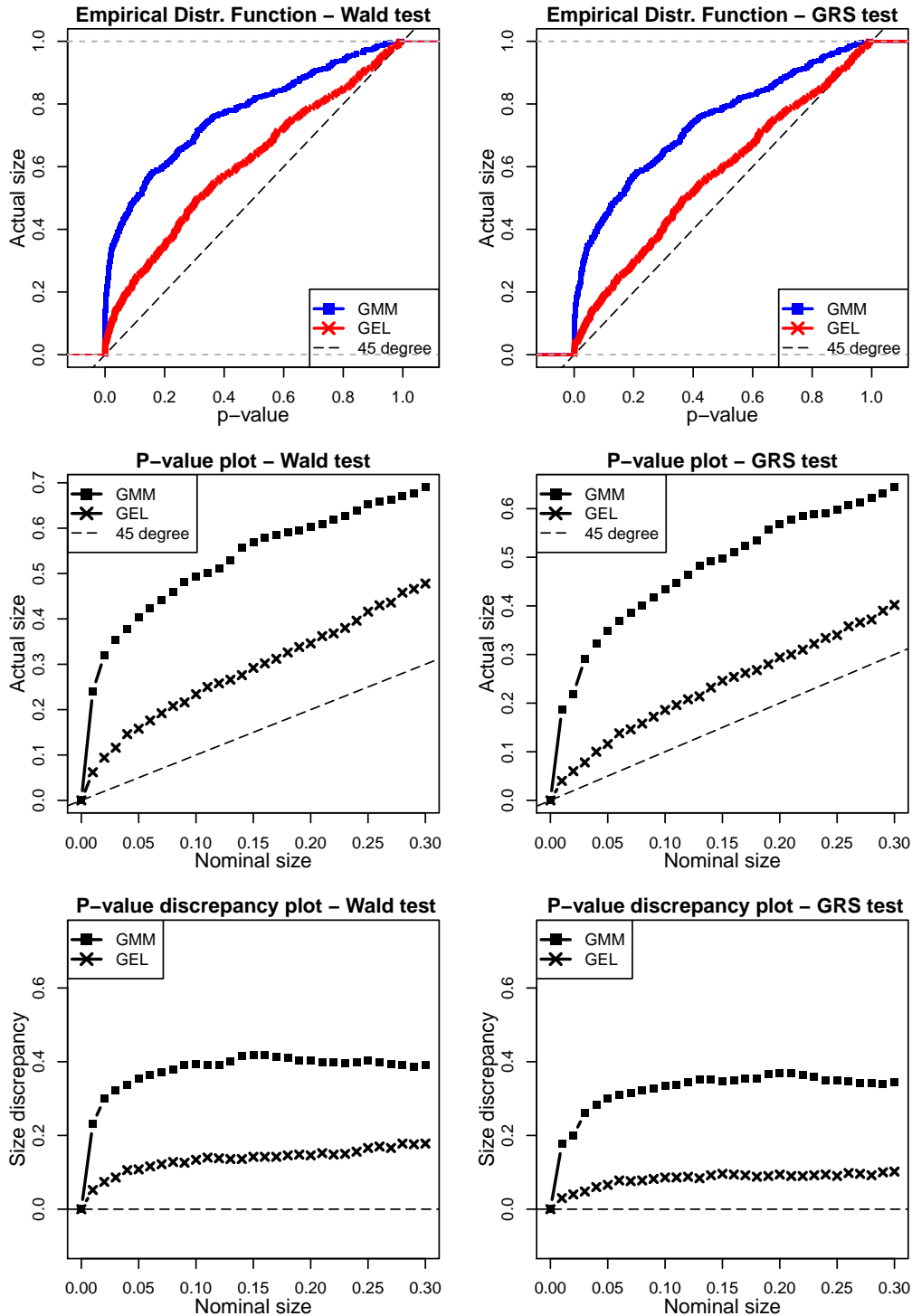
$$\mathbb{1}_{p_j \leq x_i} = \begin{cases} 1 & , \text{ if } p_j \leq x_i \\ 0 & , \text{ if } p_j > x_i \end{cases} \quad (5.7)$$

where p_j^* , for the purpose of this study, is the p-value of the J tests, i.e., $p_j^* \equiv p(J^*)$

Davidson and MacKinnon (1998) propose as first and simplest graphic what they called *P-value plot*. This graph plots $\hat{F}(x_i)$ against x_i , and if the distributions of the tests J_{Wald} and J_{GRS} used to calculate p-values p_j^* are correct; then, each p_j^* must be distributed as a Uniform (0, 1). This implies that the $\hat{F}(x_i)$ chart against x_i should be as close as possible to a 45° line. Therefore, one can say that with *P-value plot* it is possible to quickly evaluate statistical tests that systematically over-reject, under-reject or those that reject about the right proportion of the time.

For situations where the tests statistics being studied behave close to the way they should, i.e., with graphs being close to the 45° line, the authors proposed the *P-value discrepancy plot*. This chart plots $\hat{F}(x_i) - x_i$ against x_i . According to the authors, there are advantages and disadvantages in this representation. Among the advantages of this chart, we have that it presents more information than *P-value plot* when statistics of the tests are well behaved. However, this information may be spurious, being just a result of the experiment randomness conducted. Furthermore, there is no natural scale for the vertical axis, which could cause some difficulty in interpretation. For the *P-value discrepancy plot*, if the distribution is correct, then each p_j^* must be distributed as a Uniform (0, 1) and the graph of $\hat{F}(x_i) - x_i$ against x_i should be near the horizontal axis.

Figure 10 – Simulated scenario 1 with Gaussian innovations ($\hat{\varepsilon}_{i,t}^{\text{Sim}^*} = \hat{\xi}_{i,t}^{\text{Sim}1}$) in Wald and GRS tests (Model=Fama-French, N=6, T=120, 500 simulations)



All three left panels are simulations for J_{Wald} test, while the three right panels are simulations for J_{GRS} test. The two top panels are the EDF graphics of the p-values obtained via GMM and GEL for both tests. The two central panels are the *P-value plot*, while the two bottom panels are the *P-value discrepancy plot*. In order to facilitate visualization, it is included in the EDF and *P-value plot* charts a dashed line representing a 45° line. For the *P-value discrepancy plot* the dashed line represents the x-axis.

