Portfolio efficiency tests with conditioning information - Comparing GMM and GEL estimators

Caio A. Vigo Pereira and Márcio Poletti Laurini
FEA-RP - University of São Paulo

2016

Abstract

We evaluate the use of Generalized Empirical Likelihood (GEL) estimators in portfolio efficiency tests for asset pricing models in the presence of conditional information. Estimators from GEL family present 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 with the number of moment conditions included, which is expected in conditional efficiency analysis. By means of Monte Carlo experiments, we show that GEL estimators has better performance in the presence of data contaminations, especially under heavy tails and outliers. We also see that efficiency tests using GEL generate lower estimates compared to tests using the standard approach with GMM.

Keywords: Portfolio Efficiency. Conditional Information. Efficiency Tests. GEL. GMM.
JEL Classification: C12, C13, C58, G11, G12
Encontro Brasileiro de Econometria - Área de submissão - Finanças

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. 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 better statistical properties, such as robustness to outliers and heavy tails distributions, and better finite 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 GMM estimators do 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.

Firstly, we 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 evidence 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.

In the second part of the work, we evaluate how efficiency tests based on GEL and GMM estimations can lead to different decisions using real datasets. In order to better analyze this issue, we compare 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 indicate 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.

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 paper is as follows. Next section 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 3 provides the simulations experiments to evaluate tests robustness under both methods of estimation. Section 4 presents the datasets (portfolios, factors and instruments) and empirical results. Finally, section 5 concludes. Additional results, tables and figures are presented in Appendix.

2 Methodology

This study aims to evaluate an alternative econometric to the GMM 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 methods. Thus, it is expected that estimators have better statistical properties, such as robustness to data contaminations and better finite sample properties. One essential property of GEL estimators is the absence of some bias components existing in GMM, that impairs the finite sample performance in the presence of a large number of moment conditions. As the conditional portfolio analysis is based on the incorporation of information in the form of additional instruments, this is the setup where the better finite sample properties of GEL estimators can be optimal relative to the suboptimal performance of GMM due to the bias components in this estimator.

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. This properties are explained in the next section.

2.1 Incorporating Conditional Information

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

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

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 (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 (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., $R_{t+1} \equiv \dfrac{x_{t+1}}{p_t}$. Thus, follows that the pricing models can be represented in accordance with the fundamental valuation equation:

$$E(m_{t+1}R_{t+1}|Z_t) = 1. \tag{2}$$

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:

$$E_t(m_{t+1}R_{t+1} \otimes \tilde{Z}_t) = 1 \otimes \tilde{Z}_t, \tag{3}$$

where $\otimes$ represents a Kronecker product. If we take the unconditional expectation in (3) we get:

$$E(m_{t+1}R_{t+1} \otimes \tilde{Z}_t) = E(1 \otimes \tilde{Z}_t). \tag{4}$$

For the equation (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 (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.

2.2 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 (1) we can take unconditional expectations to get:

$$\begin{aligned} & p_t = E_t(m_{t+1}x_{t+1}) \\ \Rightarrow \quad & p_t = E(m_{t+1}x_{t+1}|Z_t) \\ \Rightarrow \quad & E(p_t) = E(m_{t+1}x_{t+1}). \end{aligned} \tag{5}$$

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 (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:

$$E(m_{t+1}x_{t+1} - p_t) = 0, \tag{6}$$

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

$$E(m_{t+1}R_{t+1} - 1) = 0. \tag{7}$$

For *managed portfolios*, the unconditional moments conditions are easily derived from equation (7):

$$E[(m_{t+1}R_{t+1} - 1) \otimes \tilde{Z}_t] = 0 \Rightarrow E[m_{t+1}(R_{t+1} \otimes \tilde{Z}_t) - (1 \otimes \tilde{Z}_t)] = 0. \tag{8}$$

Thus, the sample means of u_t are defined as:

$$g_T = \frac{1}{T}\sum_{t=1}^{T} u_t = \frac{1}{T}\sum_{t=1}^{T}[m_{t+1}R_{t+1} - 1], \tag{9}$$

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

$$g_T = \frac{1}{T}\sum_{t=1}^{T} u_t = \frac{1}{T}\sum_{t=1}^{T}[m_{t+1}(R_{t+1} \otimes \tilde{Z}_t) - (1 \otimes \tilde{Z}_t)]. \tag{10}$$

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 (7), instead of equation (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:

$$\hat{\theta}_T(\hat{W}) \equiv arg\min_{\hat{\theta}} g_T(\hat{\theta})'\hat{W}_T g_T(\hat{\theta}), \tag{11}$$

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

2.2.1 Empirical Likelihood Estimation

Smith (1997) and Owen (2001) recently introduced a new family of estimators denominated *Generalized Empirical Likelihood* (GEL), that just as GMM, can also be based on moment conditions. According to Anatolyev and Gospodinov (2011), this is a non-parametric method that has the important attractive of optimal asymptotic and finite samples properties. 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), consider a system of restrictions on unconditional moments, such as:

$$E[g(w, \theta_0)] = 0, \tag{12}$$

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 $\boldsymbol{p} = (p_1, p_2, \ldots, p_n)$ be a collection of probability weights assigned to each sample observation. Thus, we have the following *empirical likelihood* problem:

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

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

$$\max_{\theta\in\Theta} \min_{\lambda} \frac{1}{n} \sum_{i=1}^{n} -log(1+\lambda' g(w_i,\theta)), \tag{13}$$

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 (13) by an arbitrary criterion that is subject to certain shape conditions, it can be obtained the *GEL estimator*. To do so, let $\rho(\upsilon)$ be a strictly concave smooth function which satisfies $\rho(0)=0$, $\partial\rho(0)/\partial\upsilon = \partial^2\rho(0)/\partial\upsilon^2 = -1$. 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:

$$\min_{\theta\in\Theta} \sup_{\lambda\in\Lambda_n} \sum_{i=1}^{n} \rho(\lambda' g(w_i,\theta)), \tag{14}$$

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

$$\sqrt{n}(\hat{\theta}-\theta_0) \xrightarrow{d} N(0,\Omega_\theta). \tag{15}$$

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 and 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 (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 $\rho(\upsilon) = ln(1-\upsilon)$. Kitamura and Stutzer (1997) developed the estimator *exponential tilting* (ET), where $\rho(\upsilon) = -exp(\upsilon)$. Finally, one can still highlight the *continuous updated estimator* (CUE), where $\rho(\upsilon)$ 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.

2.2.2 HAC Estimation

The equation (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:

$$G_T(\hat{\theta}_T)' W_T g_T(\hat{\theta}_T) = \mathbf{0}, \tag{16}$$

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 asset returns 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 (11), as follows:

$$W^* = \{lim_{n\to\infty} Var(\sqrt{T}\bar{g}_T(\theta_0)) \equiv \Omega(\theta_0)\}^{-1}, \tag{17}$$

where,

$$\Omega(\theta_0) = \sum_{j=-\infty}^{\infty} \gamma(j), \tag{18}$$

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

$$\mathbf{\Omega}(\boldsymbol{\theta}_0) = \sum_{j=-\infty}^{\infty} \mathbf{\Gamma}_j = \mathbf{\Gamma}_0 + \sum_{j=1}^{\infty}\left(\mathbf{\Gamma}_j + \mathbf{\Gamma}_j'\right) = \mathbf{\Lambda}\mathbf{\Lambda}', \tag{19}$$

where $\mathbf{\Lambda}$ is a lower triangular matrix given by the Cholesky decomposition of $\mathbf{\Omega}(\boldsymbol{\theta}_0)$, and $\mathbf{\Gamma}_j$ is the j-th order autocovariance matrix defined by:

$$\mathbf{\Gamma}_j = E(\boldsymbol{g}_t \boldsymbol{g}'_{t-j}), \qquad j = 0, \pm 1, \pm 2, \ldots \tag{20}$$

An important estimator class for $\mathbf{\Omega}(\boldsymbol{\theta}_0)$ matrix are the non-parametric estimators. As the autocovariances are unknown; then, they may be replaced by sample autocovariances:

$$\begin{aligned}
&\hat{\mathbf{\Gamma}}_j = T^{-1} \textstyle\sum_{t=j+1}^{T} \hat{\boldsymbol{g}}_t \hat{\boldsymbol{g}}'_{t-j}, && j = 0, 1, \ldots, T-1 \\
&\hat{\mathbf{\Gamma}}_j = T^{-1} \textstyle\sum_{t=j+1}^{T} \boldsymbol{g}(w_t, \hat{\boldsymbol{\theta}}_T) \boldsymbol{g}(w_{t-j}, \hat{\boldsymbol{\theta}}_T)', && \text{for } j \geq 0, \\
&\hat{\mathbf{\Gamma}}_j = \hat{\mathbf{\Gamma}}'_{-j}, && \text{for } j < 0.
\end{aligned} \tag{21}$$

While it is possible to estimate the corresponding samples, the estimator (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:

$$\hat{\mathbf{\Omega}}_{HAC}(\boldsymbol{\theta}_0) = \sum_{j=-(T-1)}^{T-1} k(j/b) \hat{\mathbf{\Gamma}}_j, \tag{22}$$

where $k(\cdot)$ is a kernel function $k : R \to R$ 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 $\hat{\mathbf{\Omega}}_{HAC}(\boldsymbol{\theta}_0)$ be consistent requires that the bandwidth b grows in a lower rate when compared to the sample size, so that $b \to \infty$ and $b/T \to 0$ when $T \to \infty$.

Some options have been proposed for the kernel function, as well as the bandwidth in (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 $\hat{\mathbf{\Omega}}_{HAC}(\boldsymbol{\theta}_0)$ should be used in the estimation of the parameters via GMM methods. Therefore, the GMM estimator $\hat{\theta}_T$ may be defined as:

$$\hat{\boldsymbol{\theta}}(\hat{\mathbf{\Omega}}_{HAC}(\boldsymbol{\theta}_0)) \equiv arg \min_{\hat{\theta}} g_T(\hat{\theta})' \hat{\mathbf{\Omega}}^{-1}_{HAC}(\boldsymbol{\theta}_0) g_T(\hat{\theta}). \tag{23}$$

2.3 Tests of Efficiency

To test portfolios efficiency within the mean-variance framework, we need to predefine the asset pricing model in use. This study uses two linear pricing factors models which are the most commonly used: the CAPM from Sharpe (1964) and Lintner (1965), in addition to the three-factor of 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.

