

Identifying outliers in asset pricing data with a new weighted forward search estimator*

Alexandre Aronne¹

 <https://orcid.org/0000-0002-2049-0751>

Email: alexandre.aronne@ibmec.edu.br

Luigi Grossi²

 <https://orcid.org/0000-0003-3671-7348>

Email: luigi.grossi@univr.it

Aureliano Angel Bressan³

 <https://orcid.org/0000-0002-9333-3394>

Email: aureliano@ufmg.br

¹ Ibmec BH, Departamento de Ciências Contábeis, Belo Horizonte, MG, Brazil

² Università degli Studi di Verona, Dipartimento di Scienze Economiche, Verona, Italy

³ Universidade Federal de Minas Gerais, Faculdade de Ciências Econômicas, Centro de Pós-Graduação e Pesquisas em Administração, Belo Horizonte, MG, Brazil

Received on 02.18.2019 – Desk acceptance on 03.18.2019 – 3rd version approved on 09.06.2019 – Ahead of print on 01.31.2020

Associate Editor: Fernanda Finotti Cordeiro Perobelli

ABSTRACT

The purpose of this work is to present the Weighted Forward Search (FSW) method for the detection of outliers in asset pricing data. This new estimator, which is based on an algorithm that downweights the most anomalous observations of the dataset, is tested using both simulated and empirical asset pricing data. The impact of outliers on the estimation of asset pricing models is assessed under different scenarios, and the results are evaluated with associated statistical tests based on this new approach. Our proposal generates an alternative procedure for robust estimation of portfolio betas, allowing for the comparison between concurrent asset pricing models. The algorithm, which is both efficient and robust to outliers, is used to provide robust estimates of the models' parameters in a comparison with traditional econometric estimation methods usually used in the literature. In particular, the precision of the alphas is highly increased when the Forward Search (FS) method is used. We use Monte Carlo simulations, and also the well-known dataset of equity factor returns provided by Prof. Kenneth French, consisting of the 25 Fama-French portfolios on the United States of America equity market using single and three-factor models, on monthly and annual basis. Our results indicate that the marginal rejection of the Fama-French three-factor model is influenced by the presence of outliers in the portfolios, when using monthly returns. In annual data, the use of robust methods increases the rejection level of null alphas in the Capital Asset Pricing Model (CAPM) and the Fama-French three-factor model, with more efficient estimates in the absence of outliers and consistent alphas when outliers are present.

Keywords: asset pricing, outlier detection, robust estimation, forward search.

Correspondence address

Alexandre Aronne

Ibmec BH, Departamento de Ciências Contábeis
Rua Rio Grande do Norte, 300 – CEP 30130-130
Funcionários – Belo Horizonte – MG – Brazil

*Paper presented at the XLI ANPAD Conference, São Paulo, SP, Brazil, October of 2017.



1. INTRODUCTION

The Capital Asset Pricing Model (CAPM) introduced by Sharpe (1964) and Lintner (1965) represents a path-breaking milestone in the history of financial theory. The publication of these seminal papers led to the development of a large body of research in various areas of finance, both from normative and positive points of view. From a positive standpoint, the model has been used, for example, to explain the cross-section of expected returns (Fama & MacBeth, 1973) and the performance of mutual funds (Jensen, 1967). From a normative view, the model has been used in the context of capital budgeting decisions and portfolio management (Sharpe, 1963).

Nonetheless, the linear risk-return relationship posited by the model did not suffice to explain the cross-section of expected return and a number of departures from the model (usually called anomalies) were revealed (see Fama and French [2008]). There is, nowadays, “mounting evidence against it based on the cross section of stock returns” (Da, Guo & Jagannathan, 2012).

As evidences against the model started to appear, more sophisticated models, such as the Arbitrage Pricing Model (APT) (Ross, 1976), were developed. A model that is able to capture many of the anomalies not explained by the CAPM is the Fama and French (1992, 1993) three-factor model.

More interesting, however, is the fact that the evidences against the CAPM did not prevent it from becoming the most used model for the determination of the cost of capital in the context of capital budgeting decisions (Graham & Harvey, 2001). Moreover, there is empirical support in favor of the CAPM for this purpose (Da et al., 2012).

The choice of the ordinary least squares (OLS) method for the estimation of the CAPM parameters is natural, given that it is the best linear unbiased estimator (BLUE) under the normality assumptions posed by the CAPM theory. However, an often neglected issue related to stocks and market returns when opting for the OLS estimator is the overwhelming evidence that these returns are not normally distributed (see Mandelbrot [1963] and Merton [1976]) and exhibit fat tailed empirical distributions, i.e., the distributions of stock returns contain outliers, observations that do not belong with the majority of the (normally distributed) data. In the context of regression methods, outliers are, according to Rousseeuw and Leroy (1987), data points (observations) that deviate from the linear relation followed by the majority of the data, taking into consideration both the explanatory variables (X)

and the response variable (Y) simultaneously. Therefore, extreme values both in the Y and X variables are not considered outliers, as long as they conform to the linear relation of the bulk of the data.

It is widely known that the OLS method is extremely sensitive to the presence of outliers in the data (both in the X and in Y variables). One can prove that its breakdown point – “the smallest fraction of bad observations that may cause an estimator to take on arbitrarily large aberrant values” (Huber & Ronchetti, 2009, p. 8) – is equal to 0%, indicating that a single “bad” observation can cause massive distortions in the parameter estimate (Rousseeuw & Leroy, 1987).

Knez and Ready (1997) argue that outliers are not necessarily to be viewed as observations to be discarded or deleted, nor that they are irrelevant observations. On the contrary, they are viewed as “precious”, since they may provide a lot of information about the data generating process and a proper model specification.

Therefore, the existence of outliers in the regression variables motivates the following research question: can the application of new robust statistical methods for the analysis and estimation of the parameters of asset pricing models allow for the detection and treatment of the outlying data in financial returns, providing more reliable estimates of alphas and betas (parameters of the asset pricing model specified in equation 11) of equity portfolios?

This work consists in the application of the forward search (FS), a robust method, in the context of asset pricing models. More specifically, its aim is to assess the impact of outliers on parameter estimation in these models and to test the performance of a new weighted FS (FSW) estimator in the estimation of asset pricing models. In order to accomplish these goals, we conduct a series of Monte Carlo simulation experiments and compare the performance of the FSW with OLS and least trimmed squares (LTS).

Moreover, we apply the FSW estimator on time series regressions of the 25 Fama and French (1993) portfolios. It is, to our knowledge, the first application of estimators with high efficiency and high-breakdown point in this context, as previous research, such as Knez and Ready (1997) and Bailer (2005), has focused on the impact of outliers in cross-section regressions using methods which are either efficient or robust.

The paper is organized as follows. On Section 2, a literature review is conducted. On Section 3, the new FSW estimator is presented and defined. On Section 4, the

new method is applied to simulated and real market data and the estimation results are presented. Lastly, Section

5 presents our final remarks, limitations of the research, contributions, and possible extensions.

2. LITERATURE REVIEW

In this section, a literature review on asset pricing – CAPM and multifactor models – and on the use of robust estimators in this context is presented.

2.1 CAPM

The theoretical foundations of the CAPM were laid down by the seminal works of Sharpe (1964) and Lintner (1965). While both authors developed the mathematical groundwork of the model based on assumptions about markets, asset returns, and utility functions of investors, they did not develop empirical studies or applications of the model.