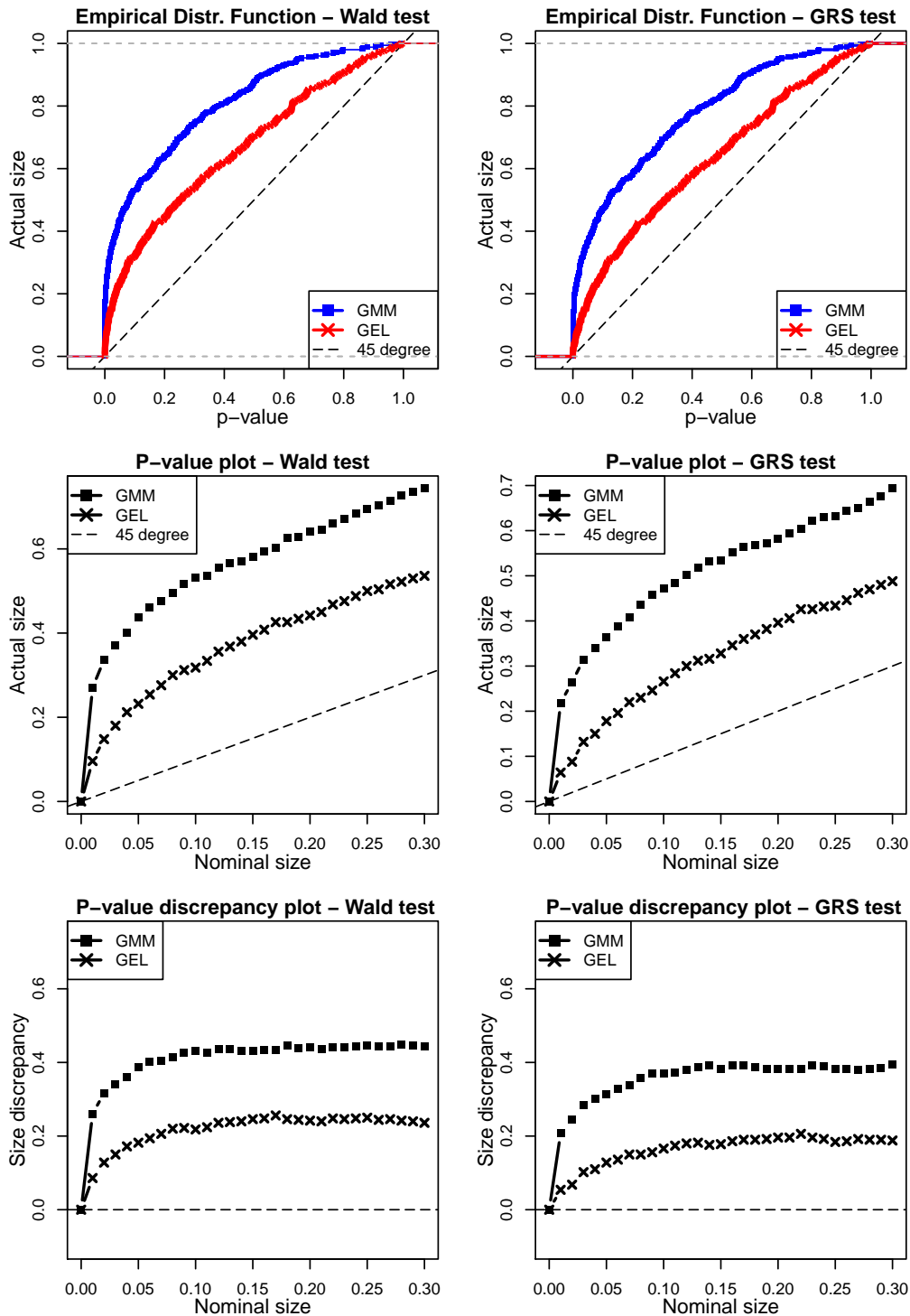
The results for the first simulated scenario derived from a Normal shock can be seen in figure 10. The panels which are in the left column, represents the Wald test, while the right ones refer to the GRS test. The panels in the first line are the EDF of the tests p-values, the panels on the center are the *P-value plot*, while those on the bottom are the *P-value discrepancy plot*. In the first two types of charts, the straight 45° line is represented by a dashed line. For *P-value discrepancy plot* the dashed line represents the abscissa axis.

Analyzing the *P-value plot* we note that GEL provides better p-values than GMM, for both, Wald and GRS tests under the null hypothesis. It can be seen that both, GEL and GMM over-reject for any nominal size. For instance, taking 5% nominal size, for Wald test, GMM shows an actual size test of 40.36%, while GEL has less than half of that (15.8%). For the same 5% nominal size, the GRS test derived for finite samples actually has better performance for both GMM and GEL. However, GEL still has a better performance. Regarding the *P-value discrepancy plot*, one can see similar results. Based on these graphs, it is possible to notice a superiority by GEL compared to GMM for estimating parameters, for J_{Wald} and J_{GRS} tests when Gaussian shocks exists.

In Appendix B can be found the table 19 presenting tabulated rejection proportions for the most common nominal sizes, both tests and all four Monte Carlo experiments scenarios.

The results for the second scenario with shocks from a t distribution are presented in figure 11. The structure of the panels remained constant. In this scenario, by inserting a shock from a t , we investigate the tests robustness for data with distribution having heavy tail characteristics. Easily one can note that tests based on GMM estimations performs badly in finite samples for distributions with long tail. For 5% nominal size, the Wald test using the GMM has an actual size of 43.68%, while GEL has slightly more than half of it (23.2%). For GRS test, both estimators have performance improved. For the same 5% nominal size, GMM has actual size of 36.47 and GEL 17.8. However, while one can say that GMM has poor performance in finite samples with heavy tails compared to GEL, this result can not hide the fact that both estimators generally over-reject under these circumstances. Even if we consider that GEL performs better, having an actual

Figure 11 – Simulated scenario 2 with shocks from a t distribution ($\hat{\varepsilon}_{i,t}^{\text{Sim}^*} = \hat{\nu}_{i,t}^{\text{Sim}2}$) in Wald and GRS tests (Model=Fama-French, N=6, T=120, 500 simulations)



All three left panels are simulations for J_{Wald} test, while the three right panels are simulations for J_{GRS} test. The two top panels are the EDF graphics of the p-values obtained via GMM and GEL for both tests. The two central panels are the P -value plot, while the two bottom panels are the P -value discrepancy plot. In order to facilitate visualization, it is included in the EDF and P -value plot charts a dashed line representing a 45° line. For the P -value discrepancy plot the dashed line represents the x-axis.