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

$$m_{t+1} = a + bR^W_{t+1}, \tag{24}$$

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-factor model the SDF is defined as:

$$m_{t+1} = a + b_1 Mkt_{t+1} + +b_2 SMB_{t+1} + +b_3 HML_{t+1}, \tag{25}$$

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 (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 (2) without incorporating conditional information and apply the decomposition of covariance,

$$\begin{aligned} & 1 = E(m_{t+1}R_{i,t+1}) \\ \Rightarrow \quad & 1 = E(m_{t+1})E(R_{i,t+1}) + Cov(m_{t+1}, R_{i,t+1}). \end{aligned} \tag{26}$$

Thus,

$$\begin{aligned} & E(R_{i,t+1}) = \frac{1}{E(m_{t+1})} - \frac{Cov(m_{t+1}, R_{i,t+1})}{E(m_{t+1})} \\ \Rightarrow \quad & E(R_{i,t+1}) = \frac{1}{E(m_{t+1})} + \left(\frac{Cov(m_{t+1}, R_{i,t+1})}{Var(m_{t+1})}\right)\left(-\frac{Var(m_{t+1})}{E(m_{t+1})}\right) \\ \Rightarrow \quad & E(R_{i,t+1}) = R^f + \beta_{i,m}\lambda_m, \end{aligned} \tag{27}$$

so that we get the beta representation, where $R^f = \dfrac{1}{E(m_{t+1})}$ denotes a risk-free rate (or zero-beta rate) if present, $\beta_{i,m} = \dfrac{Cov(m_{t+1}, R_{i,t+1})}{Var(m_{t+1})}$, i.e., the regression coefficient of the return $R_{i,t+1}$ on m (SDF), and $\lambda_m = -\dfrac{Var(m_{t+1})}{E(m_{t+1})}$. 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 slope 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 ($R^e_{i,t+1} = R_{i,t+1} - R^f_{t+1}$), 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:

$$R_{i,t} = \alpha + \beta_i \mathbf{f}_t + \varepsilon_t, \quad t = 1, \dots, T \quad ; \text{and} \quad i = 1, \dots, N \tag{28}$$

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:

$$\begin{aligned} &\boldsymbol{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} \boldsymbol{f}_t + \boldsymbol{\varepsilon}_t \\ \\ &E[\boldsymbol{\varepsilon}_t] = 0 \\ &E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'_t] = \boldsymbol{\Sigma} \\ &Cov[\boldsymbol{f}_t, \boldsymbol{\varepsilon}'_t] = \mathbf{0}, \end{aligned} \tag{29}$$

where $\mathbf{R}_t$, $\boldsymbol{\alpha}$ e $\boldsymbol{\varepsilon}_t$ have $N \times 1$ dimension; while $\boldsymbol{f}_t$ has $K \times 1$ dimension, and $\boldsymbol{\beta}$ is a $N \times K$ matrix. Defining $\boldsymbol{\Sigma}$ as the Variance-Covariance of the disturbances ε_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:

$$E(\boldsymbol{R}) = \boldsymbol{\beta} E(\boldsymbol{f}). \tag{30}$$

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:

$$\begin{aligned} &H_0 : \boldsymbol{\alpha} = \mathbf{0} \\ &H_A : \boldsymbol{\alpha} \neq \mathbf{0}, \end{aligned} \tag{31}$$

whereas the test statistic is given by:

$$J_{Wald} = \hat{\boldsymbol{\alpha}}' [\boldsymbol{Cov}(\hat{\boldsymbol{\alpha}})]^{-1} \hat{\boldsymbol{\alpha}}, \tag{32}$$

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 $\boldsymbol{\Sigma}$ 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 $\hat{\boldsymbol{\alpha}}$ 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 $\hat{\boldsymbol{\Sigma}}$. However, the test requires that the errors are normally distributed, homoskedastic and uncorrelated. This test is defined by:

$$J_{GRS} = \frac{T-N-K}{N}\left(1 + E_T(\boldsymbol{f})'\hat{\boldsymbol{\Omega}}^{-1}E_T(\boldsymbol{f})\right)^{-1}\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\hat{\boldsymbol{\alpha}}, \tag{33}$$

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

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

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

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 $\hat{\boldsymbol{\alpha}}$ has normal distribution, and $\hat{\boldsymbol{\Sigma}}$ has Wishart distribution, more precisely:

$$\begin{aligned} &\hat{\boldsymbol{\alpha}} \sim N\left(\boldsymbol{\alpha}, \frac{1}{T}\left[1 + E_T(\boldsymbol{f})'\hat{\boldsymbol{\Omega}}^{-1}E_T(\boldsymbol{f})\right]\boldsymbol{\Sigma}\right) \\ &T\hat{\boldsymbol{\Sigma}} \sim W_N\left(T-2, \boldsymbol{\Sigma}\right), \end{aligned} \tag{34}$$

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

2.4 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 with no the conditional information, it is conducted the estimation of the pricing equation given in (7). Now, when the information regarding the state of the economy is included, it is used the *managed portfolios* structure presented in section 2.1, in which the moment conditions are derived from equation (4). Notice that in this situation the instruments used, which are lagged variables, are only used in a multiplicative form in the sample moments.

Firstly, we start the assessment using a simulation approach. In order to evaluate robustness, we analyze 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 to examine the tests robustness to disturbances in the data series when both estimators are used. In particular, we analyze the presence of heavy tails and outliers.

After this analysis, we empirically assess both estimations methodologies. For each of these methods, estimations are performed for the same sample and period of time. It is evaluated different sample sizes T of the time series, in such a way that various time intervals are analyzed. In addition, the analysis is done for both asset pricing models already mentioned: the CAPM and the Fama-French three-factor. Having calculated the estimations for each scenario, it is conducted efficiency tests according to section 2.3.

3 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 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 financial literature already discussed that contaminations such as measurement errors, heavy tail and outliers may be present in this type of 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 a portfolio with six assets under the *managed portfolios* approach, for 120 months period (N = 6, T = 120), and the Fama-French three-factor model. With Monte Carlo experiments, we expect to replicate a factor model with parameters extracted from the real dataset.

To construct the dataset of artificial returns we used the following model:

$$R_{i,t} = \beta_{i,1} Mkt_t + +\beta_{i,2} SMB_t + \beta_{i,3} HML_t + \varepsilon_{i,t}, \qquad t = 1, \dots, T \quad ; \quad i = 1, \dots, N. \tag{35}$$

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. Since all parameters are based in the empirical estimations, we decided to present the complete information regarding the portfolio with 6 assets, the instruments, and the factors in the next section.

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:

$$R_{i,t}^{\text{Sim}*} = \hat{\beta}_{i,1}^{OLS} Mkt_t + +\hat{\beta}_{i,2}^{OLS} SMB_t + \hat{\beta}_{i,3}^{OLS} HML_t + \hat{\varepsilon}_{i,t}^{\text{Sim}*}, \qquad t = 1, \dots, 120 \quad ; \quad i = 1, \dots, 6. \tag{36}$$

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:

$$\begin{aligned} \hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\xi}_{i,t}^{\text{Sim1}}, \quad t = 1, \ldots, 120 \quad ; \quad i = 1, \ldots, 6 \\ \hat{\xi}_{i,t}^{\text{Sim1}} \sim N(0, \hat{\sigma}_i^{2\ \text{OLS}}). \end{aligned} \tag{37}$$

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:

$$\begin{aligned} \hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\nu}_{i,t}^{\text{Sim2}}, \quad t = 1, \ldots, 120 \quad ; \quad i = 1, \ldots, 6 \\ \hat{\nu}_{i,t}^{\text{Sim2}} \sim t(4). \end{aligned} \tag{38}$$

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 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 $\hat{\varepsilon}_{i,t}$:

$$\begin{aligned} \hat{\varepsilon}_{i,t}^{\text{Sim}*} = \mathbb{1}_{t=T/2}(\hat{\kappa}_{i,t}^{\text{Sim3}}), \quad t = 1, \ldots, 120 \quad ; \quad i = 1, \ldots, 6 \\ \mathbb{1}_{t=T/2}(\kappa_{i,t}^{\text{Sim3}}) = \begin{cases} -\hat{\kappa}_{i,t}^{\text{Sim3}} & \text{, if } t = T/2 \\ 0 & \text{, if } t \neq T/2 \end{cases} \\ \hat{\kappa}_{i,t}^{\text{Sim3}} \sim N(0, 5\hat{\sigma}_i^{2\ \text{OLS}}). \end{aligned} \tag{39}$$

Scenario 4 - Outlier with 5% probability: The fourth scenario takes another direction to simulate outliers. We use 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:

$$\begin{aligned} \hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\xi}_{i,t}^{\text{Sim4}} - \mathbb{1}_{\hat{p}_{i,t}<0.05}(\hat{\kappa}_{i,t}^{\text{Sim4}}), \quad t = 1, \ldots, 120 \quad ; \quad i = 1, \ldots, 6 \\ \mathbb{1}_{\hat{p}_{i,t}<0.05}(\kappa_{i,t}^{\text{Sim4}}) = \begin{cases} \hat{\kappa}_{i,t}^{\text{Sim4}} & \text{, if } \hat{p}_{i,t}^{\text{Sim4}} < 0.05 \\ 0 & \text{, if } \hat{p}_{i,t}^{\text{Sim4}} \geq 0.05 \end{cases} \\ \hat{p}_{i,t}^{\text{Sim4}} \sim \text{unif}(0, 1) \\ \hat{\xi}_{i,t}^{\text{Sim4}} \sim N(0, \hat{\sigma}_i^{2\ \text{OLS}}) \\ \hat{\kappa}_{i,t}^{\text{Sim4}} \sim N(0, 5\hat{\sigma}_i^{2\ \text{OLS}}). \end{aligned} \tag{40}$$

3.0.1 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 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:

$$\begin{aligned} \hat{F}(x_i) &\equiv \tfrac{1}{N}\textstyle\sum_{j=1}^{N} \mathbb{1}_{p_j \le x_i} \\ \mathbb{1}_{p_j \le x_i} &= \begin{cases} 1 & \text{, if } p_j \le x_i \\ 0 & \text{, if } p_j > x_i, \end{cases} \end{aligned} \tag{41}$$

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.