Black, Jensen, and Scholes (1972) developed one of the two benchmark frameworks for testing asset pricing models, namely the time-series regression approach. The authors proposed a simple test of the model: estimate alphas of a large number of securities and assess whether the estimates are statistically equal to 0, as predicted by the theory.

Even though Black et al. (1972) had developed cross-sectional tests of the CAPM, the most widely used cross-sectional regression approach until today is the one developed by Fama and MacBeth (1973). The so-called Fama and MacBeth (1973) procedure proposed by the authors consists of three steps:

- i. Time-series OLS regression of equation 1 obtaining estimates $\hat{\beta}_i$ for each portfolio i .

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t} \quad \boxed{1}$$

where T is the number of observations for each portfolio, N is the number of portfolios, $E_T(f)$ and $\hat{\sigma}(f)$ are, respectively, the sample mean and standard deviation of the factor, $\hat{\alpha}$ is a vector of the estimated intercepts, and $\hat{\Sigma}$ is the variance-covariance matrix of the residuals resulting from the N regressions. The test can be easily extended to consider more than one factor, i.e., to test multifactor asset pricing

where $r_{i,t} = R_{i,t} - R_{f,t}$ and $r_{M,t} = R_{M,t} - R_{f,t}$ are, respectively, the excess returns of asset and the excess return of the market (a traded aggregate wealth index) over the risk free rate $r_{f,t}$ at month t , α_i and β_i are the alpha and beta of stock i , and $\varepsilon_{i,t}$ is the zero-mean, constant-variance error term.

- ii. For each month t , run a cross-sectional regression of equation 2 using the beta estimates obtained in step (i) as independent variables.

$$r_{i,t} = \gamma_{0,t} + \gamma_{1,t} \hat{\beta}_{i,t-1} + \eta_{i,t} \quad \boxed{2}$$

- iii. As a result of step (ii), one obtains a time-series of regression coefficients estimates $\hat{\gamma}_{0,t}$ and $\hat{\gamma}_{1,t}$ and computes time-series averages and t tests, assuming $\eta_{i,t}$ – is the zero-mean, constant-variance error term – is independent of the regressors.

While Black et al. (1972) and Fama and MacBeth (1973) were all interested in testing the CAPM using a large number of portfolios, their conclusions were based on univariate portfolio-specific t statistics. In light of the limitations of these statistical tests, Gibbons, Ross, and Shanken (1989) proposed a multivariate statistic to test whether all the intercepts are jointly equal to 0. The authors have showed that the so-called Gibbons Ross Shanken (GRS) statistic has an F distribution with degrees of freedom N and $T - N - 1$:

$$GRS = \frac{T - N - 1}{N} \left[1 + \left(\frac{E_T(f)}{\hat{\sigma}(f)} \right)^2 \right]^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} \sim F_{N, T-N-1} \quad \boxed{3}$$

models (Cochrane, 2001, p. 217). The GRS statistic is still today the standard test of asset pricing models.

In spite of the empirical evidence provided by Fama and MacBeth (1973) of the linear risk-return relationship posited by the theory, the literature of asset pricing shortly turned towards the development of multifactor models, which are discussed in the next section.

2.2 Multifactor Models

Ross (1976) developed the arbitrage pricing theory (APT), which explicitly takes into account the possibility that stock returns may be generated by a multifactor model of the form:

$$r_{i,t} = r_{f,t} + \beta_{i,1}f_{1,t} + \dots + \beta_{i,k}f_{k,t} + \varepsilon_{i,t} \quad \boxed{4}$$

While the APT represents a generalization of the CAPM, it falls short in determining or providing evidence of which factors are (or should be) considered by investors.

Roll (1988) compared the explanatory power of the CAPM and a five-factor APT model in explaining individual stock returns. His results indicate that the multifactor model provided higher average R^2 (adjusted for degrees of freedom) than the single-factor model, but for his disappointment, overall average R^2 was only about 0.20 for daily returns and 0.35 for monthly returns.

Fama and French (1992, 1993) extended Roll's (1988) analysis to portfolios returns. Fama and French (1992) used size, market beta, leverage, earning/price, and book-to-market equity (BE/ME) as explanatory variables of the cross section of average stock returns. The results obtained by Fama and French (1992) based on the two-step Fama and MacBeth (1973) procedure indicate that used combined, two variables – size and BE/ME – seem to explain the cross-section of average returns, absorbing the explanatory power of other variables, such as market beta, leverage, and earning/price. Additionally, their results do not support the central idea of the CAPM, that average returns are positively related to market beta. In fact, they argue that this relationship is not present in the 1963-1990 period, and is very weak in the longer 1941-1990 period.

In a subsequent paper, Fama and French (1993) concentrated on the identification of common risk factors in the returns of stocks and bonds. Differently from Fama and French (1992), their analysis is based on the time-series regression approach of Black et al. (1972) instead of the Fama and MacBeth (1973) procedure. Regarding the stock returns analysis, the authors test the explanatory power of three risk factors (market, size, and BE/ME) on the returns of 25 size-BE/ME sorted portfolios. Their main results indicate that the size and BE/ME factors can explain the differences in average returns across stocks, but the difference between the average returns on stocks and one-month bills is explained by the market factor.

Despite the strong evidences provided by Fama and French (1992, 1993) against the CAPM and in favor of the three-factor model, their results were received with relative skepticism by other researchers. Kothari, Shanken, and Sloan (1995) argued that the BE/ME premium in Fama

and French (1992) was overstated, due to a survivorship bias in the Compustat data used, which was likely to include distressed firms that survived and to exclude those that went bankrupt. Moreover, they provide empirical evidence in favor of the CAPM, as their results indicate that there is a statistically significant market risk premium when betas are computed on annual, instead of monthly, returns.

Fama and French (1996) tested the hypothesis that the BE/ME premium was spurious by applying the three-factor model to various data sets. Their results reconfirmed the existence of the BE/ME premium and its statistical significance. Additionally, the authors provide evidence that the three-factor model explains many of the patterns in stock returns – the so-called anomalies –, which are not captured by the CAPM. Nonetheless, the three-factor model is not able to capture the short-term return continuation anomaly, currently known as the momentum anomaly, which was later analyzed by Carhart (1997).

2.3 Robust Estimation of Asset Pricing Models

The application of robust regression methods in the context of asset pricing models dates back to the work of Sharpe (1971), in which the mean absolute deviation (MAD) method was applied to estimate parameters of the CAPM. His results show that the two methods give similar beta estimates, but quite different alfa estimates when applied to stocks or non-diversified portfolios. Nonetheless, the author concludes that the gains of the MAD over OLS are modest.

Cornell and Dietrich (1978) have also developed a comparative analysis of MAD *versus* OLS, in order to test the stability of the beta coefficients across time. In line with Sharpe's (1971) conclusions, the authors find it "disappointing ... that the MAD technique fails to improve on OLS".

Chan and Lakonishok (1992) compared the performance of various robust estimation methods on both simulated and market data. The authors present the performance of each method under the null case where there are no outliers in the data and also when the stock returns are heavy-tailed. Their results reconfirm the poor performance of MAD. On the other hand, the authors reported that the use of trimmed regression quantile estimators results in loss of efficiency of only about 10% under the null case, and in efficiency gains of up to 80% under the alternative where the dependent variable is heavy-tailed, providing strong evidence in favor of the application of robust methods for beta estimation. The authors focus their analysis on beta estimation, leaving

aside the performance of the methods with respect to alpha estimation.

