Enrica Bolognesi, Mattia Pividori

Fundamental Indexation in Europe: New Evidence

(doi: 10.12831/85432)

Journal of Financial Management, Markets and Institutions (ISSN 2282-717X)
Fascicolo 2, luglio-dicembre 2016

Ente di afferenza:
()

Fundamental Indexation in Europe: New Evidence

Journal of Financial Management Markets and Institutions

Enrica Bolognesi
University of Udine, Department of Economics and Statistics

Mattia Pividori
Eurizon Capital S.A.

Abstract

Fundamental Indexation proposes an index construction methodology based on several metrics able to describe a firm's «economic footprint». We compare the risk-return profile of fundamental-weighted indexes with their related cap-weighted benchmarks, focusing on the European markets during the period 1999-2013. We provide evidence of the superiority of this approach but, only: 1) if the weighting scheme is based on firm income; 2) during the Internet Bubble Burst; 3) during the realignment phase of stock prices vs firm fundamentals. Using the factor analysis, we provide evidence of high correlation between the selected fundamentals, suggesting the opportunity to simplify the index-weighting scheme.

Keywords: Fundamental indexation; Index design; Financial crises; European equity markets; Value investing; Factor analysis.
JEL Codes: G10; G11.

1 Introduction

Since the 60s, the insights stemming from the Capital Asset Pricing Model (Markowitz, 1952; Sharpe, 1964; Lintner, 1965; Mossin, 1966) has led to the consideration of the market portfolio (a basket of securities based on a market-capitalization weighting scheme) as the mean-variance optimal portfolio. Thus, the market portfolio is considered the most desirable benchmark in asset management, being the most representative of the overall market, easily replicable and designed following objective rules. Consequently, the most popular stock market indexes are weighted proportionally to their constituents' market cap, assuming that the stock market price represents the best estimate of the company's fair value. Consistently, Cap-Weighted (CW) portfolios are desirable when the Efficient Market Hypothesis (Malkiel and Fama, 1970) is verified.

Although the market portfolio has been commonly accepted as the best proxy of the market over the last few decades many critiques have been put forward that highlight the weaknesses of the CAPM. Among others, we recall Roll's critique about the impossibility of creating fully diversified portfolios (Roll, 1977) and the Noisy Market Hypothesis,

Corresponding author: Enrica Bolognesi, enrica.bolognesi@uniud.it, +39.339.4104403, via Tomadini 30/A, 33100 Udine.
Mattia Pividori, mattia.pividori@eurizoncapital.lu, +352.4949301, 8, Avenue de la Liberté, 1930, Luxembourg.

ISSN 2282-717X

claiming that prices are not the best estimate of the firm's fair value due to the presence of market participants who trade regardless of stock price evaluations (Siegel, 2006). Consistently, stock market prices tend to deviate from their fair values creating a mispricing. Hsu (2006) argues that stock prices are inefficient, meaning that underpriced stocks show a smaller market cap with respect to the fair value and, vice versa, over priced stocks gain a larger capitalization. In other words, a cap-weighting scheme leads to a sub-optimal portfolio strategy because portfolio weights are driven by market prices; as such, more weights are allocated to overvalued stocks and less weight to undervalued stocks (Arnott and Hsu, 2008).

Therefore, a growing number of alternative index construction methodologies have been promoted in order to design a more efficient market portfolio in terms of a higher degree of adherence to the real economy, as well as a higher diversification and a better risk-return profile (for a survey see Chow, Hsu, Kalesnik and Little, 2011 and Amenc, Goltz, and Lodh, 2012). Among these, the methodology developed by Arnott, Hsu, and Moore (2005) suggests the creation of indexes based on companies accounting data rather than their market-cap. The authors, acknowledged as the pioneers of Fundamental Indexation (FI), argue that this methodology aims to catch the intrinsic value of the stocks and overcomes the above-mentioned portfolio misallocation caused by the «noise trading» of irrational investors. In particular, they design Fundamental Weighted (FW) Indexes, each based on the five-years averages of fundamental measures (i.e. employment, sales, revenues, book value, operating income and dividends) and a Composite Index, which equally weights four of these six metrics. Their analysis is based on a sample of 1,000 US stocks, over the period from 1962 to 2004, and provides evidence of a yearly average excess return of the Composite Index over the S&P500 Index equal to 1.91%, associated with a similar risk profile. Thereafter, other studies provide evidence of the FI superiority, focusing on different equity markets (see Section 2 for a literature review).

This study aims to test whether the properties of the FI are verified, focusing on the European equity market since the adoption of the Euro and during a period (1999-2013) characterized by three of the toughest financial crises: the Internet Bubble Burst (2000-2001), the Global Financial Crisis (2007-2008) and the European Sovereign Debt Crisis (2010-2011). Unlike most of the previous studies on this methodology, we construct fundamental-based indexes from the perspective of an asset manager, choosing to rebalance the constituents twice a year and relying only on the information disclosed at the time of the reweighting. This methodological choice provides two advantages, allowing a consistent comparison with the benchmark (as it follows the same rebalancing dates) as well as avoiding any possible look-ahead bias.

A further contribution of our study is the comparison of FW and CW indexes during the phases of alignment and of misalignment of their composition, in order to verify the indexes' profile when stock market prices approach or deviate from their fundamentals. Finally, we focus on the design of the Composite Index, which is based on equal weighting of four fundamental metrics. This rather simple weighting scheme is particularly suitable to the use of factor analysis in an attempt of index construction optimisation. Factor analysis, in fact, focuses on the relationships between the accounting metrics and assigns alternative coefficients to each measure.

Our main findings are as follows. We confirm the superiority of the FI but only during the bursting of the Internet Bubble, a result that can be largely explained by the intrinsic value bias associated with this methodology. However, focusing on the whole time frame and observing each fundamental metric, we show that only the weighting scheme based on income is able to outperform the reference CW portfolio. Moreover, dividing the observation period in different market phases, we show the underperformance of the Composite Index during bull markets while we do not register substantial differences during bear markets. Furthermore, we observe significant excess returns of the Composite Index during the alignment phases of the two indexes' components meaning that FI outperforms when stock market evaluations rewards fundamentals. A similar result is provided when the degree of overlap between the two indexes' composition is related to the excess returns of the fundamental-based index on a semi-annual basis: we show that the FI outperformance is followed by the narrowing of the gap between the two indexes compositions, meaning that markets reward value stocks when they are undervalued. Finally, we argue that the high correlation of the fundamental measures chosen in the weighting scheme suggests the opportunity to select a lower number of metrics allowing for less operating expenditure for asset managers.

The remainder of the paper is organised as follows. Section 2 reviews the wide literature on this subject, highlighting the lively academic debate about the fundamental-based approach. Section 3 describes the dataset and the index construction methodology; Section 4 presents the results; the final Section concludes.

2 Fundamental Indexation in literature

Since the publication of the original work of Arnott *et al.* in 2005, a wide number of empirical works have aimed to test the FI superiority. The outperformance provided by the fundamental-based indexes is confirmed by a broad literature based on several equity markets over different time frames: the German market between 1988 and 2007 (Mihm and Locarek-Junge, 2010); the Portuguese market over the period 1995-2012 (Ribeiro, 2013); the Emerging markets between 1996 and 2010 (Hsieh, 2013) and between 1994 and 2009 (Arnott and Shepherd, 2010), the South African market from 1996 to 2009 (Ferreira and Krige, 2011); the Australian market from 1985 to 2010 (Basu and Forbes, 2013); global equity markets (MSCI World and FTSE World Indexes) between 1988 and 2005 (Shimizu and Tamura, 2005), from 1984 to 2004 (Hsu and Campollo, 2006) and from 1982 to 2008 (Walkshäusl and Lobe, 2010).

Focusing on the Euro area, Hemminki and Puttonen (2008) analyse the FW version of the DJ Euro Stoxx50 Index from 1996 to 2006 and show an annual excess return equal to 1.74% (for the same risk). Stotz, Döhnert, and Wanzenried (2010) examine the DJ Euro Stoxx 600 Index between 1993 and 2007 determining an annual excess return of 1.90%, again showing the same standard deviation. The advantage of the FI on the US market has been recently verified by Chen, Dempsey, and Lajbcygier (2015), over the period 1962-2009, showing FW indexes' excess returns of 28.26% and 46.14% with respect to the DJ Industrial Average Index and the Russell 1000 Index, respectively.

Further studies aim to improve the index design methodology focusing on the analysis of the risk-return profile of alternative fundamental-based indexes to the original one proposed by Arnott *et al*. (2005). For example, Neukirch (2008) overcomes the criticisms arising from the use of backtesting methodology, designing FW portfolios of Exchange Traded Funds, where the members' weights are defined in an ex-ante perspective. Blitz, Van der Grient, and Van Vliet (2010) aimed to verify if the positive results of the FI are influenced by the timing of the rebalancing and they demonstrate that the best results are achieved when the reweighting takes place in March rather than in January. Moreover, Fisher, Shah and Titman (2015) design an enhanced index, modifying the market capitalization weights using a fundamental-based factor: this modified market index tilts towards value exhibiting a higher weight for stocks with good fundamentals and *vice versa*.

Another strand of research claims that FI is merely a variant of a value strategy (Asness, 2006; Jun and Malkiel, 2008; Asness, Frazzini, Israel and Moskowitz, 2015) due to the fact that excess returns are explained by the well-known value premium rather than by the Noisy Markets Hypothesis (and the related cap drag). Lakonishok, Shleifer, and Vishny (1994) argue that value strategies yield higher returns exploiting the suboptimal behaviour of some investors rather than a higher risk. Furthermore, still concerning investment style, other authors consider FI as an active investment style or, at least, a quasi-active strategy (Ginis and Schoenfeld, 2006), which is in contrast with the idea of its proponents who consider FI as a passive strategy. Perold (2007) refuses the notion that cap weighting imposes intrinsic drag on performance asserting that FI is largely considered as a strategy of active security selection through investing in value stocks. Arnott and Markowitz (2008) return these critiques reiterating that the assumption of mean–variance efficiency of CW portfolios is deeply flawed. In this debate, Tabner (2012) confirms the idea that the FI is an active investment strategy, claiming that any excess return may be explained by style, size, momentum or other biases. In addition, Graham (2012) recalls that stock's fair value is not observable, providing evidence that, on average, CW indexes do not experience lower returns due to the systematic over-pricing of large cap securities, as assumed by Arnott and Hsu (2008).