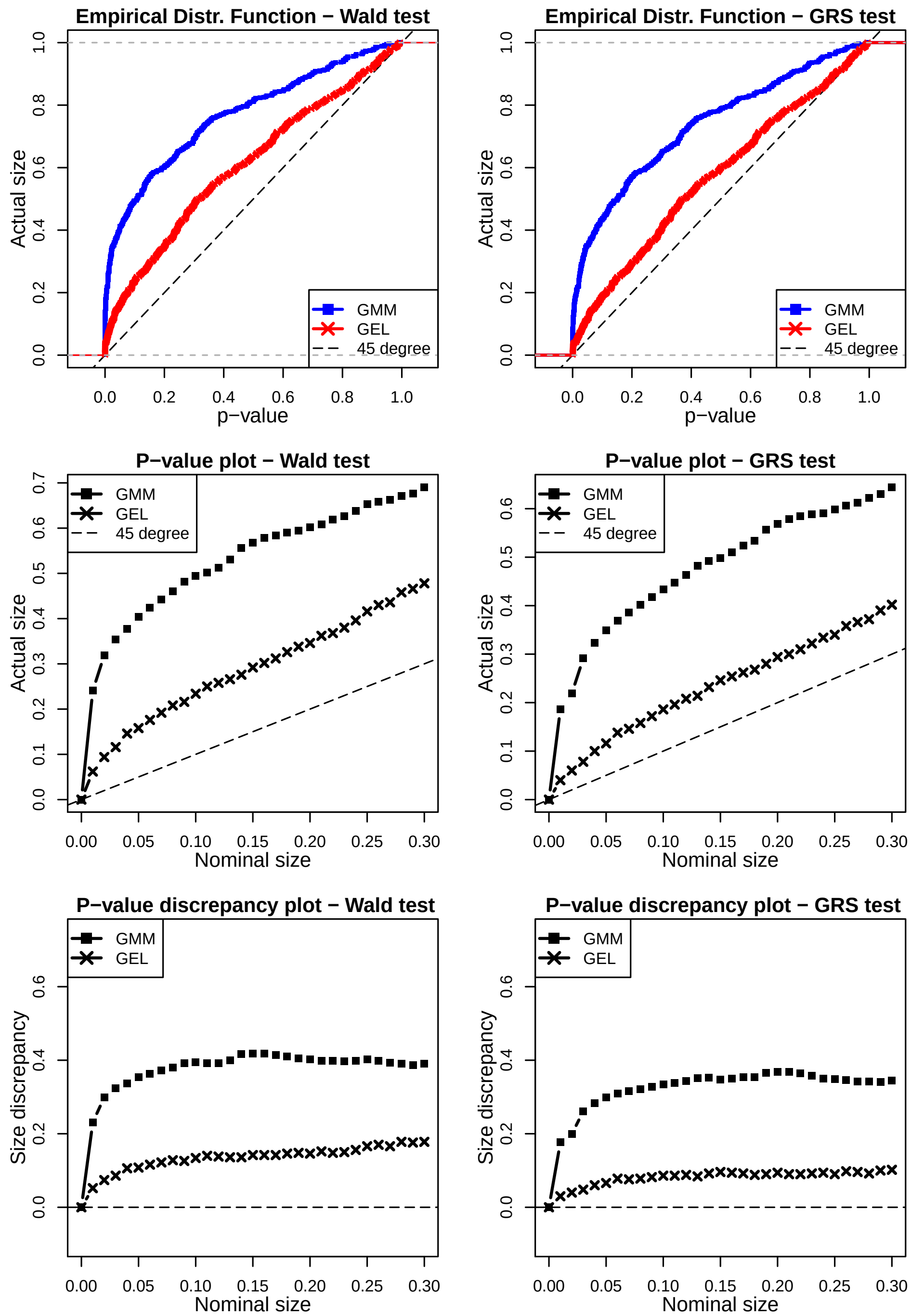
For situations where the tests statistics being studied behave close to the way they should, i.e., with graphs being close to the 45° line, the authors proposed the *P-value discrepancy plot*. This chart plots $\hat{F}(x_i) - x_i$ against x_i. According to the authors, there are advantages and disadvantages 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.

The results for the first simulated scenario derived from a Normal shock can be seen in figure 1. The left panels represent the Wald test, while the right ones refer to the GRS test. The panels in the first row 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.

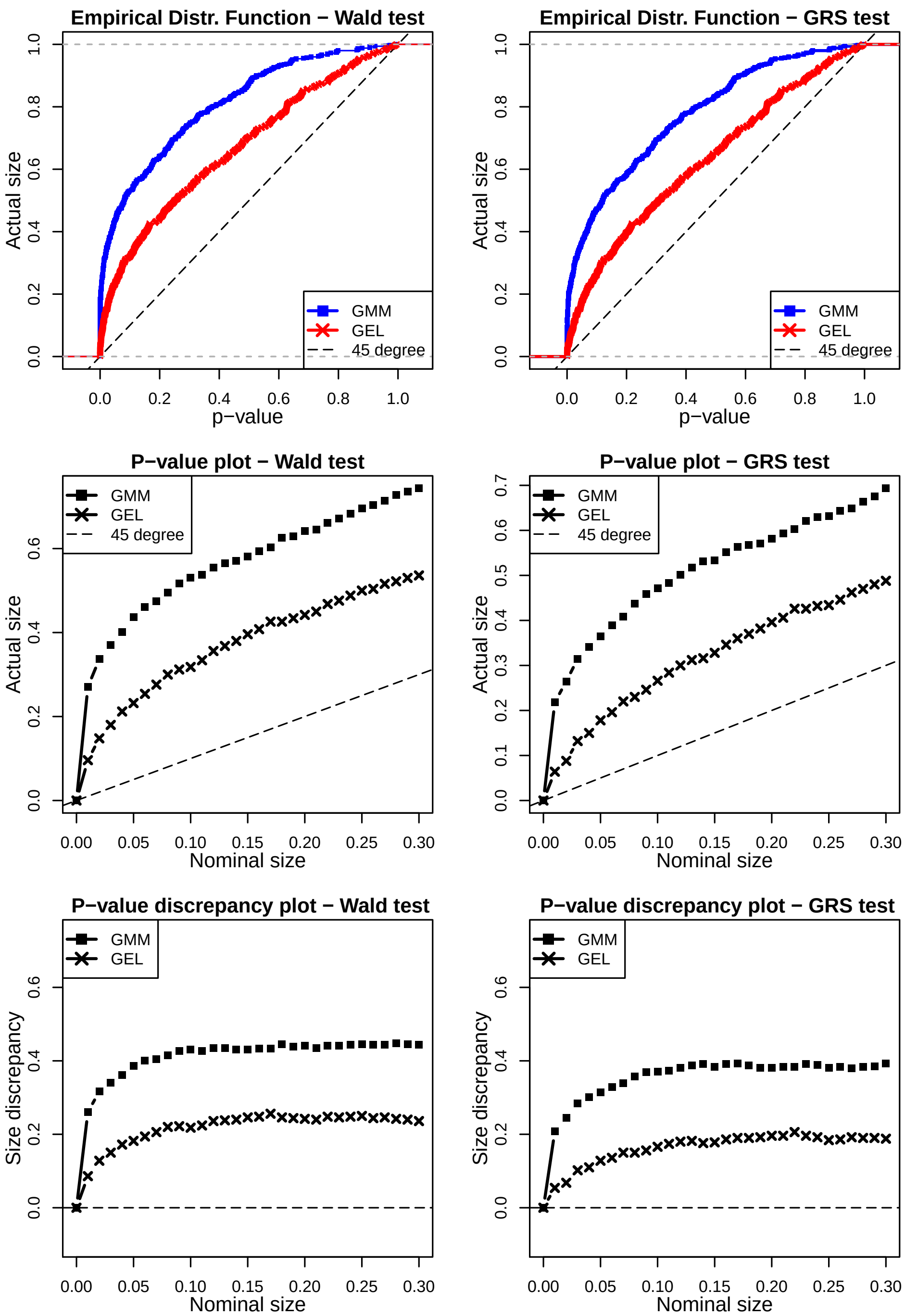
Table 12 in Appendix C presents tabulated rejection proportions for the most common nominal sizes, both tests and all four Monte Carlo scenarios.

Figure 1 – Simulated scenario 1 with Gaussian innovations ($\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\xi}_{i,t}^{\text{Sim1}}$) in Wald and GRS tests (Model=Fama-French three-factor, 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 2 – Simulated scenario 2 with shocks from a t distribution ($\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\nu}_{i,t}^{\text{Sim2}}$) in Wald and GRS tests (Model=Fama-French three-factor, 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 second scenario with shocks from a t distribution are presented in figure 2. 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 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 3 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 evidence is 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 4 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%.

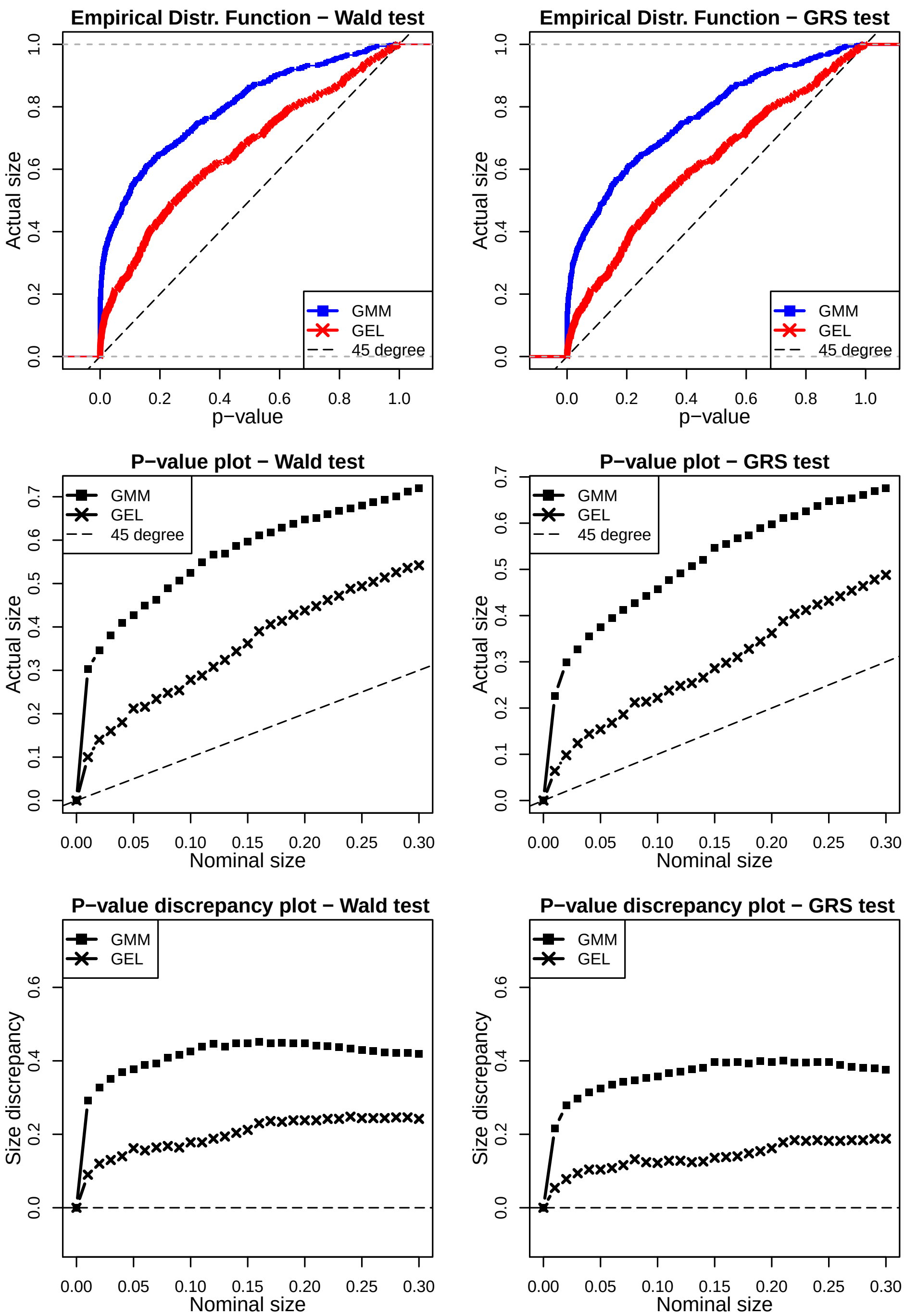
In summary, analyzing all the results presented in this section, it was possible to observe evidence that efficiency tests in finite samples with GEL estimations tend to have better performance compared to estimations via GMM. Furthermore, tests using GEL are more robust to the presence of heavy tails. 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.