Bowie and Bradfield (1998) extended the work of Chan and Lakonishok (1992) by assessing the relative performance of a wider range of robust estimators when applied for beta estimation of securities listed on the Johannesburg Stock Exchange. Their results, based on jackknife measures of efficiency, indicate that robust methods are less sensitive than OLS to model misspecification – such as extreme excess market returns –, and that the superior efficiency of the robust estimators was caused by non-normality in the distribution of residuals.

To our knowledge, Knez and Ready (1997) were the first to study the impacts of applying such techniques on cross-sectional regressions (second step of the Fama and MacBeth [1973] procedure). The authors apply the LTS on the cross-sectional regression data used in Fama and French (1992) and analyze the risk premia on size and book-to-market factors. The authors show that the negative relationship between average returns and size obtained by Fama and French is caused by only a few influential firms. In fact, their results indicate that trimming 1% of the most extreme observations each month leads to a positive relationship between average returns and size. The authors restrict their analysis to size and book-to-market factors and do not use market betas as an explanatory variable for the cross section of average returns. Additionally, the authors do not apply the LTS for the estimation of “pre-ranking” and “post-ranking” betas.

Bailer (2005) further extended Knez and Ready’s (1997) analysis in at least four directions. First, he uses the MM-estimator instead of LTS. Second, he introduces market betas (as well as size and book-to-market) as an explanatory variable of average returns. Third, he applies the robust methods both in the first and second steps of the Fama and MacBeth (1973) procedure, as well as in the time-series averages of cross-sectional estimates. Fourth, he applies the methodology to more recent time periods. The author concludes that OLS alphas tend to be over-biased and classical betas are highly sensitive to outliers, while robust alphas and betas are superior predictors. The author also finds that the beta and size risk premiums found to be respectively flat and negative in Fama and French (1992) turn out to be flat or negative for beta and positive for size when only 1 to 3% of the data are rejected, reconfirming Knez and Ready’s (1997) results.

Despite the vast literature on the robust estimation of asset pricing models, we are not aware of the application of methods with high-breakdown and high efficiency properties – such as the FSW and FSI – in this context, as all the methods previously mentioned either present high efficiency (MM-estimator) or high-breakdown (LTS).

In the next section, the FS – a high efficiency and high-breakdown robust estimator – is presented.

2.4 The FS

The FS described by Atkinson and Riani (2000) is a robust method that provides useful plots, which allows one to understand the real structure of the data being analyzed and assess the agreement between the data and the model. Differently from backward methods, the FS is immune to the well-known masking and swamping effects (Atkinson & Riani, 2000).

The basic concepts of the FS algorithm date back to the work of Hadi (1992), where the idea of fitting a model to subsets of increasing sizes was introduced. Hadi and Simonoff (1993) used it in a regression framework, while Atkinson (1994) and Hadi (1994) applied it to multivariate data. Atkinson and Riani (2000) and Atkinson, Riani, and Cerioli (2004) published books that discuss deeply how the FS can be applied in the regression and multivariate analysis contexts, respectively.

The FS is composed of three steps:

- i. choice of the initial subset;
- ii. addition of observations through the search;
- iii. monitoring of key quantities during the search.

The first step is designed to identify a subset of the data, which is free of outliers – a clean data set (CDS). This is accomplished by the use of a high-breakdown robust estimator, such as the least median of squares or the LTS. The LTS estimator is given by (Rousseeuw and Leroy, 1987):

$$\text{Min}_{\hat{\beta}} \sum_{i=1}^h (e^2)_{i:n} \quad \boxed{5}$$

where $(e^2)_{1:n} \leq \dots \leq (e^2)_{n:n}$ are the ordered squared residuals. The estimated parameter vector $(\hat{\beta})$ is, therefore, the vector that minimizes the sum of the h (out of n) smallest squared residuals.

In the initial step of the FS, the model is fitted to $m_0 = p$ observations and $h = \frac{(n+p+1)}{2}$, resulting in the highest breakdown point (50%) that can be achieved by the LTS method, in which p represents the number of parameters to be estimated.

The minimization in equation 5 is performed only approximately by searching over a large number (usually 10,000 or higher) of subsets of size p chosen at random. This procedure provides a subset that is free of outliers. Thus, the initial subset is the subset of size p that yields the

minimum value of the sum in equation 5. The parameter estimate that minimizes equation 5 is $\hat{\beta}_{m_0}$.

The second step is designed to successively add observations to the initial subset by their closeness to the bulk of the data, as the search evolves. Given a subset $S^{(m)}$ of dimension $m \geq p$, the search moves forward to finding subset $S^{(m+1)}$ by selecting the $m + 1$ observations with the smallest squared scaled residuals $e_{i,S^{(m)}}^2$, $i = 1, \dots, n$. Residuals $e_{i,S^{(m)}}$ are computed as

$$e_{i,S^{(m)}} = y_i - x_i' \hat{\beta}_m \quad \boxed{6}$$

where $\hat{\beta}_m$ is obtained by applying OLS on the observations that form subset $S^{(m)}$, for $m > m_0$.

In most moves from m to $m + 1$, only one observation joins the subset, but the method allows the inclusion of more than one observation as one or more leave $S^{(m)}$. Step 2 is repeated until $m = n$, and all observations are included in $S^{(m)}$.

The third step is the monitoring of quantities of interest. During the search, various quantities are monitored and recorded, so that informative plots can be produced and analyzed.

An important aspect is that estimates of σ^2 are not constant during the search. For each subset $S^{(m)}$, an estimate $s_{S^{(m)}}^2$ is produced, and as m increases, $s_{S^{(m)}}^2$ increases smoothly if there are no outliers in the data. An abrupt change in the trajectory of $s_{S^{(m)}}^2$ at $m = m'$ is an indication that an outlier has joined the subset.

An interesting property of the FS is that it is insensitive to the initial subset – provided that it is free of outliers – and the trajectories of the monitored quantities during the search converge, so that (roughly) the last one third of the observations to enter the search are the same irrespective of the initial subset (Atkinson, Riani, & Cerioli, 2006).

Even though the method was originally developed for data diagnostics, it has been recently extended, becoming an automatic robust procedure.

Riani, Atkinson, and Cerioli (2009) developed a hard rejection method based on the FS described by Atkinson and Riani (2000) algorithm where outlier detection is

based on (simulated or approximated) minimum deletion residual envelopes, according to the following rules:

- i. in the central part of the search, three consecutive values of $r_{\min}(m, n)$ exceed the 99.99% envelope or one exceeds the 99.999% bound;
- ii. in the final part of the search, two consecutive values of $r_{\min}(m, n)$ exceed 99.9% and one exceeds the 99% bound;
- iii. $r_{\min}(n - 2, n)$ exceeds the 99.9% envelope;
- iv. $r_{\min}(n - 1, n)$ exceeds the 99% envelope and, in this case, a single outlier is detected and the procedure terminates.

The final part of the search is defined as $m \geq n - [13 \left(\frac{n}{200}\right)^{0.5}]$.

The authors consider the break of any of the 4 rules as a signal, which is later reconfirmed by the superimposition of minimum deletion residual envelopes. The procedure is intended to provide a nominal size of 1%, meaning that the method should identify, on average, a signal once in every 100 outlier-free samples.

Grossi and Laurini (2009) develop a soft weighting robust estimator based on the FS described by Atkinson and Riani (2000). Their method is based on simulation envelopes, where the studentized residuals obtained during each stage of the search are compared with simulated envelope bounds. If the studentized residual lies outside the envelope, the distance between the value of the residual and the closest envelope bound is computed and used to calculate the weight of that observation, which will be used in a weighted regression. The envelopes are calculated at each stage of the search, based on parameter estimates, which are computed using the observations within the CDS.