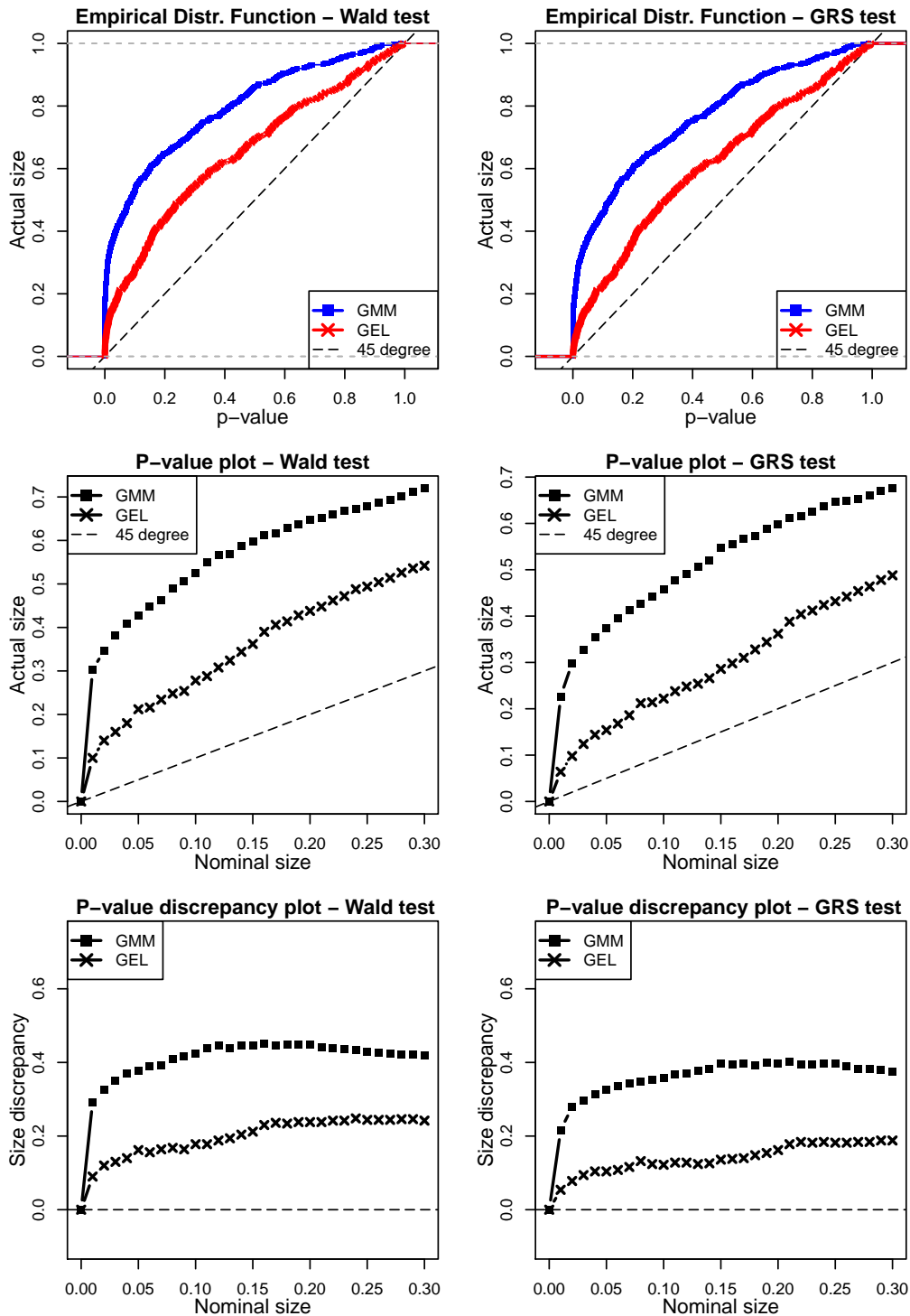
size of nearly 5 times the 5% nominal size for the Wald test, and an actual size of more than 3 times the 5% nominal size for GRS test, one can not necessarily conclude that its performance is satisfactory.

In figure 12 can be found the results for the third scenario, with great magnitudes shocks in the middle of the sample. The aim is to check robustness in the presence of outliers. Here, again the evidences are in the same direction, demonstrating that the GMM has worse performance compared to GEL under the efficiency null hypothesis imposed. Note that both estimators always over-reject when we insert a random shock with 5 standard deviations in middle of the sample.

Finally, in figure 13 we have the results for the fourth scenario, in which we also seek to evaluate robustness to outliers. In this scenario, innovation comes from a Gaussian distribution, being added a shock with 5 standard deviations from a Normal, when it is drawn from a Uniform $(0, 1)$ a value lower than 0.05. Here we have interesting results that differ from earlier ones. The J_{Wald} and J_{GRS} tests based on GMM estimations show better results than via GEL for any nominal size we take. However, note that this superiority is tenuous, being more discernible for nominal values below 10%. Taking 5% nominal size, the Wald test with GMM has an actual size of 90.34%, while GEL has 95.6%. For GRS test, assuming the same 5% nominal size, GMM has 86.5% and GEL 93%. Analyzing the *P-value discrepancy plot* we can note similar pattern with an important feature: for both tests, both GMM and GEL estimations tend to consistently improve performance after reaching a peak of discrepancy around nominal size of 5%.

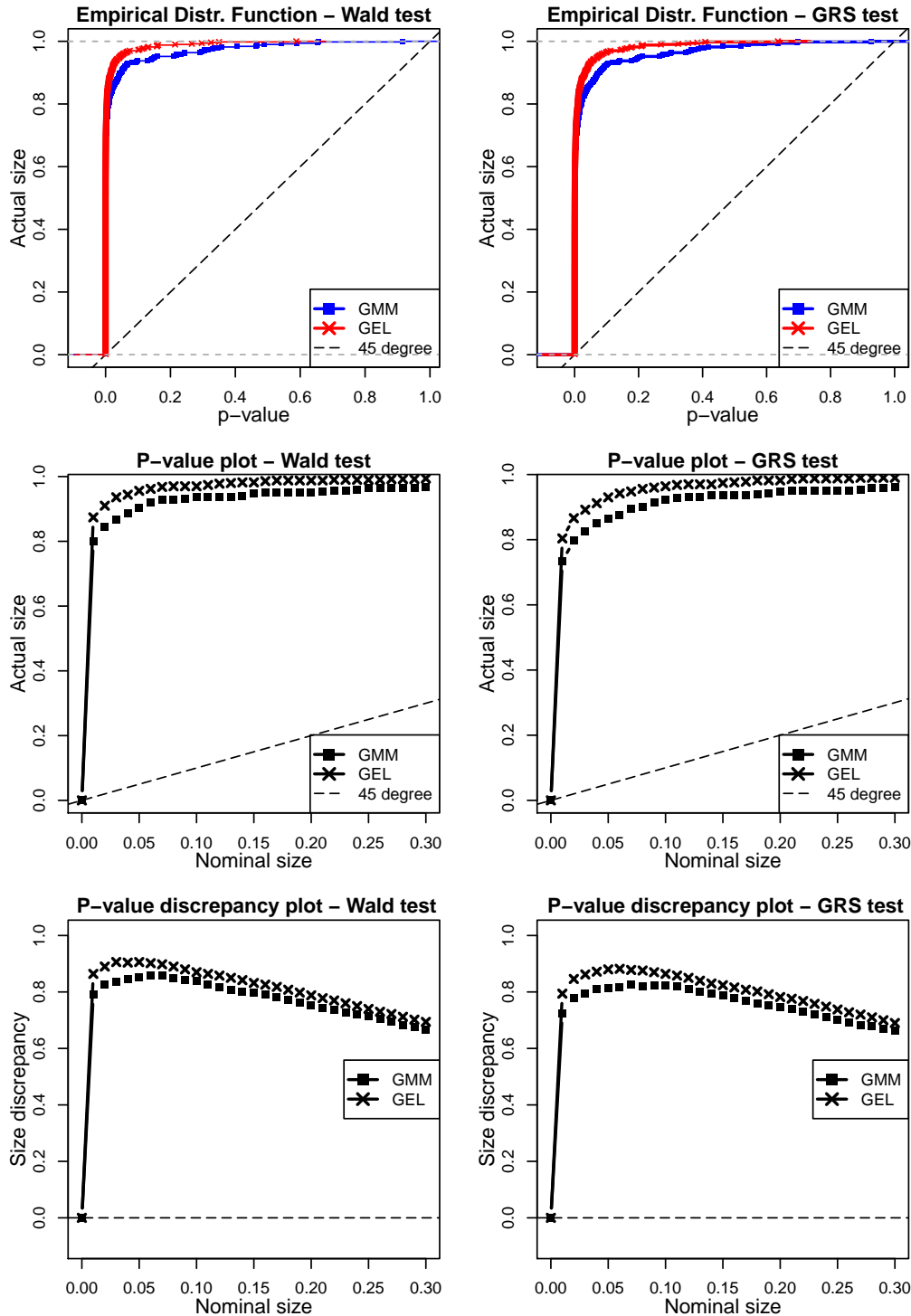
In summary, analyzing all the results presented in this section, it was possible to observe evidences that efficiency tests in finite samples with GEL estimations tend to have better performance compared to estimations via GMM. Furthermore, tests using GEL are more robust to the presence of heavy tails. To assess the robustness for outliers, depending on the generating process we assume, both GMM and GEL can be advantageously compared with each other. However, these results also demonstrate that, whichever the estimator and the test we evaluate, in general, the Wald and GRS tests have a tendency to over-reject.