A further critical position towards FI is that of Kaplan (2008a, 2008b) which challenges one of the key assumptions of the FI (the «independence assumption» described by Hsu, 2006 implies constant fair value multiples across stocks or market valuations completely unrelated to the fair values). In this debate, Hsu (2008) contends that, in an economy where market prices are affected by a pricing error that mean-reverts to zero, the weighting schemes that would be successful versus CW are those in which weights are uncorrelated with the pricing errors, such as the FI.

To sum up, the skepticism about the FI clashes with the aforementioned results prevailing in the literature that demonstrates the superiority of FW Indexes in different markets and on different time horizons. In particular, the highest extra return of FW Indexes is recorded during the Internet Bubble Burst as demonstrated by Arnott *et al*. (2005) which show an annual average over performance equal to 9.44% in the period 2000-2004. This result suggests that FI has been established in order to prevent the occurrence of the «irrational exuberance» experienced during the Internet Bubble, where stock prices were completely unrelated to their fundamentals.

3 Data Sample and Fundamental-based Index design

Our study aims to compare the FW and CW Index weighting schemes, focusing on the equity market of the Euro area during the period between July 1999 and June 2013 (14 years). We create several FW portfolios using a sample of the top 500 stocks of the Eurozone, that is to say the constituents of our benchmark, the Bloomberg European 500 (BE500) Index. The BE500 Index is weighted by free-float market capitalisation and reviewed on a semi-annual basis. The dataset used in our analysis is provided by Bloomberg Finance L.P. and includes: the list and the relative weights of the BE500 Index constituents and the selected accounting measures for each component[1] besides their stock prices and their total returns (gross dividends). The frequency of the data collection is semi-annual. In accordance with the methodology designed by Arnott *et al*. (2005), we adopt the following fundamental measures which are closely related to company size: 1) book value (BV); 2) trailing five-year average operating income (INC); 3) trailing five-year average revenue (REV); 4) trailing five-year average gross sales (SAL); 5) trailing five-year average gross dividends (DIV); 6) total employment (EMP). In particular, we calculate trailing five-year measures when we deal with measures published more frequently in interim records (REV, SAL, DIV and INC), in order to consider a full business cycle of the firms (Campbell and Shiller, 1988).

We firstly construct six FW indexes based on the aforementioned measures. We rebalance these indexes twice each year, at the beginning of January and July[2], in order to implement a reweighting methodology closer to the operational needs of asset managers. In reference to the symbols used by Stotz *et al*. (2010), the weight of a stock i in a FW Index, at time t $x_{i,t}^{FW}$, is defined as:

$$x_{i,t}^{FW} = \frac{F_{i,t-1y}}{\sum_{i=1}^{N} F_{i,t-1y}}, \quad (1)$$

where $F_{i,t-1y}$ is the metric of the stock i shown in the financial statement of the fiscal year preceding time t, $(t - 1y)$. Moreover, we design a Composite Index (COMP), which is composed by equally weighting four of the six measures mentioned (BV, REV, INC and DIV)[3], meaning that the weight of the stock i in the COMP $x_{i,t}^{COMP}$ is calculated as follows:

$$x_{i,t}^{COMP} = \tfrac{1}{4}\left(x_{i,t}^{BV} + x_{i,t}^{REV} + x_{i,t}^{INC} + x_{i,t}^{DIV}\right). \quad (2)$$

[1] The dataset begins in 1996 to be able to calculate the trailing five-year average of each measure.

[2] In particular, the fundamental metrics available on annual basis (BV and EMP) are kept constant in the two consecutive rebalancing dates that occur between the publications of financial statements; for the data published more frequently (INC, SAL, DIV, REV) we use the updated values provided by Bloomberg on January and July, depending on the timing of rebalancing.

[3] As discussed in Arnott *et al*. (2005), EMP and SAL are excluded from the Composite index: the first because the information is often unavailable, the latter because SAL and REV are closely related and, therefore, redundant.

It is important to highlight that the COMP portfolio merely includes the stocks whose four fundamental measures are available at the rebalancing date, meaning that the number of its constituents is steadily lower than that of the BE500 Index. In order to preserve the highest number of constituents, we construct a further index, namely Partial Composite (PC), that includes, with respect to the COMP, even the stocks for which only two or three fundamentals are available. The difference in the number of components between the PC and the COMP is, on average, equal to 27 stocks (Max = 42; Min = 16). The full description of the number of components of each mentioned Index is presented in Appendix A.

Furthermore, we construct a benchmark portfolio, the Reference Index (REF) for each FW Index (BV, REV, SAL, DIV INC, EMPL, COMP, PC) being composed of the same constituents of the corresponding FW Index, but constructed using the CW methodology.

Finally, we calculate the monthly total return for each index $R^{Index}_{t\to t+1m}$ as follows:

$$R^{Index}_{t\to t+1m} = \left(\sum_{i=1}^{N} x_{i,t} R_{i,t\to t+1m}\right), \tag{3}$$

where $R_{i,t\to t+1m}$ is the monthly total return of the stock i and $x_{i,t}$ is stock's weight in t.

In order to measure the extra returns of the FW Indexes we also calculate the Jensen's alpha, based on the CAPM and defined as the difference between a portfolio's excess return over the risk-free rate and the return explained by the market model:

$$R^{FW}_t - R^F_t = \alpha_{JEN} + b(R^{CW}_t - R^F_t) + \varepsilon_t, \tag{4}$$

where α_{JEN} is the Jensen's alpha, R^{FW}_t is the return of a FW Index (i.e. the COMP, PC, EMPL, BV, SAL, REV, INC and DIV), R^{CW}_t is the return of its Reference CW Index and R^F_t is the return on a risk-free asset. α_{JEN} provides an estimate of the risk-adjusted return, assuming that b is an appropriate measure for the systematic risk. The standard errors for the time series are consistent both in the case of heteroscedasticity and in the case of serial autocorrelation of residuals (Newey-West standard errors).

The consensus in finance academia, and among practitioners is that the simple one-factor model is not entirely effective in capturing the cross section of expected stock returns (Amenc *et al*., 2009). The parameter alpha can also lead to misleading considerations in comparing different portfolios when they are invested in securities with different characteristics, such as low-beta stocks, small stocks or value stocks (Fama and French, 2004). As such, we perform the Fama-French (1992) three-factor regression analysis. The aim of this is to verify whether the difference in performance between FW and CW portfolios could be explained by common risk factors, such as value and small-cap exposures. Thus, we run the following regression:

$$R^{FW}_t - R^F_t = \alpha + b(R^{CW}_t - R^F_t) + sSMB_t + hHML_t + \varepsilon_t, \tag{5}$$

where R_t^{FW} is the return of a FW Index, R_t^{CW} is the return of its Reference Index, R_t^{F} is the return on a risk-free asset, SMB is the small-cap factor and HML is the value factor. In particular, SMB is a portfolio that is long small cap stocks and short large stocks while HML is a portfolio that is long high book-to-price stocks (value stocks) and short low book-to-price stocks (growth stocks)[4]. Using this model, Fama and French (1993 & 1996) capture much of the variation in average return for portfolios formed on size, book-to-market equity and other price ratios that cause problems for the CAPM. Furthermore, Fama and French (1998) show that an international version of the model performs better than an international CAPM in describing average returns on portfolios formed on scaled price variables for stocks in 13 major markets.

4 Results

4.1 FI: a first comparison on asset allocation

As already mentioned, likewise the CW approach, FW Indexes favour large cap stocks, as the selected fundamental measures are proxies of the investable securities size and, therefore, highly correlated with the stock market capitalisation. It follows that the two methodologies are comparable in terms of liquidity and investment capacity of the constituents. Unlike the CW method however, FW portfolios do not take into account the level of equity prices, thus trying to avoid the overweighting of overestimated stocks and the underweighting of undervalued stocks. In particular, the FW methodology favours value stocks, excluding young companies as well as growth stocks.

Figure 1 shows the dynamics of the concentration of FW and CW Indexes (COMP versus its Reference portfolio, REFCOMP) on their 20 largest constituents. The REFCOMP registers a sum of the top 20 stocks weights ranging between 26% and 39% during the observation period, whereas for the COMP this sum ranges between 27% and 31%. This means that the FW Index concentration level on the top stocks is lower and more stable with respect to its comparable CW version[5].

Focusing on the sectorial allocation, we analyse the portfolios based on the following sectorial classifications provided by Bloomberg: Financial, Industrial, Communications, Basic Materials, Utilities, Consumer Cyclical, Consumer Non-Cyclical, Energy, Diversified and Technology. Figure 2 shows the different sector composition dynamics of the COMP and REFCOMP Indexes. The FW Index presents a more stable sectorial allocation than its CW version. This evidence, according to Arnott *et al.* (2005), refers to the weaker anchorage of the FW methodology to the investors' preferences in an attempt to better represent the regular growth of the real economy with a gradual change in the sectorial allocations.

[4] In our analysis, the small-cap factor is measured by means of the excess return of the S&P Small Cap Eurozone TR Index over the DJ Euro Stoxx 50 TR Index while the value factor is measured as the excess return of the S&P Eurozone Value TR Index over the S&P Eurozone Growth TR Index.

[5] Observing the other fundamental Indexes however, only INC and DIV show a higher concentration than the comparable CW Indexes during the period.

