Journal of Finance and Investment Analysis, Vol. 10, No. 2, 2021, 53-70
ISSN: 2241-0998 (print version), 2241-0996 (online)
https://doi.org/10.47260/jfia/1023
Scientific Press International Limited

Fundamentals-weighting vs. Capitalization-weighting: An Empirical Comparison

Guido Abate[1], Tommaso Bonafini and Pierpaolo Ferrari[2]

Abstract

Following the criticism surrounding capitalization-weighting, both academic and practitioner communities have developed alternative approaches to portfolio construction. We analyze one of these approaches, fundamentals-based weighting, which identifies the weights of portfolio constituents on the basis of their market multiples and accounting ratios. Our analysis is carried out on four fundamentals-weighted portfolios (FW) based on four different weighting variants, the capitalization-weighted portfolio (CW), and the equally-weighted (EW) portfolio, from January 2004 to December 2020, and in two subperiods (2004–2011 and 2011–2020). We find that in the first subperiod, the EW portfolio shows the highest risk-adjusted performance, followed by the FW portfolios. In contrast, in the second subperiod and in the period as a whole, the CW portfolio outperforms the other portfolios in terms of risk-adjusted performance. Overall, we conclude that both FW portfolios and the EW portfolio do not exhibit superior results when compared with the classic CW portfolio. Therefore, we have shown that FW and EW techniques provide superior risk-adjusted performance only during a period of exceptional financial turmoil. However, under normal conditions, they cannot be recommended as a rational investment strategy.

JEL classification numbers: G11, G14.
Keywords: Fundamental weighting, Capitalization weighting, Equal weighting, Value investing, Indexed investing.

---

1 Assistant Professor of Financial Markets and Institutions, Department of Economics and Management, University of Brescia, Brescia, Italy.
2 Affiliate Professor of Financial Markets and Institutions, Banking and Insurance Department, SDA Bocconi School of Management, Milan, Italy; Full Professor of Financial Markets and Institutions, Department of Economics and Management, University of Brescia, Brescia, Italy.

Article Info: *Received:* March 11, 2021. *Revised*: April 2, 2021.
Published online: April 21, 2021.

1. Introduction

Indexed investment management aims to replicate the performance of a benchmark index. Therefore, the weighting scheme employed plays a crucial role in investment performance. Traditionally, capitalization-weighted indices have been employed as benchmarks for passive investments (Goltz and Le Sourd, 2011). Proponents of these investments refer to financial theory to substantiate their policies and, in particular, to the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Lintner, 1965). Another advantage of value-weighted portfolios is their greater liquidity. In fact, the capitalization and, in particular, the free float of a company are highly correlated to the liquidity of its securities (Hsu, 2006).

However, value-weighting is not immune to criticism; therefore, both academic and practitioner communities have developed alternative approaches. Our research is focused on the empirical evaluation of one such alternative method, namely, fundamentals-based weighting developed by Arnott, Hsu, and West (2008).

The idea behind this weighting scheme is that value-weighted indices are efficient only if market prices are efficient, that is, if the Efficient Markets Hypothesis (EMH) holds true (Fama, 1970). Otherwise, this portfolio construction process systematically overweights overvalued stocks and underweights undervalued stocks.

To conduct our analysis, in addition to the fundamentals-weighted portfolio (FW), following the methodology of Arnott, Hsu and West (2008), we develop three other portfolios based on the use of different fundamental indicators. Both capitalization-weighted (CW) and equally-weighted (EW) portfolios are employed for comparison.

The remainder of this paper is organized as follows. Section 2 reviews the scientific literature on value investing and FW portfolios.

Section 3 describes the methodology of our empirical analysis, and the results are presented and discussed in Section 4. Finally, Section 5 presents our main findings and conclusions.

2. Literature Review

Alternative weighting techniques trace their origins to the critiques of the CAPM. Roll (1977) points out that the true market portfolio cannot be observed because it should take into account all risky assets and not only the listed securities. An index that is representative of a financial market does not consider either durable goods, real estate, human capital, or unlisted stocks. In essence, indices are not the true market portfolio, but only an approximation of it. Consequently, even if the market portfolio is efficient, the same cannot be said for an index.

One of the assumptions of the CAPM is the ability to borrow without limit at the risk-free rate and the ability to sell short without limit. If this last assumption is restricted, the property that a combination of mean-variance efficient portfolios is itself mean-variance efficient no longer follows, and the market portfolio is inefficient. Markowitz (2005) notes that an investor cannot incur unlimited debt at the risk-free rate, and concludes that the efficiency of the market portfolio does not always hold true. In fact, this conclusion by CAPM is based on the choice of

assumptions, which are unlikely to occur in the real world (Markowitz, 1983).
Moreover, a strand of literature has focused on the implications of the mispricing of assets, which occur if the EMH does not hold true empirically. These studies are known under the name of Noisy Market Hypothesis (NMH), because they assume the presence of "noise," that is, temporary market price deviations from the true intrinsic value of companies.

Treynor (2005) notes that the distribution of pricing errors can be symmetrical, with overvalued stocks counterbalancing undervalued stocks, but the lack of symmetry in market values caused by mispricing implies a potential underperformance for value-weighted portfolios. If market prices are inefficient, in the sense that they do not fully reflect a firm's situation, undervalued firms will have a market capitalization smaller than their true value, while overvalued firms will have a capitalization larger than their true value. Therefore, a CW portfolio will weight overvalued firms more and undervalued firms less. If we assume that the departure from true intrinsic value is not persistent over time, then the price correction will be tied to a negative performance of CW portfolios relative to capitalization-indifferent portfolios. It follows that it is preferable to use a weighting method that relies on proxies for a firm's intrinsic value, since CW portfolios are suboptimal.