Figure 12 – Simulated scenario 3 with shocks at $t=T/2$ defined by $\hat{\varepsilon}_{i,t}^{\text{Sim}^*} = \mathbb{1}_{t=T/2}(\hat{k}_{i,t}^{\text{Sim}3})$ in Wald and GRS tests (Model=Fama-French, $N=6$, $T=120$, 500 simulations)



All three left panels are simulations for J_{Wald} test, while the three right panels are simulations for J_{GRS} test. The two top panels are the EDF graphics of the p-values obtained via GMM and GEL for both tests. The two central panels are the P -value plot, while the two bottom panels are the P -value discrepancy plot. In order to facilitate visualization, it is included in the EDF and P -value plot charts a dashed line representing a 45° line. For the P -value discrepancy plot the dashed line represents the x-axis.

Figure 13 – Simulated scenario 4 with shocks defined by $\varepsilon_{i,t}^{\text{Sim}^*} = \hat{\xi}_{i,t}^{\text{Sim}4} - \mathbb{1}_{\hat{p}_{i,t} < 0.05}(\hat{k}_{i,t}^{\text{Sim}4})$ in Wald and GRS tests (Model=Fama-French, N=6, T=120, 500 simulations)



All three left panels are simulations for J_{Wald} test, while the three right panels are simulations for J_{GRS} test. The two top panels are the EDF graphics of the p-values obtained via GMM and GEL for both tests. The two central panels are the *P-value plot*, while the two bottom panels are the *P-value discrepancy plot*. In order to facilitate visualization, it is included in the EDF and *P-value plot* charts a dashed line representing a 45° line. For the *P-value discrepancy plot* the dashed line represents the x-axis.

6 Conclusions

The purpose of this study was to evaluate the behavior of the estimators from GEL class in tests of portfolios efficiency. Comparing the results for (a) different portfolios sizes, (b) different composition methods, as well as (c) increasing periods of time, it can be seen from the analysis that (i) in general, efficiency tests using GEL generate lower estimates compared to tests using the standard approach with GMM; and (ii) when the sample may be characterized as finite, with low n and T , we note that the results are conflicting among the methodologies. These results are significant when the hypothesis of efficiency is evaluated for both models, the CAPM and the Fama-French. Under the unconditional efficiency structure these findings also apply for either fixed weights, or conditional information with instruments in the multiplicative approach.

These results may be an evidence that estimators from GEL class really performs differently in small samples. In addition, they may show that tests based on GMM, or even by maximum likelihood estimation, have a tendency to over-reject the null hypothesis.

On the other hand, if we use the Vogelsang-Kiefer method in which one would expect more robust tests for finite samples, in order to circumvent problems of over-rejection, we obtained results for Wald and GRS tests very similar with the case of the standard long-run covariance HAC matrix estimation via GMM. However, for T low, in particular for $T = 120$, decisions based on tests of efficiency may differ between using conventional HAC and VK method, wherein with the latter there is a strong evidence in favor of rejecting the null hypothesis of efficiency. For GEL, the Vogelsang-Kiefer correction provided no evidence and nothing can be concluded about the use of this method for GEL estimators in efficiency tests.

Finally, in order to assess the robustness of the tests with the use of GMM and GEL estimators in a finite sample context, this study sought through Monte Carlo experiments to examine the effects that distortions in the data series may cause on tests of efficiency, and consequently, in decisions based on these results. In general, it was noted that GEL

has better performance when heavy tails are present. Depending on the DGP we choose to use, both GMM and GEL may have better robustness to outliers compared among them. However, under the null hypothesis, the Wald and GRS tests using both estimators have a tendency to over-reject the hypothesis of efficiency in finite samples.

Bibliography

- ALMEIDA, C.; GARCIA, R. Empirical likelihood estimators for stochastic discount factors. In: *EFA 2008 Athens Meetings Paper*. [S.l.: s.n.], 2008. 21
- ALMEIDA, C.; GARCIA, R. Assessing misspecified asset pricing models with empirical likelihood estimators. *Journal of Econometrics*, Elsevier, v. 170, n. 2, p. 519–537, 2012. 21
- ANATOLYEV, S. GMM, GEL, serial correlation, and asymptotic bias. *Econometrica*, Wiley Online Library, v. 73, n. 3, p. 983–1002, 2005. 30
- ANATOLYEV, S.; GOSPODINOV, N. *Methods for Estimation and Inference in Modern Econometrics*. [S.l.]: CRC Press, 2011. 20, 28, 29, 30
- ANDREWS, D. W. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, JSTOR, p. 817–858, 1991. 32, 33
- CAMPBELL, J. Y. et al. *The Econometrics of Financial Markets*. [S.l.]: Princeton University Press, 1997. 62
- CHAUSSÉ, P. Computing generalized method of moments and generalized empirical likelihood with R. *Journal of Statistical Software*, v. 34, n. 11, p. 1–35, 2010. 19
- CHEN, Z.; KNEZ, P. J. Portfolio performance measurement: Theory and applications. *Review of Financial Studies*, Soc Financial Studies, v. 9, n. 2, p. 511–555, 1996. 23
- CHRISTOPHERSON, J. A.; FERSON, W. E.; GLASSMAN, D. A. Conditioning manager alphas on economic information: Another look at the persistence of performance. *Review of Financial Studies*, Soc Financial Studies, v. 11, n. 1, p. 111–142, 1998. 23
- COCHRANE, J. H. *Asset Pricing*. [S.l.]: Princeton University Press, 2009. 20, 25, 38
- DAVIDSON, R.; MACKINNON, J. Graphical methods for investigating the size and power of hypothesis tests. *The Manchester School of Economic & Social Studies*, University of Manchester, v. 66, n. 1, p. 1–26, 1998. 77, 78
- FAMA, E. F. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, Wiley Online Library, v. 25, n. 2, p. 383–417, 1970. 35, 36
- FAMA, E. F.; FRENCH, K. R. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, Elsevier, v. 25, n. 1, p. 23–49, 1989. 43
- FAMA, E. F.; FRENCH, K. R. The cross-section of expected stock returns. *The Journal of Finance*, Wiley Online Library, v. 47, n. 2, p. 427–465, 1992. 58
- FAMA, E. F.; FRENCH, K. R. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, Elsevier, v. 33, n. 1, p. 3–56, 1993. 35, 45, 47
- FERSON, W. E.; FOERSTER, S. R. Finite sample properties of the generalized method of moments in tests of conditional asset pricing models. *Journal of Financial Economics*, Elsevier, v. 36, n. 1, p. 29–55, 1994. 20