4 Empirical Analysis

4.1 Instruments and Factors

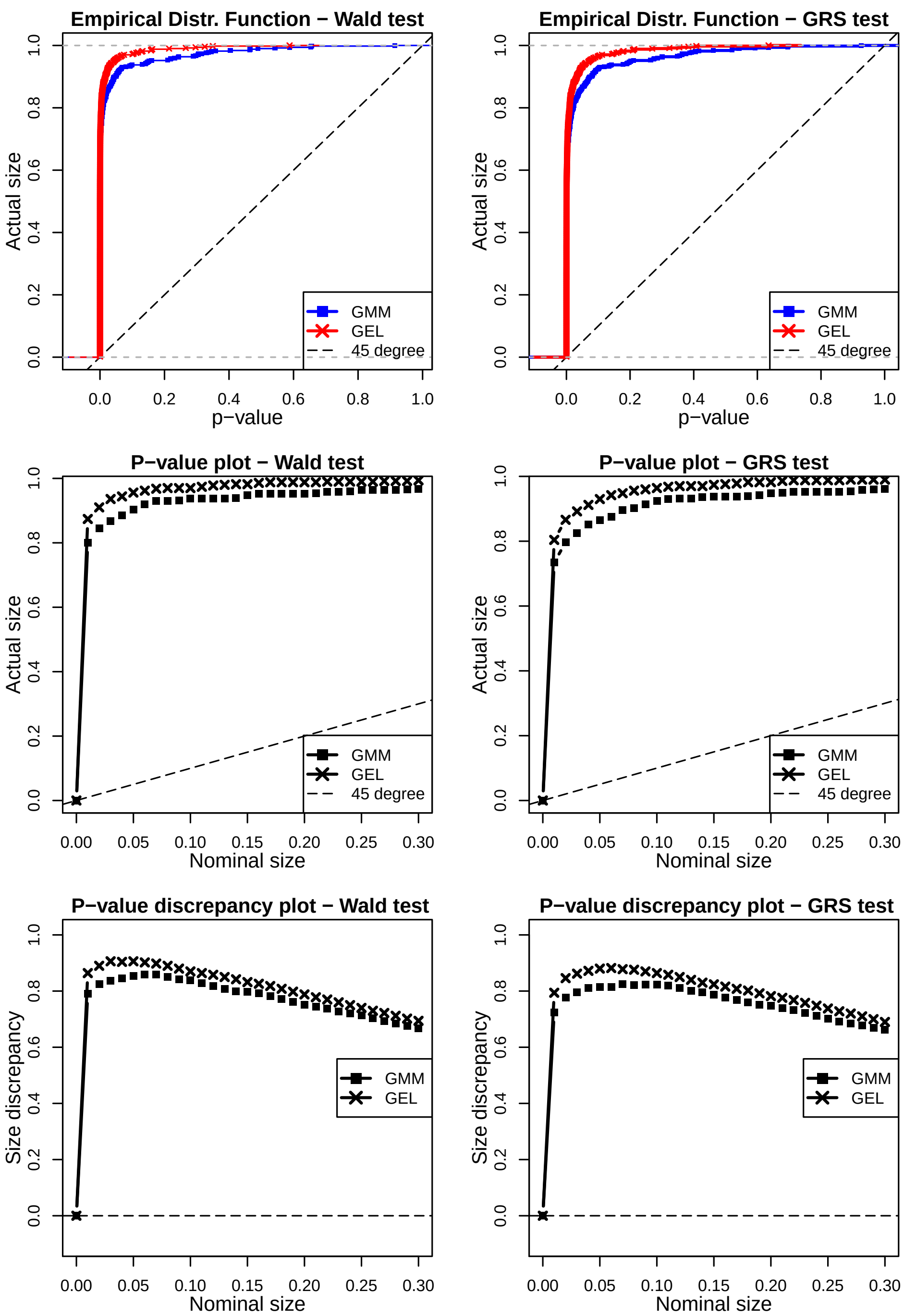
Following previous studies, we selected a limited number of instruments from those commonly used to measure the state of the economy. One can say that the lagged variables chosen are part of a standard set of instruments for this purpose.

Figure 3 – Simulated scenario 3 with shocks at t=T/2 defined by $\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \mathbb{1}_{t=T/2}(\hat{\kappa}_{i,t}^{\text{Sim3}})$ in Wald and GRS tests (Model=Fama-French three-factor, 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 4 – Simulated scenario 4 with shocks defined by $\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\xi}_{i,t}^{\text{Sim4}} - \mathbb{1}_{\hat{p}_{i,t}<0.05}(\hat{\kappa}_{i,t}^{\text{Sim4}})$ in Wald and GRS tests (Model=Fama-French three-factor, 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 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-factor 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 listed on the NYSE, AMEX and NASDAQ [1]. 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. French[2].

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
Lagged Variables					
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
Factors					
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).

Figures 5 and 6 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] $http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html$

Figure 5 – Historical series of the instruments for 720 months (from jan-1955 to dec-2014)

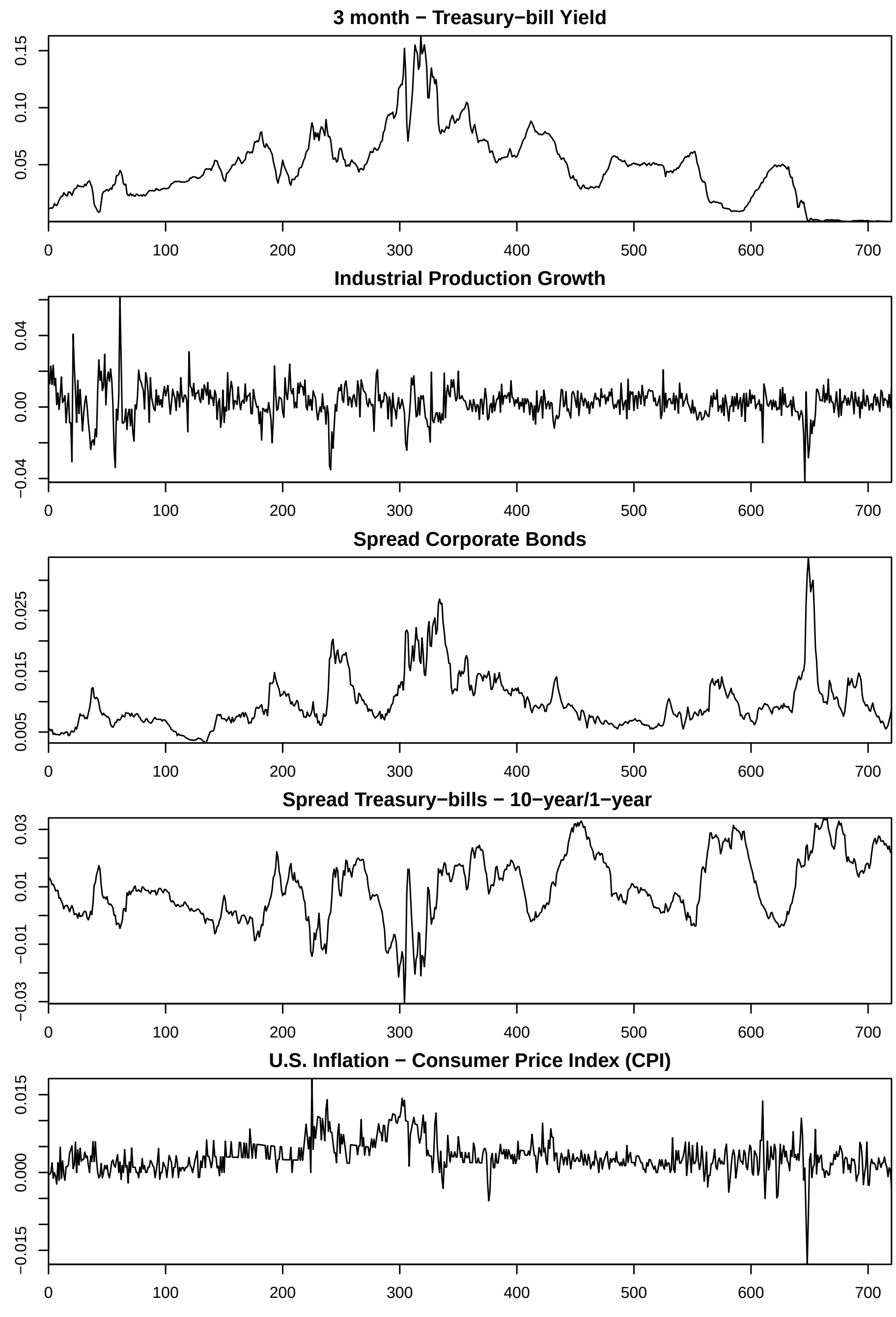
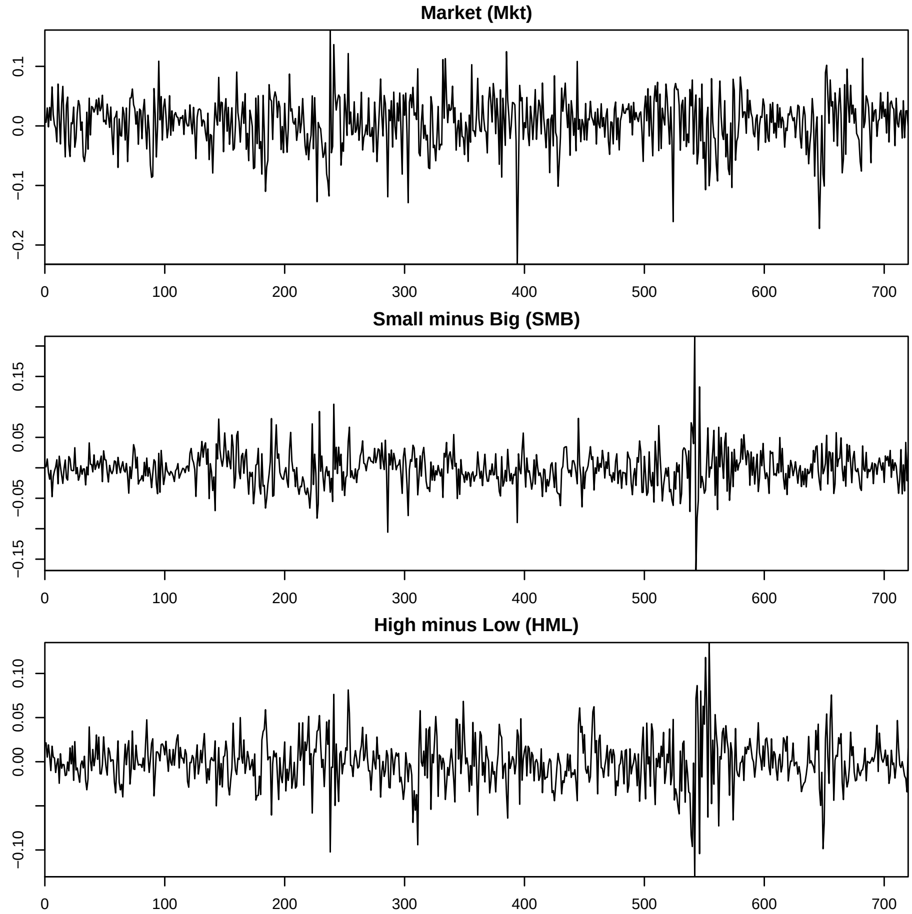