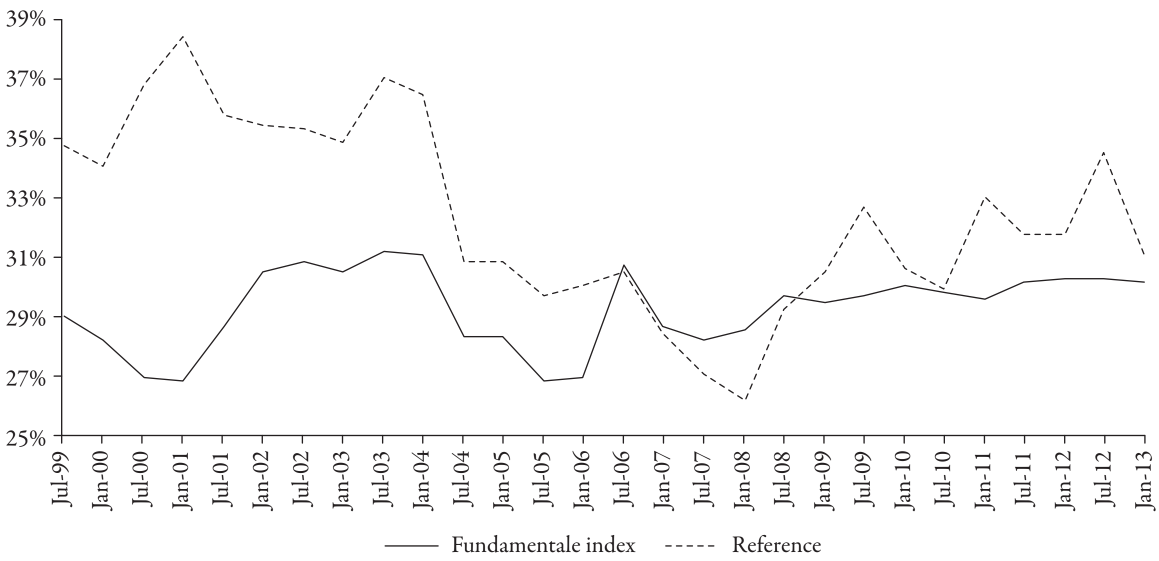

Figure 1: Cumulative weight of the 20 largest caps of COMP and REFCOMP.

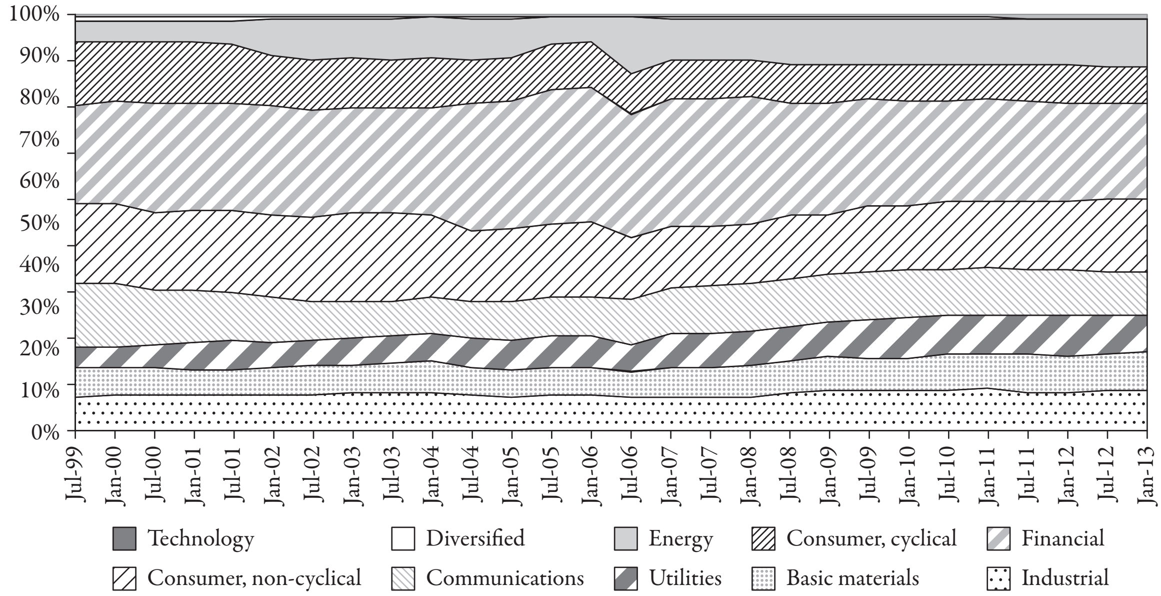

Figure 2a: Fundamental Index (COMP): Weights by Industry Sector.

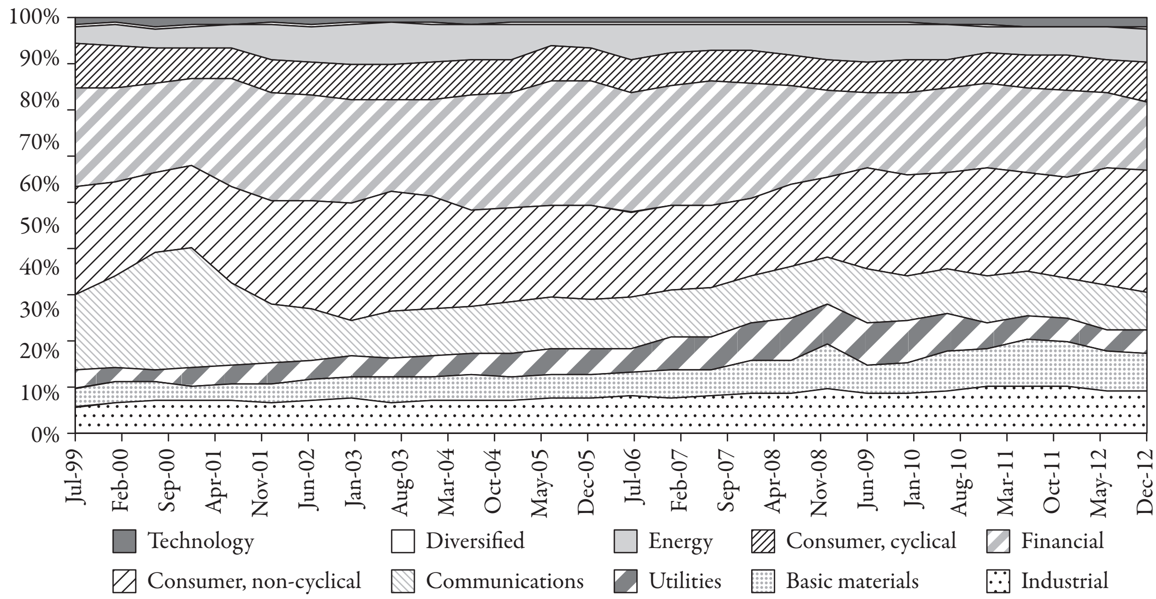

Figure 2b: Cap-weighted Index (REFCOMP): Weights by Industry Sector.

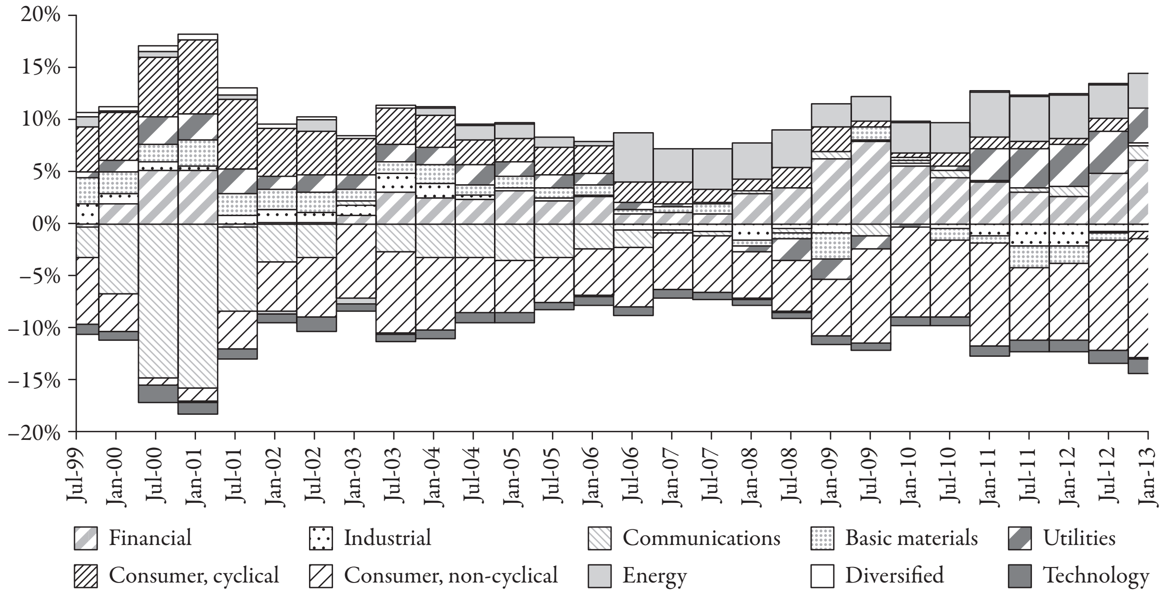

Figure 3: Differences in Sector Weights (COMP – REFCOMP).

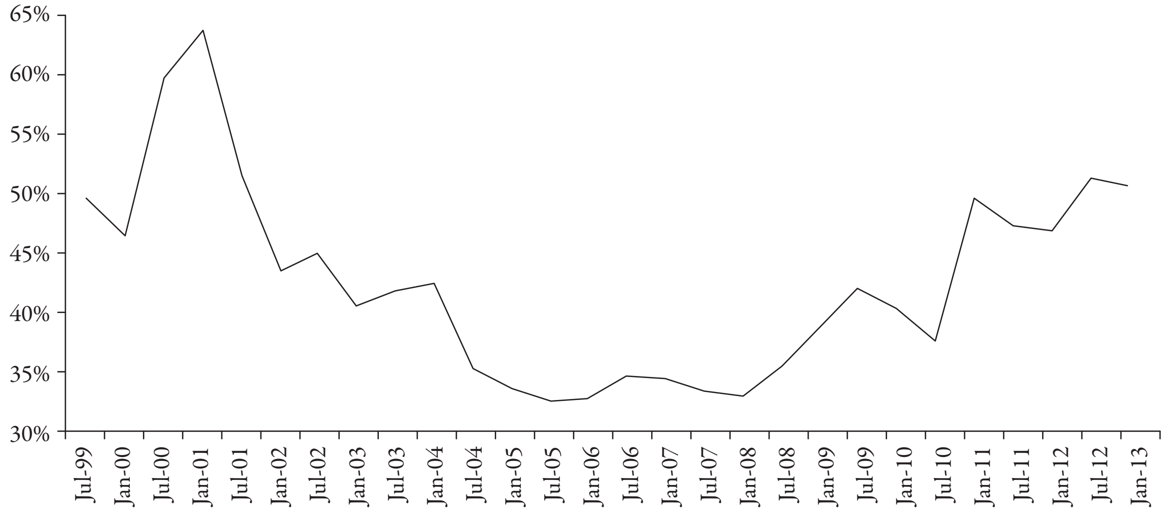

Figure 4: Cumulative sum of imbalances between COMP and REFCOMP.

