Size Matters, Book-To-Market Does Not!
The Fama-French Empirical CAPM Revisited**

by

M. LAMBERT[a] and G. HÜBNER[*b]

This version: 11th June 2014

* *Corresponding author*
[a] Assistant Professor, HEC Management School, University of Liège, Belgium. Mailing address: HEC-Management School of the University of Liège, 14 rue Louvrex, B-4000 Liège, Belgium. Tel.: (+32) 4 232 74 32. Email: marie.lambert@ulg.ac.be.
[b] The Deloitte Professor of Financial Management, HEC Management School, University of Liège, Belgium; Associate Professor, School of Business and Economics, Maastricht University, the Netherlands; Chief Scientific Officer, Gambit Financial Solutions, Belgium. Mailing address: HEC-University of Liège, Rue Louvrex 14 – Bldg N1, B-4000 Liège, Belgium. Tel: (+32) 4 2327428. Fax: (+32) 4 2327240. E-mail: g.hubner@ulg.ac.be.
** This paper has benefitted from comments by Antonio Cosma, Dan Galai, Aline Muller, Pierre Armand Michel, Patrick Navatte, Christian Wolff, Jeroen Derwall as well as the participants to the European Financial Management Association 2010 (Aarhus, Denmark), the French Finance Association 2010 (St-Malo, France), the 1st World Finance Conference (Viana do Castelo, Portugal), the Maastricht-Liège seminar day (Liège, September 2010), the University of Bologna seminar participants (September 2011) and the Luxembourg School of Finance seminar participants. Georges Hübner acknowledges financial support of Deloitte. All remaining errors are ours.

Size Matters, Book-to-Market Does Not!
The Fama & French Empirical CAPM Revisited

Abstract

The Fama and French factors do not reliably estimate the size and book-to-market effects. Our aim is to demonstrate that these factors have been under- and over-estimated, respectively, in the US market. We replaced Fama and French's independent rankings with the conditional ones introduced by Lambert and Hübner (2013), with some additional modifications designed to improve the sorting procedure. We have been able to highlight a much stronger size effect than has conventionally been documented. As a significant related outcome, the alternative risk factors have been found to deliver less specification errors when used to price passive investment indices.

Keywords: Size; Book-to-market; Momentum; Mimicking portfolios

JEL Classification: G11, G12

Pricing anomalies on the US stock markets have been documented since the early 1980s. Banz (1981) revealed a small size effect: firms with low market capitalization tend to outperform large cap stocks. Research conducted by Basu (1983), Rosenberg, Reid and Lanstein (1985) and Fama and French (1992, 1998) also reveals that value stocks (i.e. stocks with high book equity value in comparison to their market value) outperform growth stocks (i.e. stocks with low book-to-market ratio) over various sample periods. Finally, Jegadeesh and Titman (1993) point out a significant momentum effect in the US stock market by showing that significant gains can be realized from long positions in persistent winner stocks, and conversely from short positions in loser stocks.

Although these effects are well established in relation to both risk and mispricing, the influential work of Fama and French (1993) holds these first two market anomalies in proxy for liquidity risk and for market distress, respectively. Their paper develops a set of heuristics enabling the inference of size and book-to-market effects in the US market. The resulting so-called "Fama & French 3-factor model" has become a core version of empirical asset pricing models taught at many levels in many business schools.

Yet, while introducing the Fama and French model to MBA students (for example) many instructors throughout the world have been confronted with a rather uncomfortable feeling. When asking the typical question: "*After controlling for their beta and their size, which stocks' average returns do you believe is greater: those of value stocks, or those of growth stocks?*", many students would intuitively reply "*growth stocks, of course!*". Some time may be required to explain that this appears not to be the case. But are we really sure that there *is* a genuine value premium after removing the market and size effects? Or do most teachers simply try to post-rationalize what is merely the result of a methodological bias? This is the question that our paper attempts to address.

A number of recent studies have already started to fuel this debate. While the factor construction method developed by Fama and French (1993) has become the standard means by which to construct both size and value (i.e. book-to-market) premiums, some more recent studies suggest that the premiums obtained with the Fama and French technique could be misspecified. According to a study by Cremers, Petajusto and Zitzewitz (2010), value premium is overestimated in the Fama and French framework because this methodology does not distinguish the differential impact of value effects on small and larger sized portfolios (value effect has a greater impact on smaller stock portfolios). In addition, Huij and Verbeek (2009) indicated that mimicking portfolios calculated using the Fama and French methodology could suffer from an overestimation of value premium and an underestimation of momentum factors. According to Brooks, Li and Miffre (2008), the size premium could even capture some part of the value premium.

These sorts of issues pertaining to the Fama and French factor construction method could provide a realistic explanation for these apparent pollution effects, but this conjecture needs to be rigorously scrutinized. This scrutiny has not formed the primary focus of the aforementioned studies. Our paper revisits the way in which size and book-to-market effects translate onto risk factors, and applies this approach to the whole US market over an extended period (1980-2007). This time period has been selected in order to be able to compare our results with factor mimicking indices. We show that the Fama and French premiums are contaminated by cross-effects that are not adequately neutralized by their independent sorting procedure. This study follows the sequential methodology proposed by Lambert and Hübner (2013), used to isolate fundamental risks into portfolio returns. In their paper, Lambert and Hübner constructed risk premiums accounting for higher moments. The application of this sequential sort to our sample resulted in removal of most of the correlation in the data, thereby generating a new set of risk premiums for size, book-to-market and momentum. The objective of the study was to produce pure estimates of the returns associated with each risk exposure.

Our paper sheds new light on the relative importance of the size and book-to-market effects in the US market. In an experimental setting, we have sought to demonstrate that the independent sorting procedure creates a theoretical bias in the premiums definition: one that underestimates the size effect while overestimating value-growth effect. The new set of premiums better matches empirical observations of slight outperformance of growth stock over value stocks (using S&P500/Citigroup data) over our sample period. These results demonstrate the existence of a strong size effect over the period, but an insignificant value-growth effect.

The use of this modified Fama and French methodology has enabled us to deliver a new set of risk premiums that better price passive benchmark indices. The alphas of the 4-factor Carhart model proved largely insignificant across all the regressions. We also documented the superior accuracy of our alternative premiums for pricing individual stocks. Through a series of robustness checks, we were also able to further decompose the incremental explanatory power of the various modifications applied to the original Fama and French portfolio construction in our model. Amongst the incremental added value brought by each modification to the premium construction process, the choice of a sequential sort stands out as the primary and decisive source of improvement.

The rest of the paper is organized as follows. Section 1 addresses the problems related to the independent sorting procedure performed in the original Fama and French methodology. Section 2 presents the alternative methodology for constructing mimicking portfolios based on size, book-to-market, and momentum. Section 3 describes an analysis of the properties of the sequential (modified) and the independent (original) Fama and French samples of empirical risk factors. Section 4 details comparative tests regarding the specification power of each pair of premiums. Section 5 concludes.

1. Background: Correlation bias in the Fama and French (1993) methodology

The Fama and French (1993) three-factor model and its extension for momentum (authored by Carhart (1997)) have become the benchmark in performance evaluation. Using a dataset from the Center for Research in Security Prices (CRSP), Fama and French consider two methods of scaling US stocks, i.e. an annual two-way sort on market equity and an annual three-way sort on book-to-market according to New York Stock Exchange (NYSE) breakpoints (quantiles). They then construct six value-weighted (two-dimensional) portfolios at the intersections of the annual rankings (performed each June of year *y* according to the fundamentals displayed in December of year *y-1*). The size or *SMB* factor ("*Small minus Big*") measures the return differential between the average small cap and the average big cap portfolios, while the book-to-market or *HML* factor ("*High minus Low*") measures the return differential between the average value and the average growth portfolios. The authors have made these two factor series available online[i]. Carhart (1997) completes the Fama and French three-factor model by computing, along a similar method, a momentum (i.e. a 1-year prior-return) or *UMD* ("*Up minus Down*") factor that reflects the return differential between the highest and the lowest prior-return portfolios. Using his online data library, French has calculated a similar momentum premium by replacing book-to-market with the momentum risk dimension. The set of 2x3 size/momentum-sorted portfolios is rebalanced on a monthly basis.

1.1. Theoretical framework

In order to group together US stocks with small/large market capitalization and with low/high book-to-market ratios, Fama and French performed two independent rankings on market capitalization and on book-to-market.

Under independent sorting, the six portfolios will have approximately the same number of stocks only if size and book-to-market are unrelated characteristics; that is, if there is no significant correlation between the risk fundamentals. However, market capitalization and book-to-market are correlated. The study of Fama and French (1993) even points out that "*using independent size and book-to-market sorts of NYSE stocks to form portfolio means that the highest book-to-market/market equity quintile is tilted toward the smallest stocks*" (Fama and French, 1993, pp. 12). The disproportion between these portfolios indicates that the size effect cannot be equivalently diversified across book-to-market sorted portfolios, and therefore the size effect cannot be eliminated by difference. The use of NYSE breakpoints in the Fama and French approach also favors this over-representation of small stocks in the portfolios. The second consequence, as already noted by Cremers, Petajusto and Zitzewitz (2010), is that the Fama and French methodology underestimates the size effect by making calculations based on the return spread between 1) large cap stocks and 2) those displaying small and intermediate level of capitalization, instead basing calculations on the return spread of pure small cap stocks.

1.2 Numerical experiment evidence

This section sets out to qualify the existence and magnitude of the above-mentioned theoretical bias in the Fama and French method in the presence of correlated risk characteristics.

We performed this numerical test by generating random samples of stock returns and constructing theoretical size and book-to-market premiums according to the Fama and French method. We contrasted two scenarios: one with and one without correlation between the ranking based on company size and on book-to-market ratios. Specifically, we simulated the two-dimensional ranking on company size and book-to-market ratios of 70 stocks, as well as their corresponding return, and constructed the size and book-to-market premiums under both scenarios. Through this method of construction, the simulated premiums can be expected to display descriptive statistics close to the input parameters of the model and should not display significant differences in descriptive statistics. In the case that the Fama and French methodology proved unable to deal with the correlated rankings, we expected a significant deviation between the statistical properties of these sets of premiums under the two scenarios.

We proceeded as follows:

(a) We simulated the rankings of 70 stocks along the company size and the book-to-market criteria over a 120-month sample period. As in Fama and French (1993), stocks were ranked on a scale of 1 to 2 for their market capitalization, and on a scale of 1 to 3 for their book-to-market ratio. We adopted two scenarios for generating the simulated two-way rankings along the size dimension and the three-way ranking along the book-to-market dimension: 1) with correlation between the rankings made on size and book-to-market ratios, and 2) without correlation between the rankings made on size and book-to-market ratios. In order to run the Monte Carlo simulation, we used a uniform law for both rankings and modeled the correlation among those rankings by applying the historical average correlation between the ratios, i.e. 41%.

(b) We simulated the monthly returns of these 70 stocks conditional on their rankings defined in stage (a), i.e. defining one stochastic model per category (related to the six possible rankings) of stocks. Monthly returns of the 70 stocks that made up our sample were modeled using a multivariate Gaussian distribution.

The average and volatility of the return distribution of the six characteristic portfolio types as well as the correlation matrix were then used as the input of our subsequent analysis. Parameters of the Gaussian distribution were estimated based on historical data (from January 1980 to December 2007): we considered the average return and volatility of the six categories of stocks considered in stage (a). The value-weighted average return of each category (made available on French's website) enabled us to directly incorporate the market value weight in our simulation.

Table 1 displays historical statistics for the 2×3 portfolios composing the size and book-to-market premiums.

< Insert Table 1 >

We finally incorporated the correlations between these six stock categories into our simulation. Table 2 displays the correlation matrix between the 2×3 characteristic-sorted portfolios.

< Insert Table 2 >

(c) Based on these random samples of stock returns and rankings, we formed 2×3 portfolios and constructed the size and book-to-market factors along the original Fama and French methodology. Contrary to Fama and French, however, the factors were updated via a monthly rebalancing.