Figure 6 – Historical series of the factors for 720 months (from jan-1955 to dec-2014)



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.

Table 1 remarks some 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 for 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

The portfolio chosen has six assets selected with equal weights by size book-to-market[3]. 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

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

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

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).

The evaluated periods begin with 120 months prior to 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).

4.3 Tests of Efficiency Using GEL and GMM for N=6

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

In order to better fulfill this analysis, the assessment is made by comparing the test results for different sample sizes, as well as for two asset pricing models (CAPM and the Fama-French three-factor).

First we start analyzing the portfolio consisting of the six portfolios formed on size and book-to-market. In Appendix A we also performed the same tests for different types and sizes of portfolios.

For all portfolios, 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.

4.3.1 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.

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-factor 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.

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 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 evidence to reject efficiency. One possible explanation for these results is that the DGP changed over time.

4.3.2 *Managed Portfolios*

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**			
GMM	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
GEL	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

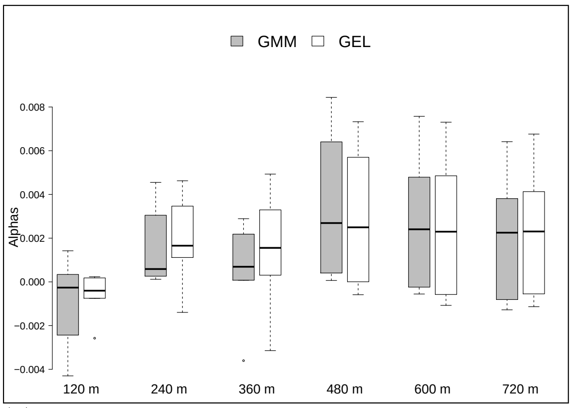
Tests of portfolio efficiency using 6 portfolios formed on size and book-to-market (2x3) for 6 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years) and T=720 (60 years). The tests are evaluated using conditioning information when instruments 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-factor 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 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.

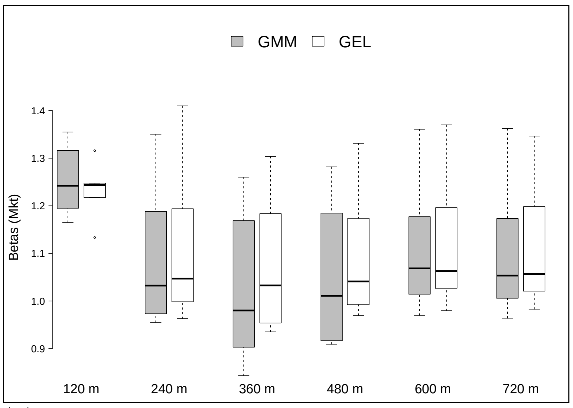
As before we have compelling evidence to reject efficiency for all intervals above 120 months. This result is true for the estimates using either GMM or GEL; and for both models under analysis, CAPM and Fama-French. Note that the longest period in which we can perform the tests is 720 months due to possess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

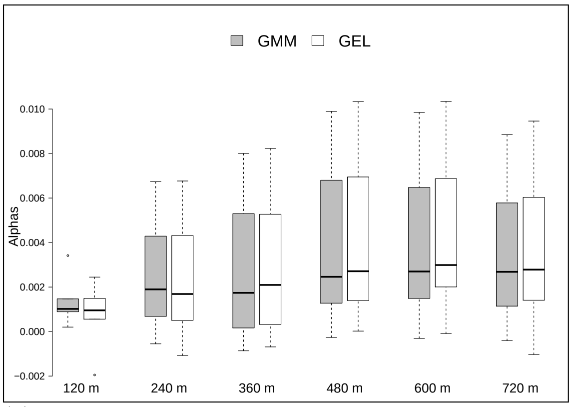
Figure 7 – Boxplots for comparison of estimations by GMM and GEL using 6 portfolios formed on size and book-to-market under *scaled returns* for selected periods of time



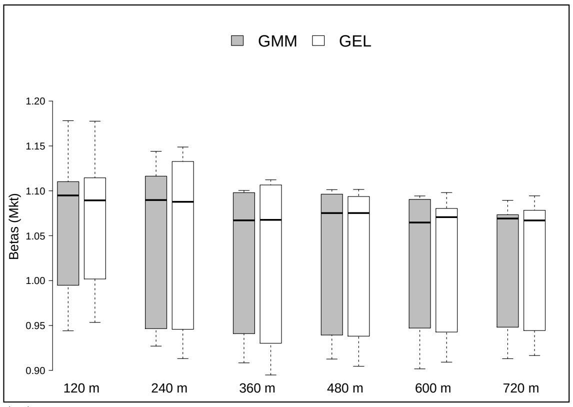
(a) CAPM Model: Alphas estimated by GMM and GEL



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



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

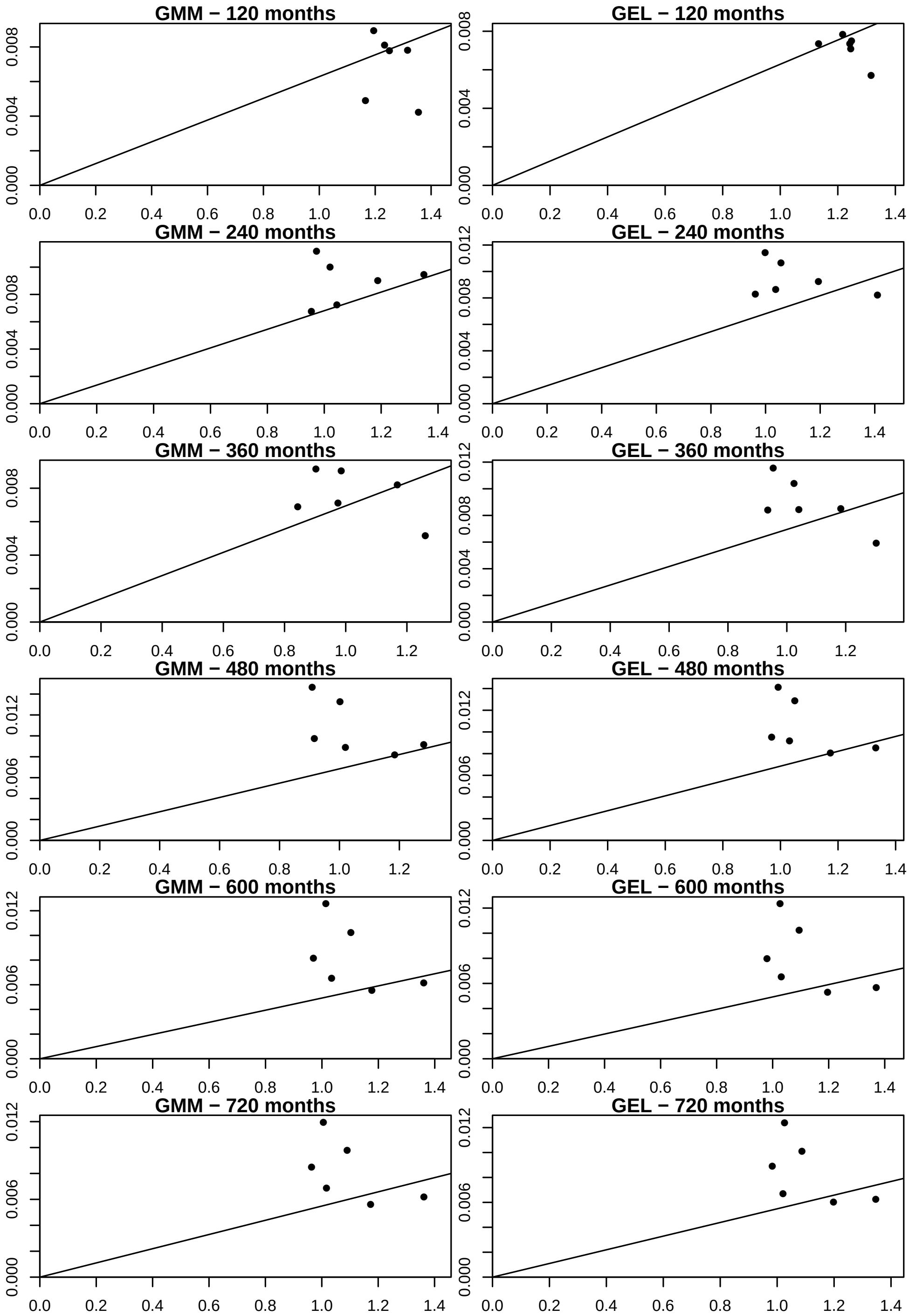


(d) Fama-French Model: $Betas_{Mkt}$ estimated by GMM and GEL

In all 4 panels, the x-axis represents the time intervals, starting with 120 months before December 2014 to 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.