Figure 3 shows the difference of the two indexes' exposure to each sector (COMP minus REFCOMP) highlighting: 1) the constant underweighting on IT sectors (Technology and Communications) in the COMP index is explained by its structural underweight in growth stocks; 2) the stable underweighting of Consumer Non-Cyclical in the COMP index is explained by a size effect (notoriously this sector is composed of large cap stocks); 3) the constant overweighting of value sectors (such as Energy and Financials) inside the COMP index.

Finally, we focus on the degree of overlap between the two Indexes. Figure 4 shows the dynamics of the cumulative difference between the weights of each constituent within the two Indexes: the sum of the weight differences ranges between 32% and 63%, whereas

Figure 5: Cumulative monthly log-returns of COMP, REFCOMP, BE500 Indexes.

the highest level of overlap occurs during the Global Financial Crisis while the lowest occurred during the bursting of the Internet Bubble.

4.2 Risk-Adjusted Performance analysis

Figure 5 shows the performance of the observed indexes during the period. The constant alignment between REFCOMP and the BE500 Indexes should not come to a surprise, as both are CW: the only difference lies in the number of components which the REFCOMP omits, in accordance with the COMP, relating to all the stocks that do not present all the measures required.

Table 1 shows the results of the Risk-Adjusted Performance (RAP) analysis for each portfolio during the 14 years between 1999 and 2013. Among the FW Indexes considered, the best performer is INC, showing a cumulative total return of 112.58% and an annualized rate of return of 5.53% (the others range from 3.99% and 4.67% in the case of DIV and COMP respectively). Moreover, each FW Index outperforms its CW version: the annualized excess returns range from 1.34% (SAL) to 3.11% (INC), the latter being the only statistical significant result at 1% level); COMP and PC outperform their Reference portfolios by 1.46% and 1.52% respectively. These statistics confirm the results of previous studies on the European stock market, showing the outperformance of the FW over the related CW Indexes of nearly 2% (Hemminki and Puttonen, 2008; Stotz *et al.*, 2010). In our analysis, however, statistical significance occurs only in the case of INC and EMPL.

On the other hand, moving on to risk profile, the standard deviation registered by the REFCOMP (16.45%) index is the lowest one with respect to all the FW portfolios (ranging from 16.74% to 19.32%). Focusing on the beta parameter, the recorded values lie within the range 0.96-1.09, revealing a similar risk attitude for all the analysed Indexes.

Jensen's alpha expresses the incremental return of a portfolio over the market return: all the FW indexes show positive coefficients, which means that the excess returns are not fully explained by a higher level of risk but, at least partially, by portfolio composition (i.e. security selection). The Calmar ratio relates to the average returns of the index for each year and the measure of maximum drawdown recorded during the period considered. The lower the result (as in the case of REFCOMP), the worse the risk adjusted performance registered is. Conversely, the higher this value is, as in the case of the FW Indexes (particularly the COMP), the better the risk-return profile of the portfolio tends to be. The advantage of a fundamental approach is also confirmed by other statistics: Sharpe, Ω, Sortino and Kappa. In particular, Ω (calculated as the ratio between the average of portfolio returns exceeding a certain threshold rate, and the average of portfolio returns not exceeding the same threshold) registers values greater than one for both the FW and CW indexes: this indicates an overall positive performance during the period and, even in this case, INC shows the highest value. Finally, we focus on Tracking Error Volatility (TEV) and Information Ratio (IR) as measures used to identify an active portfolio management style. TEV registers values relatively low while IR ranges between 0.3 and 0.8: these values, according to the metric used by Grinold and Kahn (1995), are associated with a rating of «good» or «very good» in judging the performance of an active asset manager.

4.3 Transaction costs

We estimate the rebalancing costs that must be incurred when a FW strategy is implemented. These costs are directly related to the annual turnover of the fundamentals based portfolios that is, in our case, on average equal to 12.8%[6]. This result is lower than the 15%-30% range calculated by Dash *et al.* (2010) for the US equity Indexes. Moreover, assuming negotiation fees equal to 10 bps for stock trading, the average transaction costs that must be considered are limited to nearly 1 bps per year. Furthermore, aiming to repeat the exercise provided by Arnott *et al.* (2005), we estimate what the level of trading fees able to compensate the excess return registered by the FW Indexes should be. In our case, this threshold is equal to 14%.

4.4 CAPM and the Fama-French three-factor model

Table 2 shows the results of the CAPM and Fama-French three-factor regression analyses. Panel A refers to the one-factor model results. Our findings highlight that FW portfolios have positive coefficients but are not significantly different from zero, except for INC. The alpha generated by the fundamental-based indexes range between 13 and 25 bps per month, among which COMP generates an extra return of 12 bps per month.

[6] The detailed turnover data for all FW indexes are as follows: 13.69% (EMPL); 13.71% (BV); 11.50% (SAL); 11.18% (REV); 12.99% (INC); 14.53% (DIV); 12.65% (COMP); 12.17% (CP).

Table 1: Risk-Adjusted Performance Analysis

Index	Cumulative Return	Geom. Return (yearly)	ER vs Reference	St. Dev. (yearly)	Beta	Alpha	R2	Calmar	Sharpe	Omega	Sortino	Kappa	Treynor	TEV	IR	t-value for ER	
BE500	38.45%	2.35%	–	16.24%	–	–	–	–	–	–	–	–	–	–	–	–	
EMPL	82.88%	4.41%	**2.22%**	18.77%	1.06	2.42%	0.93	0.079	5.66%	1.19	0.094	0.195	0.0029	4.93%	0.48	(1,88)	*
BV	79.05%	4.25%	1.59%	18.90%	1.09	1.76%	0.95	0.074	5.40%	1.19	0.093	0.186	0.0027	4.49%	0.41	(1,59)	
SAL	74.50%	4.06%	1.34%	19.32%	1.09	1.56%	0.93	0.071	5.15%	1.18	0.087	0.178	0.0026	5.46%	0.29	(1,17)	
REV	73.43%	4.01%	1.58%	19.23%	1.10	1.84%	0.93	0.071	5.08%	1.18	0.086	0.176	0.0026	5.43%	0.34	(1,36)	
INC	112.58%	5.53%	**3.11%**	16.81%	0.97	3.07%	0.95	0.104	7.56%	1.26	0.124	0.257	0.0038	3.80%	0.78	(2,94)	***
DIV	73.00%	3.99%	1.67%	16.74%	0.96	1.67%	0.93	0.073	5.02%	1.18	0.087	0.178	0.0025	4.46%	0.34	(1,35)	
COMP	89.43%	4.67%	1.46%	17.59%	1.05	1.52%	0.96	0.085	6.09%	1.21	0.102	0.209	0.003	3.76%	0.41	(1,60)	
PC	75.07%	4.08%	1.52%	17.97%	1.04	1.62%	0.95	0.073	5.15%	1.18	0.089	0.179	0.0026	3.98%	0.4	(1,58)	
REFCOMP	55.58%	3.21%	–	16.45%	1	–	1	0.06	3.70%	1.14	0.068	0.138	0.0018	–	–	–	

* Statistically significant at the 10 percent level.
** Statistically significant at the 5 percent level.
*** Statistically significant at the 1 percent level.

Table 2: CAPM and Fama and French three-factor model. HAC standard errors, bandwidth 4 (Bartlett kernel). No. observations: 168

	α		*b*		*s*		*h*	
Panel A: One-factor model								
COMP	0.0013		**1.044**					
	(1.13)		(35.99)	***				
PC	0.0013		**1.043**					
	(1.16)		(32.26)	***				
EMPL	0.0020		**1.063**					
	(1.39)		(29.23)	***				
BV	0.0015		**1.093**					
	(1.37)		(28.64)	***				
SAL	0.0013		**1.085**					
	(0.87)		(25.67)	***				
REV	0.0015		1.098					
	(1.00)		(25.89)	***				
INC	**0.0025**		**0.972**					
	(2.07)	**	(33.06)	***				
DIV	0.0014		**0.960**					
	(0.99)		(26.42)	***				
Panel B: Three-factor model								
COMP	0.0006		**1.036**		–0.003		**0.528**	
	(1.03)		(80.78)	***	(–0.14)		(14.41)	***
PC	0.0006		**1.039**		–0.003		**0.564**	
	(1.05)		(77.64)	***	(–0.15)		(15.52)	***
EMPL	0.0008		**1.068**		**0.115**		**0.581**	
	(0.89)		(48.45)	***	(–4.23)	***	(10.74)	***
BV	0.0008		**1.088**		–0.002		**0.501**	
	(1.12)		(48.99)	***	(–0.06)		(7.73)	***
SAL	0.0003		**1.082**		0.004		**0.734**	
	(0.38)		(49.12)	***	(0.11)		(14.86)	***
REV	0.0006		**1.094**		0.003		0.717	
	(0.71)		(49.17)	***	(0.09)		(14.03)	***
INC	**0.0019**		**0.969**		–0.004		**0.513**	
	(3.18)	***	(82.84)	***	(–0.18)		(8.96)	***
DIV	0.0006		**0.957**		–0.027		**0.641**	
	(1.10)		(77.33)	***	(–1.31)		(11.62)	***

* Statistically significant at the 10 percent level.
** Statistically significant at the 5 percent level.
*** Statistically significant at the 1 percent level.