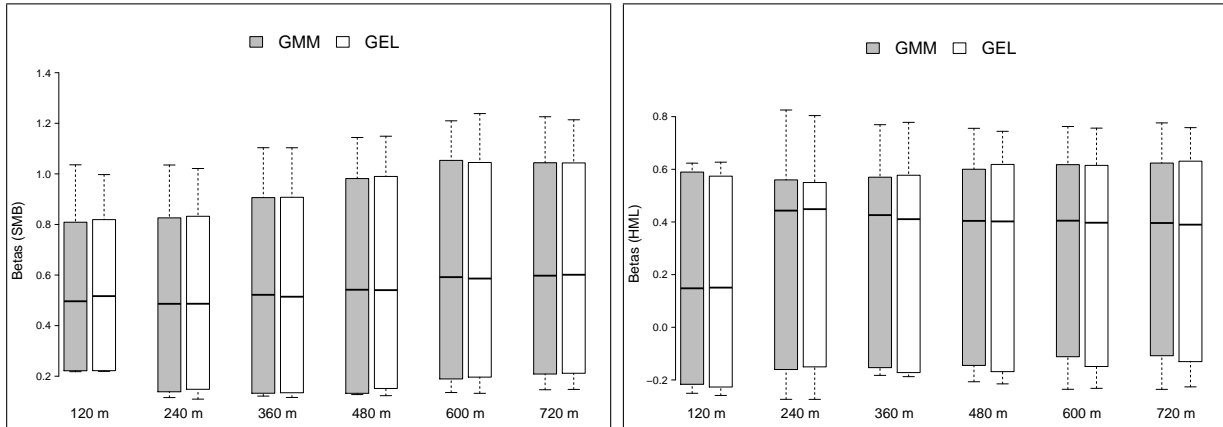
- FERSON, W. E.; QIAN, M. Conditional performance evaluation, revisited. *Research Foundation Monograph of the CFA Institute*, Citeseer, 2004. 43
- FERSON, W. E.; SCHADT, R. W. Measuring fund strategy and performance in changing economic conditions. *Journal of Finance*, JSTOR, p. 425–461, 1996. 23
- FERSON, W. E.; SIEGEL, A. F. The efficient use of conditioning information in portfolios. *The Journal of Finance*, Wiley Online Library, v. 56, n. 3, p. 967–982, 2001. 23
- FERSON, W. E.; SIEGEL, A. F. Testing portfolio efficiency with conditioning information. *Review of Financial Studies*, Oxford University Press (OUP), v. 22, n. 7, p. 2735–2758, Jul 2009. ISSN 1465-7368. 19, 23, 24, 26, 40, 43, 58, 61
- GIBBONS, M. R.; ROSS, S. A.; SHANKEN, J. A test of the efficiency of a given portfolio. *Econometrica*, JSTOR, p. 1121–1152, 1989. 39, 65
- HANSEN, L. P.; HEATON, J.; YARON, A. Finite-sample properties of some alternative GMM estimators. *Journal of Business & Economic Statistics*, Taylor & Francis Group, v. 14, n. 3, p. 262–280, 1996. 31
- HANSEN, L. P.; RICHARD, S. F. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, JSTOR, p. 587–613, 1987. 23
- HANSEN, L. P.; SINGLETON, K. J. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica*, JSTOR, p. 1269–1286, 1982. 39
- JENSEN, M. C. The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, Wiley Online Library, v. 23, n. 2, p. 389–416, 1968. 23, 27
- JOBSON, J. D.; KORKIE, B. Potential performance and tests of portfolio efficiency. *Journal of Financial Economics*, Elsevier, v. 10, n. 4, p. 433–466, 1982. 40, 65
- KEIM, D. B.; STAMBAUGH, R. F. Predicting returns in the stock and bond markets. *Journal of Financial Economics*, Elsevier, v. 17, n. 2, p. 357–390, 1986. 43
- KIEFER, N. M.; VOGELSANG, T. J. Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. *Econometrica*, JSTOR, p. 2093–2095, 2002. 35
- KIEFER, N. M.; VOGELSANG, T. J. Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size. *Econometric Theory*, Cambridge Univ Press, v. 18, n. 06, p. 1350–1366, 2002. 34, 35, 68
- KIEFER, N. M.; VOGELSANG, T. J. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory*, Cambridge Univ Press, v. 21, n. 6, p. 1130, 2005. 34
- KITAMURA, Y.; STUTZER, M. An information-theoretic alternative to generalized method of moments estimation. *Econometrica*, JSTOR, p. 861–874, 1997. 30
- LINTNER, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, JSTOR, p. 13–37, 1965. 35

- MARKOWITZ, H. Portfolio selection. *The Journal of Finance*, Wiley Online Library, v. 7, n. 1, p. 77–91, 1952. 19, 22
- MÜLLER, U. K. A theory of robust long-run variance estimation. *Journal of Econometrics*, Elsevier, v. 141, n. 2, p. 1331–1352, 2007. 33
- NEWBY, W. K.; SMITH, R. J. Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica*, Wiley Online Library, v. 72, n. 1, p. 219–255, 2004. 20, 29, 30, 31
- NEWBY, W. K.; WEST, K. D. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, v. 55, n. 3, p. 703–708, 1987. 32, 33
- OWEN, A. B. *Empirical Likelihood*. [S.l.]: Chapman and Hall, 2001. 28, 30
- PEÑARANDA, F.; SENTANA, E. Spanning tests in return and stochastic discount factor mean–variance frontiers: A unifying approach. *Journal of Econometrics*, Elsevier, v. 170, n. 2, p. 303–324, 2012. 61
- SHARPE, W. F. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, Wiley Online Library, v. 19, n. 3, p. 425–442, 1964. 19, 22, 35
- SHARPE, W. F. Mutual fund performance. *Journal of Business*, JSTOR, p. 119–138, 1966. 23
- SMITH, R. J. Alternative semi-parametric likelihood approaches to generalised method of moments estimation. *The Economic Journal*, JSTOR, p. 503–519, 1997. 28
- SORTINO, F. A.; PRICE, L. N. Performance measurement in a downside risk framework. *The Journal of Investing*, Institutional Investor Journals, v. 3, n. 3, p. 59–64, 1994. 23
- TREYNOR, J. L. How to rate management of investment funds. *Harvard Business Review*, Harvard Business School Publishing Corporation, v. 43, n. 1, p. 63–75, 1965. 23