More recently, Crosato and Grossi (2017) extended the FSW procedure of Grossi and Laurini (2009), developing a new approach for the identification of outliers in dependent data, more specifically in generalized autoregressive conditional heteroskedasticity (GARCH) models.

3. A NEW FSW ESTIMATOR

The FSW combines the soft trimming concept used by Grossi and Laurini (2009) and Crosato and Grossi (2017) with the “early stopping” concept proposed by Riani et al. (2009). While the former allows flexibility in down weighing observations, the latter guarantees the high breakdown of the method, allowing weights to be computed before the inclusion of outliers in the subset. Additionally, the FSW is based on a modified version of the simulation envelopes proposed by Atkinson (1981), which are constructed for every subset $S^{(m)}$, subject to $\frac{(n+p+1)}{2} \leq m \leq m^*$. When $m = m^*$, i.e., when any of the 4 rules used in the hard trimming method developed by Riani et al. (2009) is broken, the search is interrupted and weights are calculated before any outlier is included in the subset.

3.1 Simulation Envelopes

Simulation envelopes have long been used for the detection of outliers (see Atkinson [1981] and Flack and Flores [1989]). In the context of the FS, the simulation envelopes reflect the distribution of the studentized residuals at a specific subset size m of the search.

The envelopes proposed by Atkinson (1981) are generated in 4 steps:

- i. Simulate M vectors Z of dimension $(n \times 1)$ from the standardized normal distribution;
- ii. regress each of these vectors on the X matrix and obtain M simulated vectors of studentized residuals $R_{z(i)}$;
- iii. order the elements of each simulated studentized residual vector;
- iv. for each $i = 1, \dots, n$, select $l_i = \min r_{z(i)}$ and $u_i = \max r_{z(i)}$. These lower and upper values for the i th-order statistic of the M simulated residual vectors form the lower and upper bounds of the diagnostic envelopes, respectively.

Setting $M = 19$, one obtains envelope bounds corresponding to the 5th and 95th percentiles of the distribution of the i th-order statistic of the externally studentized residual vector, given X .

The methodology above is immediately applicable when OLS regressions are run on the full set of observations. However, given that the FS starts with subset size $m = p$,

the envelope is obtained by simulating M vectors Z and running a FS for each of these M vectors on matrix X , such as in Atkinson and Riani (2006), allowing the envelopes to be independent of the initial subset of the FS.

The envelopes are constructed according to the following steps:

- i. simulate M vectors Z of dimension $(n \times 1)$ from the standardized normal distribution;
- ii. conduct a FS of each of these vectors on the X matrix and obtain for each subset size m , M simulated vectors of studentized residuals r_z ;
- iii. for each subset size, order the elements of each simulated studentized residual vector, obtaining M simulated vectors of ordered studentized residuals r_z^* . Grouping all these vectors, one obtains, for each subset m , a matrix R with dimension $(M \times n)$, populated with elements $r_{(i,j)}$;
- iv. for each subset size m , sort each column of matrix R , such that the lowest value of the residuals in that column is allocated in the 1st row and the highest value is allocated in the last row, obtaining matrix R^* , with elements $r_{(i,j)}^*$;
- v. in order to obtain 5% and 95% envelope bounds select for each subset size, elements $l_i = r_{(i,0.05M)}^*$ and $u_i = r_{(i,0.95M)}^*$, for $i = 1, \dots, n$. These upper and lower values for the i th-order statistic of the M simulated residual vectors form the upper and lower bounds of the diagnostic envelopes, respectively.

3.2 Weighing Observations

For every $m \geq h$, the envelopes are constructed and the distance of the studentized residuals to the envelopes is computed.

The distance for observation i at subset size m is calculated by:

$$\pi_m^{(i)} = \begin{cases} 0, & \text{if } l_i \leq r_i \leq u_i \\ (l_i - r_i), & \text{if } r_i < l_i \\ (r_i - u_i), & \text{if } r_i > u_i \end{cases} \quad \boxed{7}$$

At the end of the search, the average distance of each residual to the envelope is used to determine the weights of each observation in the weighted regression. The overall average distance of observation i , calculated when the search is interrupted, is given by:

$$\pi_i = \frac{\sum_{m=\frac{(n+p+1)}{2}}^{m^*} \pi_m^{(i)}}{m^* - \frac{(n+p+1)}{2}} \quad \boxed{8}$$

The weight attributed to observation i is then calculated as:

$$w_i = e^{-\pi_i} \quad \boxed{9}$$

Finally, parameters are estimated through a weighted regression:

$$\hat{\beta} = (X'WX)^{-1}X'WY \quad \boxed{10}$$

Notice from equations 8 and 10 that, even if the search is interrupted before the inclusion of all observations in the subset (i.e., $m^* < n$), all observations are used for the estimation of $\hat{\beta}$.

4. DATA ANALYSIS

In this section, the performance of the FSW is assessed with simulated data, and the new estimator is applied to market data.

4.1 Application to Simulated Data

In this section, we conduct Monte Carlo simulations in order to assess the properties of the FSW estimator in a number of experiments specially designed to reproduce usual assumptions and stylized facts about the models and data used in the context of asset pricing. We also compare the performance of the FSW with OLS and LTS estimators both for outlier-free and for contaminated data. The LTS is set to trim 30% of the observations of the data.

We simulate pairs of asset and factors returns based on the true (beta) parameters and then estimate the parameters of the model from the simulated data.

In order to assess the bias, average discrepancy, and comparative efficiency of each estimator, we compute summary statistics of the cross-sectional distribution (across the N replications). The statistics are the mean estimated intercept (α) and mean estimated slopes (β_k), along with their cross-sectional average discrepancy (root-mean-square deviation [RMSE]) of estimated parameters away from their true values. We also compute the relative efficiency ($r.e$) of each estimator, determined as the squared ratio between the RMSE of the OLS estimates and the RMSE of the robust procedure $\left(\frac{RMSE_{OLS}}{RMSE_{Robust}}\right)^2$ in similar fashion to the analysis performed by Chan and Lakonishok (1992). Additionally, we present Diebold-Mariano (DM) statistics for the null hypothesis that the robust method and the OLS have equal forecast accuracy in forecasting the true parameters used across the N simulations.

All our experiments are based on the following general return-generating process:

$$\begin{aligned} r_{i,t} &= \alpha + \beta_1 f_{1,t} + \dots + \beta_k f_{k,t} + \varepsilon_{i,t}; \\ i &= 1, \dots, N; \\ t &= 1, \dots, T; \\ k &= 1, \dots, K. \end{aligned} \quad \boxed{11}$$

We set $N = 1,000$ simulations and report results for $T = 60, 180, 300$.

The Monte Carlo experiment consists of sampling random numbers for each of the K factors and for the error term of equation 11. The distributions from which the random numbers are sampled – as detailed below – are designed to replicate the mean and the SD of the factors and the error term.

In each experiment, we follow Grossi and Laurini (2011) and randomly replace 30% of the values of the dependent and independent variables with values drawn from a distribution with higher variance. This procedure results in fat-tailed distributions for both the dependent and independent variables, which are consistent with returns on stocks or portfolios. The variance of the contaminated data has been set at five times the original variance of the variable.