Panel B shows the results of the three-factor model, which takes into account the exposure to value and small cap factors. The exposure to the value premium (coefficient *h*) is positive and statistically significant at the 1% level for each index. Looking at the exposure to the small-cap factor (SMB), the coefficient *s* is positive and statistically significant only for EMPL. Overall, it can be noted that, when the style factor is considered in the regression model, abnormal returns are considerably lower than when the one-factor model is applied: in the case of COMP, on average, the monthly alpha is reduced to less than 1 bp per month. Our findings confirm the prevailing view in the literature (Asness, 2006; Blitz and Swinkels, 2008; Asness *et al.*, 2015) namely that the strong value tilt accounts for most of the outperformance of FW portfolios.

4.5 FI during bull and bear markets

As in Arnott *et al.* (2005), we proceed with the comparison of the indexes during bull and bear market periods, defining a bull market as 20% of upside from the previous low and as a bear market as a 20% decline from the previous high. The authors' results reveal an outperformance of the FW indexes both during bull and bear markets of 55 bps and 640 bps, respectively. On the contrary, our results presented in Table 3 show an underperformance of each FW Index during bull markets and a slight out- or under-performance during bear markets, depending on which FW index is examined. In particular, in the case of bull markets, REFCOMP shows a consistent over performance of 352 bps with respect to the average of the FW Indexes and of 56 bps in the case of bear markets. Focusing on the comparison between COMP and REFCOMP, our results show an underperformance (363 bps) only in the case of bull markets while we record an extra return of 84 bps during bear markets.

4.6 FI during financial crisis

We continue the analysis of the FW Indexes focusing on the three financial crises that occurred during the observation period and we calculate the maximum drawdown recorded by our indexes. We refer to these financial crises as follows: Internet Bubble Burst (28/04/00-31/03/03), Global Financial Crisis (31/05/07-27/02/09) and European Sovereign Debt Crisis (29/04/11-30/09/11).

The aim is to verify whether the methodology based on fundamentals is able to mitigate the losses during market downturns as widely claimed in the literature. Table 4 shows interesting results. Not surprisingly, we register a strong outperformance of the fundamental-based portfolios during the Internet Bubble Burst (in the 10%-21% range). It is worth remembering that the collapse of TMT (technology, media and telecom) stocks demonstrated the need to prevent a vicious circle, where portfolio managers were led to keep buying stocks already overvalued due to the well-known return drag in CW portfolios. On this issue, Amenc *et al.* (2012) argue that the various types of fundamentals-based indexes, which have appeared after the bursting of the Internet Bubble, can be seen as an attempt to avoid drawing investors into being exposed to this type of crisis.

Focusing on the financial crises, our results highlight that during both the Global Financial Crisis and the European Sovereign Debt Crisis, FW Indexes underperform their benchmark; in the case of COMP, the underperformance is in the 144-283 bps range (among the fundamental-based indexes, INC continues to be the best performer). Our results confirm the findings of Amenc *et al.* (2012) claiming that the fundamentals-based indexes have not been able to protect investors from being exposed to the financial crisis and the ensuing sovereign risk crisis, which heavily affected banking stocks. Likewise, Chen *et al*. (2015) argue that the poor performance of the FI during the Global Financial Crisis was due to the overweight position in the financial sector (value stocks) whose stock prices had dropped sharply.

Table 3: FI during bull and bear markets

	Bull markets			Bear markets		
	Geometric Return (%)	Volatility (%)	Sharpe Ratio	Geometric Return (%)	Volatility (%)	Sharpe Ratio
BE500	29.55	13.27	1.95	−25.63	19.51	−1.61
REFCOMP	30.93	13.00	2.05	−24.80	20.19	−1.49
COMP	27.30	14.02	1.77	−23.96	21.38	−1.36
PC	27.35	14.06	1.76	−25.42	21.53	−1.44
EMPL	27.67	14.47	1.69	−26.37	22.73	−1.43
BV	29.75	14.24	1.92	−27.04	22.52	−1.46
SAL	27.67	14.84	1.68	−26.75	23.29	−1.42
REV	27.07	14.95	1.65	−26.70	23.21	−1.42
INC	27.50	13.76	1.83	−21.84	20.70	−1.27
DIV	24.78	13.70	1.70	−23.45	19.80	−1.45
Avg (ex Comp.)	27.41	14.33	1.74	−25.36	22.04	−1.41

Table 4: Indexes performance during financial crisis

	Internet Bubble Burst (28.04.00-31.03.03)		Global Financial Crisis (31.05.07-27.02.09)		EU Sovereign Debt Crisis (29.04.11-30.09.11)	
	Geom. Return (%)	ER vs Reference (%)	Geom. Return (%)	ER vs Reference (%)	Geom. Return (%)	ER vs Reference (%)
EMPL	−42.77	12.27	−55.93	−1.66	−23.72	−3.62
BV	−41.62	10.25	−57.56	−3.33	−24.66	−4.48
SAL	−41.48	11.09	−56.98	−2.72	−24.82	−4.74
REV	−41.16	12.47	−56.87	−2.90	−24.87	−4.74
INC	−32.06	21.57	−52.89	1.08	−20.62	−0.47
DIV	−34.76	20.60	−54.70	−1.20	−20.29	−0.41
COMP	−37.83	12.65	−54.78	−1.44	−21.80	−2.38
PC	−40.17	13.35	−55.54	−1.77	−22.74	−2.83

Table 5 exhibits the outcomes of the one-factor and three-factor regression models related to the three financial crises. The results shown in Panel A reconfirm the significance of the outperformance of FI during the Internet Bubble Burst. During the Global Financial Crisis, however, the performance difference between the two indexes appears to not be significant while, during the European Sovereign Debt Crisis, an underperformance of FI is significant at the 5% level.

Panel B shows the results of the Fama-French model. Focusing on the Internet Bubble Burst, we show a positive coefficient α, statistically significant at the 5% level. Both the beta coefficient and the value factor are highly significant: in particular, b is equal to 1.0798, while h is equal to 0.5174. The small-cap factor s is positive but not significant. These results indicate that the excess return of COMP is only partially explained by the aforementioned value tilt attributed to the fundamental methodology. Even during the second financial crisis the coefficients b (1.0245) and h (0.2785) are highly significant. Finally, focusing on the European Sovereign Debt Crisis, we find rather different results. In this case, α is negative and significant at the 5% level; the beta coefficient is highly significant and slightly lower than 1; the coefficient s is negative and significant at the 10% level; the value factor h is not statistically significant. Although based on a few observations, these latter results lead to interesting considerations. They suggest that,

Table 5: Fundamental Indexation during financial crises

Panel A: One-factor model					
HAC SE bandwidth 4	*α*	*b*	*s*	*h*	No. Obs.
(Bartlett kernel)					
Internet Bubble Burst (28/04/2000-31/03/2003)					
COMP	**0.0071** (3.88) ***	**1.027** (23.10) ***			36
COMP	0.0004 (0.27)	**1.044** (100.70) ***			22
COMP	**–0.0043** (–3.64) **	**1.024** (59.30) ***			6
Panel B: Three-factor model					
Internet Bubble Burst (28/04/2000-31/03/2003)					
COMP	**0.0027** (2.35) **	**1.080** (54.89) ***	0.032 (0.69)	**0.517** (12.16)***	36
Global Financial Crisis (31/05/2007 – 27/02/2009)					
COMP	0.0014 (1.36)	**1.025** (108.56) ***	–0.002 (–0.07)	**0.279** (6.13) ***	22
COMP	**–0.0059** (–4.59) **	**0.993** (55.48) ***	**–0.253** (–4.24)*	–0.425 (–2.02)	6

* Statistically significant at the 10 percent level.
** Statistically significant at the 5 percent level.
*** Statistically significant at the 1 percent level.

during the last financial crisis, the value factor is not able to explain the performance of the FW Index, while the size factor has registered a negative effect.

4.7 FI: comparing stock market prices and stock fundamentals

Our latest analysis focuses on the dynamics of the indexes composition and, in particular, on the degree of overlap between the constituents of the two indexes. We define ‘fundamental abnormal weight’ as the difference between the stock’s fundamental index weight (i.e. the COMP’s weights) and the stock’s market index weight (i.e. the REFCOMP’s weights). We calculate, on a semiannual basis, the fundamental abnormal weight for each constituent in order to aggregate them: this latter figure is the degree of imbalance between the two Indexes.

Going back to the graph of Figure 4, we note that the degree of imbalance is rather different during the observation period: descending in the first period and, then, ascending during the second. Therefore, we decide to analyze the performance of the two Indexes over two sub periods: during the first period (01/01/01-31/12/05) the compositions gradually realign (with a strong convergence after the Internet Bubble) from 2007 on-

Table 6: Stock market evaluations and stock fundamentals

HAC SE bandwidth 4	α	*b*	*s*	*h*	No. Obs.
Panel A: One-factor model					
(Bartlett kernel)					
Alignment phase (01/01/2001-31/12/2005)					
COMP	**0.0031** (3.00) ***	**1.037** (34.77)***			60
Misalignment phase (31/12/2005-30/06/2013)					
COMP	0.0000 (0.09)	**1.103** (37.43)***			90
Panel B: Three-factor model					
Alignment phase (01/01/2001-31/12/2005)					
COMP	0.0005 (0.81)	**1.070** (66.47)***	**0.054** (1.82) *	**0.428** (8.83) ***	60
Misalignment phase (31/12/2005-30/06/2013)					
COMP	0.0007 (0.25)	**1.039** (70.45)***	−0.027 (−0.31)	**0.485** (7.02) ***	90

* Statistically significant at the 10 percent level.
** Statistically significant at the 5 percent level.
*** Statistically significant at the 1 percent level.

wards, however, the two indexes start to move away again in their composition. In other words, the fundamental abnormal weights decrease when the stocks' evaluations return to reward the fundamentals (alignment phase) and vice versa increase when investors' preferences deviate from stocks' fundamentals (misalignment phase).