Table 3 displays the results of the Monte Carlo simulation conducted over 100 runs. Descriptive statistics for the simulated size and book-to-market premiums can be seen for each of the two scenarios considered. The table also analyzes the properties of the two series, defined as the differences between the two sets of factors.

< Insert Table 3 >

The data presented in Table 3 reveals significant differences in the statistical properties of the two sets of simulated premiums under the two scenarios of correlation for both the company size and book-to-market dimension. T-stats of the differences between the simulated distributions (S_SMB^C-S_SMB and S_HML^C-S_HML) are significantly negative in the case of the size premium, and positive in the case of the book-to-market premium. This suggests an undervaluation of the size premium, but an overvaluation of the book-to-market premium under the scenario of correlation between the rankings. Additionally, as for the size effect, the premium defined under the no-correlation scenario likewise displayed statistical properties very similar to the original simulation input.

This simulation experiment delivers strong numerical evidence that the independent sorting procedure – due to the correlation between the rankings and the disproportionate weights between portfolios – causes spurious estimates of the returns related to size and value-growth effects. As highlighted by Cremers, Petajusto and Zitzewitz (2010), this analysis shows that the Fama and French method creates an imbalance between the base portfolios that yield the premiums. A disproportionate weight is placed on small value stocks. The independent ranking procedure appears to only sub-optimally diversify sources of risk other than the one to be priced. It does not sufficiently take into account the correlations across risk dimensions.

2. An alternative to the Fama and French procedure: the sequential sorting technique

Another approach would be to replace the independent rankings with a sequential sorting procedure. We have sought to demonstrate that such a technique would lead to a purification of risk factors by ensuring the homogeneity of each constructed portfolio on all three fundamental risk dimensions (i.e. book-to-market, momentum and size).

2.1. Principle

The modified factor construction approach differs from the original Fama and French methodology on a number of points. Firstly, the modified methodology comprises a comprehensive framework that analyzes the three empirical risk dimensions (size, book-to-market, and momentum) together. Furthermore, each form of risk is considered with identical weighting. In addition, the modified methodology proposes a consistent and systematic sorting of all listed stocks, while Fama and French perform a heuristic split according to NYSE stocks only.

Secondly, the monthly rebalancing of the portfolios more realistically captures the returns associated with some time-varying dimensions of risk, such as liquidity issues or market distress. Finally, our sequential sort avoids spurious cross-effects in risk factors due to any correlation between the rankings underlying the construction of the benchmarks.

The following subsections detail the construction of the sequential premiums.

2.1.1. Sequential sorting procedure

In designing the sorting procedure, our objective was to detect whether, when controlling for two out of the three risk dimensions, there is still enough return variation related to the third risk criterion. Therefore, we substituted the Fama and French "independent sort" with a "sequential" or "conditional sort", i.e. a multi-stage sorting procedure. More specifically, we performed three sorts successively. The first two sorts operated on "control risk" dimensions, followed by the risk dimension to be priced.

The sequential sorting produced 27 portfolios capturing the return related to a low, medium, or a high level on the risk factor, conditional on the levels registered on the two control risk dimensions. Taking the simple average of the differences between the portfolios scoring, respectively, high and low on the risk dimension to be priced, but scoring at the same levels for the two control risk dimensions, we were able to obtain the return variation related to the risk under consideration.

This procedure is similar to that of Lambert and Hübner (2013). To obtain the risk premium corresponding to dimension *X,* after sequentially controlling for dimensions *Y* and *Z*, the factor can be computed as:

$$X_{Y,Z,t} = \frac{1}{9}\left[\sum_{b=H,M,L}\sum_{c=H,M,L} R_t\left(HX|bY|cZ\right) - \sum_{b=H,M,L}\sum_{c=H,M,L} R_t\left(LX|bY|cZ\right)\right] \quad (1)$$

where $R_t\left(aX|bY|cZ\right)$ represents the return of a portfolio of stocks ranked *a* on dimension *X*, among the basket of stocks ranked *b* on dimension *Y*, themselves among the basket of stocks ranked *c* on dimension *Z*. Dimensions *X*, *Y* and *Z* stand for size, book-to-market and momentum (in any order) while *H*, *M* and *L* stand for high, medium and low, respectively.

In contrast to an independent sorting, this sequential sorting ensures the same number of stocks in all 27 portfolios. All portfolios therefore provide the same level of diversification.

2.1.2. Three-way sort

We split the sample according to three levels of size, book-to-market (BTM), and momentum[ii]. Two breakpoints (1/3rd and 2/3rd percentiles) were used for all fundamentals. Instead of the original six portfolios, this method leads to a set of 27 baskets of stocks. The breakpoints are based on all US markets, not only on NYSE stocks. The finer size classification also contributes to balance the proportion between the small/value, small/growth, large/value and large/growth portfolios. It also provides a better distinction between small and large cap stocks.

2.1.3. Monthly rebalancing

To apply a monthly rebalancing strategy, we assumed market participants to have referred to the last quarterly reporting to form their expectations about each stock. On this basis we therefore employed a linear interpolation to transpose annual debt and asset values into quarterly data, as this is the usual publishing frequency on the US markets:

$$D_{ik} = D_{i,y-1} + \frac{k}{12}(D_{i,y} - D_{i,y-1}) \tag{2}$$

$$A_{ik} = A_{i,y-1} + \frac{k}{12}(A_{i,y} - A_{i,y-1}) \tag{3}$$

for k = 3,6,9,12, i.e. k^{th} month of year y. Secondly, we ignored unrealistic values[iii] of BTM for the US markets (i.e. higher than 12.5) in line with the study of Mahajan and Tartaroglu (2008).

2.2. Data

The sample used in this paper comprised all NYSE, AMEX, and NASDAQ stocks collected from *Thomson Financial Datastream* for which the following information was available[iv]: company annual total debt, company annual total asset[v], the official monthly closing price adjusted for subsequent capital actions, and the monthly market value. We only recorded monthly observations of returns and market values[vi] when stock return did not exceed 100% and whose market values were strictly positive. This step was implemented in order to avoid outliers that could result from errors in the data collection process. We define the book value of equity as the net accounting value of the company assets, i.e. the value of the assets net of all debt obligations.

From a total of 25,463 dead and 7,094 live stocks available as of August 2008, we retained 6,579 dead and 4,798 live stocks matching all the criteria previously identified for the period ranging from February 1973 to June 2008. The usable sample for the risk premiums ranges from May 1980 to April 2007 due to some missing accounting data. The analysis covers 324 monthly observations. The market risk premium corresponds to the value-weighted return on all US stocks minus the one-month T-Bill rate.

We can illustrate our methodology with the *HML* factor construction. We started by breaking up the NYSE, AMEX, and NASDAQ stocks into three groups according to the market capitalization criterion. We then successively scaled each of the three size-portfolios into three classes according to their 2-12 prior return. Splitting each of these nine portfolios once again to form three new portfolios according to their book-to-market fundamentals, we ended up with 27 value-weighted portfolios. The rebalancing was performed on a monthly basis. For each month *t*, each stock was ranked on the selected risk dimensions. An analogy could be made to a cubic construction: each month, any stock integrates one slice, then one row, then one cell of a cube and thus enters one and only one portfolio. The stock specific value-weighted return in the month following the ranking was then related to the reward gained through the risks incurred in this portfolio.

Amongst the 27 portfolios inferred from the sequentially sorted risk factors, we retrieved only the 18 that scored at a high or a low level on the risk dimension, i.e. value/growth. The nine self-financing portfolios were then created from the difference between high- and low-scored portfolios displaying the same ranking on the size and momentum dimensions (used as control variables). Finally, the *HML* risk factor was computed as the arithmetic average of these nine portfolios.

3. Properties of the sequential factors and the original Fama and French factors

Section 1 of this paper presented only preliminary evidence regarding the correlation bias inherent in the Fama and French factor construction. This section outlines our investigation into the empirical impact of the methodological changes introduced above, the first of which was the sequential sort procedure. To this end, we performed a systematic analysis of the pricing properties of the competing sets of factors.

Table 4 details the descriptive statistics for size (SMB' and SMB_{ff} stand respectively for the alternative or original Fama-French factors), book-to-market (HML' and HML_{ff} stand respectively for the alternative or original Fama-French factors), and momentum (UMD' and UMD_{ff} stand respectively for the alternative or original Fama-French factors) premiums over the period May 1980-April 2007.

< Insert Table 4 >

All Fama and French premiums displayed positive average returns over the period studied, but only the HML_{ff} and UMD_{ff} premiums proved significant over the period (at the usual significance levels). The momentum strategy displayed the strongest returns, with an average value more than five times higher than that displayed by the size premium, and almost twice that of the HML_{ff} strategy. The momentum premium also proved more volatile. The *HML'* premium yielded a very small, insignificant negative average return over the total period. The descriptive statistics confirmed our previous claim that, in the presence of correlations between the rankings, the independent sorting procedure advocated by Fama and French (1993) underestimates the size premium while overvaluing the value-growth premium.

We compared the dynamics of the *HML'* premium with the S&P 500/Citigroup Growth and Value Indexes over the same period. Our results match empirical data. The S&P 500/Citigroup Growth Index slightly outperforms the value of the S&P 500/Citigroup Value Index over this period.[vii] Note that the magnitude of the *SMB'* premium proved similar to that of the momentum strategy in the alternative framework; they produce approximately the same (significant) positive average return over the period. The *UMD'* premium displays very similar characteristics to the corresponding Fama and French premium.

In order to analyze the impact of our modifications on the original Fama and French method and the documented differences in descriptive statistics, we examined the nine return spreads resulting from each of our three-stage sequential sorting procedures, as well as the return spreads resulting from the Fama and French construction.

Table 5 reports descriptive statistics for the three sets of nine return spreads related respectively to the *SMB'*, *HML'*, and *UMD'* factors. For each panel, the ordering sequence ends up with the dimension to be priced, as explained in the methodological section. We closely examined the correlations between these three sets of nine portfolios and the *SMB'*, the *HML'*, and the *UMD'* factors[viii].

< Insert Table 5 >

Panel A shows that each of the nine return spreads related to the *SMB'* factor evaluates equivalently the premium related to the size effect. All portfolios were found to offer comparable levels of mean returns and volatilities. The coefficient of variation for the series of average returns across portfolios was quite low (i.e. CV=0.26/0.88 or 0.30). In addition, the portfolios displayed strong correlations with the *SMB'* factor, but weak correlations (less than 30%) with the *HML'* and *UMD'* factors.

Panel B shows the nine differences to be correlated at, on average, 54.77% with the *HML'* factor, but display only weak correlations (less than 30%) with other types of risk. The book-to-market risk premium proved to be highest in portfolios formed of stocks of low (resp. medium) market capitalizations and delivering low (resp. medium) levels of prior returns. The table shows very large variations within the series of mean returns across the different book-to-market spreads. This preliminary descriptive analysis does not confirm the presence of a book-to-market effect in our sample. This might indicate that the sort on BTM captures only noisy returns that cannot be related to a source of risk priced on the market. Several papers point out the possibility of a mispricing as an explanation for the positive return spread between value and growth stock, the latter being even considered riskier than the former. Research conducted by Mohanram (2004) and Michou (2007) shows that the distinction between growth and value stocks could help in distinguishing winner from loser stocks. The book-to market effect has also been presented as being the strongest in low capitalized stocks (Griffin and Lemmon, 2002), or even as being partially explained by the size effect (Brooks, Li and Miffre, 2008). If these conjectures are valid, we could reasonably expect that after having controlled for the size and the momentum effects, the book-to-market effect would be seriously mitigated, if not entirely removed.

Panel C shows that the momentum effect decreases with market capitalization. The momentum spreads tend to be highest in stocks presenting small or medium levels of market capitalization.

Table 6 outlines the same analysis conducted on the two spread portfolios leading to the HML_{ff} and UMD_{ff} factors respectively, and the three spread portfolios forming the SMB_{ff} factor.