Arnott (2005) measures the cost of "noise" in market prices. The idiosyncratic risk of a stock includes both company-related factors and "noise." Assuming that only one-eighth of idiosyncratic risk is due to mispricing, a CW portfolio performs 2% per year worse than portfolios that take into account the presence of "noise."

Based on historical data and assuming that only 5% of the variation in market prices is due to "noise," Hsu (2006) calculates that value-weighted portfolios have annual returns that are 1.78% lower than portfolios that reflect fundamentals. This means that, given a total transaction cost of 2%, a portfolio built to reflect fundamentals should have a turnover 89% higher than that of a CW portfolio to negate its advantage.

In contrast to NMH, Perold (2007) argues that CW portfolios are not characterized by lower returns than portfolios constructed using alternative methods. Perold criticizes NMH's thesis that higher returns can be achieved even without knowing the true values of securities, because market capitalization does not reveal whether a stock is undervalued or overvalued; thus, mispricing does not lead CW portfolios to invest more in overvalued stocks. The debate that ensued after this article has not yet reached a consensus on NMH (Ennis et al., 2008).

2.1 Value investing

One of the first studies to empirically identify the superiority of value investments is Basu (1977), who analyzes the period between 1957 and 1971 and finds that portfolios composed of securities with low price-earnings (P/E) ratios exhibit higher absolute returns and risk-adjusted returns than portfolios composed of securities with high P/E ratios. This behavior is inconsistent with that of EMH.

Oppenheimer (1984) focuses on portfolios formed by New York Stock Exchange (NYSE) and American Stock Exchange (AMEX) stocks whose issuers have the

following:
(1) a P/E ratio at least equal to twice the yield of AAA bonds;
(2) a book value higher than total debt;
(3) a dividend yield at least equal to two-thirds the AAA bond yield.
This study tests three different portfolios for both the NYSE and AMEX. The first is constructed using criteria (1) and (2), the second using criteria (2) and (3), and the third using all three criteria. These portfolios show a return in excess of the market portfolio in both markets.
Fama and French (1992) and Lakonishok, Shleifer and Vishny (1994) examine the returns of stocks characterized by a high book-to-market ratio (value stocks) relative to stocks with a low ratio (growth stocks). Both studies find that, over the past 30 years, the returns of value stocks are significantly higher than the returns of growth stocks. According to Fama and French (1992), the difference between the returns of value and growth stocks stems from the higher risk in the former. Lakonishok, Shleifer, and Vishny (1994) find that value strategies performed better than growth strategies in most recessions over the past 30 years. Therefore, they attribute the higher returns of value stocks to temporary undervaluation.
Arnott and Hsu (2008) show that the anomalies due to value and size factors are not linked to greater risk, but are derived from noise in the market prices of securities. Therefore, the factors identified by Fama and French (1992) can be easily explained by information inefficiency in stock prices. In addition, Arnott, Hsu, Liu, and Markowitz (2015) state that value stocks are more likely to exhibit negative noise. Therefore, these stocks are undervalued and have higher expected returns than can be justified by their level of risk. The same is true for stocks of thinly capitalized companies.
Value investing, as defined by the high minus low (HML) factor of Fama and French's model, has performed worse than growth investing since 2007. Arnott, Harvey, Kalesnik and Linnainmaa (2021) associate this unsatisfactory performance to two causes. First, the valuation of value stocks relative to growth stocks has plummeted in recent years, but an investment style can weaken in certain periods and thus show disappointing results. In addition, the HML factor does not capture the growing importance of intangible assets in companies, because the book-to-market ratio actually fails to capture a firm's investment in intangible capital (Amenc, Goltz and Luyten, 2020). As a result, many stocks are classified as growth stocks, even though they are not. To account for this, Arnott, Harvey, Kalesnik and Linnainmaa (2021) modify Fama and French's HML factor so that intangible capital is added to equity. To this end, they capitalize 100% of research and development (R&D) expenditures and 30% of administrative expenses. The factor thus modified, called iHML, performs significantly better than the HML factor.
Fama and French (2020) test whether the value factor premium has declined or even disappeared after the publication of Fama and French (1992). They measure that the mean value premium is higher in the July 1963–June 1991 subperiod of their analysis than in the July 1991–June 2019 subperiod.

However, even though the value premium has decreased in the last 28 years, the mean value premium remains positive in the full time frame.

2.2 Fundamental weighting

The hypothesis by Arnott, Hsu, and West (2008) is that CW indices are efficient only if market prices are efficient. If this does not hold true, then building and weighting an index based on fundamentals produces higher returns than capitalization weighting, which systematically overweights overvalued stocks and underweights undervalued stocks. To test their idea, they compare the S&P 500 with an index based on companies' total sales revenues. They find that this index has had a 2.5% higher annual return than the S&P 500 for more than 30 years.

Given that a portfolio formed using only one fundamental would not be representative of the market as a whole, Arnott, Hsu and West (2008) choose to focus on a combination of fundamental measures: sales revenues, net cash flow, dividends, and shareholders' equity. They are all measured as a mean over the previous five years, with the exception of shareholders' equity, which is measured at the date of portfolio construction. In their empirical analysis, they find that all FW portfolios beat CW indices since 1962.

However, Blitz and Swinkels (2008) criticize fundamental indexation, claiming that it is actually an active investment strategy disguised as an index. In support of this assertion, they argue that it is unclear who owns an FW portfolio in the market equilibrium condition, that indexing by fundamentals is not a buy-and-hold strategy, and that its construction requires subjective choices about the fundamentals to be used. Similarly, Amenc, Goltz and Le Sourd (2008) associate the stronger performance of FW portfolios with value tilt; therefore, they regard them as a cheaper version of an active strategy exploiting the value factor.

Balatti, Brooks and Kappou (2017) compare the risk-adjusted performances, net of transaction costs, of CW, EW, and FW indices over the period from January 1989 to September 2014. The three FW portfolios, one of which is constructed following Arnott, Hsu and West (2008), beat the CW portfolio. However, the EW portfolio shows better performance than the FW portfolio built following Arnott, Hsu and West (2008).