As it was possible to perform the estimation under both methodologies, for both models, and all periods of time; figure 7 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

Figure 8 – 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.

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 B.

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 8 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 Conclusions

The purpose of this study was to evaluate the behavior of the estimators from GEL class in tests of portfolios efficiency. 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.

Comparing the results for (a) different portfolios sizes (see Appendix A for more details), (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.

Finally, 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.

Bibliography

ALMEIDA, C.; GARCIA, R. Empirical likelihood estimators for stochastic discount factors. In: *EFA 2008 Athens Meetings Paper*. [S.l.: s.n.], 2008. 3

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. 3

ANATOLYEV, S. GMM, GEL, serial correlation, and asymptotic bias. *Econometrica*, Wiley Online Library, v. 73, n. 3, p. 983–1002, 2005. 7

ANATOLYEV, S.; GOSPODINOV, N. *Methods for Estimation and Inference in Modern Econometrics*. [S.l.]: CRC Press, 2011. 2, 6, 7, 8

ANDREWS, D. W. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, JSTOR, p. 817–858, 1991. 9

CAMPBELL, J. Y. et al. *The Econometrics of Financial Markets*. [S.l.]: Princeton University Press, 1997. 38

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. 2

COCHRANE, J. H. *Asset Pricing*. [S.l.]: Princeton University Press, 2009. 2, 4, 11

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. 15

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. 21

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. 34

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. 9, 21, 23

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. 2

FERSON, W. E.; QIAN, M. Conditional performance evaluation, revisited. *Research Foundation Monograph of the CFA Institute*, Citeseer, 2004. 21

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. 2, 3, 4, 21, 34, 36

GIBBONS, M. R.; ROSS, S. A.; SHANKEN, J. A test of the efficiency of a given portfolio. *Econometrica*, JSTOR, p. 1121–1152, 1989. 12

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. 8

HANSEN, L. P.; SINGLETON, K. J. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica*, JSTOR, p. 1269–1286, 1982. 12

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. 6

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. 21

KITAMURA, Y.; STUTZER, M. An information-theoretic alternative to generalized method of moments estimation. *Econometrica*, JSTOR, p. 861–874, 1997. 8

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. 9

MARKOWITZ, H. Portfolio selection. *The Journal of Finance*, Wiley Online Library, v. 7, n. 1, p. 77–91, 1952. 2

NEWEY, 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. 2, 7, 8

NEWEY, 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. 9

OWEN, A. B. *Empirical Likelihood*. [S.l.]: Chapman and Hall, 2001. 6, 8

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. 36

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. 2, 9

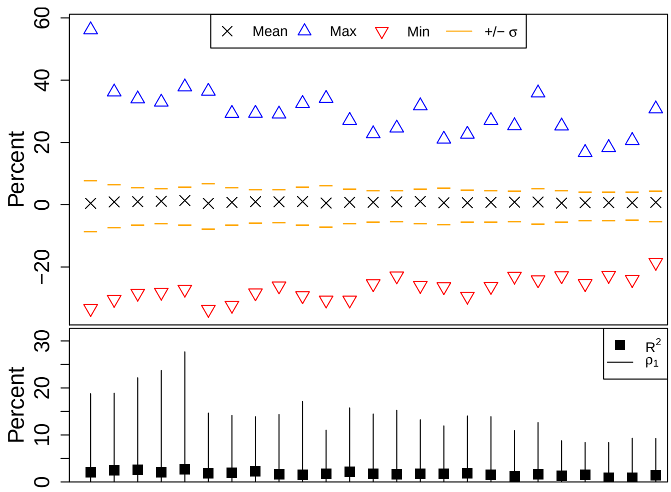
SMITH, R. J. Alternative semi-parametric likelihood approaches to generalised method of moments estimation. *The Economic Journal*, JSTOR, p. 503–519, 1997. 6

APPENDIX A – Results for Different Types and Sizes of Portfolios

1 Data - Portfolios with 25 and 49 assets

In order to examine the estimations behavior for various sample sizes, we selected other three portfolios with increasing number of assets. We sought not to be limited solely on a single portfolio composition methodology. Among the selected portfolios, two of them have their composition based on size and book-to-market, while one is composed with categories derived from industries classification according to the business segment. All observations 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.

Figure 9 – 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



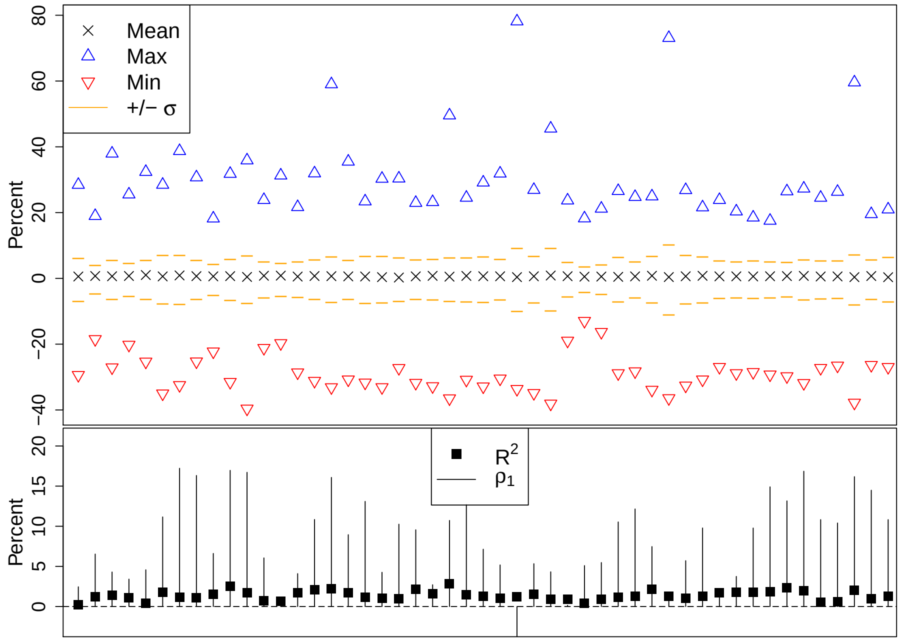
The top panel shows the sample mean statistics (represented by "X"), Max (represented by a blue triangle), Min (represented by an upside down red triangle), and the distance between the two horizontal bars represent the range of $\pm\sigma$ for a 720 months period. The bottom panel shows the first-order 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.

We selected a portfolio composed with 25 assets selected with equal weights by size and book-to-market [4]. Figure 9 presents remarks the main 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%.

Finally, the last chosen portfolio consists of 49 assets representing industrial portfolios selected with equal weights [5]. Figure 10 presents the descriptive statistics of the 720 monthly returns. Notice that both mean and standard deviation are similar to previous portfolio, while the maximum and minimum returns are magnified. Note that

[4] *25 Portfolios Formed on Size and Book-to-Market (5 x 5)*
[5] *49 Industry Portfolios*

Figure 10 – 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.

most of the first order autocorrelation is lower than 20% and there is only one asset with negative value. The R^2 keeps low values, with adjustments coefficients no higher than 5%.

2 Results - Tests of Efficiency Using GEL and GMM for Portfolios with 25 and 49 assets

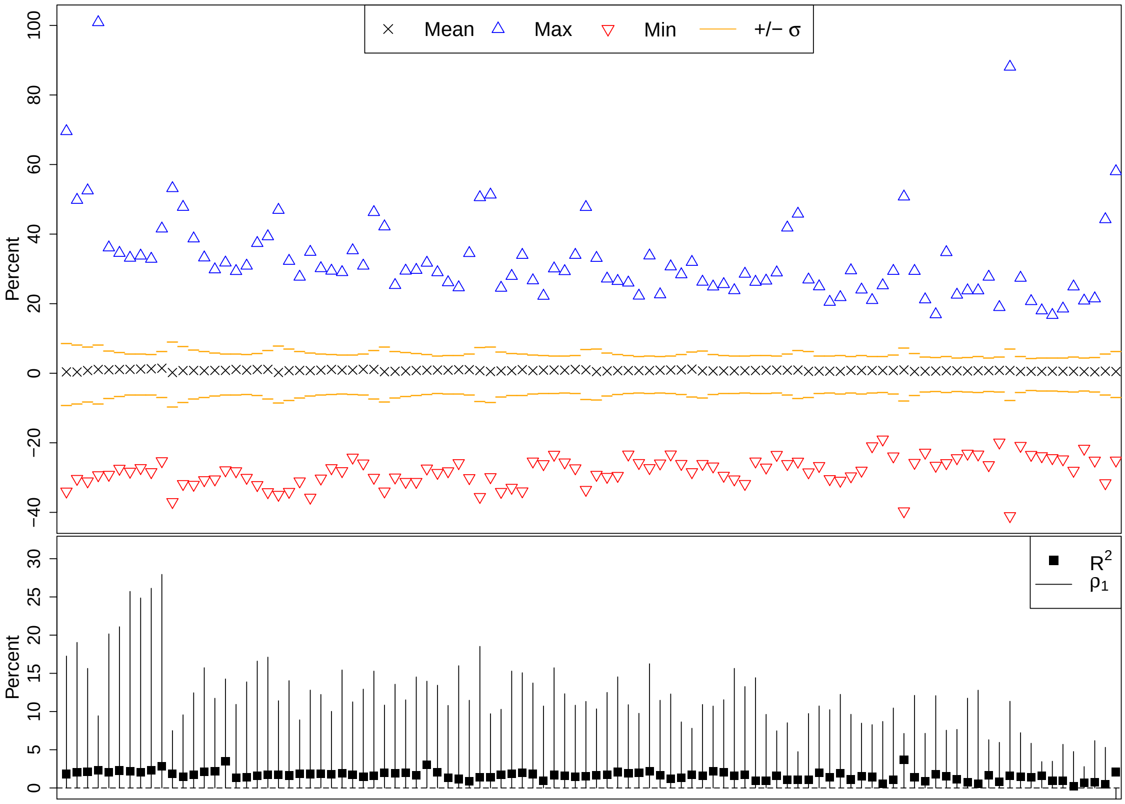
2.1 25 Portfolios Formed on Size and Book-to-Market

2.1.1 No Conditional Information

We present here the tests of efficiency using GEL and GMM estimations and how they can lead to different decisions with a portfolio with higher amount of assets. Firstly, we use the 25 portfolios formed on size and book-to-market. Table 5 shows the results of estimations 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. 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

Figure 11 – 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.