< Insert Table 6 >

The size spreads forming the SMB_{ff} factor displayed in Table 5 are all proved strongly correlated with the SMB_{ff} factor but, contrary to the return spreads forming the *SMB'* factor, they also displayed substantial correlations with the HML_{ff} factor (greater than 30% for all three portfolios). While our specification delivers portfolios that are quite homogeneous with regards to the return spreads related to size, in this case the low book-to-market-sorted portfolios display a very different average size spread compared to those of the other two portfolios. The coefficient of variation even increases from 0.30 to 2.14 (i.e. CV=0.30/0.14).

This evidence suggests that the Fama and French empirical size factor is contaminated by a book-to-market effect; this is also indicated by the values taken by the cross-correlations between the size return spreads and the HML_{ff} factor. This even results in a negative size return spread in the low book-to-market portfolio (where the reward associated with the book-to-market effect is in fact negative).

As previously mentioned, our size factor has been constructed on the basis of the return differential between portfolios of extremely small caps and portfolios of big stocks. By considering all the NYSE, NASDAQ, and AMEX stocks, our breakpoints are tilted towards small caps compared to the Fama and French premium. This could explain the larger average spread observed for this premium.

Similarly, the two book-to-market spreads forming the HML_{ff} factor display strong correlation with the HML_{ff} factor, but still present moderate levels of correlation with the SMB_{ff} factor. The characteristics of the book-to-market return spread portfolios confirm evidence that the book-to-market effect is highest in low size portfolios. Finally, the momentum factor constructed according to our modified specification only halves the level of volatility compared to the Fama and French UMD_{ff} factor. Substantial variation in returns related to the Fama and French momentum risk between small and big capitalizations was observed. The returns proved more stable across the nine different portfolios resulting from the sequential technique, showing that the size effect has been eliminated. The average coefficient of variation for the series of cross-sectional mean returns equaled 0.80 (0.63/0.79) while remaining a moderate 0.55 (0.5/0.91) when considering the sequential sorts.

In conclusion, the sequential construction method appears to induce a large correlation of the post-formation spread portfolios with the priced factor, while simultaneously isolating the effects of the other two sources of risk. The Fama and French factors do not seem to adequately price the returns attached to the size and book-to-market effects respectively, but rather appear to be contaminated with correlated sources of risk. The book-to-market premium (insignificant in the sequential framework) might be responsible for this contamination effect. Following Cremers, Petajusto and Zitzewitz. (2010), we argue that the value-growth effect could be exaggerated; our analysis even suggests that the book-to-market effect does not capture any kind of systematic risk priced on the stock market.

Table 7 displays the correlation matrix of these two sets of premiums.

< Insert Table 7 >

The bottom-left corner displays the cross-correlations between the two sets of premiums. The *SMB'* and *HML'* factors are correlated at 67.16 % and 68.25% with their Fama and French counterparts, respectively. These levels indicate that, although the original and the modified size and value premiums are intended to price the same risk, approximately one third of their variation provides different information. The analyses detailed in Tables 4 and 5 highlight the potential reasons for this difference. The momentum premium displays a higher correlation with the UMD_{ff} factor. Contrary to the SMB_{ff} and HML_{ff} factors, French's downloadable momentum premium does not follow the Fama and French (1993) methodology exactly: the premium is rebalanced monthly rather than annually. It differs from our momentum premium only with regard to the breakpoints used for the rankings and the sequential sorting. The bottom-right corner presents the intra-correlations among the Fama and French premiums. The SMB_{ff} and HML_{ff} factors are highly negatively correlated over the period (-40.83%). The UMD_{ff} premium also displays a negative correlation with the HML_{ff} factor, but a positive correlation with the SMB_{ff} factor. This evidence contrasts with the top-left corner, which presents the intra-correlations among the sequential premiums. The signs are the same as those displayed by the Fama and French premiums, but the correlation levels are considerably lower; this is consistent with our objective of designing uncorrelated premiums. The intra-correlations among the Fama and French premiums are all statistically significant, whereas the correlations among our alternative factors are only significant (at an inferior level) between the SMB_{ff} and HML_{ff} factors.

4. Specification tests

Two types of specification tests conducted are outlined in this section. We first performed a basic efficiency test (similar to that performed by Cremers, Petajusto and Zitzewitz (2010)) on the empirical asset pricing model in order to evaluate whether the original and the modified Fama and French specifications are able to price passive indexes and passive investment portfolios without specification errors. Following this, we carried out a direct and rigorous comparison of the competing models. The procedure featured a test of non-nested models on individual stocks. The outcome of this test delivers the proportions of stock return series for which there is a statistical dominance of one specification over the other.[ix]

4.1. Factor efficiency test

We evaluated the specification errors displayed by both a modified and original four-factor Carhart analysis on the set of 2x3 Fama and French portfolios and on a set of passive benchmark indices.

4.1.1. The 2x3 set of Fama and French portfolios

These portfolios were constructed on the basis of a two-way sort into size and a three-way sort into book-to-market. The time series are downloadable from French's website.

We considered the following multivariate linear regression to test the values of the alphas:

$$R_{pt} = \alpha_p + \beta_m R_{mt} + \beta_{SMB} R_{SMBt} + \beta_{HML} R_{HMLt} + \beta_{UMD} R_{UMDt} + \varepsilon_{pt} \quad for \;\; p=1,\ldots,N \qquad (4)$$

Instead of testing N univariate t-statistics based on each equation, we used the Gibbons, Ross and Shanken (1989) (GRS) test on the joint significance of the estimated values for α_p across all N equations:

$$H_0 : \alpha_p = 0 \qquad\qquad for\ p=1,...,N \tag{5}$$

Following Gibbons, Ross and Shanken (1989), under the null hypothesis (H_0) that α_p is equal to 0 for all N portfolios, the statistics $[T/(1+\overline{R}_F^{'}\hat{\Omega}^{-1}\overline{R}_F)]\alpha_p^{'}\sum^{-1}\alpha_p$ follows a central F distribution with degrees of freedom N and (T-N-L), where $\overline{R}_F$ is a vector of sample means for the L factors ($\widetilde{R}_{Ft}$), $\hat{\Omega}$ is the sample variance-covariance matrix for $\widetilde{R}_{Ft}$, $\sum$ is the variance-covariance matrix of the residuals, $\hat{\alpha}_p$ is a vector of the least squares estimates of the α_p across the N equations.

We applied the Gibbons, Ross and Shanken (GRS) methodology to the set of size and book-to-market-sorted portfolios using returns from May 1980 to April 2007. We considered the case where L = 4 (i.e.the market index, the *SMB*, the *HML*, and the *UMD* factor) and N = 6 for the six independent portfolios. The F statistic to test hypothesis (4) when using the set of Fama and French premiums is 0.0597, so we cannot reject efficiency of the Fama-French model at the usual levels of significance. When using the sequential premiums, the F statistic is reduced even further to 0.0000272. In this way, both sets of premiums seem to efficiently explain stock returns, with a slight advantage to the sequential approach. In other words, the different changes performed on the original Fama and French methodology do not seem to affect the efficiency of the factors.

4.1.2. Passive benchmark indices

Following the study of Cremers, Petajusto and Zitzewitz (2010) we applied a four-factor Carhart model to a set of passive indexes. We evaluated both the original Fama and French factors and the modified versions developed in this paper. We considered the following passive benchmarks (All, Growth and Value): Russell 1000, Russell 2000, Russell 3000, S&P500, S&P MidCap, S&P SmallCap over the period April 1997-April 2007, which represents the common sample period for all benchmarks.

< Insert Table 8 >

Table 8 shows that the original four-factor model of Fama and French and Carhart produces significant levels of specification errors (alphas of the model) for almost all passive benchmark indices. This result is fully consistent with what Cremers, Petajusto and Zitzewitz (2010) demonstrate in their study conducted over the period 1980-2005. The modifications to the Fama and French methodology enabled us to deliver a new set of risk premiums that better prices passive benchmark indices. Indeed, alphas of the four-factor Carhart model are mostly insignificant across all the regressions.[x]

4.2. Non-nested models

This sub-section attempts to identify the potential superiority of one set of empirical premiums (i.e. either those of Fama and French or our updated premiums) over the other. Our investigation follows the literature on model specification tests against non-nested alternatives (MacKinnon, 1983; Davidson and MacKinnon, 1981, 1984). Such tests have already been used in both financial and macroeconomics literature[xi].

We considered the following two models:

1. M1 or the *F&F* model:

$$R_{it} = \alpha_{1,i} + \beta_i \mu_t + \delta_i X_t^{'} + \varepsilon_{it} \tag{6}$$

2. M2 or the sequential model :

$$R_{it} = \alpha_{2,i} + \beta_i \mu_t + \gamma_i Z_t^{'} + \varepsilon_{it} \quad (7)$$

where R_i stands for the excess return on asset *i,* μ for the market premium, *X'* for the Fama and French premiums, and *Z'* for the sequential risk premiums.

We carried out a joint test of the Fama-French model and the sequential model. First, the model to be tested was designated as M1, and the alternative model M2. To test the model specification, we set up a composite model within which both models are nested. The composite model (M3) appears as:

$$R_{i,t} = \alpha_{3,i} + \beta_i \mu_t + (1-\theta_{i,1})\delta_i X_t^{'} + \theta_{i,1}\gamma_i Z_t^{'} + \varepsilon'_{it} \quad (8)$$

Under the null hypothesis $\theta_{i,1} = 0$, M3 reduces to M1; if $\theta_{i,1} \neq 0$, M1 is rejected. Tests were conducted on the value of $\theta_{i,1}$. Davidson and MacKinnon (1981, 1984) prove that under H_0, $\hat{\gamma}$ can be replaced by its ordinary least square (OLS) estimate from M2 so that $\theta_{i,1}$ and δ_i (and $\alpha_{3,i}, \beta_i$) are estimated jointly. This procedure is called the "J-test". We define $\delta_i^{*} = (1-\theta_{i,1})\delta_i$ so that M3 can be rewritten as follows:

$$R_{i,t} = \alpha_{3,i} + \beta_i \mu_t + \delta^{*}{}_i X_t^{'} + \theta_{i,1}\hat{\gamma}_i Z_t^{'} + \varepsilon'_{it} \quad (9)$$

To test M2, we reversed the roles of the two models. We constructed model M4:

$$R_{i,t} = \alpha_{4,i} + \beta_i \mu_t + \theta_{i,2}\hat{\delta}_i X_t^{'} + \gamma^{*}{}_i Z_t^{'} + \varepsilon''_{it} \quad (10)$$

We replaced δ_i by its estimate along M1 ($\hat{\delta}_i$) and estimated $\gamma^{*}{}_i$ (and $\alpha_{4,i}, \beta_i$) jointly with $\theta_{i,2}$. If $\theta_{i,2}$ = 0, M4 reduces to M2; if $\theta_{i,2} \neq 0$, M2 is rejected. Tests were conducted on the value of $\theta_{i,2}$.

We evaluated the goodness-of-fit of the two alternative asset pricing models on both a set of 11,377 individual stocks and on the Fama and French 2x3 portfolios sorted by size and book-to-market. The following hypotheses were jointly tested on all the individual test assets:

Hypothesis I: $H_0 : \theta_{i,1} = 0$ *against* $H_1 : \theta_{i,1} \neq 0$;

Hypothesis II: $H'_0 : \theta_{i,2} = 0$ *against* $H'_1 : \theta_{i,2} \neq 0$

Each θ_j followed a normal distribution with mean $\bar{\theta}_j$ and volatility σ_j. Therefore, under the null hypothesis, the statistics θ_j / σ_j followed a Student distribution with 315 degrees of freedom – the number of observation in each time-series (i.e. 324) minus the number of factors in each regression (i.e. 9: the constant, the market portfolio, the 2 sets of 3 empirical premiums, and the θ estimate). Among the four possible scenarios, we considered two cases:

- (H_0, H'_1), M1 is not rejected but M2 is;
- (H'_0, H_1), M2 is not rejected but M1 is[xii].