Miziolek and Zaremba (2017) analyze the performance of FW portfolios in the three largest European emerging markets, namely, Russia, Poland, and Turkey, over the period from 2002 to 2015, finding higher returns than CW indices after accounting for transaction costs.

Chow, Hsu, Kalesnik and Little (2011) measure that CW portfolios have an annual turnover of 8.4% when considering the global market (1987–2009 period) and 6.69% when considering the U.S. market (1964–2009 period), while FW portfolios have a 14.9% and 13.6% turnover, respectively. Li, Chow, Pickard and Garg (2019) find a similar turnover for FW portfolios. Therefore, the annual turnover of FW portfolios is significantly lower than that of active strategies, which is often around 100%.

3. The Methodology of the Empirical Analysis

Since the mid-2000s, the performance of value stocks relative to growth stocks has collapsed. Therefore, we aim to verify whether the superiority of FW portfolios over CW portfolios is still detectable in this context. Specifically, we will analyze the U.S. Large Cap market, both in the period from January 2004 to December 2020 and in two subperiods. The first subperiod begins in January 2004 and ends in December 2011, encompassing the 2008 crisis and the subsequent recovery, while the second subperiod begins in January 2012 and ends in December 2020. The second subperiod is characterized by strong growth in the stock market.

According to Arnott, Hsu and West (2008), FW portfolios perform worse than CW portfolios during periods of strong equity market growth. This is because the former have a value tilt relative to the latter. As a result, FW portfolios do not perform well during positive trends for growth stocks. For this reason, we would expect to find that, at least with respect to returns from 2012 onwards, CW portfolios perform better than FW portfolios.

To implement our analysis, we construct four FW portfolios using the constituents of the S&P 500 index as the investable universe. The four portfolios are defined as follows:

- Arnott: We follow the methodology of Arnott, Hsu and West (2008), using net revenues from sales, net cash flow, book value, and dividends to identify portfolio weights. We use the latest book value (BV) available at the time of portfolio construction, while for the other fundamental indicators, we employ their mean values over the previous five years in order to smooth out annual fluctuations and thus maintain portfolio weights at a stable level.
- Arnott Variant 1: Similar to the previous portfolio, it is built using a five-year average of the net revenues from sales, net cash flow, and dividends. However, we use a modified book value, calculated by adding research and development expenses (R&D) and selling, general, and administrative expenses (SG&A) to the latest book value. We add these costs to account for the increasing importance of intangible assets in firms during the last decades, as suggested by Arnott, Harvey, Kalesnik and Linnainmaa (2021).
- Arnott Variant 2: This portfolio is constructed using the latest modified book value (as defined above) and a five-year average of dividends and net income.
- Value: To identify the weights of this portfolio, we use the five-year average of dividends, the earnings-to-price ratio, and the book-to-price ratio. The earnings-to-price ratio is calculated using the company's five-year mean net income as the numerator, and the company's latest market value as the denominator. The book-to-price ratio is calculated using the latest modified book value as the numerator and the company's latest market value as the denominator. Thus, the companies with the highest weight in the Value portfolio are those with the highest dividends, earnings-to-price, and book-to-price ratios.

For comparison purposes, we also construct a CW portfolio and an EW portfolio using the same components as the FW portfolios. Therefore, the difference in the performance of the portfolios is only due to the calculation method of their weights and not their constituents.

Table 1 summarizes the methodology employed for each portfolio.

Portfolios were reconstructed once a year on December 31. Therefore, portfolio weights in the first month of year y+1 are calculated using the data available on December 31 of year y, and then they are allowed to drift following a buy-and-hold strategy for the remainder of year y+1. After testing different rebalancing frequencies, Arnott, Hsu and West (2008) found that rebalancing over shorter periods has almost no effect on FW portfolio performance, but results in a higher turnover.

Given that we evaluate the performance of the U.S. stock market between January 2004 and December 2020, and given that five years of data are required to calculate portfolio weights, our time series starts on December 31, 1999. The data were retrieved from Datastream and performances are total returns, that is, they take into account the reinvestment of dividends.

Table 1: Methodology of portfolio construction

<table>
<tr><th>Portfolio</th><th>Fundamental indicators</th><th>Weights calculation</th></tr>
<tr><td>Cap-weighted</td><td>- Market value (MV)</td><td>$$w_{Ci} = \frac{MV_i}{\sum_{i=1}^{N} MV_i}$$ where N is the number of portfolio components.</td></tr>
<tr><td>Equally-weighted</td><td>- Not applicable</td><td>$$w_{Ei} = \frac{1}{N}$$</td></tr>
<tr><td>Arnott</td><td>- Book value
- Net revenues from sales
- Dividends
- Net cash flow</td><td rowspan="4">$$w_{Fi} = \begin{cases} \dfrac{\sum_{j=1, j\neq d}^{M} w_{ij} + w_{id}}{M + 1} & if \;\; w_{id} > 0 \\ \dfrac{\sum_{j=1, j\neq d}^{M} w_{ij}}{M} & if \;\; w_d = 0 \end{cases}$$ with: $$w_{ij} = \frac{IND_{ji}}{\sum_{i=1}^{N} IND_{ji}}$$ w_{ij} is the weight of the company *i* with regard to the *j*-th indicator; *M* is the number of the indicators excluding dividends, w_{id} is the weight of company i with respect to dividends, and IND_{ji} is the value of the *j*-th fundamental indicator for company *i*. Short selling is not allowed; thus, negative weights are set equal to zero. The weights are then normalized to sum to one.</td></tr>
<tr><td>Arnott Variant 1</td><td>- Modified book value
- Net revenues from sales
- Dividends
- Net cash flow</td></tr>
<tr><td>Arnott Variant 2</td><td>- Modified book value
- Net income
- Dividends</td></tr>
<tr><td>Value</td><td>- Dividends
- Net income/MV
- Modified book value/MV</td></tr>
</table>

4. The Results of the Empirical Analysis

Table 2 provides the descriptive statistics of the portfolios' total monthly returns. Among the FW portfolios, only the mean return of the Value portfolio is higher than that of the CW portfolio, while the EW portfolio has the highest mean return among all portfolios. However, the Value and EW portfolios are characterized by a higher standard deviation. It is also evident that the distribution of the returns of each portfolio is asymmetrical. All portfolios show negative skewness and leptokurtosis. The EW and Value portfolios are most subject to excess kurtosis.