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)[6].

From the formulas in (32) and (33), 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

6 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**			
GMM	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
GEL	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

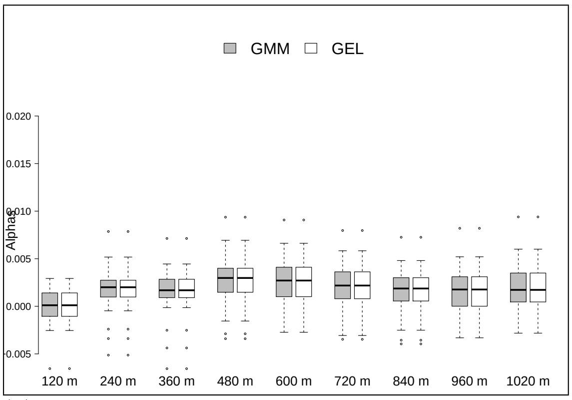
Tests of portfolio efficiency using 25 portfolios formed on size and book-to-market (5x5) for 9 selected periods of time: T=120 (10 years), T=240 (20 years), T=360 (30 years), T=480 (40 years), T=600 (50 years), T=720 (60 years), T=840 (70 years), T=960 (80 years) and T=1020 (85 years). The tests are 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-factor 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.

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.

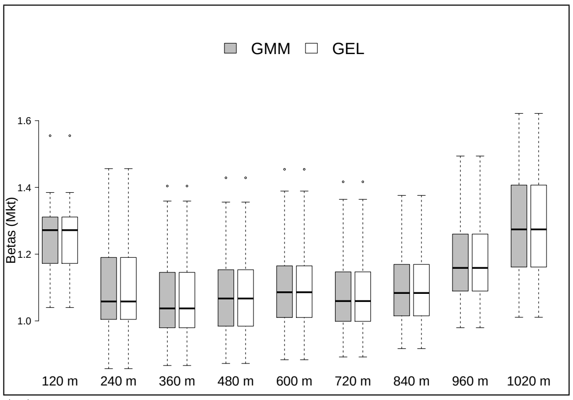
In Figure 12 can be seen the boxplots of the estimated coefficients of α and β_{Mkt} for CAPM and Fama-French models. Each boxplot summarizes the 25 estimated values using GMM or GEL for each time interval. Note that for every T estimation was possible to be performed. Even if the test for 120 months 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 B.

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 they decrease, and finally, increase again.

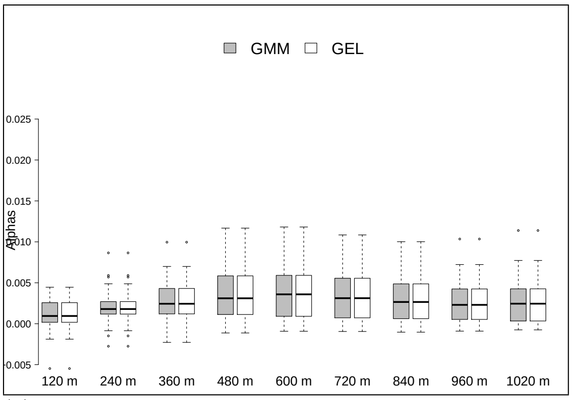
Figure 12 – Boxplots for comparison of estimations by GMM and GEL using 25 portfolios formed on size and book-to-market for selected periods of time



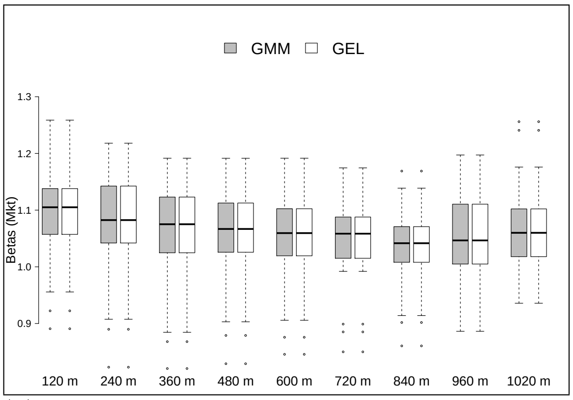
(a) CAPM Model: Alphas estimated by GMM and GEL



(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.

In Appendix B 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.

2.1.2 *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 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

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-factor 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.

covariance matrix $\hat{\mathbf{\Omega}}(\boldsymbol{\theta}_0)$ 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 evidence to reject the null hypothesis.

2.2 49 Industry Portfolios

2.2.1 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 evidence 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-factor 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.

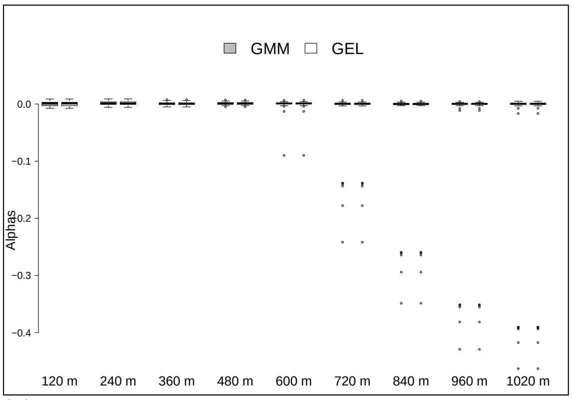
When we analyze the results with the use of GEL class estimators, the first obvious observation is that the estimates for all periods with Fama-French model were not able to be computed. This fact occurred because with GEL methodology was not feasible to estimate the coefficients when $n = 49$, for any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

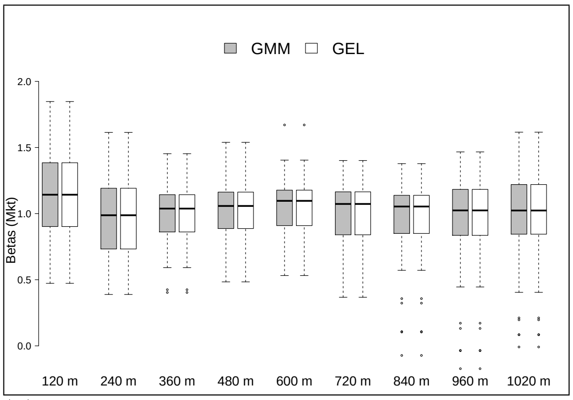
Examining specifically 360 and 480 months periods, and comparing to the GMM estimation, we see that the results are in line: both provide evidence to reject the null hypothesis. However, one can note that the p-values from GEL are higher for both tests. For instance, the GRS test using GEL obtained p-values 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.

Figure 13 – Boxplots for comparison of estimations by GMM and GEL for CAPM model using 49 industry portfolios for selected periods of time



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



(b) CAPM Model: $Betas_{Mkt}$ 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.

As only the estimation of CAPM model was possible using both GMM and GEL, in figure 13 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.

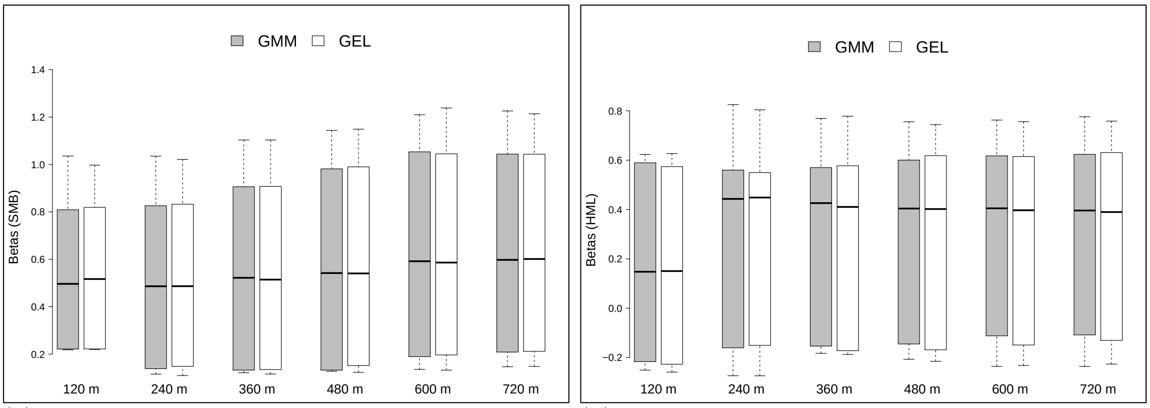
A point that worth a note is the existence of outliers in alpha estimates, occurring mainly for longer periods, from 600 months or higher. In addition, notice also that the estimates for both, $\hat{\alpha}$, as well as for $\hat{\beta}_{mkt}$ in CAPM model change marginally as time intervals increases.

In appendix B 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.