Setting $K = 3$, i.e., the return-generating process is the Fama and French (1992, 1993) three-factor model, where factor 1, 2, and 3 stand for, respectively, the excess market, size (small minus big [SMB]), and BE/ME (high minus low [HML]) returns. In each simulation i , excess market returns are drawn from a normal distribution with mean 0.43% and SD 4.54% – these values were taken from Fama and French (1993) –, and residuals ε_i are drawn from a normal distribution with mean 0 and SD 2.5%, and were designed to reproduce the average R^2 of approximately 75% obtained by the same authors. Factor 2 is drawn from a normal distribution with mean 0.27% and SD 2.89% and factor 3 is drawn from a normal distribution with

average 0.40% and SD 2.54%. Lastly, α , β_1 , β_2 , and β_3 are set to 0 and 1, 1.5, and 0.5, respectively.

Table 1 presents the results obtained when data is not contaminated: all the methods are unbiased and the FSW is highly efficient when data is not contaminated, showing efficiencies of at least, in contrast to the LTS, whose efficiency is approximately only 40%. The results

of the DM test indicate that one should reject, at the 5% significance level, the null hypothesis that the FSW and the OLS methods have equal forecast accuracy when $T = 180$ and $T = 300$. When $T = 60$, one cannot reject that the FSW and OLS have equal forecast accuracy. Nonetheless, one should reject the null hypothesis that the LTS and OLS have equal forecast accuracy for every T .

Table 1

Monte Carlo simulation – Three-factor model: Clean data results using Monte Carlo simulation on three alternative sample sizes

Estimation method	Performance statistic	T = 60				T = 180				T = 300			
		α	β_1	β_2	β_3	α	β_1	β_2	β_3	α	β_1	β_2	β_3
OLS	Bias	-0.01%	0.08%	0.52%	-0.41%	0.01%	-0.29%	0.02%	0.22%	0.00%	-0.13%	-0.13%	0.14%
	RMSE	(0.00)	(0.08)	(0.11)	(0.14)	(0.00)	(0.04)	(0.07)	(0.07)	(0.00)	(0.03)	(0.05)	(0.06)
	r.e.	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	DM	-	-	-	-	-	-	-	-	-	-	-	-
LTS (70%)	Bias	0.01%	-0.26%	0.23%	0.08%	0.00%	-0.26%	0.00%	-0.38%	0.00%	-0.19%	-0.41%	0.19%
	RMSE	(0.01)	(0.12)	(0.17)	(0.21)	(0.00)	(0.06)	(0.10)	(0.12)	(0.00)	(0.05)	(0.07)	(0.09)
	r.e.	0.42	0.41	0.43	0.44	0.43	0.43	0.39	0.40	0.44	0.45	0.50	0.43
	DM	13.43	7.79	12.86	10.79	16.07	14.39	14.79	15.64	15.40	15.89	16.31	15.62
FSW	Bias	0.00%	0.07%	0.50%	-0.44%	0.01%	-0.29%	0.01%	0.22%	0.00%	-0.13%	-0.13%	0.14%
	RMSE	(0.00)	(0.08)	(0.11)	(0.14)	(0.00)	(0.04)	(0.07)	(0.07)	(0.00)	(0.03)	(0.05)	(0.06)
	r.e.	0.98	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	DM	-2.26	-2.00	-1.42	-2.60	0.89	-0.68	-0.57	0.40	-0.82	0.18	-0.66	-0.91

DM = Diebold-Mariano; FSW = weighted forward search; LTS = least trimmed squares; OLS = ordinary least squares; r.e. = relative efficiency; RMSE = root-mean-square deviation; T = number of monthly simulated returns.

Source: Elaborated by the authors.

Table 2 presents the results obtained when data is contaminated with outliers: the FSW and LTS offer protection against the outliers introduced in the data, as they present much lower RMSE than OLS. However, the FSW presents higher efficiency than the LTS. Once again, all the methods provide unbiased estimates both

for the intercept and for slopes. At the 5% significance level, the DM statistics indicate the rejection of the null hypothesis that the LTS and FSW methods have the same forecast accuracy of the OLS method; the DM statistics indicating higher rejection levels for the FSW than for the LTS.

Table 2

Monte Carlo simulation – Three-factor model: Contaminated data results using Monte Carlo simulation on three alternative sample sizes

Estimation method	Performance statistic	T = 60				T = 180				T = 300			
		α	β_1	β_2	β_3	α	β_1	β_2	β_3	α	β_1	β_2	β_3
OLS	Bias	0.00%	-0.03%	0.14%	-0.05%	0.03%	-0.24%	0.07%	0.53%	0.00%	-0.11%	-0.07%	0.45%
	RMSE	(0.01)	(0.08)	(0.12)	(0.14)	(0.00)	(0.04)	(0.07)	(0.08)	(0.00)	(0.03)	(0.05)	(0.06)
	r.e.	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	DM	-	-	-	-	-	-	-	-	-	-	-	-
LTS (70%)	Bias	0.01%	0.05%	0.00%	-0.03%	0.01%	-0.09%	0.26%	0.26%	0.00%	-0.01%	-0.12%	0.03%
	RMSE	(0.01)	(0.06)	(0.08)	(0.10)	(0.00)	(0.03)	(0.04)	(0.05)	(0.00)	(0.02)	(0.03)	(0.04)
	r.e.	2.54	1.88	2.13	1.97	2.43	2.18	2.30	2.33	2.50	2.16	2.39	2.06
	DM	12.67	6.39	9.55	8.43	13.22	10.57	11.21	12.13	13.18	11.37	13.06	11.11

Table 2

Cont.

Estimation method	Performance statistic	T = 60				T = 180				T = 300			
		α	β_1	β_2	β_3	α	β_1	β_2	β_3	α	β_1	β_2	β_3
FSW	Bias	0.00%	-0.03%	0.09%	-0.02%	0.00%	-0.04%	0.22%	0.23%	-0.01%	-0.06%	-0.07%	0.04%
	RMSE	(0.01)	(0.05)	(0.08)	(0.09)	(0.00)	(0.02)	(0.04)	(0.04)	(0.00)	(0.02)	(0.03)	(0.03)
	r.e.	2.78	2.51	2.57	2.56	3.52	3.42	3.57	3.48	3.72	3.59	3.65	3.16
	DM	-14.99	-14.71	-14.34	-15.23	-15.46	-15.14	-16.25	-15.17	-15.33	-14.40	-15.02	-15.19

DM = Diebold-Mariano; FSW = weighted forward search; LTS = least trimmed squares; OLS = ordinary least squares; r r.e. = relative efficiency; RMSE = root-mean-square deviation; T = number of monthly simulated returns.

Source: Elaborated by the authors.

The results of the Monte Carlo simulations support the use of FSW for the estimation of parameters of asset pricing models. In the next section, these estimators are applied on real market data.

4.2 Application to Financial Data

In this section, we apply the FSW to the time series regressions framework developed by Black et al. (1972) and used by Fama and French (1993, 1996).

The dependent variables used in our tests are the excess returns of the well-known Fama-French 25 portfolios applied to the United States of America equity market (Fama & French, 1993), whereas the independent variables considered are the SMB, HML, and excess market returns. We perform GRS tests of whether the robustly and non-robustly estimated intercepts of all the portfolios are jointly equal to 0 and present results based on the CAPM and the three-factor models.