Table 2: Descriptive statistics of monthly returns (January 2004–December 2020)

Portfolio	Mean	Standard deviation	Maximum	Minimum	Median	Percentile		Skewness	Excess kurtosis
						5	95		
CW	0.87%	4.10%	12.81%	-16.19%	1.39%	-6.76%	7.05%	-0.63	1.90
EW	0.95%	4.70%	17.04%	-19.40%	1.22%	-7.34%	7.26%	-0.64	3.13
Arnott	0.83%	4.45%	14.14%	-17.22%	1.28%	-6.95%	7.41%	-0.65	2.55
Variant 1	0.85%	4.42%	14.42%	-17.21%	1.28%	-6.88%	7.41%	-0.63	2.54
Variant 2	0.83%	4.24%	12.30%	-16.27%	1.37%	-6.37%	7.13%	-0.73	2.40
Value	0.91%	4.72%	17.53%	-19.10%	1.27%	-7.25%	7.55%	-0.69	3.52

Correlations

	CW	EW	Arnott	Variant 1	Variant 2	Value
CW	1.0000					
EW	0.9727	1.0000				
Arnott	0.9823	0.9807	1.0000			
Variant 1	0.9834	0.9818	0.9999	1.0000		
Variant 2	0.9875	0.9742	0.9974	0.9974	1.0000	
Value	0.9657	0.9897	0.9909	0.9910	0.9854	1.0000

4.1 January 2004-December 2020

Table 3 shows the annualized mean returns, risk measures, and typical risk-adjusted performance measures of the portfolios for the period 2004–2020 (Basile and Ferrari, 2016). The CW portfolio is used as a benchmark for relative risk and performance measures. Table 3 also shows the final investment amount of 100 dollars in each portfolio at the beginning of the period. The cumulative performance of these investments is shown in Figure 1.

Table 3: Performance and risk measures (January 2004–December 2020)

Portfolio	Portfolio returns			Portfolio risk measures			
	Growth of $100	Arithmetic return	Geometric return	Standard deviation	Downside deviation	Beta	Tracking error volatility
CW	494.53	10.46%	9.86%	14.21%	9.64%	1.0000	0.00%
EW	546.89	11.37%	10.51%	16.27%	11.05%	1.1137	4.11%
Arnott	444.47	10.01%	9.17%	15.40%	10.62%	1.0645	3.03%
Variant 1	459.24	10.18%	9.38%	15.32%	10.50%	1.0597	2.91%
Variant 2	447.69	9.94%	9.22%	14.70%	10.18%	1.0216	2.34%
Value	499.90	10.86%	9.93%	16.36%	11.28%	1.1120	4.54%

Portfolio	Risk-adjusted performance measures				
	Sharpe ratio	Modigliani RAP	Sortino ratio	Information ratio	Treynor ratio
CW	0.6504	10.46%	0.9583	-	0.0924
EW	0.6245	10.09%	0.9197	0.2237	0.0912
Arnott	0.5707	9.32%	0.8280	-0.1484	0.0826
Variant 1	0.5857	9.54%	0.8542	-0.0933	0.0847
Variant 2	0.5936	9.65%	0.8572	-0.2195	0.0854
Value	0.5895	9.59%	0.8553	0.0893	0.0868

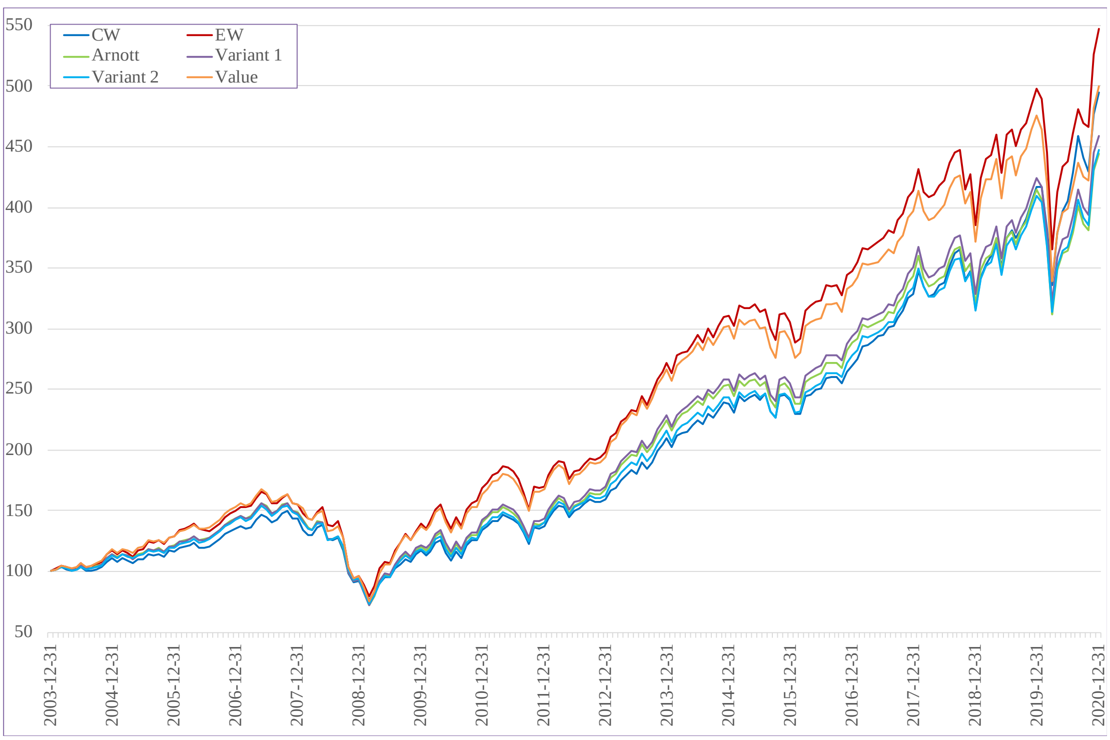

Figure 1: Cumulative performance (January 2004–December 2020)