Table 9 presents the results of the tests over the significance of θ_1 and θ_2 across both sets of assets for different confidence levels. The tests about the value of θ_1 and θ_2 have the form:

$$H_0\left|\frac{\theta_{1i}}{\sigma_{1i}}\right| \le t_{stat}, H_1\left|\frac{\theta_{1i}}{\sigma_{1i}}\right| > t_{stat} \quad and \quad H'_0\left|\frac{\theta_{2i}}{\sigma_{2i}}\right| \le t_{stat}, H'_1\left|\frac{\theta_{2i}}{\sigma_{2i}}\right| > t_{stat} \tag{11}$$

< Insert Table 9 >

The table displays, for different levels of significance, the frequency of non-rejections of the Fama and French model, i.e. H_0 (resp. of the sequential model, H'_0) while rejecting the sequential premiums i.e. H'_1 (resp. of the F&F premiums, H_1). It also reports the frequency of assets for which M1 and M2 were both either rejected or not rejected. The first quarter of the table reflects the performance of the sequential model (Not reject M2 & Reject M1), while the second quarter identifies the frequency of dominance of the original Fama and French model. The modified version proved less frequently rejected than Fama and French premiums for individual stocks. The gap proved largest at the 10% significance level, where the test lead to the non-rejection of the sequential premiums 6.54% more often than for Fama and French premiums.

Overall, the non-nested econometric analysis performed on individual assets shows that in most cases neither the original Fama and French nor the sequential models were rejected when compared to the augmented model. This result does not imply that any of the models provides a good fit, as this judgment lies outside the scope of this test when performed on a database of individual stocks. For a limited subset of stocks (up to ca. one third) we can, however, discriminate between these models. Our alternative premiums seem to outperform the Fama and French specification. The extent of this superiority is economically quite important, as the adoption of Fama and French factors instead of the sequential ones would be (statistically) an incorrect choice for almost 4,000 individual stocks.[xiii]

5. Conclusions

This paper revisits the size and book-to-market effect in the US market over the 1980-2007 sample period. Our study has demonstrated a strong size effect but an insignificant book-to-market effect over the sample period. Our results challenge the evidence previously presented by Fama and French suggesting the presence of a stronger book-to-market than size effect in the US market. Fama and French's size and book-to-market premiums are indeed shown to be, respectively, insignificant and significantly positive over the analyzed period[xiv]. Their evidence is partly supported in the standard construction methodology itself, as more weight is attributed to the ranking according to the book-to-market dimensions (Fama and French, 1993).

We propose an alternative way to construct the empirical risk factors of Fama and French (1993), avoiding contamination of the premiums from the correlation structure of the data. Indeed, this paper aims to address some of the drawbacks identified in this heuristic approach to the construction of risk factors. We have focused on the potential misevaluation of the size and book-to-market effect implicit in the way the Fama and French methodology was constructed (Cremers et al., 2010; Huij and Verbeek, 2009; Brooks et al., 2008). The original Fama and French method performs a 2x3 sort of US stocks on market capitalization and book-to-market, forming six two-dimensional portfolios at the intersections of the two independent rankings. The premiums are defined as the spread between the average low- and high-scoring portfolios. Our main argument motivating the modifications to the original Fama and French method is that the independent sorting procedure underlying the formation of the six Fama and French two-dimensional portfolios distorts the way stocks are ranked into portfolios by placing disproportionate weights between the portfolios.

Following the methodology of Lambert and Hübner (2013), we applied a generalized Fama and French technique to infer the size, book-to-market and momentum factors from the US stock market over the sample period of 1980-2007. The main innovations of our premiums reside in a monthly rebalancing of the portfolios (underlying the construction of the risk premiums) in order to capture the time-varying dimensions of risk. Our model also featured a finer size classification and a conditional sorting of stocks into portfolios. We considered three risk dimensions. The conditional sorting procedure addresses the question of whether return variation related to the third risk criterion still exists even after having controlled for two other risk dimensions. The sorting procedure involves performing a sequential sort in three stages: the first two sorts were performed on control risks, followed by the risk dimension to be priced.

Echoing the findings of Cremers, Petajusto and Zitzewitz (2010) and Huij and Verbeek (2009), our paper demonstrates that the book-to-market premium of Fama and French is overvalued. We performed several asset pricing tests to check the validity and pricing power of our alternative premium specification. Compared to the Fama and French method, our factor construction method more accurately captures the return spread associated with the source of risk to be priced. It maximizes the dispersion in the related source of risk while keeping minimal dispersion in correlated sources of risk. The conditional sorting and the finer size classification both contribute to better balance the weightings placed on the small/large value/growth portfolios. The most significant improvement engendered in the new method lies in the reduction of specification errors when pricing passive benchmark investment portfolios. Additionally, the modified technique is, without loss of significance, neater and leads to risk premiums that may not necessarily be used jointly in a regression-based model. This contrasts with the original Fama and French factors, whose risk exposures are highly sensitive to the inclusion of the other Fama and French risk factors in the regression.

Our paper generally supports Lambert and Hübner's (2013) previous evidence that a sequential sorting procedure could be more appropriate to take into consideration the contamination effects between the premiums. We show that the premiums constructed in this way deliver more consistent risk properties while reaching at least the same specification level as the Fama and French premiums. Given the critical stance of our paper, it has been necessary to explore in depth the origins of the improvements of the proposed sequential procedure associated with various methodological variations, over the original Fama and French method. The robustness checks deliver clear insights with regards to the key drivers of the alternative approach’s pricing performance. As predicted, the replacement of an independent sort by a sequential one seems to make the most significant difference.

Appendix A

Robustness checks

Appendix A checks the robustness of our empirical results. First, we analyzed the specification of the sequential premiums for different variations from the original Fama and French setup in an isolated way. Second, we tested the superiority of our sequential technique over the value-weighting of the sorted portfolios proposed by Cremers et al. (2008). Finally, we tested whether the approach is still valid when reduced to a two-dimensional framework. This approach simplifies the construction of the factors, especially in a small market with a considerably lower number of stocks and where a 27 portfolio grid results in very low diversified portfolios.

1. Methodological choices under the modified Fama and French approach: marginal analysis

The impacts of the following variations with respect to the original Fama and French setup were separately tested:

- the sequential approach relies on a finer firm size classification than the Fama and French model: US stocks are split into groups of small or big stocks in the Fama and French framework, while our modified approach considers small, medium and big caps;

- the sequential approach defines the sorting breakpoints over the whole sample and not only with the NYSE data as in Fama and French (1993);

- the sequential approach consists in sorting stocks according to the interest pricing variable conditionally on the levels of their control variable(s). The original Fama and French approach, in contrast, sorts stocks independently according to size, momentum and book-to-market dimensions;

- finally, the sequential approach rebalances sorting portfolios monthly whereas the original Fama and French method relies on an annual rebalancing.

We considered the sequential sorting and the finer size classification jointly with the use of whole-sample breakpoints. Indeed, the sequential sorting of stocks into portfolios does not ensure a balanced repartition between the portfolios unless whole-sample breakpoints are used. Besides, whole-sample breakpoints alone do not ensure that the large capitalization group contains only big stocks; the joint consideration of finer and whole-sample breakpoints, however, contributes to balance the portfolios. The test of the whole-sample breakpoints was carried out when considering the finer size classification.

The following tests were performed:

- (T.1) test of the sequential sort over an independent sort: we constructed new size, book-to-market, and momentum premiums based on the following methodological choices: an independent sorting, a finer size classification, a three-dimensional sort, whole-sample breakpoints and a monthly rebalancing. We compared the efficiency of this model (M3) over our initial sequential model (M2), whose premiums depend on the following methodological choices: sequential sorting, a finer size classification, a three-dimensional sort, whole-sample breakpoints, and a monthly rebalancing.

- (T.2) test of finer size classification (small, mid and big caps) and whole-sample breakpoints versus NYSE breakpoints with two-dimensional size classification: we constructed new size and book-to-market premiums based on the following methodological choices: independent sorting, a finer size classification, a two-dimensional sort, whole-sample breakpoints and annual rebalancing. We tested a new model comprising the new size (SMB_{M4}) and book-to-market premiums (HML_{M4}) and the Fama and French momentum factor. We compare the efficiency of this model (M4) over the Fama and French model (M1), whose premiums depend on the following methodological choices: independent sorting, Small/Big size classification, a two-dimensional sort, NYSE breakpoints and annual rebalancing.

- (T.3) test of monthly rebalancing against annual rebalancing: we constructed new size and book-to-market premiums based on the following methodological choices: independent sorting, a finer size classification, a two-dimensional sort, whole-sample breakpoints and monthly rebalancing. We tested a new model made of the new size (SMB_{M5}) and book-to-market premiums (HML_{M5}) and the Fama and French momentum factor. We compared the efficiency of this model (M5) over the M4 model, whose premiums depend on the following methodological choices: independent sorting, finer size classification, 2-dimensional sort, NYSE breakpoints and annual rebalancing.

Table A.1 provides descriptive statistics for the SMB_{M3}, HML_{M3}, $\mathrm{UMD}_{\mathrm{M3}}$, SMB_{M4}, HML_{M4}, SMB_{M5}, and HML_{M5}. Tables A.2 and A.3 display the correlation matrix for these premiums.[xv]

< Insert Tables A.1 to A.3 >

The three size premiums displayed descriptive statistics similar to our sequential premium. The only difference resides in a higher average return and/or a higher level of positive skewness. As in the Fama and French original framework, all our *HML* premiums displayed positive return. Because we considered different methodological choices and because the different versions of the premium stayed positive over the period, we are confident in relating the negative average return of the modified premium to our sequential procedure. Except for HML_{M3}, the book-to-market premiums presented strong correlations with the other *HML* premiums, including the Fama and French original and sequential premiums. Replacing the Fama and French independent procedure with an independent sorting coupled with whole-sample breakpoints do not sufficiently take into account the cross-size effects as shown from the high levels of correlation between the premium and the size factor.

1.1. Test of a sequential sorting

To test the relevance of implementing a sequential approach, we re-ran the analysis conducted in Section 3 for M3. We compared the results obtained to those yielded by the sequential approach (M2).

We first tested whether the M3 model was able to price passive portfolios or indexes such as the Russell without specification error. We interpreted these results by comparing them to Table 8 on the Fama and French modified model.

< Insert Table A.4 >

The analysis did not allow us to discriminate between the sequential and the independent approaches. The levels of specification errors on passive indexes produced by a four-factor Carhart model using M3 premiums is very close to those arising from the sequential approach. For the M3 model, six out of 18 indices presented significant alphas in both specifications, while the modified Fama and French model delivered only four. Moreover, a Gibbons et al. (1989) test rejected the null hypothesis that M3 could price passive indexes with no specification error at the usual significance level (F_{stat} = 6.97). We did not consider the Gibbons et al efficiency test with regard to the F&F 2x3 portfolios sorted on size and book-to-market, as it gives a clear advantage to the original Fama and French premiums.

Table A.5 reports a direct comparison of both sets of premiums on individual stocks and also on the set of 2x3 F&F portfolios sorted on size and book-to-market.

< Insert Table A.5 >

The analysis conducted on individual assets demonstrates the superiority of the sequential approach over the independent approach.

1.2. Test of finer size classification and whole-sample breakpoints

We the proceeded to test the finer size classification and the whole-sample breakpoints. To conduct this test, we compared the results of the specification error tests and of the non-nested models for M4 to our previous results regarding the original Fama and French model (M1). Note that because only SMB_{M4} and HML_{M4} were recomputed for the purposes of this model, the Fama and French momentum factor is also included in the model.

Table A.6 shows our analysis of the levels of specification errors produced by M4 on a set of passive indexes.

< Insert Table A.6 >

M4 produced significant levels of specification errors for eight indices out of 18. The Fama and French model, however, displayed significant level of specification errors for 16 out of 18 indices. This indicates the superiority of the finer size classification and the whole-sample breakpoints. Additionally, the Gibbons et al. (1989) test could not reject the null hypothesis that M4 could price passive indexes with no specification error at the usual significance level ($F_{stat} = 0.02$).

Table A.7 tests the superior specification of M4 and M1 in order to test the superiority of a finer size classification with whole-sample breakpoints over a Small/Big size classification with NYSE breakpoints.