Aiming to detect the performance of the two indexes during these sub periods, we calculate the excess returns between COMP and REFCOMP and test their statistical significance. Table 6 presents our results. During the alignment phase the fundamentals-based index registers a monthly over performance of 0.30% that is highly statistically significant. Performing the Fama-French three-factor regression analysis, we detect that this excess return is explained by a strong style effect (i.e. values stocks over perform growth stocks during the alignment phase) and a size effect – this is statistically significant at the 10% level. On the contrary, during the second phase, we do not register any significant excess return. These results suggest the detachment between stock market values and stock fundamentals should be carefully observed. When this level is high it means that investors base their assessments on different criteria with respect to stock fundamentals, sometimes leading to irrational movements of the markets, as seen during the Internet Bubble.

In order to further emphasise this result, we finally observe the evolution of the fundamental abnormal weights after each semi-annual rebalancing date. By relating the change in the overlap level of the two indexes (deriving from rebalancing) to the extra return registered by COMP during the previous six months, we verify that each outperformance of the fundamental-based index corresponds to a decrease in the fundamental abnormal weights (after rebalancing), and *vice versa*.

Figure 6 shows this inverse relationship ($R^2 = 0.75$) providing evidence that, when market rewards the best stocks (from a fundamental perspective), the gap in the composi-

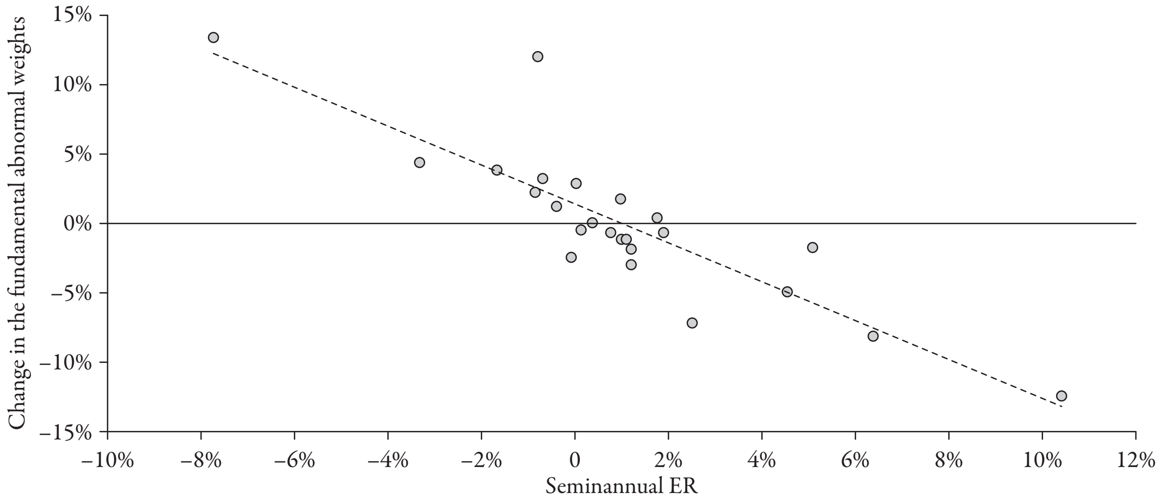

Figure 6: Linear interpolation between the semi-annual excess returns of COMP over COMPREF, and the changes in the sum of fundamental abnormal weights.

tion of the two indexes tends to narrow, as the weight of best stocks will have risen in the price sensitive CW Index. On the other hand, a widening of the fundamental abnormal weights follows an underperformance of COMP.

As known, COMP summarises four accounting data (BV, REV, INC, DIV) and it is calculated as an equally weighted average of the constituents of these indexes. The choice of this specific index links to its representativeness of the entire category of fundamental indexes and to its function of synthesis, which helps correct any possible distortion in corporate and performance evaluation. Theoretically, this means that it implicitly assumes no correlation between the four accounting variables or even a lack of consideration of the relationships between them. Therefore, we claim that the arithmetic mean (while helping to mitigate potential data errors of the selected fundamentals) could also be a misleading simplification in researching an indexing methodology able to mirror the dynamics of the real economy.

4.8 FI: an attempt of optimisation using the factor analysis

With the goal of providing a more realistic footprint of each index constituent, we employ the factor analysis to explore the relationships between the four accounting metrics (BV, REV, INC, DIV) and assign alternative coefficients to each fundamental measure in the Composite Index. Operationally, for each constituent, we calculate four coefficients considering the dynamics of the correlations between the metrics over time. We use factor analysis, which is a statistical method of multivariate analysis whose goal is to synthesise the information included in a set of correlated variables, to determine a number of latent dimensions (factors) that are not directly observable. Therefore, the general purpose is to determine one or more factors, each one representative of a linear combination of the original variables: that is to say that a limited number of independent components are found and identified as to represent the proportion of variance in common among the variables[7].

A further step in the analysis involves the calculation of factor scores: these are standardised scores associated with the original variables for each of the factors identified by the analysis[8]. The common goal is, once the number of factors underlying a dataset has been identified, to use the information about the factors in subsequent analyses (Gorsuch, 1983). The common principle behind all these calculation methods (especially valid for refined ones) is to obtain factors as linear combinations of the original variables (X_1,..., X_p), which consider both shared variance and error term variance:

$$F_1 = c_{1,1}X_1 + \ldots + c_{1,i}X_i + \ldots \ c_{1,p}X_p, \tag{6}$$

where $c_{k,i}$ is the factor score coefficient which represents the weight of the i-th standardised variable in determining the single factor F_1. It is worth highlighting how similar the weights definitions are for a generic factor and for COMP itself; in fact, the latter is a linear combination of accounting variables:

$$x_{i,t}^{COMP} = \alpha_t x_{i,t}^{BV} + \beta_t x_{i,t}^{REV} + \gamma_t x_{i,t}^{INC} + \delta_t x_{i,t}^{DIV}, \tag{7}$$

where all four coefficients of the original model (α_t, β_t, γ_t, δ_t) always assume the same standard value of 0.25 and $x_{i,t}$ represents the weight of the i-th firm for each fundamental index at time t. Thus, our optimisation aims to obtain, date by date, different values for each of the four coefficients used in calculating $x_{i,t}^{COMP}$. Unlike originally intended by Arnott *et al*. (2005), these coefficients become unknown variables instead of constants: our inputs are semi-annual time series of the normalised weights (i.e. varying between 0 and 1) we computed and assigned to each constituent firm when calculating BV, REV, INC and DIV indexes.

From some preliminary tests, we decide to employ the factor analysis with maximum likelihood extraction, using the first factor as a cut-off threshold. This choice of a single factor has been confirmed either by scree test, by Kaiser criterion or by the portion of variance explained – in all cases this was close to (or exceeded) the 70% threshold. Moreover, this was found to still be the only choice consistent with the target of the optimisation, given the ideal overlap between the first factor extracted and the COMP index itself.

Table 7 exhibits the un-rotated factor loadings taken from the factor matrix extracted that represents the correlations between the observed variables and the single factor (for the un-rotated solution) and potentially ranging from –1 to +1.

Statistical significance is generally high in each semester as shown by the eigenvalues associated with the individual factors, each greater than one. The high levels of total

[7] The general rule of thumb is that, if the absolute value of the standardised loading is greater than 0.3, the variable is relevant for the factor, so it can be considered: the variable with the highest loading will be the most significant relatively to the factor considered.

[8] Concretely, there are many possibilities to compute these factor scores, among which the two main classes are refined (e.g. multiple regression, Bartlett's approach and Anderson-Rubin method) and non-refined methods (e.g. sum scores).

variance explained confirm statistical significance[9]. The loadings are rotation-indifferent meaning that they do not change after any rotation. The results highlight a strong correlation between the four measures and the factor that according to the transitive property indicates a high correlation between the variables themselves.

The next step is to adopt factor scores as proxies for the coefficients (α_t, β_t, γ_t, δ_t) in the calculation of the optimised COMP. As already hinted at, there are different ways to create factor scores: the method to use depends on many issues such as the goal of the project, the nature of the work, and even issues such as the individual researcher's knowledge of methodology, statistical techniques, and software (DiStefano *et al.*, 2009). The most common refined methods use standardised information to create standardised factor scores. Generally, these methods are preferable because they aim to maximise validity (by producing factor scores that are highly correlated with a given factor) and obtain unbiased estimates of the true factor scores (Gorsuch, 1983). In addition, they attempt to preserve the relationships among factors. In particular, the multiple regression method may seem the most appropriate technique for our optimisation because of the two additional advantages of taking into account correlation among observed variables (as well as the correlation among factors, and the correlation between factors and observed variables) and due to the use of its underlying model to virtually produce optimal factor scores. The main risks of this procedure are that scores could be not univocal, biased or correlated. However, the most important matter concerns its application to our optimisation model, since the definition of a standardised score implies that the average of all factor scores will be zero[10]. To overcome these issues, our proposal is to use a non-refined approach, such as the «weighted sum score» method. In general, non-refined factor scores are thought to be more stable across samples than refined methods (Grice and Harris, 1998). This advantage means that the obtained results do not heavily depend on the particular sample used even if, as a drawback, this method may produce scores that are correlated (Glass and Maguire, 1966).

In order to preserve the information provided by the factor analysis about the relationships among the accounting variables, we decide to approximate the coefficients on the basis of the factor loadings. Consistent with this for convenience, we suppose hypothetically identical (unitary) items for all fundamental metrics, so that new coefficients (α_t, β_t, γ_t, δ_t) are calculated transforming at each date the four factors loadings in index weights for the construction of a new Composite (COMPAF) Index to be compared to the original COMP.

Results are presented in Table 7. In column 6-9, it emerges clearly that, except in few cases, the deviations from the 25% of the original COMP are not very large. This result can be explained firstly by the fact that the correlation coefficient between the accounting variables tends to zero, meaning that it is not possible to operate any synthesis.