We also estimate the model on annual data, as according to Kothari et al. (1995), there are at least three reasons for using longer measurement interval returns in asset pricing tests: (i) the CAPM does not provide explicit guidance on the choice of interval for assessment of the explanatory power of beta; (ii) the use of longer interval returns mitigates biases in the beta estimates due to trading frictions and non-synchronous trading; (iii) using annual data is one way of bypassing statistical complications created by seasonality in monthly returns.

Given the lower availability of annual data (i.e., smaller sample size), our tests are based on the longer period of 1927-2012, yielding time series of 86 observations. Estimates on monthly data are based on the period July/1963-December/1991, as in Fama and French (1993).

The estimated CAPM and three-factor models follow the specifications in equations 12 and 13, respectively:

$$\begin{aligned}
 r_{i,t} - r_{f,t} &= a_i + b_i[r_{M,t} - r_{f,t}] + \varepsilon_{i,t}; \\
 i &= 1, \dots, 25; \\
 t &= 1, \dots, T.
 \end{aligned}
 \tag{12}$$

$$\begin{aligned}
 r_{i,t} - r_{f,t} &= a_i + b_i[r_{M,t} - r_{f,t}] + s_iSMB_t + h_iHML_t + \varepsilon_{i,t} \\
 i &= 1, \dots, 25; \\
 t &= 1, \dots, T.
 \end{aligned}
 \tag{13}$$

where $T = 86$ for annual data and $T = 342$ for monthly data.

In the single-factor model, the FSW intercept estimates obtained for each individual portfolio with monthly data are similar to those obtained by Fama and French (1993) with the OLS. The signs of the 25 intercept estimates are identical in both methods and all significant ($|t - statistic| > 2$) intercepts obtained with OLS are significant when estimated with the FSW.

The three-factor FSW estimates are also similar to those obtained with OLS. All the 4 $t - statistics$ with absolute values higher than 2 obtained with OLS also presented absolute values higher than 2 when estimated with FSW.

Moreover, the absolute values of the $t - statistics$ of two of these portfolios are considerably higher when estimated with FSW. Besides that, one additional portfolio presents $t - statistics$ with absolute value higher than 2 when FSW is used. Overall, results based on independent $t - tests$ suggest a worse explanatory power of the three-factor model when robust estimates are used.

Figures 1 and 2 show additional outputs obtained when applying the FSW on annual data, namely, the weights attributed to each observation of the dataset and the studentized residuals of the observations along with the simulation envelopes.

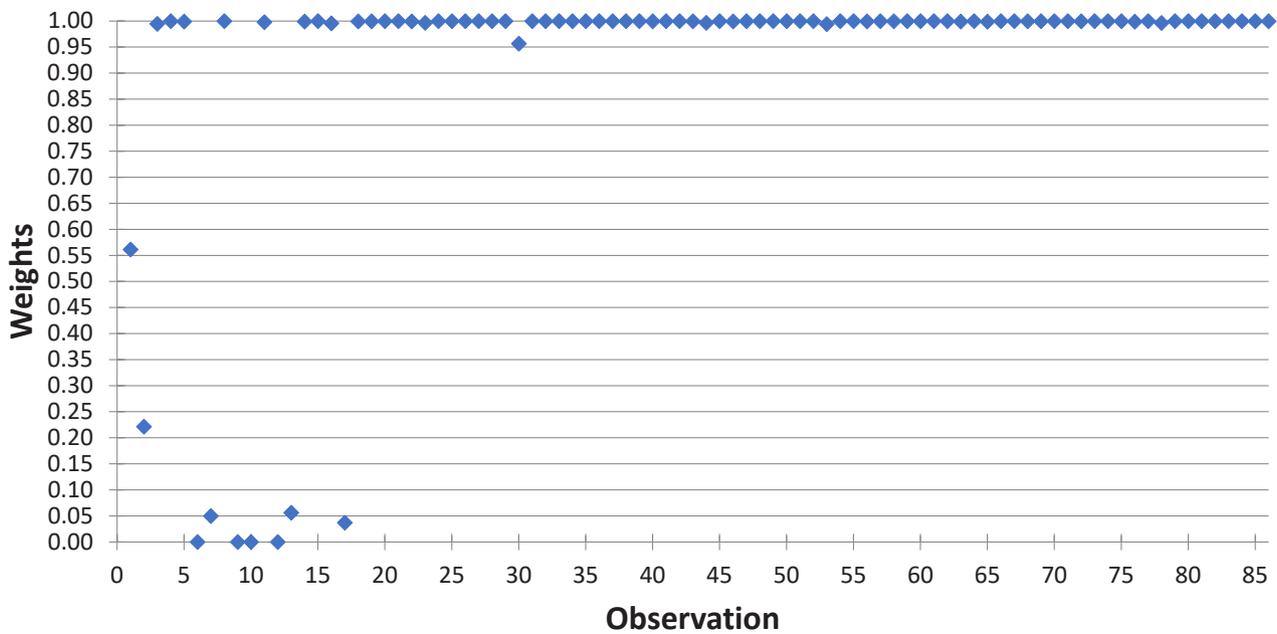


Figure 1 Weights attributed by the weighted forward search (FSW): Portfolio 2, three-factor model
Source: Elaborated by the authors.

Figure 1 presents the weight w_i – defined in equation 9 – of each observation, obtained as a result of the weighted estimation procedure. The results show that 7 observations, namely observations 6, 7, 9, 10, 12, 13, and 17, were severely down weighted by the FSW, while 2 others, observations 1 and 2, were attributed weights

between 0.2 and 0.6. The observed downweighting results from these observations lying outside the bounds of the simulation envelopes during the search.

Figure 2 shows the distribution of sorted studentized residuals when $m^* = 77$, i.e., when the search is interrupted.

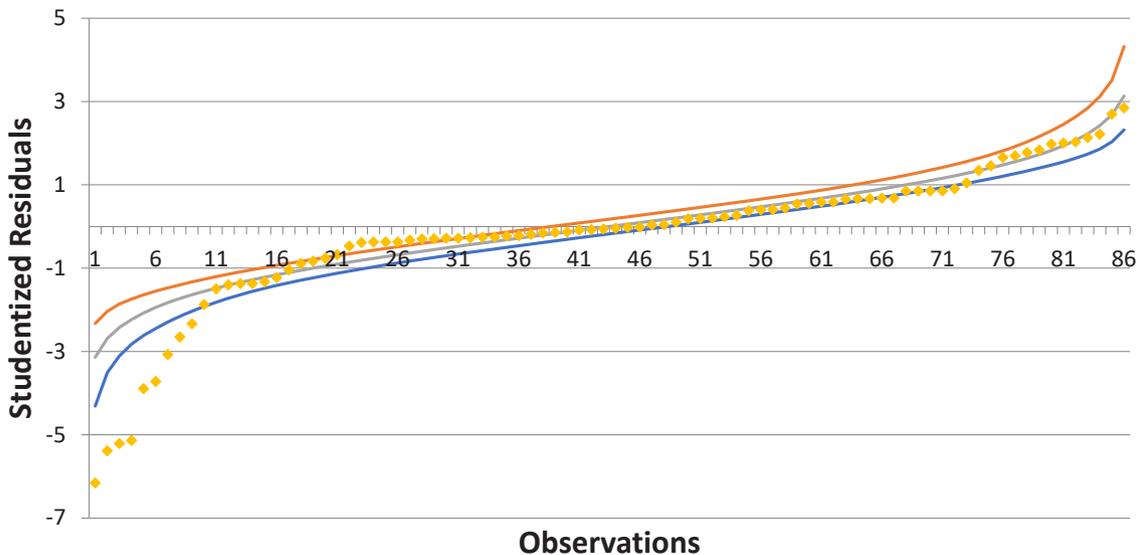


Figure 2 Sorted studentized residuals and simulation envelopes
Source: Elaborated by the authors.