< Insert Table A.7 >

The table shows the superiority of the finer size classification and whole-sample breakpoints over the Fama and French methodological choices. The results on the portfolios are contradictory to the individual assets analysis. However, we argue that the test conducted on portfolios is misspecified because of the strong factor structure of the characteristic-sorted portfolios on size and book-to-market. There is a bias in the direction of the acceptance of the Fama and French model.

1.3. Test of monthly rebalancing

Table A.8 shows our test of the specification of M5 on passive indexes.

< Insert Table A.8 >

Compared to M4, M5 delivered higher specification errors. Therefore, the analysis does not confirm the superiority of the monthly rebalancing. However, a Gibbons et al. (1989) test could not reject the null hypothesis that M5 could price passive indexes with no specification error at the usual significance level (F_{stat} = 0.019).

Table A.9 outlines the non-nested econometric analyses on M5 and M4.

< Insert Table A.9 >

The non-nested analysis does demonstrate the superiority of the monthly rebalancing over the annual rebalancing. The analysis conducted on the portfolios, however, did not prove conclusive.

1.4. Summary

Our results support the relevance of both a sequential approach and of a finer size classification together with the use of whole-sample breakpoints. The monthly rebalancing appears to be only slightly responsible for the improved specification of the modified Fama and French approach.

Finer size classification together with whole-sample breakpoints alone improves the Fama and French specification. The addition of a sequential approach serves to virtually eliminate the cross-effects between empirical factors by ensuring the same number of stocks per portfolio when constructing the model factors. For this reason, we argue that the sequential approach provides the largest improvement over the Fama and French approach.

2. The modified Fama and French model versus the Cremers et al. (2010) model

The Fama-French independent sort yielded disproportionate weights between portfolios. Cremers et al. (2010) propose to value-weight the Fama and French 2x3 portfolios when forming the premiums; they propose the replacement of the "independent/equally-weighted" methodological choice of Fama and French with an "independent/value-weighted" portfolio sort. Our modified Fama and French premiums deal with this issue by replacing the independent sorting with a sequential sorting.

For this reason it has been necessary to test the superiority of a sequential/whole-sample breakpoints/equally-weighted portfolios approach (as in the modified Fama and French) over an independent/value-weighting of stocks into portfolios as proposed by Cremers et al. (2010), i.e. M6 (also with whole-sample breakpoints).

Descriptive statistics over SMB_{M6} and HML_{M6} are displayed at Table A.10.

< Insert Table A.10 >

Interestingly, when value-weighted, the average return on the *HML* independent premiums became negative, as for the sequential definition of the premium. In addition, premium HML_{M3} is defined on an independent sorting with equally-weighted portfolios, whereas HML_{M6} is based on an independent sorting with value-weighted portfolios. Contrarily to HML_{M3}, HML_{M6} displayed strong levels of correlation with the original (70%) and sequential (82%) Fama and French *HML* premiums[xvi]. This supports the evidence supplied by Cremers et al. that value weighting of the portfolio could catch part of the cross-size effects.

First, the specification errors on model M6 with passive indices were tested.

< Insert Table A.11 >

Table A.11 indicates that M6 produced less specification error than the Fama and Frenchmodel. This result is consistent with the evidence provided by Cremers et al. (2010). Nevertheless, the sequential model still delivered the least specification errors when pricing passive indexes. Additionally, a Gibbons et al. (1989) test rejected the null hypothesis that M6 could price passive indexes with no specification error at the usual significance level (F_{stat} = 229.11).

Table A.12 details the econometric comparative tests between M6 and M2 (the modified F&F model).

< Insert Table A.12 >

Non-nested econometric model analysis supported our main finding that a sequential methodological choice outperforms an independent-value-weighting choice for dealing with disproportion among portfolios. The analysis performed on portfolios revealed an inability to disentangle the relative added value of both models.

3. A two-dimensional approach

Our last robustness check tested whether the cubic sequential approach could be reduced to a two-dimensional approach. In order to perform this test, we constructed new size and book-to-market factors based on the following methodological choices: sequential sorting, a finer size classification, a two-dimensional sort, whole-sample breakpoints and monthly rebalancing. We compared the efficiency of a four-factor Carhart model using these size and book-to-market premiums and the modified Fama and French momentum (M7) over the modified Fama and French model (M2), whose premiums depend on the following methodological choices: sequential sorting, finer size classification, 3-dimensional sort, whole-sample breakpoints and monthly rebalancing.

Table A.13 shows descriptive statistics over the SMB_{M7} and HML_{M7} premiums issued from model M7.

< Insert Table A.13 >

SMB_{M7} displayed very similar descriptive statistics to the modified *SMB* premium, except for a higher skewness and a higher average return. The premium displayed very high correlation with the Fama and French modified *SMB* premium (91%) and the original Fama and French *SMB* premium (75%). Contrary to the modified Fama and French *HML* premium, HML_{M7} presented a positive average return over the period. The premium also displayed moderate levels of correlation with the modified (55%) and the original (60%) Fama and French premiums.

Table A.14 details a four-factor Carhart M7 model performed using a set of passive indexes.

< Insert Table A.14 >

The model delivered higher level of significant specification errors than a sequential portfolio. However, M7 still provided less significant alpha on passive indexes than the Fama and French version of the four-factor Carhart model. A Gibbons et al. (1989) test could not reject the null hypothesis that M7 could price passive indexes with no specification error at the usual significance level (F_{stat} = 0.021).

Table A.15 displays the results of non-nested econometric tests for the superiority of the M7 model over the sequential model (M2).

< Insert Table A.15 >

From the test on individual assets, the 2-dimensional approach appeared to outperform the sequential 4-factor Carhart model. At the 5%-significance level, the 2-dimensional sequential premiums correctly priced 24.5% of the individual securities, while rejecting the modified Fama and French 4-factor Carhart model. The sequential model, however, only priced 17.7% at the same significance level. From the portfolio analysis, both models were rejected for most portfolios. Our analysis suggests that the sequential approach could easily be translated to a two-dimensional model sharing the same set of assumptions.

References

Al-Muraikhi, H., Moosa, I. A., 2008, "The role of technicians and fundamentalists in emerging financial markets: A case study of Kuwait", *International Research Journal of Finance and Economics,* 13, 77-83.

Banz, R. W., 1981, "The relationship between return and market value of common stocks", *Journal of Financial Economics*, 9 (1), 3-18.

Basu, S., 1983, "The Relationship between earnings yield, market value and return for NYSE common stocks: Further evidence", *Journal of Financial Economics* 12,129-156.

Bernanke, B., Bohn, H., Reiss, P. C., 1986, "Alternative nonnested specification tests of time-series investment models", *Working Paper*, NBER Research Paper 890.

Brooks, C., Li, ., Miffre, J., 2008, "The value premium and time-varying idiosyncratic risk", *Working Paper*, ICMA Centre-University of Reading, EDHEC Business School.

Carhart, M. M., 1997, "On persistence in mutual fund performance", *Journal of Finance*, 52(1), 57-82.

Cremers, M., Petajusto, A., Zitzewitz, E., 2010, "Should benchmark indices have alpha? Revisiting performance evaluation", *Working Paper*, Yale School of Management, AFA 2010 Atlanta Meetings Paper.

Davidson, R., MacKinnon, J. G., 1981, "Several tests for model specification in the presence of alternative hypotheses", *Econometrica*, 49, 781-793.

Davidson, R., MacKinnon, J. G., 1984, Model specification tests based on artificial linear regressions, *International Economic Review,* 25, 485-502.

Elyasiani, E., Nasseh, A., 1994, "The appropriate scale variable in the U.S money demand: An application of nonnested tests of consumption versus income measures", *Journal of Business and Economic Statistics*, 12(1), 47-55.

Elyasiani, E., Nasseh, A., 2000, "Nonnested procedures in econometric tests of asset pricing theories", *Journal of Financial Research*, 23(1), 103-128.

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 on stocks and bonds", *Journal of Financial Economics*, 33(1), 3-56.

Fama, E. F., French, K. R., 1998, "Value versus growth: The international evidence", *Journal of Finance*, 53, 1975-1999.

Gibbons, M. R., Ross, S. A., Shanken, J., 1989, "A test of the efficiency of a given portfolio", *Econometrica*, 57(5), 1121-1152.

Griffin, J. M., Lemmon, M. L., 2002, "Book-to-market equity, distress risk, and stock returns", *Journal of Finance*, 57(5), 2317-2336.

Huij, J., Verbeek, M., 2009, "On the use of multifactor models to evaluate mutual fund performance", *Financial Management*, 38(1), 75-102.

Jegadeesh, N., Titman, S., 1993, "Returns to buying winners and selling losers: Implications for stock market efficiency", *Journal of Finance*, 48(1), 65-91.

Jegadeesh, N., Titman, S., 2001, "Profitability of momentum strategies: An evaluation of alternative explanations", *Journal of Finance*, 56(2), 699-720.

Lambert, M. and Hübner, G., 2013. "Comoment risk and stock returns", *Journal of Empirical Finance*, 23, 191-205.

MacKinnon, J. G., 1983, "Model specification tests against non-nested alternatives", *Econometric Reviews*, 2, 85-110.

Mahajan, A., Tartaroglu, S., 2008, "Equity market-timing and capital structure: International evidence", *Journal of Banking and Finance*, 32(5), 754-766.

Mohanram, P. S., 2004, "Separating winners from losers among low book-to-market stocks using financial statement analysis", *Working Paper*, Columbia Business School.

Michou, M., 2007, "Value investing: The use of financial statements to separate winners from losers", *Working Paper*, University of Edinburgh, Management School and Economics.

Rosenberg, B., Reid, K. and Lanstein, R., 1985, "Persuasive evidence of market inefficiency", *Journal of Portfolio Management*, 11, 9-17.

Endnotes

---

[i] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[ii] The one-year momentum anomaly for month *t* is defined as the trailing eleven-month returns lagged one month (*t*-11 to *t*-1), as previously outlined in Jegadeesh and Titman (2001) and Carhart (1997), Stocks that do not have a price at the end of month *t*-12 were not considered for that period.

[iii] We allowed a variation of up to one standard deviation around the US average BTM.

[iv] As in the Fama and French analysis, temporary unavailability of data excludes the stock from the analysis at that time.

[v] The company total debt at year *y* (*D)* concerns all interest bearing and capitalized lease obligations (long and short term debt) at the end of the year. The company total asset at year *y* (*A)* is the sum of current and long term assets owned by the company for that year. These variables have been collected on *Compustat*.

[vi] We designate by market value at month *t*, the quoted share price multiplied by the number of ordinary shares of common stock outstanding at that moment. As in Fama and French (1993), negative or zero book values that result from particular cases of persistently negative earnings are excluded from the analysis.

[vii] The indexes display an average monthly return of 0.9403% and 0.8784% for respectively the Growth and the Value Index over the period May 1980-April 2007.

[viii] Note that all correlations are significantly different from 1 at the usual significance levels.

[ix] Please refer to the appendix for robustness checks conducted on the updated Fama and French methodology.

[x] Performing the Gibbons, Ross and Shanken (1989) test on the set of passive indexes, we significantly reject, however, the joint efficiency of both models at the usual significance levels.

[xi] Bernanke, Bohn and Reiss (1986) and Elyasiani and Nasseh (1994), among others, use non-nested models to compare some model specifications regarding investment and US money demand, respectively. Elyasiani and Nasseh (2000) differentiate between the performance of the Capital Asset Pricing Model (CAPM) and of the consumption CAPM through non-nested econometric procedures. Al-Muraikhi and Moosa (2008) tested the impact of the actions of traders who act on the basis of fundamental or of technical analysis on financial prices based on non-nested models.

[xii] Note that the rejection of H_0 does not reveal anything about the validity of H_0'.

[xiii] The analysis performed on the characteristic portfolios shows that either both models are rejected, or the Fama and French model is accepted while the modified Fama and French model is rejected. It should be pointed out, however, that the characteristic portfolios possess a strong factor structure and that because they are based on a similar sorting to that underlying the formation of the Fama and French premiums, the test could be biased in the direction of the Fama and French acceptance.