The opposite interpretation, best suited to our case, is that of very high correlations between the fundamental metrics: for this reason, the equally weighted average of these

[9] Minimum values recorded for this over our time horizon belong to the first two dates but, in general, the level is high enough to ensure the validity of the analysis, with a maximum of 86.5%.

[10] Indeed, this entails the presence of both positive and negative scores but it is not permissible given the subsequent use of such scores as weighting coefficients for the fundamental weights composing the COMP: we cannot consider negative weights, as they are equivalent to short positions.

Table 7: Factor loadings and new fundamental coefficients

	Factor loadings				New fundamental coefficients				Significance
	FBV	FREV	FINC	FDIV	PBV	PREV	PINC	PDIV	σ^2 explained (%)
05/07/1999	0.383	0.734	0.958	0.886	0.129	0.248	0.324	0.299	66.77
03/01/2000	0.387	0.732	0.957	0.89	0.13	0.247	0.323	0.3	66.85
03/07/2000	0.835	0.793	0.923	0.804	0.249	0.236	0.275	0.24	77.77
08/01/2001	0.834	0.788	0.925	0.81	0.249	0.235	0.276	0.241	77.89
02/07/2001	0.829	0.771	0.914	0.857	0.246	0.229	0.271	0.254	78.45
07/01/2002	0.835	0.793	0.919	0.87	0.244	0.232	0.269	0.255	79.96
01/07/2002	0.855	0.807	0.883	0.858	0.251	0.237	0.259	0.252	79.28
06/01/2003	0.837	0.804	0.891	0.86	0.247	0.237	0.263	0.254	79.02
07/07/2003	0.844	0.804	0.871	0.855	0.25	0.238	0.258	0.253	78.38
05/01/2004	0.857	0.803	0.864	0.857	0.253	0.238	0.256	0.253	78.59
05/07/2004	0.844	0.795	0.893	0.882	0.247	0.233	0.262	0.259	79.72
03/01/2005	0.841	0.79	0.894	0.879	0.247	0.232	0.263	0.258	79.43
04/07/2005	0.845	0.739	0.944	0.891	0.247	0.216	0.276	0.261	80.14
02/01/2006	0.846	0.742	0.943	0.891	0.247	0.217	0.276	0.26	80.27
03/07/2006	0.841	0.844	0.97	0.939	0.234	0.235	0.27	0.261	85.83
08/01/2007	0.83	0.811	0.961	0.919	0.236	0.23	0.273	0.261	83.42
02/07/2007	0.856	0.831	0.967	0.924	0.239	0.232	0.27	0.258	85.14
07/01/2008	0.859	0.831	0.967	0.928	0.24	0.232	0.27	0.259	85.36
07/07/2008	0.884	0.839	0.943	0.942	0.245	0.233	0.261	0.261	86.10
05/01/2009	0.885	0.837	0.942	0.941	0.245	0.232	0.261	0.261	86.01
06/07/2009	0.87	0.824	0.936	0.954	0.243	0.23	0.261	0.266	85.40
04/01/2010	0.872	0.836	0.939	0.953	0.242	0.232	0.261	0.265	85.96
05/07/2010	0.885	0.834	0.927	0.946	0.246	0.232	0.258	0.263	85.67
03/01/2011	0.886	0.834	0.927	0.947	0.247	0.232	0.258	0.263	85.72
04/07/2011	0.875	0.843	0.955	0.936	0.243	0.234	0.265	0.259	86.36
02/01/2012	0.876	0.843	0.955	0.936	0.243	0.234	0.265	0.259	86.42
02/07/2012	0.852	0.837	0.953	0.93	0.239	0.234	0.267	0.26	85.24
07/01/2013	0.851	0.824	0.954	0.929	0.239	0.232	0.268	0.261	84.83

Table 8: Comparative analysis of the Composite Index (COMP) and the optimised version (COMPAF)

Index	Cumulative Return (%)	Annualised Average Rate (%)	Excess Return (%)	Annualised St. Dev. (%)	Beta	Alpha (%)	R2	Calmar	Sharpe
COMP	89.42	4.67	1.46	17.62	1.047	1.52	0.96	0.0852	0.0608
COMPAF	88.78	4.64	1.44	17.57	1.045	1.49	0.96	0.0848	0.0605
	Omega	Sortino	Kappa	Treynor	TEV (%)	IR	Average Ptf Turnover (%)	Max drawdown (%)	
COMP	1.208	0.101	0.208	0.00295	3.76	0.411	12.63	−54.80	
COMPAF	1.206	0.100	0.206	0.00294	3.74	0.405	12.70	−54.73	

measures in the Composite Index is redundant. Table 8 shows the performance of the two indexes COMP and COMPAF to demonstrate their substantial equivalence.

Although this index construction methodology based on the factor analysis should be tested on different markets and time frames, we can advance some operative remarks. The high correlation observed between the four metrics throughout the period suggests the opportunity to focus on a lower number of variables, contributing to a significant operational simplification.

Theoretically, the index design should be based on a single measure, such as dividends or income, showing the highest new weights in the factor analysis. A fundamental index based on dividends is supported by the view that '*dividends are the only fundamental variable that is completely objective, transparent and unable to be manipulated*' (Siegel, 2006). On the other hand, we recall from Section 4.2 that the fundamental index based on income (INC) has registered the best risk-adjusted performance.

5 Conclusions

In this paper we analyse the risk-return profile of FW indexes based on several metrics representing the «economic footprint» of firms, such as book value, income, revenues and dividends. We focus on the Euro equity market following the research methodology proposed by Arnott *et al.* (2005) but following the operational needs of an asset manager and as such, avoiding any look-head bias. We confirm the superiority of the FI, during the overall period, but only when the index weighting scheme is based on firm income. In the other cases, the outperformance is not statistically significant. Focusing on the three financial crises that occurred during our observation period, we confirm the superiority of the FI during the Internet Bubble Burst, i.e. when overvalued stocks realigned towards their fundamentals. During the other financial crises, the FW Indexes underperform their benchmark as well as during bull market phases. Furthermore, we find a significant outperformance of the Composite Index when the composition of FW and CW indexes converges after a period of strong misalignment, often explainable by irrational stock market evaluations. Coherently, we find an inverse relation between the change in the overlap level of the two indexes (deriving from the rebalancing) and the extra return registered by the fundamental-based index during the previous six months. This result means that value stocks generally outperform when they are underweighted (i.e. undervalued) in market-weighted portfolios.

Finally, in the attempt to optimise the weighting scheme of the Composite Index, we argue that the high correlation of the selected metrics permit us to focus only on one of the four metrics (income or dividends show the most convincing results), thus simplifying the index construction process.

Acknowledgements

The authors would like to thank Prof. Maurizio Polato for helpful comments. The paper benefited from the valuable comments of the discussant, Prof. Eliana Angelini, and the suggestions of the participants at the ADEIMF Piacenza 2015 Conference. The opinions expressed in the paper are those of the authors and do not necessarily represent the views of Eurizon Capital S.A.

Appendix A

Table A1: Dataset Characteristics

Date	Comp. BE500	Stocks with available accounting metrics								Stocks with a zero FI weight (in %)					
		EMPL	SAL	BV	REV	INC	DIV	PC	COMP	EMPL	SAL	BV	REV	INC	DIV
5/7/99	546	471	454	473	510	510	442	409	393	0.00	0.00	1.69	0.00	1.37	2.26
3/1/00	560	482	458	491	524	524	443	414	397	0.00	0.00	1.63	0.00	1.91	2.48
3/7/00	563	473	451	476	530	530	459	413	396	0.00	0.00	1.26	0.00	2.64	2.83
8/1/01	551	473	449	477	518	518	449	414	397	0.00	0.00	1.26	0.00	2.32	2.00
2/7/01	562	495	475	487	535	535	465	422	380	0.00	0.00	1.85	0.00	6.54	6.24
7/1/02	515	467	440	466	490	490	426	404	365	0.00	0.00	2.15	0.00	5.71	5.87
1/7/02	514	444	447	456	488	488	427	398	360	0.00	0.00	2.41	0.00	4.92	5.39
6/1/03	512	449	449	463	484	484	426	405	369	0.00	0.00	1.94	0.00	3.93	4.69
7/7/03	508	431	443	445	480	480	424	390	354	0.00	0.00	1.57	0.00	5.42	4.25
5/1/04	495	426	435	442	470	470	407	382	354	0.00	0.00	1.36	0.00	4.47	2.70
5/7/04	500	440	448	453	476	476	448	429	401	0.00	0.00	1.32	0.00	4.41	2.68
3/1/05	500	449	450	460	475	475	448	436	403	0.00	0.00	1.30	0.00	5.26	3.35
4/7/05	500	444	447	453	476	476	467	444	419	0.00	0.00	0.88	0.00	3.78	2.36
2/1/06	500	449	452	458	476	476	463	446	424	0.00	0.00	0.87	0.00	3.57	2.16
3/7/06	500	443	450	452	477	477	466	444	423	0.00	0.00	1.11	0.00	2.10	2.36
8/1/07	499	445	451	461	477	477	468	453	428	0.00	0.00	1.30	0.00	1.89	2.78
2/7/07	501	449	454	457	482	482	474	450	432	0.00	0.00	1.31	0.00	1.24	2.32
7/1/08	501	454	451	461	480	480	473	455	430	0.00	0.00	1.08	0.00	1.46	3.38
7/7/08	501	460	456	463	480	480	477	460	434	0.00	0.00	2.38	0.00	1.67	2.94
5/1/09	501	467	462	473	479	479	470	464	437	0.00	0.00	1.69	0.00	1.88	3.62
6/7/09	501	456	457	459	483	483	475	451	427	0.00	0.00	0.22	0.00	2.48	3.37
4/1/10	501	470	469	472	485	485	473	460	440	0.00	0.00	0.21	0.00	1.65	2.75
5/7/10	500	469	471	474	485	484	472	462	441	0.00	0.00	0.63	0.00	2.69	2.54
3/1/11	500	472	474	477	485	484	474	467	450	0.00	0.00	0.63	0.00	1.86	1.69
4/7/11	499	475	468	474	483	482	476	469	437	0.00	0.21	1.90	0.21	3.73	2.10
2/1/12	500	475	470	478	484	483	477	472	442	0.00	0.21	1.88	0.21	3.11	2.10
2/7/12	499	474	472	479	485	485	480	474	443	0.00	0.21	1.46	0.21	4.12	2.29
7/1/13	499	474	472	480	484	484	479	475	445	0.00	0.21	1.67	0.21	3.51	2.51
Average	512	460	456	466	489	488	458	438	411	0.00	0.03	1.39	0.03	3.20	3.07

References

Amenc N., Goltz F. and Le Sourd V. (2009) 'The Performance of Characteristics-Based Indices', *European Financial Management*, 15(2), pp. 241-278.

