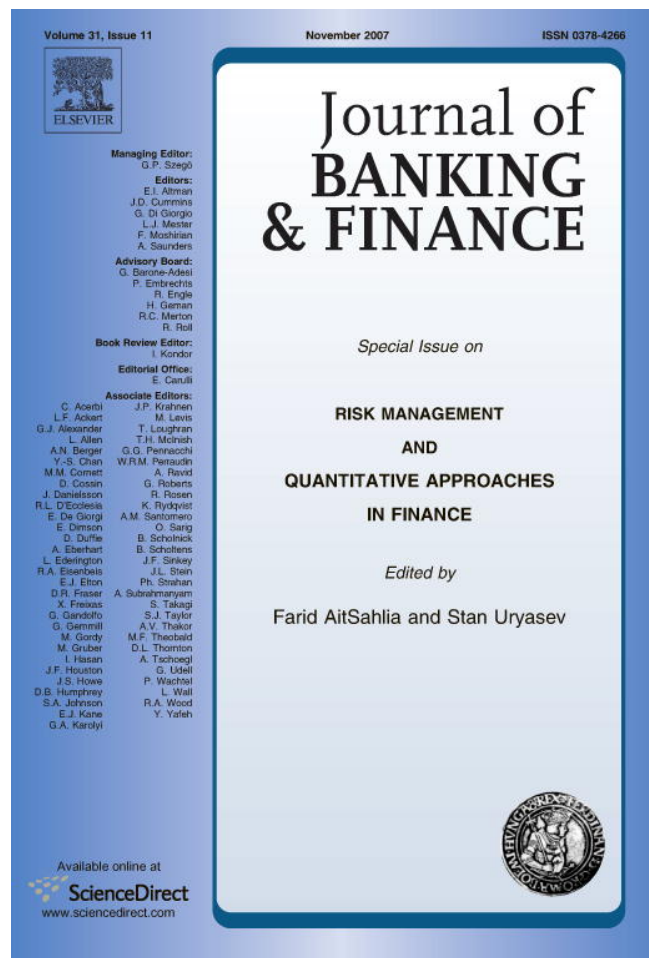




This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



# Fundamental indexation via smoothed cap weights

Chen Chen <sup>a,1</sup>, Rong Chen <sup>a,b,2</sup>, Gilbert W. Bassett <sup>c,\*</sup>

<sup>a</sup> *Department of Information and Decision Sciences, University of Illinois at Chicago, United States*

<sup>b</sup> *Department of Business Statistics and Econometrics, Peking University, PR China*

<sup>c</sup> *Department of Finance, University of Illinois at Chicago, United States*

Available online 20 April 2007

---

## Abstract

If prices of individual stocks are unbiased but noisy approximations to fundamental values, there will be a gap in returns between the standard cap-weighted market portfolio and the one based on fundamentals. The discrepancy occurs because, relative to fundamentals, cap-weights are too large (small) for stocks with positive (negative) deviations from fundamental values. It follows that the usual cap-weighted portfolio will underperform relative to the fundamental-based portfolio as long as prices revert to fundamental values. This has led Arnott et al. to propose new market indices based on a firm's fundamental size as measured by its revenues, number of employees, and so on. In this paper we follow the same principle but propose to estimate fundamental weights using a smoothed average of standard cap-weights. Since the putative excess returns of a fundamentals-weighted portfolio requires reversion to fundamental values, and because fundamental values are likely to change slowly, we can estimate current fundamentals by smoothing the time series of a stock's noisy prices. The determination of fundamental size in terms of accounting data is thereby replaced by a simple estimate based on price history. We derive expressions for expected returns of the market capitalization-based and fundamentals-based portfolios under various assumptions about (i) the random deviations from fundamental values and (ii) the change in fundamentals over time. We present empirical comparisons between portfolios and find the returns of the fundamentals-based portfolios exceed the standard indices by an amount comparable to the prior estimates that used accounting data to determine size.

© 2007 Elsevier B.V. All rights reserved.

---

\* Corresponding author.

E-mail address: [gib@uic.edu](mailto:gib@uic.edu) (G.W. Bassett).

<sup>1</sup> Chen Chen is at Allegiant Asset Management Company. The views expressed herein are those of the author and do not necessarily reflect the views of Allegiant.

<sup>2</sup> Rong Chen's research is sponsored in part by NSF grants DMS 0244541 and NIH grant R01 Gm068958.

JEL classification: G11

Keywords: Capitalization-weighted index; Smoothed capitalization-weighted index; Fundamental-based portfolio; Price inefficiency

---

## 1. Introduction

It has been long established that under the efficient market assumption and the framework of the capital asset pricing model, the “market portfolio” is mean–variance optimal. Based on the theory, portfolios with trillions of dollars in total value have been launched to track capitalization-weighted market indexes. However, there has also been ample evidence indicating that the market is not efficient. Such a violation of market efficiency makes the cap-weighted portfolio no longer mean–variance optimal. Accordingly, since the early 1980s *enhanced index funds* have been established to exploit this inefficiency. Many of the funds utilize a quantitative strategy that “tilts” the portfolio composition away from the commonly used capitalization-weighted market index to a slightly different composition expected to produce higher return for similar levels of risk. Riepe and Werner (1998) and diBartolomeo (2000) find evidence that enhanced index funds can in fact improve mean–variance efficiency. Other studies and empirical evidence have also shown that the capitalization-weighted market portfolio is inefficient so that a carefully structured portfolio can increase mean returns for comparable risk.