[xiv] With regard to the data available on the Fama and French website from May 1980-April 2007.

[xv] Note that the sample has been reduced to July 1980 to April 2007 due to 2 missing data for the M4 premiums.

[xvi] Correlation matrix for SMB_{M6} and HML_{M6} with other empirical factors is available upon requests.

Table 1

Descriptive statistics over the empirical risk premiums and its components (January 1980-December 2007)[a]

	Panel A: F&F premiums		**Panel B: 2x3 sorted portfolios**					
	SMB	*HML*	*Low/Low*	*Low/Med*	*Low/High*	*Big/Low*	*Big/Med*	*Big/High*
Mean (%)	0.108	0.378	0.858	1.406	1.504	1.088	1.159	1.197
Std. Dev. (%)	3.225	3.129	6.862	4.876	4.682	4.745	4.232	4.093
t-stat	0.614	2.21**	2.29**	5.29***	5.89***	4.202***	5.022***	5.361***
# Obs.	336	336	336	336	336	336	336	336

[a] Table 1 displays descriptive statistics for F&F size (*SMB*) and book-to-market (*HML*) premiums as well as of their 2x3 characteristic-sorted portfolios (sorted on size and book-to-market) over January 1980-December 2007: time-series mean, standard deviation, t-stat for the bilateral test of time series mean equals to 0 as well as the number of observations considered are displayed. Low/High stand for a portfolio made of stocks ranked low based on market capitalization and high on the book-to-market ratios. *,** ,and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively.

Table 2

Correlation matrix of the 2x3 characteristics portfolios (January 1980- December 2007)[b]

	Low/Low	*Low/Med*	*Low/High*	*Big/Low*	*Big/Med*	*Big/High*
Low/Low	100%					
Low/Med	92.86%	100%				
Low/High	86.30%	96.66%	100%			
Big/Low	80.35%	77.22%	71.56%	100%		
Big/Med	67.78%	77.60%	76.30%	84.97%	100%	
Big/High	59.75%	72.70%	75.90%	75.32%	90.19%	100%

[b] Table 2 reports the paired correlations (in %) among the 2x3 characteristic-sorted portfolios on size and book-to-market. Low/High stand for a portfolio made of stocks ranked low based on market capitalization and high on the book-to-market ratios.

Table 3

Descriptive statistics over the simulated empirical risk premiums[c]

	Size premiums			**Value premiums**		
	S_SMB	*S_SMB*^{C}	*S_SMB*^{C}-*S_SMB*	*S_HML*	*S_HML*^{C}	*S_HML*^{C}-*S_HML*
Mean	0.117	0.074	-0.0441	0.336	0.380	0.0439
Std. Dev.	0.120	0.128	0.178	0.165	0.161	0.222
T-stat	9.75***	5.744***	-2.47**	20.35***	23.60	1.98**
t-stat*	0.77	-2.70***	-2.47**	-2.55**	0.10	1.98**
# Obs.	100	100	100	100	100	100

[c] Table 3 displays descriptive statistics for the simulated size and book-to-market premiums under two scenarios: with correlations among the rankings (*S_SMB*^{C} and *S_HML*^{C}) and without correlation (*S_SMB* and *S_HML). The statistics describe the simulated distribution of the average S_SMB/S_SMB*^{C} and S_HML/*S_HML*^{C} . The difference between the two simulated series is also described. The table displays the time-series mean, standard deviation, t-stat for the bilateral test for the simulated premiums mean being equal to 0, as well as the t-stat* for a bilateral test for the simulated premiums mean being equal to respectively the mean of the SMB and HML premiums. *, ** ,and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively.

Table 4

Descriptive statistics over the empirical risk premiums (May 1980-April 2007)[d]

	Panel A: F&F premiums			Panel B: Sequential premiums		
	SMB_{ff}	HML_{ff}	UMD_{ff}	*SMB'*	*HML'*	*UMD'*
Mean	0.14%	0.44%	0.79%	0.88%	-0.07%	0.91%
Median	-0.06%	0.38%	0.90%	0.84%	0.01%	0.92%
Maximum	21.96%	13.85%	18.39%	12.88%	19.15%	10.65%
Minimum	-16.79%	-12.40%	-25.06%	-11.71%	-14.16%	-11.26%
Std. Dev.	3.24%	3.16%	4.26%	3.12%	3.23%	2.71%
Skewness	0.76	0.07	-0.56	0.08	0.24	-0.25
Kurtosis	11.47	5.34	9.06	5.18	8.34	5.56
Jarque-Bera	999***	74.5***	512***	64.4***	388***	91.9***
t-stat	0.83	2.13**	3.57***	4.85***	-0.35	6.21***
# Obs.	324	324	324	324	324	324

[d] Table 4 displays descriptive statistics for size (*SMB*), book-to-market (*HML*), and momentum (*UMD*) premiums over the period ranging from May 1980 to April 2007. T-tests of the significance of the different time-series are conducted. The values of the t-stats have been corrected for the presence of autocorrelation in the time-series. Panel A presents the statistics for the empirical risk premiums of F&F, while Panel B presents the statistics for the updated (sequential) premiums built along our alternative methodology. *,** ,and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively.

Table 5

Descriptive statistics over the return spread portfolios forming each sequential risk factor[e]

Panel A: 9 size spread portfolios										
	LLL-LLH	LML-LMH	LHL-LHH	MLL-MLH	MML-MMH	MHL-MHH	HLL-HLH	HML-HMH	HHL-HHH	Average / (σ)
Mean (%)	0.97	0.73	0.99	0.80	0.88	1.25	1.09	0.33	0.87	0.88/ (0.26)
Median (%)	0.62	0.57	1.19	1.05	0.72	0.87	1.17	0.22	0.81	0.80/ (0.31)
Min (%)	18.01	14.69	12.74	17.16	14.21	13.55	19.78	14.88	16.84	15.76/ (2.31)
Max (%)	-15.06	-12.82	-11.71	-24.87	-11.92	-14.12	-18.25	-13.84	-15.65	-15.36/ (4.10)
S. D. (%)	4.84	3.84	3.74	4.98	3.62	3.68	5.03	3.35	3.61	4.08 / (0.67)
Skewness	0.1687	0.2333	-0.0354	-0.5078	0.0733	0.0401	-0.1984	0.1209	0.0936	
Kurtosis	3.7854	4.4247	3.6944	6.7823	4.4983	3.9962	4.5764	5.2606	5.6633	
Jarque-Bera	9.86***	3.03***	6.58**	2.07***	3.06***	1.35***	3.57***	6.98***	9.62***	
$\rho_{SMB',spread_i}$	77.15	82.38	60.20	83.88	84.29	76.97	78.49	65.96	75.59	76.10/ (8.14)
$\rho_{HML',spread_i}$	-5.60	-10.23	-4.01	-24.65	-20.37	-25.80	-17.96	6.81	1.40	-11.16/ (11.68)
$\rho_{UMD',spread_i}$	13.04	2.27	-16.65	9.94	1.93	-7.69	9.66	5.65	-9.17	1.00/ (10.09)

Panel B: 9 book-to-market spread portfolios										
	LLH-LLL	LMH-LML	LHH-LHL	MLH-MLL	MMH-MML	MHH-MHL	HLH-HLL	HMH-HML	HHH-HHL	Average / (σ)
Mean (%)	0.54	-0.54	0.27	-0.29	0.60	0.20	0.04	0.12	-1.63	-0.08/ (0.69)
Median (%)	0.22	-0.22	0.22	-0.33	0.83	0.19	0.19	0.29	-0.62	0.09/ (0.42)
Min (%)	-13.45	-9.65	-11.75	-17.14	-9.54	-19.68	-11.05	-125.79	-21.07	-26.57/ (37.45)
Max (%)	19.76	10.95	14.52	22.37	9.07	22.83	12.62	122.18	18.21	28.06/ (35.64)
S. D. (%)	5.14	3.71	4.12	4.22	2.94	4.49	3.39	17.98	3.90	5.54/ (4.71)
Skewness	0.1472	0.0094	0.3301	0.2708	-0.1031	0.0283	-0.0933	-0.4665	-0.0846	
Kurtosis	3.7096	2.9000	3.8073	5.8770	3.5258	6.4773	4.4572	21.6317	7.2400	
Jarque-Bera	7.97***	0.14	14.68***	115.70***	4.31***	163.28***	29.14***	4698.14***	243.09***	
$\rho_{SMB',spread_i}$	-21.67	-24.44	-19.21	-21.64	-24.18	-26.01	-10.45	6.66	-27.11	-18.67/(10.70)
$\rho_{HML',spread_i}$	39.57	51.48	56.32	57.27	59.12	61.61	40.28	76.29	50.99	54.77/ (11.20)
$\rho_{UMD',spread_i}$	-6.72	-0.74	-6.75	-9.70	0.15	-6.86	8.29	-6.03	-16.13	-4.94/ (6.87)

Table 5 (continued)

Panel C: 9 momentum spread portfolios										
	LLH-LLL	LMH-LML	LHH-LHL	MLH-MLL	MMH-MML	MHH-MHL	HLH-HLL	HMH-HML	HHH-HHL	Average / (σ)
Mean (%)	1.27	1.14	0.83	0.90	0.82	0.38	0.61	1.95	0.29	0.91/ (0.50)
Median (%)	1.52	1.27	0.96	0.56	1.00	0.35	0.76	2.37	0.48	1.03/ (0.63)
Min (%)	-19.73	-15.39	-12.49	-15.91	-14.09	-14.15	-20.26	-31.33	-15.03	-17.60/ (5.75)
Max (%)	23.32	16.62	13.92	19.41	14.53	12.04	18.04	13.51	10.02	15.71/ (4.08)
S. D. (%)	4.88	4.32	3.98	4.55	3.44	3.43	5.20	4.20	3.62	
Skewness	0.2780	-0.0848	-0.1026	0.4531	-0.6048	-0.0594	-0.2038	-1.7190	-0.3773	
Kurtosis	5.7598	4.4316	3.6097	5.7932	5.9957	4.3471	4.7966	15.0953	4.4072	
Jarque-Bera	106.99***	28.05***	5.59***	116.41***	140.90***	24.69***	45.82***	2134.57***	34.42***	
$\rho_{SMB',spread_i}$	-5.03	6.68	15.07	-19.65	2.52	17.68	-24.26	22.84	13.37	3.24/ (16.57)
$\rho_{HML',spread_i}$	11.16	-14.02	-27.47	22.11	7.06	-15.46	25.21	-28.65	-12.34	-3.60/ (20.45)
$\rho_{UMD',spread_i}$	71.86	79.82	64.07	70.05	78.44	62.13	59.75	30.14	69.62	65.10 (14.77)

[e] Table 5 displays descriptive statistics for the 9 return spreads forming the *SMB'*, *HML'*, and *UMD'* factors. The correlations (in %) of each spread portfolio with the *SMB', HML',* and *UMD'* factors are reported. The last column reports the average and the standard deviation of the statistics for the different portfolios. The size (resp. book-to-market, resp. momentum) spread portfolios are formed by performing a 3-stage sequential sorting procedure on, successively, book-to-market, momentum and market capitalization (resp. market capitalization, momentum, and finally book-to-market; resp. book-to-market, market capitalization, and momentum). Each spread portfolio is defined from a difference between two portfolios defined by 3 letters describing the 3-stage sequential sorting procedure. *L* stands for a low scoring portfolio, *M* for a medium scoring portfolio, and *H* for a high scoring portfolio. *S.D. = Standard Deviation*. The row corresponding to the dimension sought after by the spread portfolios is grayed.