Appendix

APPENDIX A – Additional Graphics

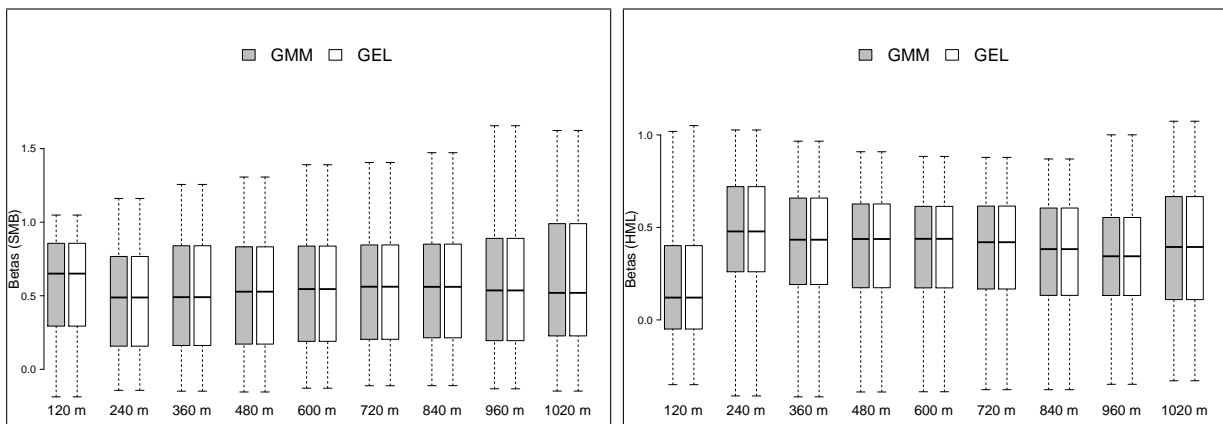
Figure 14 – Boxplots for comparison of estimations by GMM and GEL for Fama-French model using 6 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time



(a) Fama-French Model: $Betas_{SMB}$ estimated by GMM and GEL (b) Fama-French Model: $Betas_{HML}$ estimated by GMM and GEL

In both panels, the x-axis represents the time intervals, starting with 120 months before December 2014 to 720 months prior to this date. The y-axis are the estimated coefficients values by GMM and GEL. Estimation by GMM is represented by gray boxplots, while GEL estimation is represented by white boxplots.

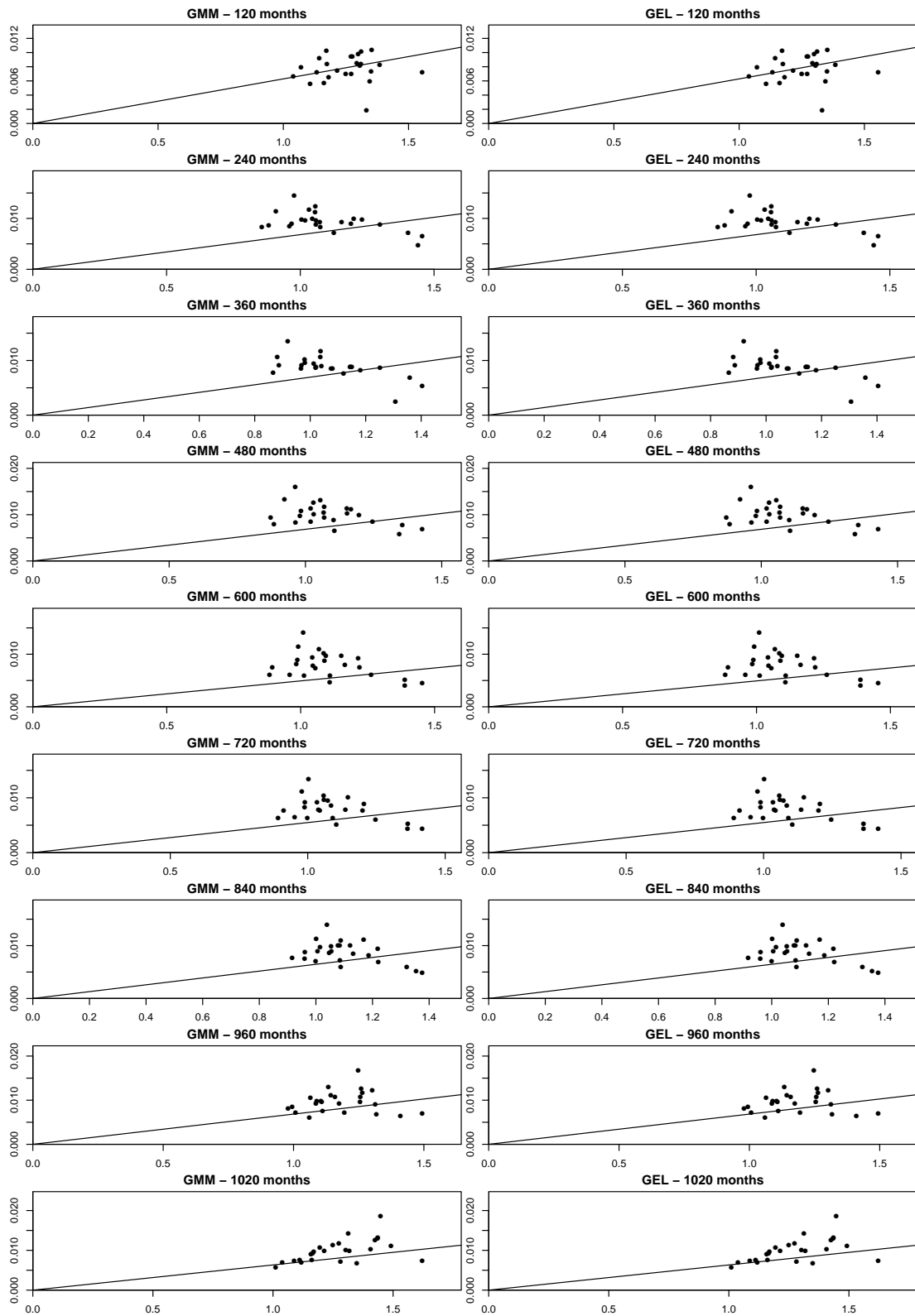
Figure 15 – Boxplots for comparison of estimations by GMM and GEL for Fama-French model using 25 portfolios formed on size and book-to-market for selected periods of time