The CW portfolio has the lowest risk in terms of standard and downside deviations. As a result, it has the highest Sharpe ratio, Modigliani's RAP, and Sortino ratio.
The EW portfolio is the best portfolio in terms of both arithmetic and geometric mean returns, followed by the Value portfolio. However, these portfolios also exhibit the highest absolute and asymmetric risk. Despite this, the EW portfolio is the second portfolio in terms of risk-adjusted performance measures owing to its mean return.
The Value portfolio has the highest mean return among FW portfolios. However, given its higher risk in terms of standard and downside deviations, it performs worse than the Arnott Variant 2 portfolio in terms of the Sharpe ratio, RAP, and Sortino ratio.
The Arnott portfolio has the lowest risk-adjusted performance measures. It also has the lowest arithmetic mean return, along with the Arnott Variant 2 portfolio. Variant 2, however, has the second lowest standard deviation and downside deviation of all portfolios. For this reason, it is the best FW portfolio in terms of the Sharpe ratio, RAP, and Sortino ratio.
Betas show that both the FW and EW portfolios are more aggressive than the CW portfolio. Among these portfolios, the Arnott Variant 2 portfolio is the least aggressive, with a beta close to 1.
The Treynor ratio is related to portfolio beta. According to this measure, the most efficient portfolio is again the CW portfolio, followed by the EW portfolio. However, using this measure of risk-adjusted performance, the ranking of the remaining portfolios changes and the Value portfolio emerges as the best of the four FW portfolios.
Analyzing tracking error volatility, the Arnott portfolio and its variants deviate less from the CW portfolio than the Value and EW portfolios. However, the Arnott portfolio and its variants have a negative information ratio, that is, they fail to produce an additional return per unit of additional risk compared to the cap-weighted portfolio. The EW and Value portfolios instead have a positive information ratio.

4.2 January 2004-December 2011

Table 4 provides the annualized returns, risk measures, and risk-adjusted performance measures of portfolios over the period from January 2004 to December 2011. In this subperiod, portfolios show lower returns and greater variability when compared to the second subperiod and the sample as a whole.

Table 4: Performance and risk measures (January 2004–December 2011)

Portfolio	Portfolio returns			Portfolio risk measures			
	Growth of $100	Arithmetic return	Geometric return	Standard deviation	Downside deviation	Beta	Tracking error volatility
CW	137.00	5.10%	4.01%	15.17%	11.20%	1.0000	0.00%
EW	169.75	8.26%	6.84%	18.02%	12.60%	1.1610	4.59%
Arnott	140.95	5.74%	4.38%	16.91%	12.26%	1.0958	3.43%
Variant 1	143.62	5.96%	4.63%	16.84%	12.14%	1.0918	3.34%
Variant 2	139.99	5.50%	4.29%	15.94%	11.75%	1.0357	2.69%
Value	167.45	8.08%	6.66%	17.95%	12.71%	1.1479	4.94%

Portfolio	Risk-adjusted performance measures				
	Sharpe ratio	Modigliani RAP	Sortino ratio	Information ratio	Treynor ratio
CW	0.2110	5.10%	0.2859	-	0.0320
EW	0.3531	7.26%	0.5048	0.6884	0.0548
Arnott	0.2269	5.34%	0.3130	0.1856	0.0350
Variant 1	0.2411	5.56%	0.3344	0.2566	0.0372
Variant 2	0.2256	5.32%	0.3060	0.1458	0.0347
Value	0.3444	7.13%	0.4865	0.6030	0.0538

In this subperiod, the CW portfolio has the lowest standard and downside deviations, but its geometric and arithmetic mean returns are lower than those of all the other portfolios. For this reason, the cap-weighted portfolio is the worst among the portfolios in terms of the Sharpe ratio, Sortino ratio, and Treynor ratio.

The EW portfolio is the best portfolio in terms of risk-adjusted performance, with the Value portfolio in second place. These two portfolios are also the riskiest. Their results in terms of risk-adjusted performance derive from their mean returns, which are significantly higher than those of the other portfolios.

Table 4 also illustrates that, over this period, all FW portfolios performed better in terms of risk-adjusted performance than the CW portfolio.

The relative risk, measured by beta and tracking error volatility, is higher than in the full period. In contrast to the full period, however, all portfolios have a positive information ratio. In particular, the information ratio of the EW portfolio is the highest, followed by the Value portfolio.

4.3 January 2012-December 2020

Table 5 provides the annualized returns, risk measures, and risk-adjusted performance measures of portfolios over the period of January 2012 to December 2020.