Amenc N., Goltz F. and Lodh A. (2012) 'Choose Your Betas: Benchmarking Alternative Equity Strategies', *Journal of Portfolio Management*, 39, pp. 88-111.

Amenc N., Goltz F. and Martellini L. (2011) 'Improved Beta? A Comparison of Index-Weighting Schemes', *Journal of Indexes*, 14, pp. 10-19.

Arnott R.D. and Hsu J.C. (2008) 'Noise, CAPM and the Size and Value Effects', *Journal of Investment Management*, 6, pp. 1-11.

Arnott R.D., Hsu J.C. and Moore P. (2005) 'Fundamental Indexation', *Financial Analysts Journal*, 61, pp. 83-99.

Arnott R.D. and Markowitz H. (2008) 'Fundamentally Flawed Indexing: Comments', *Financial Analysts Journal*, 64, pp. 12-14.

Arnott R.D. and Shepherd S.D. (2010) 'The Fundamental Index Concept in Emerging Markets', Working Paper Research Affiliates.

Asness C.S. (2006) 'The Value of Fundamental Indexing', *Institutional Investor*, 40, pp. 94-99.

Asness C.S., Frazzini A., Israel R. and Moskowitz T. (2015) 'Fact, Fiction, and Value Investing', *The Journal of Portfolio Management*, 42, pp. 34-52.
Basu A.K. and Forbes B. (2014) 'Does Fundamental Indexation Lead to Better Risk-Adjusted Returns? New Evidence from Australian Securities Exchange', *Accounting & Finance*, 54, pp. 699-728.
Blitz D. and Swinkels L. (2008) 'Fundamental Indexation: An Active Value Strategy in Disguise', *Journal of Asset Management*, 9, pp. 264-269.
Blitz D., Van der Grient B. and Van Vliet P. (2010) 'Fundamental Indexation: Rebalancing Assumptions and Performance', *Journal of Index Investing*, 1, pp. 82-88.
Campbell J.Y. and Shiller R.J. (1988) 'Stock Prices, Earnings, and Expected Dividends', *Journal of Finance*, 43, pp. 661-676.
Chen D., Dempsey M. and Lajbcygier P. (2015) 'Is Fundamental Indexation Able to Time the Market? Evidence from the Dow Jones Industrial Average and the Russell 1000', *Journal of International Financial Markets, Institutions & Money*, 37, pp. 162-177.
Chow T., Hsu J.C., Kalesnik V. and Little B. (2011) 'A Survey of Alternative Equity Index Strategies', *Financial Analysts Journal*, 67, pp. 37-57.
Dash S., Guarino D. and Zeng L. (2010) 'Equal Weight Indexing: Seven Years Later', *Standard & Poor's indices: Research and Design*.
DiStefano C., Zhu M. and Mindrila D. (2009) 'Understanding and Using Factor Scores: Considerations for the Applied Researcher', *Practical Assessment, Research & Evaluation*, 14, pp. 1-11.
Fama E.F. and French K.R. (1992) 'The Cross-Section of Expected Stock Returns', *Journal of Finance*, 47, pp. 427-465.
Fama E.F. and French K.R. (1993) 'Common Risk Factors in the Returns on Stocks and Bonds', *Journal of Financial Economics*, 33, pp. 3-56.
Fama E.F. and French K.R. (1996) 'Multifactor Explanations of Asset Pricing Anomalies', *Journal of Finance*, 51, pp. 55-84.
Fama E.F. and French K.R. (1998) 'Value versus Growth: the International Evidence', *Journal of Finance*, 53, pp. 1975-1999.
Fama E.F. and French K.R. (2004) 'The Capital Asset Pricing Model: Theory and Evidence', *Journal of Economic Perspectives*, 18, pp. 25-46.
Fama E.F. and French K.R. (2012) 'Size, Value, and Momentum in International Stock Returns', *Journal of Financial Economics*, 105, pp. 457-472.
Ferreira R. and Krige J.D. (2011) 'The Application of Fundamental Indexing to the South African Equity Market for the Period 1996 to 2009', *Investment Analysts Journal*, 40, pp. 1-12.
Fisher G.S., Ronnie S. and Sheridan T. (2015) 'Decomposing Fundamental Indexation', Working Paper.
Ginis R. and Schoenfeld S.A. (2006) 'Are Alternatively Weighted Indexes Worth their Weight?', *Journal of Indexes*, 8, pp. 10-21.
Glass G.V. and Maguire T.O. (1966) 'Abuses of Factor Scores', *American Educational Research Journal*, 3, pp. 297-304.
Gorsuch R.L. (1983), *Factor Analysis*. Hillsdale, NJ: Erlbaum Associates.
Graham J.P. (2012) 'Comment on the Theoretical and Empirical Evidence of Fundamental Indexing', *Journal of Investment Management*, 10(1), p. 90.
Grinold R.C. and Kahn R.N. (1995), *Active Portfolio Management*. Chicago: Probus.
Grice J.W. and Harris J.R. (1998) 'A Comparison of Regression and Loading Weights for the Computation of Factor Scores', *Multivariate Behavioral Research*, 33, pp. 221-247.
Haugen R.A. and Baker N.L. (1991) 'The Efficient Market Inefficiency of Capitalization-Weighted Stock Portfolios', *Journal of Portfolio Management*, 17, pp. 35-40.
Hemminki J. and Puttonen V. (2008) 'Fundamental Indexation in Europe', *Journal of Asset Management*, 8, pp. 401-405.

Heng-Hsing H. (2013) 'Unlocking the Secrets of Fundamental Indexes: Size Effect or Value Effect? Evidence from Emerging Stock Markets', *Investment Management and Financial Innovations*, 10, pp. 48-63.
Hsu J.C. (2006) 'Cap-Weighted Portfolios are Sub-Optimal Portfolios', *Journal of Investment Management*, 4, pp. 1-10.
Hsu J.C. (2008) 'Why Fundamental Indexation Might – or Might Not – Work: A Comment', *Financial Analysts Journal*, 64, pp. 17-18.
Hsu J.C. and Campollo C. (2006) 'An Examination of Fundamental Indexation. New Frontiers in Index Investing', *Journal of Indexes*, 58, pp. 32-37.
Jun D. Malkiel B.G. (2008) 'New Paradigms in Stock Market Indexing', *European Financial Management*, 14, pp. 118-126.
Kaplan P.D. (2008a) 'Why Fundamental Indexation Might – or Might Not – Work', *Financial Analysts Journal*, 64, pp. 32-39.
Kaplan P.D. (2008b) 'Why Fundamental Indexation Might – or Might Not – Work: Author response', *Financial Analysts Journal*, 64, p. 18.
Lakonishok J., Shleifer A. and Vishny R.W. (1994) 'Contrarian Investment, Extrapolation, and Risk', *Journal of Finance*, 49, pp. 1541-1578.
Lintner J. (1965) 'The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets', *Review of Economics and Statistics*, 47, pp. 13-37.
Malkiel B.G. and Fama E.F. (1970) 'Efficient Capital Markets: A Review of Theory and Empirical Work', *Journal of Finance*, 25, pp. 383-417.
Markowitz H.M. (1952) 'Portfolio Selection', *Journal of Finance*, 7, pp. 77-91.
Mihm M. and Locarek-Junge H. (2010) 'Empirical Examination of Fundamental Indexation in the German Market', *Advances in Data Analysis, Data Handling and Business Intelligence*, Springer.
Mossin J. (1966) 'Equilibrium in a Capital Asset Market', *Econometrica*, 34, pp. 768-783.
Neukirch T. (2008) 'Alternative Indexing with the MSCI World Index', Working Paper.
Perold A.F. (2007) 'Fundamentally Flawed Indexing', *Financial Analysts Journal*, 63, pp. 31-37.
Ribeiro A.S. (2013) 'Índices Fundamentais: Exposição Metodológica e Avaliação da Aplicação ao Mercado Português', Working paper.
Roll R. (1977) 'A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory', *Journal of Financial Economics*, 4, pp. 129-176.
Sharpe W.F. (1964) 'Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk', *Journal of Finance*, 19, pp. 425-442.
Shimizu Y. and Tamura H. (2005) 'Global Fundamental Indices: Do They Outperform Market-Cap Weighted Indices on a Global Basis?', *Security Analysts Journal*, 43, pp. 32-46.
Siegel J. (14 June 2006) 'The Noisy Market Hypothesis', *Wall Street Journal*: A14.
Stotz O., Döhnert K. and Wanzenried G. (2010) 'Do Fundamental Indexes Produce Higher Risk-Adjusted Returns Than Market Cap Weighted Indexes? Evidence for European Stock Markets', *Financial Markets and Portfolio Management*, 24, pp. 219-243.
Summers L.H. (1986) 'Does the Stock Market Rationally Reflect Fundamental Values?', *Journal of Finance*, 41, pp. 591-601.
Tabner I. (2012) 'In Defense of Capitalization Weights: Evidence from the FT100 and S&P500 Indices', *European Financial Management*, 18, pp. 142-161.
Treynor J. (2005) 'Why Market-Valuation-Indifferent Indexing Works', *Financial Analysts Journal*, 61, pp. 65-69.
Treynor J. (2008) 'Fundamentally Flawed Indexing: Comments', *Financial Analysts Journal*, 64, p. 14.
Walkshäusl C. and Lobe S. (2010) 'Fundamental Indexing Around the World', *Review of Financial Economics*, 19, pp. 117-127.