Table 6

Descriptive statistics over the return spread portfolios forming each F&F risk factor[f]

	Mean (%)	Median (%)	Min (%)	Max (%)	S.D. (%)	Skewness	Kurtosis	J-B	$\rho_{SMB_{ff},spread_i}$	$\rho_{HML_{ff},spread_i}$	$\rho_{UMD_{ff},spread_i}$
Panel A: Size											
Spread 1. Low BTM	-0.2	-0.33	27.75	-22.54	4.17	0.6418	11.0458	896.17	93.89	-42.08	4.38
Spread 2. Mid BTM	0.29	0.14	19.94	-14.3	3.11	0.7586	10.0509	702.23	94.36	-32.91	12.19
Spread 3. High BTM	0.33	0.25	18.29	-13.71	3.12	0.5545	8.0572	361.87	91.7	-37.94	15.13
Average / (σ)	0.14	0.02	21.99	-16.85	3.47				93.32	-37.64	10.57
	(0.30)	(0.31)	(5.05)	(4.94)	(0.61)				(1.42)	(4.59)	(5.56)
Panel B: Book-to-market											
Spread 1. Low size	0.7	0.57	13.53	-17.1	3.68	-0.206	5.5328	88.90***	-49.29	93.37	-7.17
Spread 2. High size	0.17	0.1	14.91	-10.39	3.16	0.324	4.8724	53.00***	-24.12	90.84	-17.54
Average / (σ)	0.44	0.34	14.22	-13.75	3.42				-36.71	92.11	-12.36
	(0.37)	(0.33)	(0.98)	(4.74)	(0.37)				(17.80)	(1.79)	(7.33)
Panel C: Momentum											
Spread 1. Low size	1.23	1.36	20.84	-26	4.2	-0.7975	12.0118	1130.72***	10.29	-10.48	94.13
Spread 2. High size	0.34	0.65	19.23	-24.08	4.79	-0.3104	6.1806	141.77***	10	-13.66	95.55
Average / (σ)	0.79	1.01	20.04	-25.04	4.50				10.15	-12.07	94.84
	(0.63)	(0.50)	(1.14)	(1.36)	(0.42)				(0.21)	(2.25)	(1.00)

[f] Table 6 displays descriptive statistics for the return spreads forming the SMB_{ff} (Panel A), HML_{ff} (Panel B), and UMD_{ff} (Panel C). The correlation (in %) of each spread portfolio with the SMB_{ff}, HML_{ff}, and UMD_{ff} factors are reported. The last line in each panel reports the average and the standard deviation of the statistics for the different portfolios. The size spread portfolios are formed from the return spreads between small and big caps for 3 levels of book-to-market. The book-to-market (resp. momentum) spread portfolios are formed from the return spreads between high and low levels of book-to-market (resp. momentum) for two levels of market capitalization. Each spread portfolio is defined from a difference between two portfolios formed at the intersection of a two-way sort of stocks on size and a three-way sort on book-to-market or on momentum. *S.D. = Standard Deviation. J-B = Jarque-Bera.* The column corresponding to the dimension sought after by the spread portfolios is grayed.

Table 7

Correlation matrix of the empirical risk premiums (May 1980-April 2007)[g]

	SMB'	*HML'*	*UMD'*	*SMB$_{ff}$*	*HML$_{ff}$*	*UMD$_{ff}$*
SMB'	1					
HML'	-15.50***	1				
UMD'	2.56	-3.19	1			
SMB$_{ff}$	67.16***	-34.46***	3.24	1		
HML$_{ff}$	-18.87***	68.25***	2.35	-40.83***	1	
UMD$_{ff}$	9.61*	-12.03**	82.63***	10.66*	-12.85**	1

[g] Table 7 reports the paired correlations (in %) among the modified (sequential) and the original F&F empirical risk premiums, as well as across these two sets of factors. Tests over the significance of the pair-wise correlations are performed: *,** ,and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively.

Table 8

Specification errors of passive investment indexes[h]

	Panel A			Panel B		
	4-Factor Carhart Model :F&F specification			4-Factor Carhart Model: F&F modified Specification		
	All	**Growth**	**Value**	**All**	**Growth**	**Value**
Russell 1000	-0.0010***	0.0002	-0.0021**	0.0000	-0.0015	0.0024
Russell 2000	-0.0043***	-0.0043***	-0.0040***	-0.0040	-0.0065*	-0.0002
Russell 3000	-0.0013***	-0.0002	-0.0023**	-0.0003	-0.0020	0.0021
S&P 500	-0.0031***	-0.0021*	-0.0047***	-0.0021**	-0.0037**	-0.0008
S&P Mid Cap	-0.0033*	-0.0034*	-0.0035**	-0.0009	-0.0039	0.0021
S&P Small Cap	-0.0060***	-0.0062***	-0.0065***	-0.0044	-0.0070**	-0.0025

[h] Table 8 performs a 4-factor Carhart analysis on a set of passive indexes: Russell 1000/2000/3000 and the S&P 500/MidCap/SmallCap: only the levels of specification errors (alphas) and their significance are displayed. For each index, the results for the composite, the growth and the value index are presented.

Table 9

Tests over the value of θ_1 and θ_2 in the Nested Models M1 and M2 [i]

Panel A – Individual assets

	H'_0 and H_1: Not reject M2 & Reject M1			H_0 and H'_1: Not reject M1 & Reject M2			H_1 and H'_1: Reject M1 & M2			H_0 and H'_0: Not reject M1 & M2		
signif	Test	#	%	Test	#	%	Test	#	%	Test	#	%
10%	$\|t'\| \le 1.65$ and $\|t\| > 1.65$	3786	34.15	$\|t\| \le 1.65$ and $\|t'\| > 1.65$	3061	27.61	$\|t\| > 1.65$ and $\|t'\| > 1.65$	1890	17.05	$\|t\| \le 1.65$ and $\|t'\| \le 1.65$	2350	21.20
5%	$\|t'\| \le 1.97$ and $\|t\| > 1.97$	3431	30.95	$t \le 1.97$ and $\|t'\| > 1.97$	2884	26.01	$\|t\| > 1.97$ and $\|t'\| > 1.97$	997	9.00	$\|t\| \le 1.97$ and $\|t'\| \le 1.97$	3775	34.05
1%	$\|t'\| \le 2.59$ and $\|t\| > 2.59$	2275	20.52	$\|t\| \le 2.59$ and $\|t'\| > 2.59$	2017	18.19	$\|t\| > 2.59$ and $\|t'\| > 2.59$	265	2.39	$\|t\| \le 2.59$ and $\|t'\| \le 2.59$	6530	58.90

Panel B – 2x3 F&F portfolios

	H'_0 and H_1: Not reject M2 & Reject M1			H_0 and H'_1: Not reject M1 & Reject M2			H_1 and H'_1: Reject M1 & M2			H_0 and H'_0: Not reject M1 & M2		
signif	Test	#	%	Test	#	%	Test	#	%	Test	#	%
10%	$\|t'\| \le 1.65$ and $\|t\| > 1.65$	0	0	$\|t\| \le 1.65$ and $\|t'\| > 1.65$	2	1/3	$\|t\| > 1.65$ and $\|t'\| > 1.65$	4	2/3	$\|t\| \le 1.65$ and $\|t'\| \le 1.65$	0	0
5%	$\|t'\| \le 1.97$ and $\|t\| > 1.97$	0	0	$t \le 1.97$ and $\|t'\| > 1.97$	3	1/2	$\|t\| > 1.97$ and $\|t'\| > 1.97$	3	1/2	$\|t\| \le 1.97$ and $\|t'\| \le 1.97$	0	0
1%	$\|t'\| \le 2.59$ and $\|t\| > 2.59$	0	0	$\|t\| \le 2.59$ and $\|t'\| > 2.59$	3	1/2	$\|t\| > 2.59$ and $\|t'\| > 2.59$	3	1/2	$\|t\| \le 2.59$ and $\|t'\| \le 2.59$	0	0

[i] Table 9 estimates the models M1 (testing the F&F model) and M2 (testing the modified F&F model) for the 11,377 individual assets (only 11,087 were available for the analysis) in Panel A and on the 2x3 F&F portfolios sorted on size and book-to-market on Panel B. It jointly tests the significance of θ_1 and θ_2 using Equation (10). The table reports, for different levels of significance, the number of assets (and the frequency) for which both models are "accepted", rejected, or accepted while the other one rejected.

Table A.1

Descriptive statistics over M3, M4 and M5 factors (July 1980-April 2007)[i]

	Panel A: M3			Panel B: M4		Panel C : M5	
	SMB_{M3}	HML_{M3}	UMD_{M3}	SMB_{M4}	HML_{M4}	SMB_{M5}	HML_{M5}
Mean	1.16%	1.27 %	0.67	0.87%	0.14%	1.14%	0.09%
Median	0.95%	0.95%	0.85	0.76%	0.001%	0.91%	0.07%
Maximum	14.34%	31.32%	10.26	13.12%	9.87%	15.49%	10.66%
Minimum	-10.47%	-16.13%	-14.81	-11.75%	-13.76%	-9.49%	-14.94%
Std. Dev.	3.49%	5.16%	3.15	3.30%	2.60%	3.69%	2.81%
Skewness	0.1719	1.6125	-0.6996	0.2224	-0.5962	0.419%	-0.4618
Kurtosis	4.3625	12.2657	6.0610	4.5063	7.2957	4.4347	7.6302
Jarque-Bera	26.4938***	1291.390***	151.97***	33.0970***	266.65***	37.0453***	99.09***
# Obs.	322	322	322	322	322	322	322

[a] Table A.1 displays descriptive statistics for SMB_{M3}, HML_{M3},UMD_{M2}, SMB_{M4}, HML_{M4}, SMB_{M5}, and HML_{M5}. over the period ranging from July 1980 to April 2007. *,** ,and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively

Table A.2

Correlation matrix for SMB_{M3}, SMB_{M4} and SMB_{M5} (July 1980-April 2007)[j]

	SMB_{M3}	SMB_{M4}	SMB_{M5}
SMB'	70.72	71.92	70.41
SMB_{ff}	93.48	88.26	90.60
SMB_{M3}	100.00	90.44	97.30
SMB_{M4}	90.44	100.00	89.37
SMB_{M5}	97.30	89.37	100.00

[b] Table A.2 reports the paired correlations (in %) among the F&F original (SMB_{ff}) and sequential (SMB') empirical risk premiums and SMB_{M3}, SMB_{M4} and SMB_{M5}.

Table A.3

Correlation matrix for HML_{M3}, HML_{M4} and HML_{M5} (July 1980-April 2007)[k]

	HML_{M3}	HML_{M4}	HML_{M5}
HML'	24.52	81.94	83.61
HML_{ff}	20.11	68.57	71.85
HML_{M3}	100.00	22.49	28.66
HML_{M4}	22.49	100.00	91.60
HML_{M5}	28.66	91.60	100.00

[c] Table A.3 reports the paired correlations (in %) among the F&F original (HML_{ff}) and sequential (HML') empirical risk premiums and HML_{M3}, HML_{M4} and HML_{M5}.

Table A.4

Specification errors of passive investment indexes for M3[l]

	Independent sorting/ finer size classification S/M/B 3-dimensional/ whole-sample breakpoints/monthly		
	All	**Growth**	**Value**
Russell 1000	0.0004	-0.0007	0.0021
Russell 2000	-0.0063**	-0.0092***	-0.0026
Russell 3000	-0.0017	-0.0015	0.0017
S&P 500	-0.0015*	-0.0011	-0.0023
S&P Mid Cap	-0.0020	-0.0042	-0.0001
S&P Small Cap	-0.0062**	-0.0053*	-0.0077**

[d] Table A.4 performs a 4-factor Carhart analysis using M3 premiums on a set of passive indexes: Russell 1000/2000/3000 and the S&P 500/MidCap/SmallCap: only the levels of specification errors (alphas) and their significance are displayed. For each index, the results for the composite, the growth and the value index are presented.