Table 5: Performance and risk measures (January 2012–December 2020)

Portfolio	Portfolio returns			Portfolio risk measures			
	Growth of $100	Arithmetic return	Geometric return	Standard deviation	Downside deviation	Beta	Tracking error volatility
CW	360.97	15.21%	15.33%	13.21%	8.11%	1.0000	0.00%
EW	322.17	14.14%	13.88%	14.57%	9.55%	1.0727	3.54%
Arnott	315.34	13.80%	13.61%	13.92%	9.01%	1.0353	2.62%
Variant 1	319.75	13.94%	13.79%	13.81%	8.89%	1.0295	2.44%
Variant 2	319.80	13.90%	13.79%	13.49%	8.65%	1.0102	1.95%
Value	298.54	13.33%	12.92%	14.87%	9.93%	1.0860	4.06%

Portfolio	Risk-adjusted performance measures				
	Sharpe ratio	Modigliani RAP	Sortino ratio	Information ratio	Treynor ratio
CW	1.1056	15.21%	1.8021	-	0.1461
EW	0.9288	12.88%	1.4168	-0.3031	0.1262
Arnott	0.9482	13.13%	1.4648	-0.5402	0.1275
Variant 1	0.9656	13.36%	1.4996	-0.5231	0.1295
Variant 2	0.9855	13.63%	1.5366	-0.6748	0.1316
Value	0.8562	11.92%	1.2814	-0.4630	0.1172

The CW portfolio emerges as the best portfolio in this subperiod. It is characterized by higher arithmetic and geometric mean returns and has a lower standard deviation and downside deviation. As a result, all measures of risk-adjusted performance dominate those of the other portfolios. Owing to this performance, the portfolio compensates for its inefficiency in the first subperiod and emerges as the best over the full time frame.

The EW portfolio has returns close to those of the Arnott portfolio and its variants, but with both higher absolute and asymmetric risk. For this reason, the Sharpe ratio, RAP, and Sortino ratio of the EW portfolio are lower than those of the Arnott portfolio and its variants. The EW portfolio is the best portfolio over the entire period in terms of absolute return only, because of its performance in the first subperiod.

The Value portfolio has the worst mean return and the highest standard and downside deviations among the portfolios from 2012 to 2020; therefore, in the second subperiod, it is the worst portfolio in terms of risk-adjusted performance. As a consequence, even though its risk-adjusted performance in the first subperiod is second only to that of the EW portfolio, over the entire time frame it performs worse than the EW, CW, and Variant 2 portfolios with regard to the Sharpe ratio, RAP, and Sortino ratio.

Both the EW and FW portfolios have a negative information ratio. Therefore, no portfolio can produce an additional return compared to the CW portfolio per unit of additional relative risk.

4.4 Hypothesis testing

We employ a one-tail paired t-test to verify the null hypothesis that the mean return of each portfolio is equal to the mean return of the CW portfolio. The results of this test are presented in Table 6.

Over the full period, the mean returns of all portfolios are not statistically different from those of the CW portfolio at a significance level of 10%. In the subperiod between January 2004 and December 2011, the mean returns of the EW portfolio and the Value portfolio are statistically higher than the mean return of the CW portfolio at a significance level of 5%. In the second subperiod, the mean returns of the FW portfolios are statistically lower than the mean return of the cap-weighted portfolio if we consider a significance level of 10%. The mean return of portfolio Variant 2 is statistically lower, even at the 5% level.

Furthermore, we employ the non-parametric Kolmogorov-Smirnov test to verify the null hypothesis that the distributions of returns of the FW and EW portfolios are statistically different from the distribution of returns of the CW portfolio. The results in Table 6 show that all portfolios have distributions that are statistically equal to that of the CW portfolio.

Table 6: T-tests and Kolmogorov-Smirnov tests on returns

Portfolio	Arithmetic return	Excess return	t-statistic	D-statistic
January 2004–December 2020				
EW	11.37%	0.92%	0.9223	0.0686
Arnott	10.01%	-0.45%	-0.6119	0.0294
Variant 1	10.18%	-0.27%	-0.3848	0.0343
Variant 2	9.94%	-0.51%	-0.9051	0.0392
Value	10.86%	0.41%	0.3682	0.0441
January 2004–December 2011				
EW	8.26%	3.16%	1.9470**	0.1250
Arnott	5.74%	0.64%	0.5250	0.0417
Variant 1	5.96%	0.86%	0.7258	0.0417
Variant 2	5.50%	0.39%	0.4125	0.0625
Value	8.08%	2.98%	1.7057**	0.0938
January 2012–December 2020				
EW	14.14%	-1.07%	-0.9094	0.0926
Arnott	13.80%	-1.42%	-1.6207*	0.0648
Variant 1	13.94%	-1.27%	-1.5694*	0.0648
Variant 2	13.90%	-1.32%	-2.0245**	0.0648
Value	13.33%	-1.88%	-1.3889*	0.0833

***, **, and * represent statistical significance levels of 1%, 5%, and 10%, respectively.

Table 7 provides the results of the robust test developed by Ledoit and Wolf (2008) to verify whether the difference between the two Sharpe ratios is statistically different from zero. Specifically, we test whether the difference between the Sharpe ratios of the FW and EW portfolios and the Sharpe ratio of the CW portfolio is

statistically significant.
If we consider a significance level of 5%, none of the portfolios show a Sharpe ratio statistically different from that of the CW portfolio, regardless of the period examined. However, using a significance level of 10%, the Sharpe ratio of the EW portfolio is statistically higher than that of the CW portfolio in the first period, while three FW portfolios (Arnott, Variant 1, Value) show Sharpe ratios statistically lower than that of the CW portfolio in the second period.