(a) Fama-French Model: $Betas_{SMB}$ estimated by GMM and GEL (b) Fama-French Model: $Betas_{HML}$ estimated by GMM and GEL

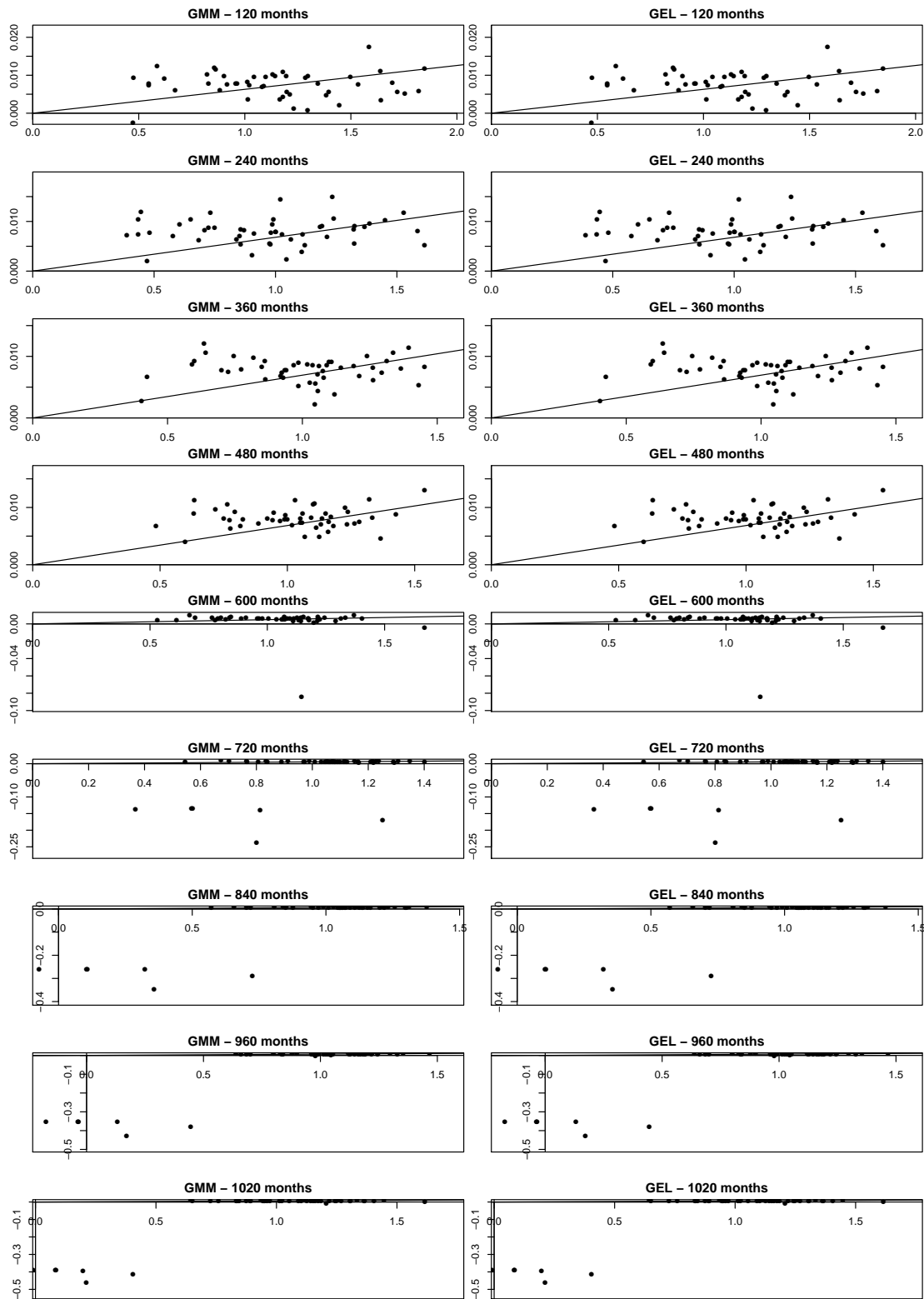
In both panels, the x-axis represents the time intervals, starting with 120 months before December 2014 to 1020 months prior to this date. The y-axis are the estimated coefficients values by GMM and GEL. Estimation by GMM is represented by gray boxplots, while GEL estimation is represented by white boxplots.

Figure 16 – CAPM Model - Comparison of GMM and GEL estimated betas against the sample mean of monthly excess returns for the portfolio with 25 assets



In all panels, the estimated betas ($\hat{\beta}_{Mkt}$) are in the x-axis, and the sample mean of the monthly excess returns for each of the $n = 25$ assets in the portfolio are in the y-axis. Estimations via GMM are on the left, while via GEL are on the right. Each panel represents one of the time intervals, starting with 120 months before December 2014 to 1020 months (85 years) prior to this date.

Figure 17 – CAPM Model - Comparison of GMM and GEL estimated betas against the sample mean of monthly excess returns for the portfolio with 49 assets



In all panels, the estimated betas ($\hat{\beta}_{Mkt}$) are in the x-axis, and the sample mean of the monthly excess returns for each of the $n = 49$ assets in the portfolio are in the y-axis. Estimations via GMM are on the left, while via GEL are on the right. Each panel represents one of the time intervals, starting with 120 months before December 2014 to 1020 months (85 years) prior to this date.

APPENDIX B – Additional Tables

Table 15 – Tests of portfolio efficiency (VK method - two-step) using 6 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	996.5	0.000	156.4	0.000	124.6	0.000	19.2	0.000
	240	988.1	0.000	159.9	0.000	4,115.9	0.000	660.3	0.000
	360	4,217.0	0.000	689.2	0.000	5,874.9	0.000	954.7	0.000
	480	4,386.5	0.000	720.4	0.000	11,108.3	0.000	1,816.7	0.000
	600	5,261.8	0.000	866.7	0.000	16,045.4	0.000	2,634.1	0.000
	720	4,652.3	0.000	767.8	0.000	8,646.5	0.000	1,423.1	0.000

Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market (2x3) for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). Estimations are performed following the methodology of Vogelsang-Kiefer, in the case in which VK method uses an optimal weighting matrix extracted from a prior estimation by standard GMM. The tests are evaluated using conditioning information when instruments are incorporated to the pricing equation. The lagged variables consisting the conditioning information are: (i) 3 month Treasury-bill yield, (ii) industrial production growth, (iii) yield spreads of low-grade over high-grade corporate bonds, (iv) yield spreads of long-term over short-term Treasury-bills (10-year/1-year) and (v) U.S. inflation (CPI). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case.