The upper and lower solid curves represent the simulation envelope bounds. There are 9 observations lying outside the lower envelope bound on the bottom-left of the plot, corroborating the results presented in Figure 1.

These 9 observations are the ones with the lowest weights in Figure 1. The largest the distance of an observation from the bulk of the data – which lies within the envelope bounds – the lowest is its weight.

4.2.1 Influential observations

In this section, we analyze the observations that were down weighted by the FSW. Figure 3 shows the

average “outlyingness” – defined as one minus w_i obtained according to equation 9 – of each observation across the 25 portfolios for monthly data.

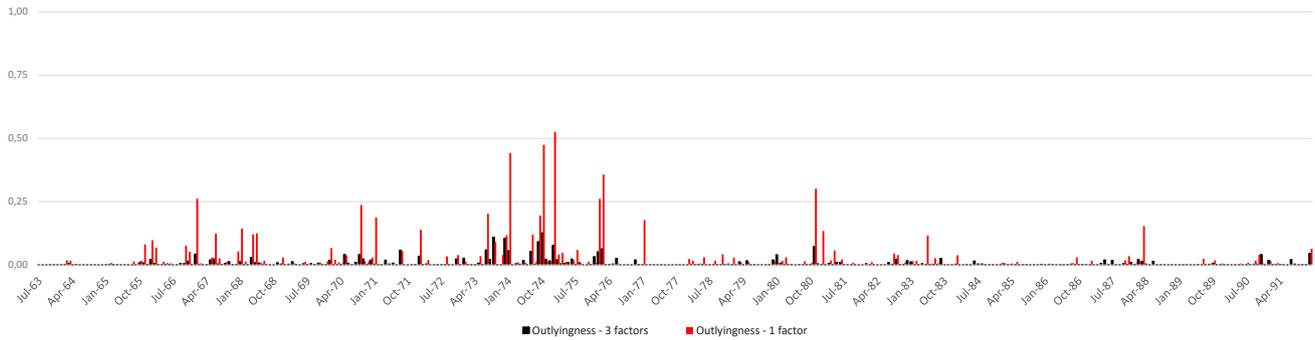


Figure 3 Average outlyingness across the 25 portfolios: Monthly data
Source: Elaborated by the authors.

The higher the bar of an observation in Figure 3, the higher its degree of outlyingness. Three conclusions can be immediately obtained from Figure 3: (i) average outlyingness is higher in the single-factor setting than in the three-factor one; (ii) the most outlying observations in the three-factor setting also show a high degree of outlyingness in the single-factor setting, i.e., there are common influential observations across the portfolios in both settings; (iii) there is a cluster of influential observations around the year of 1974.

The 4 most outlying observations (highest bars) in both setting are circumscribed in the period from September 1973 to January 1976. Interestingly, the period

corresponds to the well-known stock market crash of 1973-1974 (see Shiller [2015]).

It is worth mentioning that not all observations identified as influential are related to extreme values of a specific factor (dependent variable). For instance, February 1976, which is the 4th and 7th most outlying observation, respectively, in the single-factor and three-factor setting, corresponds only to an excess market return of 0.32%. Additionally, the most extreme value of the excess market return – which is 23.24% and corresponds to October 1987 – is not among the 10 most influential observations.

Figure 4 shows the average outlyingness of each observation across the 25 portfolios for annual data.

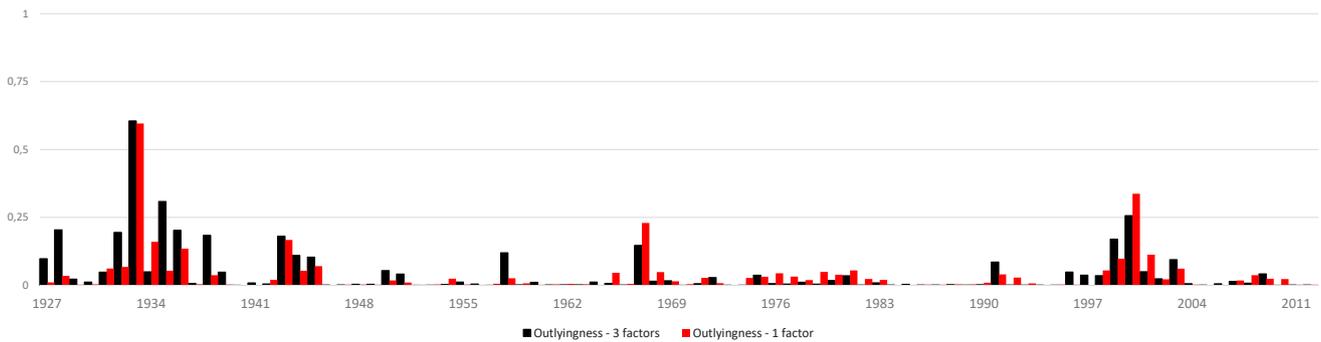


Figure 4 Average outlyingness across the 25 portfolios: annual data
Source: Elaborated by the authors.

In contrast to Figure 3, now the three-factor model shows a higher average outlyingness than in the single-factor setting. Figure 4 reveals the presence of clusters of influential observations: around 1933 and around 2000, which correspond to, respectively, the end of the 4-year recession post the 1929 stock market crash (1935 and 1936 correspond to the implementation of Franklin Roosevelt’s

New Deal Policy) and the collapse of the dot-com bubble. A comparison of Figures 3 and 4 suggests that the SMB and HML factors are able to explain extreme returns, which are considered outliers in a single-factor setting for monthly data, but not for annual data.

While the year of 1933, the most outlying observation, is in fact the most extreme excess return of factor 1, the

year of 1954 – the 2nd most extreme return – does not stand out as an outlying year.

Our results also corroborate the findings of Knez and Ready (1997) and Bailer (2005) that January is, indeed, an influential month. In the single-factor setting, the average outlyingness of the month of January across the 25 portfolios is 3.6 times the average outlyingness of October, the 2nd most outlying month, whereas in the three-factor

setting, the average outlyingness of the month of January is 1.43 times the average outlyingness of September, the 2nd most outlying month.

4.2.2 GRS tests

In order to test whether all the intercepts are jointly equal to 0, we compute the GRS (1989) statistics and p-values. Results are presented in Table 3.

Table 3

Gibbons Ross Shanken (GRS) test: monthly and annual data

	Monthly				Annual			
	Single-factor		Three-factor		Single-factor		Three-factor	
	OLS	FSW	OLS	FSW	OLS	FSW	OLS	FSW
GRS statistic	2.20	2.68	1.59	2.09	2.68	2.97	2.49	2.31
p-value (%)	99.90	100.00	96.20	99.79	99.91	99.97	99.78	99.54

FSW = weighted forward search; OLS = ordinary least squares.

Source: *Elaborated by the authors.*

Results on monthly data indicate that the three-factor model – estimated with OLS – is marginally rejected at the 0.95 level, while the CAPM is rejected at the 0.99 level, in line with the results of Fama and French (1993). Our results also show that using the robust estimates obtained with the FS leads to an increase in the rejection level of both models. In particular, the

three-factor model rejection level shifts from 0.9620 to 0.9979.