Table 7: Ledoit-Wolf test on Sharpe ratios

Portfolio	Sharpe ratio	Sharpe ratio$_p$–Sharpe ratio$_c$	t-statistic
January 2004–December 2020			
EW	0.6245	-0.0258	-0.4326
Arnott	0.5707	-0.0796	-1.5017
Variant 1	0.5857	-0.0646	-1.3069
Variant 2	0.5936	-0.0567	-1.1424
Value	0.5895	-0.0608	-0.8329
January 2004–December 2011			
EW	0.3531	0.1420	1.7131*
Arnott	0.2269	0.0159	0.2578
Variant 1	0.2411	0.0301	0.4965
Variant 2	0.2256	0.0146	0.2360
Value	0.3444	0.1334	1.4698
January 2012–December 2020			
EW	0.9288	-0.1768	-1.5250
Arnott	0.9482	-0.1575	-1.7218*
Variant 1	0.9656	-0.1401	-1.6480*
Variant 2	0.9855	-0.1201	-1.6070
Value	0.8562	-0.2494	-1.7308*

***, **, and * represent statistical significance levels of 1%, 5%, and 10%, respectively.

4.5 The CAPM and Fama-French three-factor model

Table 8 shows the results of the regressions of the CAPM and Fama-French three-factor model for both the full period and the two subperiods. The proxy of the market portfolio is identified in the CW portfolio, while the small minus big (SMB) and high minus low (HML) factors were retrieved from Kenneth R. French's Data Library. Given that the residuals of the regressions show autocorrelation in some cases, we employed the Newey-West estimator, which accounts for the autocorrelation and heteroscedasticity of residuals.

Table 8: CAPM and Fama-French three-factor model

Portfolio	CAPM			Fama-French three-factor model				
	Alpha	**Beta**	**R^2_{Adj}**	**Alpha**	**Beta**	**b_{SMB}**	**b_{HML}**	**R^2_{Adj}**
January 2004–December 2020								
EW	-0.01%	1.1137***	0.946	0.06%	1.0372***	0.2657***	0.1505***	0.970
Arnott	-0.09%	1.0645***	0.965	0.01%	1.0183***	0.0121	0.2492***	0.987
Variant 1	-0.07%	1.0597***	0.967	0.02%	1.0159***	0.0152	0.2321***	0.986
Variant 2	-0.06%	1.0216***	0.975	0.02%	0.9921***	-0.0346***	0.2041***	0.991
Value	-0.05%	1.1120***	0.933	0.07%	1.0304***	0.1515***	0.3011***	0.967
January 2004–December 2011								
EW	0.22%**	1.1610***	0.953	0.19%**	1.0670***	0.3230***	0.1165**	0.97
Arnott	0.03%	1.0958***	0.966	0.06%	1.0394***	-0.0266	0.2716***	0.99
Variant 1	0.05%	1.0918***	0.967	0.08%	1.0358***	-0.0138	0.2576***	0.99
Variant 2	0.02%	1.0357***	0.972	0.06%	0.9975***	-0.0807***	0.2410***	0.99
Value	0.21%*	1.1479***	0.939	0.22%**	1.0492***	0.1302***	0.3133***	0.96
January 2012–December 2020								
EW	**-0.18%**	1.0727***	0.945	-0.04%	1.0143***	0.2085***	0.1657***	0.972
Arnott	-0.16%	1.0353***	0.966	-0.02%	0.9987***	0.0403**	0.2220***	0.989
Variant 1	-0.14%	1.0295***	0.970	-0.01%	0.9963***	0.0361**	0.2027***	0.990
Variant 2	-0.12%	1.0102***	0.979	-0.02%	0.9872***	0.0027	0.1689***	0.993
Value	-0.26%	1.0860***	0.931	-0.07%	1.0211***	0.1606***	0.2773***	0.973

***, **, and * represent statistical significance levels of 1%, 5%, and 10%, respectively.

The alpha in the CAPM regression represents the excess return of the portfolio; its values are not statistically different from zero in all periods, with two exceptions. At the 5% significance level, only the EW portfolio exhibits a statistically positive alpha between January 2004 and December 2011. Additionally, if we consider a 10% level, the alpha of the Value portfolio is also statistically significant.

The Fama-French three-factor model shows that the portfolios are all positively related to the HML factor, that is, they have positive exposure to value stocks. In contrast, the exposure to the SMB factor of the Arnott and Variant 1 portfolios is not statistically different from zero in the first subperiod and the full period, while it is significant and positive in the second subperiod. A positive value of the SMB factor coefficient indicates that the portfolios are more exposed to stocks of small-cap companies. The exposure of the Arnott Variant 2 portfolio to the SMB factor is not statistically different from zero in the second subperiod and is significant and negative in the full period and the first subperiod. The EW and Value portfolios are positively exposed to the SMB factor in all periods. Moreover, these portfolios exhibit statistically non-zero alpha over 2004–2011; that is, they exhibit excess returns not captured by the market factor and the SMB and HML factors.

Between January 2004 and December 2011, the EW and Value portfolios performed the best (Table 4). Their performances may be explained partly by their positive alpha and partly by the fact that they are the only ones with a positive exposure to the SMB factor. In this subperiod, small-cap stocks outperformed the market. The two portfolios are also positively exposed to the value factor, which shows a slightly negative trend in the subperiod; however, that does not offset the better performance of the SMB factor.

All FW portfolios that were positively exposed to the HML factor performed worse than the CW portfolio from January 2012 to December 2020. Indeed, the cumulative returns of the HML factor declined over this time frame, impairing the performance of all FW portfolios. The decrease in the HML factor returns signals that growth stocks performed better than value stocks. This is reminiscent of the debate in the scientific literature regarding the performance of FW portfolios during strong positive trends in stock prices. In fact, Arnott, Hsu, and West (2008) point out that, in these periods, portfolios with a value tilt, such as portfolios weighted by fundamentals, fail to beat CW portfolios that more heavily weight the stocks of large-cap companies.

5. Conclusion

The scope of this study was to investigate whether FW portfolios exhibited better performance than the CW portfolio during the period from January 2004 to December 2020.