Table A.5

Tests over the value of θ_1 and θ_2 in the Nested Models M2 and M3 for individual and portfolio assets[m]

signif	H'_0 and H_1: Not reject M3 & Reject M2				H_0 and H'_1: Not reject M2 & Reject M3				H_1 and H'_1: Reject M2 & M3				H_0 and H'_0: Not reject M2 & M3			
	Portfolios		Individual assets		Portfolios		Individual assets		Portfolio s		Individual assets		Portfolios		Individual assets	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
10%	1	1/6	2749	24.79	0	0	3155	28.46	5	5/6	1165	10.51	0	0	4018	36.24
5%	1	1/6	2141	19.31	0	0	2584	23.31	5	5/6	673	6.07	0	0	5689	51.31
1%	2	1/3	1029	9.28	0	0	1532	13.82	4	2/3	219	1.98	0	0	8307	74.93

[e] Table A.5 estimates the models M2 and M3 for the 11,377 individual assets (only 11,087 were available for the analysis) and for the F&F 2x3 portfolios sorted on size and book-to-market. It jointly tests the significance of θ_1 and θ_2 using Equation (10). The table reports, for different levels of significance, the number of assets (and the frequency) for which both models are "accepted", rejected, or accepted while the other one rejected.
M3: independent sorting / finer size classification / three-dimensional sorting (3x3x3)/ whole-sample breakpoints / monthly rebalancing
M2: sequential sorting / finer size classification / three dimensional sorting (3x3x3)/ whole-sample breakpoints / monthly rebalancing

Table A.6

Specification errors of passive investment indexes for M4[n]

	Independent sorting/ finer size classification S/M/B 2-dimensional/ whole-sample breakpoints/annual		
	All	**Growth**	**Value**
Russell 1000	-0.0005	0.0002	-0.0012
Russell 2000	-0.0045*	-0.0055**	-0.0030
Russell 3000	-0.0009***	-0.0003	-0.0014
S&P 500	-0.0027***	-0.0020	-0.0039***
S&P Mid Cap	-0.0026	-0.0034	-0.0021
S&P Small Cap	-0.0058**	-0.0066**	-0.0058**

[f] Table A.6 performs a 4-factor Carhart analysis using M4 premiums on a set of passive indexes: Russell 1000/2000/3000 and the S&P 500/MidCap/SmallCap: only the levels of specification errors (alphas) and their significance are displayed. For each index, the results for the composite, the growth and the value index are presented.

Table A.7

Tests over the value of θ_1 and θ_2 in the Nested Models M4 and M1 for individual and portfolio assets[o]

signif	*H*'₀ and *H*₁: Not reject M4 & Reject M1				*H*₀ and *H*'₁: Not reject M4 & Reject M1				*H*₁ and *H*'₁: Reject M4 & M1				*H*₀ and *H*'₀: Not reject M4 & M1			
	Portfolios		Individual assets		Portfolios		Individual assets		Portfolios		Individual assets		Portfolios		Individual assets	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
10%	0	0	4185	37.75	3	1/2	2337	21.08	3	1/2	828	7.47	0	0	3737	33.71
5%	0	0	3502	31.59	4	2/3	1900	17.14	2	1/3	436	3.93	0	0	5249	47.34
1%	0	0	2160	19.48	4	2/3	1068	9.63	2	1/3	115	1.04	0	0	7744	69.98

[g] Table A.7 estimates the models M1 and M4 for the 11,377 individual assets (only 11,087 were available for the analysis) and for the 2x3 portfolios sorted in size and book-to-market of F&F. It jointly tests the significance of θ_1 and θ_2 using Equation (10). The table reports, for different levels of significance, the number of assets (and the frequency) for which both models are "accepted", rejected, or accepted while the other one rejected.

M4: independent sorting / finer size classification (S/M/B) / 2-dimensional sorting (3x3) / whole-sample breakpoints / annual rebalancing

M1: independent sorting / S/B size classification / 2-dimensional sorting (2x3) / NYSE breakpoints / annual rebalancing

Table A.8

Specification errors of passive investment indexes for M5[p]

	Independent sorting/ finer size classification S/M/B 2-dimensional/ whole-sample breakpoints/monthly		
	All	**Growth**	**Value**
Russell 1000	0.0000	0.0007	0.0005
Russell 2000	-0.0073***	-0.0084***	-0.0056**
Russell 3000	-0.0007*	-0.0001	-0.0010
S&P 500	-0.0021***	-0.0015	-0.0033***
S&P Mid Cap	-0.0034*	-0.0029	-0.0043*
S&P Small Cap	-0.0086***	-0.0092***	-0.0088***

[h] Table A.8 performs a 4-factor Carhart analysis using M5 premiums on a set of passive indexes: Russell 1000/2000/3000 and the S&P 500/MidCap/SmallCap: only the levels of specification errors (alphas) and their significance are displayed. For each index, the results for the composite, the growth and the value index are presented.

Table A.9

Tests over the value of θ_1 and θ_2 in the Nested Models M4 and M5 for individual assets[q]

signif	H'_0 and H_1: Not reject M5 & Reject M4				H_0 and H'_1: Not reject M4 & Reject M5				H_1 and H'_1: Reject M4 & M5				H_0 and H'_0: Not reject M6 & M7			
	Portfolios		Individual assets		Portfolios		Individual assets		Portfolios		Individual assets		Portfolios		Individual assets	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
10%	0	0	2302	20.76	1	1/6	1791	16.15	0	0	348	3.14	5	5/6	6646	59.94
5%	1	1/6	1672	15.08	1	1/6	1130	10.19	0	0	173	1.60	4	2/3	8112	73.17
1%	3	1/2	717	6.47	0	0	399	3.60	1	1/6	58	0.52	2	1/3	9913	89.41

[i] Table A.9 estimates the models M4 and M5 for the 11,377 individual assets (only 11,087 were available for the analysis) and for the F&F 2x3 portfolios sorted on size and book-to-market. It jointly tests the significance of θ_1 and θ_2 using Equation (10). The table reports, for different levels of significance, the number of assets (and the frequency) for which both models are "accepted", rejected, or accepted while the other one rejected.
M5: independent sorting / finer size classification (S/M/B) / 2-dimensional sorting (3x3) / NYSE breakpoints / monthly rebalancing
M4: independent sorting / finer size classification (S/M/B) / 2-dimensional sorting (3x3) / NYSE breakpoints / annual rebalancing

Table A.10

Descriptive statistics over M6 factors (July 1980-April 2007)[r]

	M6	
	SMB_{M6}	*HML_{M6}*
Mean	-0.09%	-0.001
Median	-0.1320%	0.001%
Maximum	1.96%	0.98%
Minimum	-1.05%	-1.11%
Std. Dev.	0.35%	0.25%
Skewness	0.8595	-0.4249
Kurtosis	6.7666	6.6790
Jarque-Bera	229.99***	191.29***
# Obs.	322	322

[j] Table A.10 displays descriptive statistics for SMB_{M6}, HML_{M6} over the period ranging from July 1980 to April 2007. *,** ,and *** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively.

Table A.11

Specification errors of passive investment indexes for M6[s]

Independent sorting/ finer size classification S/M/B 3-dimensional/ whole-sample breakpoints/monthly / VW			
	All	**Growth**	**Value**
Russell 1000	-0.0033***	-0.0037***	-0.0033**
Russell 2000	0.0054	0.0050	0.0050
Russell 3000	-0.0027***	-0.0032***	-0.0028**
S&P 500	-0.0066***	-0.0072***	-0.0066***
S&P Mid Cap	-0.0002	0.0004	-0.0014
S&P Small Cap	0.0019	0.0018	0.0014

[k] Table A.11 performs a 4-factor Carhart analysis using M6 on a set of passive indexes: Russell 1000/2000/3000 and the S&P 500/MidCap/SmallCap: only the levels of specification errors (alphas) and their significance are displayed. For each index, the results for the composite, the growth and the value index are presented.

Table A.12

Tests over the value of θ_1 and θ_2 in the Nested Models M2 and M6 for individual assets[1]

	H'$_0$ and *H*$_1$: Not reject M6 & Reject M2				*H*$_0$ and *H'*$_1$: Not reject M2 & Reject M6				*H*$_1$ and *H'*$_1$: Reject M2 & M6				*H*$_0$ and *H'*$_0$: Not reject M2 & M6			
signif	Portfolios		Individual assets		Portfolios		Individual assets		Portfolios		Individual assets		Portfolios		Individual assets	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
10%	1	1/6	2712	24.46	1	1/6	3819	34.45	4	2/3	2777	25.05	0	0	1779	16.05
5%	1	1/6	2576	23.32	1	1/6	3853	34.75	4	2/3	1673	15.09	0	0	2985	26.92
1%	1	1/6	1783	16.08	1	1/6	2997	27.03	4	2/3	508	4.58	0	0	5799	52.30

[1] Table A.12 estimates the models M2 and M6 for the 11,377 individual assets (only 11,087 were available for the analysis) and for the F&F 2x3 portfolios sorted on size and book-to-market. It jointly tests the significance of θ_1 and θ_2 using Equation (10). The table reports, for different levels of significance, the number of assets (and the frequency) for which both models are "accepted", rejected, or accepted while the other one rejected.
M6: independent sorting / VW / finer size classification (S/M/B) / 3-dimensional sorting (3x3x3) / whole-sample breakpoints / monthly rebalancing
M2: sequential sorting / EW / finer size classification (S/M/B) / 3-dimensional sorting (3x3x3) / whole-sample breakpoints / monthly rebalancing

Table A.13

Descriptive statistics over M7 factors (July 1980-April 2007)[u]

	M7	
	SMB_{M7}	HML_{M7}
Mean	1.00%	0.05%
Median	0.73%	-0.01%
Maximum	16.68%	7.53%
Minimum	-9.63%	-7.34%
Std. Dev.	3.53	2.12%
Skewness	0.5297	0.1292
Kurtosis	4.8801	4.2850
Jarque-Bera	62.87***	23.19***
# Obs.	322	322

[m] Table A.13 displays descriptive statistics for SMB_{M7}, HML_{M7} over the period ranging from July 1980 to April 2007. *,** ,and*** indicate statistical significance at the 0.1, 0.05 and 0.01 levels, respectively.

Table A.14

Specification errors of passive investment indexes for M7[v]

sequential sorting/ finer size classification S/M/B 2-dimensional/ whole-sample breakpoints/monthly			
	All	**Growth**	**Value**
Russell 1000	0.0006	-0.0019	0.0039*
Russell 2000	-0.0080***	-0.0115***	-0.0030
Russell 3000	-0.0001	-0.0028*	0.0033
S&P 500	-0.0012	0.0007	-0.0034**
S&P Mid Cap	-0.0020	-0.0051**	0.0011
S&P Small Cap	-0.0082***	-0.0111***	-0.0060*

[n] Table A.14 performs a 4-factor Carhart analysis using M7 premiums on a set of passive indexes: Russell 1000/2000/3000 and the S&P 500/MidCap/SmallCap: only the levels of specification errors (alphas) and their significance are displayed. For each index, the results for the composite, the growth and the value index are presented.

Table A.15

Tests over the value of θ_1 and θ_2 in the Nested Models M7 and M2 for individual assets[w]

signif	*H'*₀ and *H*₁: Not reject M7 & Reject M2				*H*₀ and *H'*₁: Not reject M2 & Reject M7				*H*₁ and *H'*₁: Reject M2 & M7				*H*₀ and *H'*₀: Not reject M2 & M7			
	Portfolios		Individual assets		Portfolios		Individual assets		Portfolios		Individual assets		Portfolios		Individual assets	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
10%	0	0	3351	30.22	0	0	2340	21.11	6	1	1275	11.50	0	0	4121	37.17
5%	0	0	2718	24.52	2	1/3	1959	17.67	4	2/3	680	6.13	0	0	5730	51.68
1%	0	0	1498	13.51	2	1/3	1141	10.29	4	2/3	170	1.53	0	0	8278	74.66

[o] Table A.15 estimates the models M7 and M2 for the 11,377 individual assets (only 11,087 were available for the analysis). It jointly tests the significance of θ_1 and θ_2 using Equation (10). The table reports, for different levels of significance, the number of assets (and the frequency) for which both models are "accepted", rejected, or accepted while the other one rejected.

M7: sequential sorting / finer size classification (S/M/B) / 2-dimensional sorting (3x3) / whole-sample breakpoints / monthly rebalancing

M2: sequential sorting / finer size classification (S/M/B) / 3-dimensional sorting (3x3x3)/ whole-sample breakpoints / monthly rebalancing