Using annual data, there is no significant difference from using OLS or robust estimates in both the single-factor and the three-factor model, but the robust estimates increase the rejection level of the single-factor model and decrease the rejection level of the three-factor one.

5. FINAL REMARKS

Our aim in this paper is to assess the impact of outliers on the estimation of asset pricing models and in associated statistical tests. For this, we propose a new weighted robust estimator, which was developed and applied for the estimation of asset pricing models. Comparison of the performance of the FSW, OLS, and LTS on simulated data indicates that the FSW method provides more reliable estimates in the presence of outliers, at the same time being almost as efficient as OLS when data is outlier-free. It should be also noticed that the precision of the intercept estimates is highly increased when the FS methods are used on contaminated data.

One contribution of the research is the application of the FSW – both efficient and robust – on the Fama and French (1993) 25 portfolios, which are a benchmark frequently used in the literature. The FSW allowed us to identify that many of these portfolios contain outlying observations, both in the single and three-factor model settings. The rejection levels in the GRS tests were increased when robust estimates were used, indicating

that the marginal rejection of the three-factor model is influenced by the presence of outliers in the portfolios. This is in line with previous research, which indicates that estimates of asset pricing models are highly sensitive to a few influential returns.

Another contribution of this paper is the estimation of the Fama and French (1993) three-factor model on annual data. Our results indicate that the model does not perform as well as it does on monthly data in explaining the cross-section of expected returns. Moreover, more outlying observations are detected in the three-factor model than in the single-factor CAPM. This is in contrast to our results on monthly data. The results of the GRS tests support the rejection of the hypothesis that the intercepts of the portfolios are jointly 0, both for the CAPM and for the three-factor model.

We have also provided evidence that some observations are commonly considered outliers both in the single and three-factor models. It is possible to relate these observations to particularly relevant events in the economy, such as financial market crashes, economic

crisis, and asset price bubbles. This is one direction that should be explored further.

Limitations of the research include the assumption that beta coefficients and risk premiums are constant across time. These assumptions could be relaxed and one could explore to what extent the results of

the conditional CAPM, such as those provided by Jagannathan and Wang (1996) are influenced by the presence of outliers.

Another promising possible way for future research would be to extend the works of Jensen (1967) and Carhart (1997) on the performance of mutual funds.

REFERENCES

- Atkinson, A. C. (1981). Two graphical displays for outlying and influential observations in regression. *Biometrika*, 68(1), 13-20.
- Atkinson, A. C. (1994). Fast very robust methods for the detection of multiple outliers. *Journal of the American Statistical Association*, 89(428), 1329-1339.
- Atkinson, A. C., & Riani, M. (2000). *Robust diagnostic regression analysis*. New York, NY: Springer-Verlag.
- Atkinson, A. C., & Riani, M. (2006). Distribution theory and simulations for tests of outliers in regression. *Journal of Computational and Graphical Statistics*, 15(2), 460-476.
- Atkinson, A. C., Riani, M., & Cerioli, A. (2004). *Exploring multivariate data with the forward search*. New York, NY: Springer-Verlag.
- Atkinson, A. C., Riani, M., & Cerioli, A. (2006). Random start forward searches with envelopes for detecting clusters in multivariate data. In Zani, S., Cerioli, A., Riani, M., & Vichi M. (Eds.). *Data analysis. Classification and the forward search* (pp. 163-171). Berlin: Springer-Verlag.
- Bailer, H. M. (2005). *Robust estimation of factor models in finance* (Tese de Doutorado). University of Washington, Washington.
- Black, F., Jensen, M. C., & Scholes, M. (1972). *The Capital Asset Pricing Model: Some empirical tests. Studies in the theory of capital markets*. New York, NY: Praeger.
- Bowie, D. C., & Bradfield, D. J. (1998). Robust estimation of beta coefficients: Evidence from a small stock market. *Journal of Business Finance & Accounting*, 25(3, 4), 439-454.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82.
- Chan, L. K. C., & Lakonishok, J. (1992). Robust measurement of beta risk. *The Journal of Financial and Quantitative Analysis* 27(2), 265-282.
- Cochrane, J. H. (2001). *Asset pricing*. Princeton, NJ: Princeton University Press.
- Cornell, B., & Dietrich J. K. (1978). Mean-absolute-deviation versus least-squares regression estimation of beta coefficients. *The Journal of Financial and Quantitative Analysis*, 13(1), 123-131.
- Crosato, L., & Grossi, L. (2017). Correcting outliers in GARCH models: A weighted forward approach. *Statistical Papers*, 1-32.
- Da, Z., Guo, R.-J., & Jagannathan, R. (2012). CAPM for estimating the cost of equity capital: Interpreting the empirical evidence. *Journal of Financial Economics* 103(1), 204-220.
- Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2), 427-465.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.
- Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1), 55-84.
- Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *The Journal of Finance*, 63(4), 1653-1678.
- Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy*, 81(3), 607-636.
- Flack, V., & Flores, R. (1989). Using simulated envelopes in the evaluation of normal probability plots of regression residuals. *Technometrics*, 31(2), 219-225.
- Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57(5), 1121-1152.
- Graham, J. R., & Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, 60(2, 3)187-243.
- Grossi, L., & Laurini, F. (2009). A robust forward weighted Lagrange multiplier test for conditional heteroscedasticity. *Computational Statistics and Data Analysis*, 53(6), 2251-2263.
- Grossi, L., & Laurini, F. (2011). Robust estimation of efficient mean-variance frontiers. *Advanced Data Analysis and Classification*, 5(1), 3-22.
- Hadi, A. S. (1992). Identifying multiple outliers in multivariate data. *Journal of the Royal Statistical Society B*, 54(3), 761-771.
- Hadi, A. S. (1994). A modification of a method for the detection of outliers in multivariate samples. *Journal of the Royal Statistical Society B*, 56(2), 393-396.
- Hadi, A. S., & Simonoff, J. S. (1993). Procedures for the identification of multiple outliers in linear models. *Journal of the American Statistical Association*, 88(424), 1264-1272.
- Huber, P., & Ronchetti, E. (2009). *Robust statistics* (2a. ed.). Hoboken, NJ: John Wiley & Sons.
- Jagannathan, R., & Wang, Z. (1996). The Conditional CAPM and the cross-section of expected returns. *The Journal of Finance*, 51(1), 3-53.
- Jensen, M. C. (1967). The performance of mutual funds in the period 1945-1964. *Journal of Finance*, 23(2), 389-416.
- Knez, P., & Ready, M. (1997). On the robustness of size and book-to-market in cross-sectional regressions. *The Journal of Finance*, 52(4), 1355-1382.

- Kothari, S. P., Shanken, J., & Sloan, R. (1995). Another look at the cross-section of expected stock returns. *The Journal of Finance*, 50(185), 185-224.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1), 13-37.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36(1), 3-27.
- Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1), 125-144.
- Riani, M., Atkinson, A. C., & Cerioli, A. (2009). Finding an unknown number of multivariate outliers. *Journal of the Royal Statistical Society*, 71(2), 447-466.
- Roll, R. (1988). . *The Journal of Finance*, 43(3), 541-566.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3), 341-360.
- Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust regression and outlier detection*. New York, NY: John Wiley & Sons.
- Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science*, 9(2), 277-293.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3), 425-442.
- Sharpe, W. F. (1971). Mean-absolute-deviation characteristic lines for securities and portfolios. *Management Science*, 18(2), 1-13.
- Shiller, R. (2015). *Irrational exuberance*. Princeton, NJ: Princeton University Press.