The performance of FW portfolios, both absolute and risk-adjusted, depends on the period under consideration. On the one hand, this analysis finds that the FW portfolios have higher risk-adjusted performance than the CW portfolio in the January 2004–December 2011 subperiod, which includes the 2008 crisis and the subsequent recovery. On the other hand, this study shows that, in the second subperiod, the CW portfolio is by far the best portfolio, both in terms of risk-adjusted and cumulative performance. This can be explained by the negative mean returns of the HML factor in this subperiod, since both the FW and EW portfolios are positively exposed to it. As a result, the empirical analysis shows that, in the whole period, both the FW portfolios and the EW portfolio do not exhibit higher risk-adjusted performances than the CW portfolio. In addition, the EW portfolio achieved the highest cumulative performance only at the expense of a higher risk.

Therefore, we have shown that FW and EW techniques provide superior risk-adjusted performance only during a period of financial turmoil. Under normal conditions, however, they cannot be recommended as a rational investment strategy. Accordingly, we conclude that both FW portfolios and the EW portfolio underperform in terms of risk-adjusted performance compared to the classic CW portfolio during the period under consideration, providing an empirical validation of the investment techniques based on the CAPM.

References

[1] Amenc, N., Goltz, F. and Le Sourd, V. (2008). A Comparison of Fundamentally Weighted Indices: Overview and Performance Analysis, EDHEC, Nice.

[2] Amenc, N., Goltz, F. and Luyten, B. (2020). Intangible Capital and the Value Factor: Has Your Value Definition Just Expired?, The Journal of Portfolio Management, 46(7), pp. 83 - 99.

[3] Arnott, R.D. (2005). What Cost Noise?, Financial Analysts Journal, 61(2), pp. 10 - 14.

[4] Arnott, R.D., Harvey, C.R., Kalesnik, V. and Linnainmaa, J.T. (2021). Reports of Value's Death May Be Greatly Exaggerated, Financial Analysts Journal, 77(1), pp. 44 - 67.

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

[6] Arnott, R.D., Hsu, J.C., Liu, J. and Markowitz, H.M. (2015). Can Noise Create the Size and Value Effects?, Management Science, 61(11), pp. 2569 - 2579.

[7] Arnott, R.D., Hsu, J.C. and West, J.M. (2008). The Fundamental Index: A Better Way to Invest, John Wiley & Sons, Hoboken.

[8] Balatti, M., Brooks, C. and Kappou, K. (2017). Fundamental Indexation Revisited: New Evidence on Alpha, International Review of Financial Analysis, 51(C), pp. 1 - 15.

[9] Basile, I. and Ferrari, P. (editors) (2016). Asset Management and Institutional Investors, Springer, Cham.

[10] Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price - Earnings Ratios: A Test of the Efficient Market Hypothesis, The Journal of Finance, 32(3), pp. 663 - 682.

[11] Blitz, D. and Swinkels, L. (2008). Fundamental Indexation: An Active Value Strategy in Disguise, Journal of Asset Management, 9(4), pp. 264 - 269.

[12] Chow, T.-M., Hsu, J.C., Kalesnik, V. and Little, B. (2011). A Survey of Alternative Equity Index Strategies, Financial Analysts Journal, 67(5), pp. 37 - 57.

[13] Ennis, R.M., Arnott, R.D., Markowitz, H.M., Treynor, J., Perold, A.F., Hsu, J.C. and Kaplan, P.D. (2008). Fundamental Disagreement, Financial Analysts Journal, 64(2), pp. 12 - 18.

[14] Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, The Journal of Finance, 25(2), pp. 383 - 417.

[15] Fama, E.F. and French, K.R. (1992). The Cross-Section of Expected Stock Returns, The Journal of Finance, 47(2), pp. 427 - 465.

[16] Fama, E.F. and French, K.R. (2020). The Value Premium, Fama-Miller Center for Research in Finance, University of Chicago Booth School of Business, working paper.

[17] Goltz, F. and Le Sourd, V. (2011). Does Finance Theory Make the Case for Capitalization-Weighted Indexing?, The Journal of Index Investing, 2(2), pp. 59 - 75.

[18] Hsu, J.C. (2006). Cap-Weighted Portfolios are Sub-Optimal Portfolios, Journal of Investment Management, 4(3), pp. 44 - 53.

[19] Lakonishok, J., Shleifer, A. and Vishny, R.W. (1994). Contrarian Investment, Extrapolation, and Risk, The Journal of Finance, 49(5), pp. 1541 - 1578.

[20] Ledoit, O. and Wolf, M. (2008). Robust Performance Hypothesis Testing with the Sharpe Ratio, Journal of Empirical Finance, 15(5), pp. 850 - 859.

[21] Li, F., Chow, T.-M., Pickard, A. and Garg, Y. (2019). Transaction Costs of Factor-Investing Strategies, Financial Analysts Journal, 75(2), pp. 62 - 78.

[22] Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, The Review of Economics and Statistics, 47(1), pp. 13 - 37.

[23] Markowitz, H.M. (1983). Nonnegative or Not Nonnegative: A Question about CAPMs, The Journal of Finance, 38(2), pp. 283 - 295.

[24] Markowitz, H.M. (2005). Market Efficiency: A Theoretical Distinction and So What?, Financial Analysts Journal, 61(5), pp. 17 - 30.

[25] Miziolek, T. and Zaremba, A. (2017). Fundamental Indexation in European Emerging Markets, Romanian Journal of Economic Forecasting, 20(1), pp. 23 - 37.

[26] Oppenheimer, H.R. (1984). A Test of Ben Graham’s Stock Selection Criteria, Financial Analysts Journal, 50(5), pp. 68 - 74.

[27] Perold, A.F. (2007). Fundamentally Flawed Indexing, Financial Analysts Journal, 63(6), pp. 31 - 37.

[28] 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(2), 1977, pp. 129 - 176.

[29] Sharpe, W.F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, The Journal of Finance, 19(3), 1964, pp. 425 - 442.

[30] Treynor, J. (2005).
Why Market-Valuation-Indifferent Indexing Works, Financial Analysts Journal, 61(5), pp. 65 - 69.