Table 16 – Tests of portfolio efficiency (VK method - two-step) using 25 portfolios formed on size and book-to-market for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	2,120.7	0.000	66.4	0.000	NA	NA	NA	NA
	240	4,206.3	0.000	150.0	0.000	4,022.6	0.000	142.1	0.000
	360	7,141.5	0.000	265.0	0.000	7,423.2	0.000	273.8	0.000
	480	2,498.5	0.000	94.5	0.000	7,267.7	0.000	273.7	0.000
	600	3,437.0	0.000	131.5	0.000	11,034.2	0.000	420.8	0.000
	720	3,563.1	0.000	137.4	0.000	5,877.6	0.000	226.0	0.000
	840	3,453.8	0.000	133.9	0.000	5,568.0	0.000	215.3	0.000
	960	3,537.9	0.000	137.7	0.000	5,985.3	0.000	232.4	0.000
	1020	3,463.4	0.000	135.0	0.000	4,973.2	0.000	193.5	0.000

Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market (5x5) for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). Estimations are performed following the methodology of Vogelsang-Kiefer, in the case in which VK method uses an optimal weighting matrix extracted from a prior estimation by standard GMM. The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

Table 17 – Tests of portfolio efficiency (VK method - two-step) using 25 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
GMM	120	NA	NA	NA	NA	NA	NA	NA	NA
	240	NA	NA	NA	NA	NA	NA	NA	NA
	360	NA	NA	NA	NA	NA	NA	NA	NA
	480	45,470.8	0.000	1,720.3	0.000	81,208.6	0.000	3,058.9	0.000
	600	42,184.5	0.000	1,614.3	0.000	108,302.0	0.000	4,129.9	0.000
	720	38,143.0	0.000	1,470.6	0.000	97,003.1	0.000	3,729.2	0.000

Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market (5x5) for 6 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years) and T=720 (60 years). Estimations are performed following the methodology of Vogelsang-Kiefer, in the case in which VK method uses an optimal weighting matrix extracted from a prior estimation by standard GMM. The tests are evaluated using conditioning information when instruments are incorporated to the pricing equation. The lagged variables consisting the conditioning information are: (i) 3 month Treasury-bill yield, (ii) industrial production growth, (iii) yield spreads of low-grade over high-grade corporate bonds, (iv) yield spreads of long-term over short-term Treasury-bills (10-year/1-year) and (v) U.S. inflation (CPI). The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

Table 18 – Tests of portfolio efficiency (VK method - two-step) using 49 industry portfolios for selected periods of time

	Months	Wald Test		GRS Test		Wald Test		GRS Test	
		Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
		CAPM				FF			
<i>GMM</i>	120	NA	NA	NA	NA	NA	NA	NA	NA
	240	2,140.4	0.000	34.6	0.000	NA	NA	NA	NA
	360	2,815.1	0.000	49.5	0.000	2,011.7	0.000	35.1	0.000
	480	4,723.6	0.000	86.4	0.000	3,340.9	0.000	60.8	0.000
	600	NA	NA	NA	NA	NA	NA	NA	NA
	720	NA	NA	NA	NA	NA	NA	NA	NA
	840	NA	NA	NA	NA	NA	NA	NA	NA
	960	NA	NA	NA	NA	NA	NA	NA	NA
	1020	NA	NA	NA	NA	NA	NA	NA	NA

Tests of portfolio efficiency using 49 industry portfolios for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). Estimations are performed following the methodology of Vogelsang-Kiefer, in the case in which VK method uses an optimal weighting matrix extracted from a prior estimation by standard GMM. The tests are conducted based on both estimations methodology: via GMM are on top, while via GEL are on the bottom. Tests of efficiency under the CAPM asset pricing model are on the left, while tests under Fama-French three-factors model (represented as “FF”) are on the right. Table presents the statistic and the p-values of the Wald and the GRS tests for each case. “NA” denotes not applicable, situations in which singularities problems occur impeding the inversion of the Var-Cov matrix.

Table 19 – Rejections proportions of J_{Wald} and J_{GRS} tests for nominal sizes of 1%, 2.5%, 5% e 10% from the Monte Carlo experiments for each of the 4 scenarios (Model=Fama-French, N=6, T=120, 500 simulations)

	1%		2.50%				5%				10%					
	Wald Test		GRS Test		Wald Test		GRS Test		Wald Test		GRS Test					
	GMM	GEL	GMM	GEL	GMM	GEL	GMM	GEL	GMM	GEL	GMM	GEL				
$\hat{\epsilon}_{i,t}^{Sim*} = \hat{\xi}_{i,t}^{Sim1}$	0.2410	0.0620	0.1867	0.0400	0.3454	0.1100	0.2631	0.0720	0.4036	0.1580	0.3494	0.1160	0.4940	0.2340	0.4337	0.1860
$\hat{\epsilon}_{i,t}^{Sim*} = \hat{r}_{i,t}^{Sim2}$	0.2705	0.0960	0.2184	0.0640	0.3547	0.1580	0.2926	0.1140	0.4369	0.2320	0.3647	0.1780	0.5311	0.3180	0.4709	0.2660
$\hat{\epsilon}_{i,t}^{Sim*} = \mathbb{1}_{t=T/2}(\hat{r}_{i,t}^{Sim3})$	0.3026	0.1000	0.2265	0.0640	0.3667	0.1480	0.3126	0.1080	0.4269	0.2120	0.3747	0.1540	0.5251	0.2780	0.4569	0.2220
$\hat{\epsilon}_{i,t}^{Sim*} = \hat{\xi}_{i,t}^{Sim4} - \mathbb{1}_{\hat{r}_{i,t} < 0.05(\hat{r}_{i,t}^{Sim4})}$	0.8008	0.8740	0.7344	0.8040	0.8571	0.9260	0.8189	0.8820	0.9034	0.9560	0.8652	0.9300	0.9376	0.9700	0.9235	0.9640

Tabulated rejection proportions from the Monte Carlo experiments for nominal sizes of 1%, 2.5%, 5% e 10%. The sample data used is the 6 portfolios formed on size and book-to-market (2x3) with T=120 observations for each asset. Tests of efficiency are evaluated only for the Fama-French three-factors asset pricing model. Table presents results for 4 different scenarios with 500 simulations each. First case ($\hat{\epsilon}_{i,t}^{Sim*} = \hat{\xi}_{i,t}^{Sim1}$) assess Wald and GRS tests under Gaussian shocks. For the second scenario ($\hat{\epsilon}_{i,t}^{Sim*} = \hat{r}_{i,t}^{Sim2}$) innovations are drawn from a t-Student distribution. In the third case ($\hat{\epsilon}_{i,t}^{Sim*} = \mathbb{1}_{t=T/2}(\hat{r}_{i,t}^{Sim3})$) the DGP insert a large magnitude shock on a fixed date. Finally, the fourth scenario ($\hat{\epsilon}_{i,t}^{Sim*} = \hat{\xi}_{i,t}^{Sim4} - \mathbb{1}_{\hat{r}_{i,t} < 0.05(\hat{r}_{i,t}^{Sim4})}$) the DGP assumes that there is a probability of 5% that an outlier exists.