APPENDIX B – 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

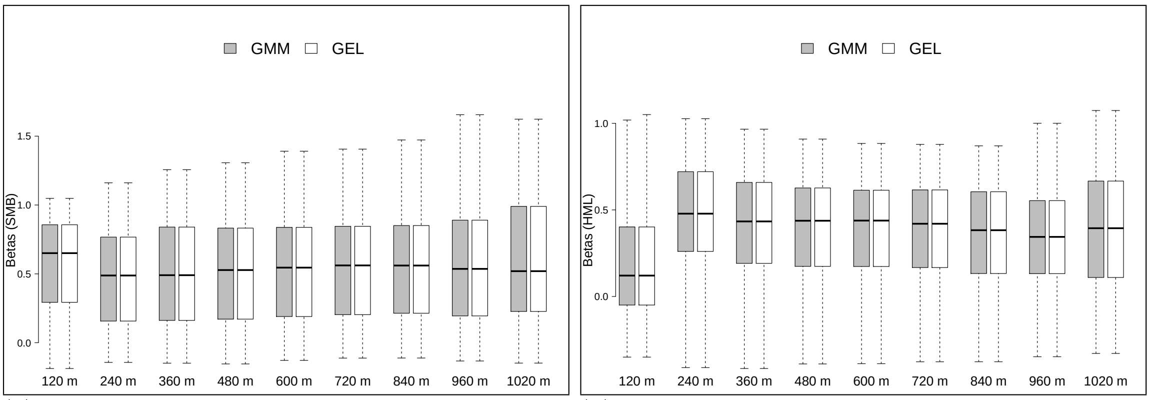


(a) Fama-French Model: $Betas_{SMB}$ estimated by GMM and GEL

(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

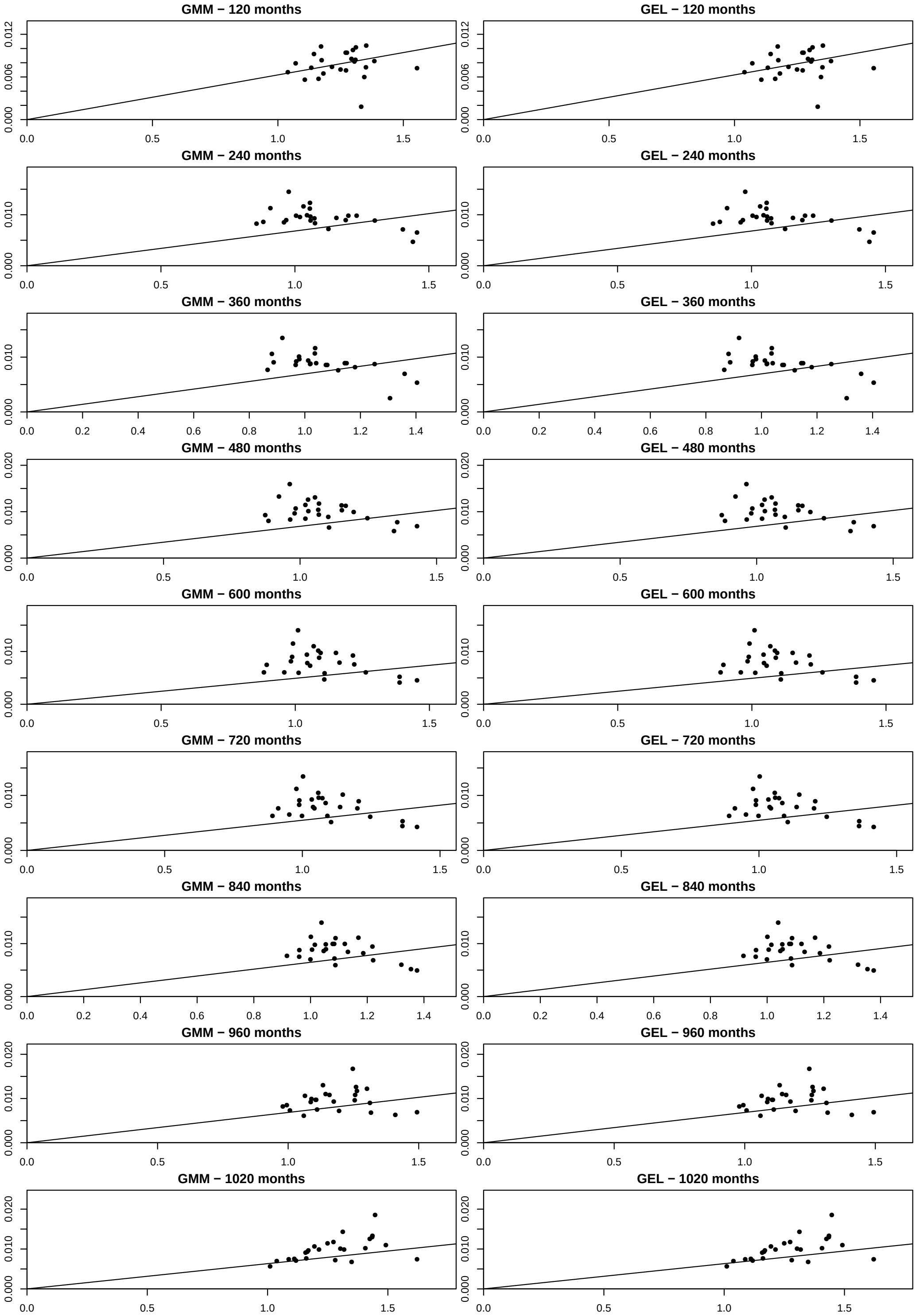


(a) Fama-French Model: $Betas_{SMB}$ estimated by GMM and GEL

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

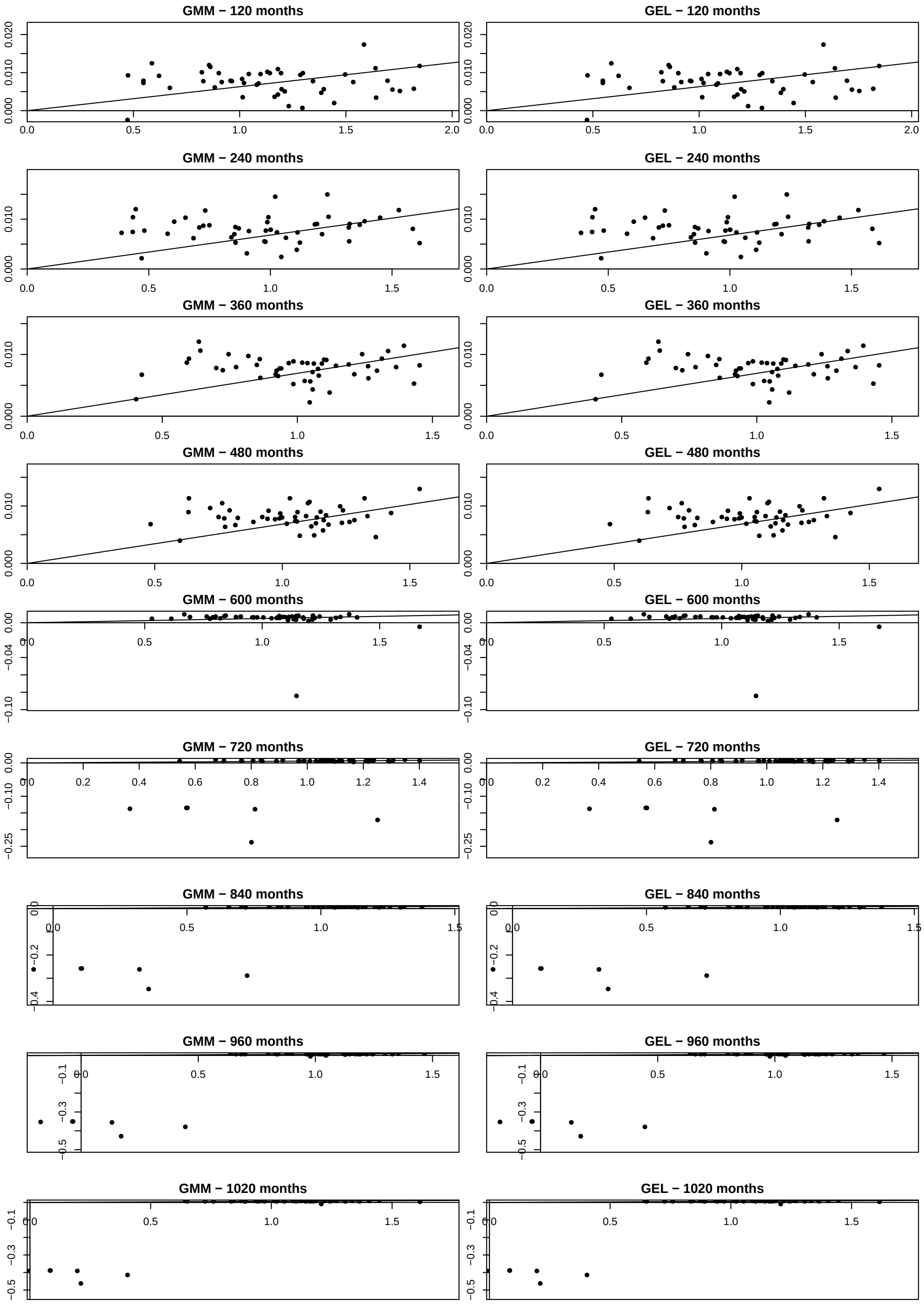
In both panels, the x-axis represents the time intervals, starting with 120 months before December 2014 to 1020 months prior to this date. The y-axis are the estimated coefficients values by GMM and GEL. Estimation by GMM is represented by gray boxplots, while GEL estimation is represented by white boxplots.

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 C – Additional Tables

Table 8 – 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-factor 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 9 – 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-factor 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 10 – 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-factor 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 11 – 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**			
GMM	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-factor 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 12 – 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		Wald Test		GRS Test	
	GMM	GEL	GMM	GEL	GMM	GEL	GMM	GEL	GMM	GEL	GMM	GEL	GMM	GEL	GMM	GEL
$\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\xi}_{i,t}^{\text{Sim1}}$	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
$\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\nu}_{i,t}^{\text{Sim2}}$	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
$\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \mathbb{1}_{t=T/2}(\hat{\kappa}_{i,t}^{\text{Sim3}})$	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
$\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\xi}_{i,t}^{\text{Sim4}} - \mathbb{1}_{\hat{p}_{i,t}<0.05}(\hat{\kappa}_{i,t}^{\text{Sim4}})$	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-factor asset pricing model. Table presents results for 4 different scenarios with 500 simulations each. First case ($\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\xi}_{i,t}^{\text{Sim1}}$) assess Wald and GRS tests under Gaussian shocks. For the second scenario ($\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\nu}_{i,t}^{\text{Sim2}}$) innovations are drawn from a t-Student distribution. In the third case ($\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \mathbb{1}_{t=T/2}(\hat{\kappa}_{i,t}^{\text{Sim3}})$) the DGP insert a large magnitude shock on a fixed date. Finally, the fourth scenario ($\hat{\varepsilon}_{i,t}^{\text{Sim}*} = \hat{\xi}_{i,t}^{\text{Sim4}} - \mathbb{1}_{\hat{p}_{i,t}<0.05}(\hat{\kappa}_{i,t}^{\text{Sim4}})$) the DGP assumes that there is a probability of 5% that an outlier exists.