In a recent paper, Arnott et al. (2005) (hereafter AHM) proposed an alternative to the standard cap-weighted market index; see also Arnott et al. (2005) and Hsu and Campollo (2005). The new index is based on a model in which a stock price is a noisy approximation to its fundamental value. Under this model, a stock whose price is (randomly) above its fundamentals will have a cap weight that is too large, and conversely for negative deviations. The AHM index replaces the cap weight with an estimate of fundamental value based on accounting measures of value. A variety of accounting-based measures are suggested, including sales, earnings, cash flow and others. Most importantly, the AHM index will outperform the standard index if temporary deviations revert to fundamentals. In fact, AHM showed that their fundamental-based returns average 170 basis points over the equivalent cap-weighted index for the 42 years of the study. This idea has attracted notice in the practitioner community and has also led to related work by academicians. For example, the basic idea of AHM has been expanded in a recent work by Arnott et al. (2005), who demonstrate how noise alone can lead to portfolios with apparent Fama French factors, that is, excess returns to small size and value.

The fundamental-based indexation proposed by AHM is based on the notion that each stock has a fundamental value, which however differs by noise from its observed price. While a vast literature exists on the topic of fundamental values, there is no universally agreed framework in theory or practice, that defines the fundamental value. Indeed, it is usually suggested that the market’s best measure of fundamentals is the current price. It is on this basis that the cap-weighted index is mean–variance efficient. Apparent deviations from fundamentals however are well known. The excess variability of prices relative to fundamentals has been well documented, see Shiller (1981) for the original paper in a long line of research. For evidence on price reversals, and over and under reaction of prices, see Kim et al. (1991), Lo and MacKinlay (1988), De Bondt and Thaler (1985, 1987) and

Chopra et al. (1992). In this context, it is reasonable to assume that while the expected market price is the fundamental value, the actual price can differ from fundamentals because of market noise. We therefore assume the observed price is the fundamental value plus a zero-mean noise. The motivation for constructing the fundamental-value index, then, is that the “market portfolio” would be mean–variance optimal if only it could be based on an accurate measure of fundamental value.

The fundamental index proposed by AHM uses economic determinants to measure fundamentals. Because of the difficulty in finding a comprehensive measurement of fundamental value using economic determinants, they resort to simple accounting data as proxies for the fundamental value.<sup>3</sup>

In this paper we show how to implement the *idea* of fundamental indexation without directly measuring fundamental values. Our approach is motivated by the intuition behind AHM, but our implementation does not require estimation of fundamentals using accounting data. AHM show that excess returns of the fundamental index depends on both an accurate estimate of fundamentals but also on the reversion of prices to fundamental values. If prices do revert and if fundamentals change slowly, then we can statistically smooth past prices to estimate fundamentals and hence create a market index that outperforms the standard cap-weighted version.

Specifically, our approach is to use statistical methods to infer fundamental values from past prices. In the model the past price data is a (noisy) representation of fundamental economic determinants that is more accurate than the current price. Assuming the underlying fundamentals change slowly and the observed price is an unbiased but noisy signal for fundamentals, a moving average of past prices provides an estimate of fundamental price. A statistical estimate based on past prices can therefore be used to represent the fundamental value. Additionally our model-based approach identifies assumptions under which the fundamental-based index outperforms a cap-weighted index. This analysis is presented in the technical part of the paper.

Empirically we find our approach works about as well as the methods using accounting data reported by AHM. However our approach is easier to implement and can be justified analytically. The outperformance evidently occurs because deviations from fundamentals are temporary and prices do revert to fundamentals.

The basic idea is illustrated with a simple example in the next section. This is elaborated upon in Section 3 in the context of a more realistic model for prices as well as how fundamentals change over time. Section 4 describes portfolio construction and presents empirical results that show our proposed weighting produces returns in excess of cap weights comparable to those reported in AHM. Section 5 is a conclusion.

## 2. Two-stock, one-period illustration

Similar to Hsu (2004), we use a two stock, one-period model to illustrate the basic approach. Suppose there are two stocks *A* and *B* with the same number of shares outstanding. The outstanding shares do not change in the holding period. Suppose that prices at time 1 and 2 are

<sup>3</sup> For additional ways to measure the economic determinants of the fundamental value, see the price-earning decomposition approach of Campbell and Shiller (1988) and Campbell and Vuolteenaho (2004).

$$\begin{aligned} P_{A,1} &= F_A + \Delta, & P_{A,2} &= F_A, \\ P_{B,1} &= F_B - \Delta, & P_{B,2} &= F_B, \end{aligned}$$

where  $F_A$  and  $F_B$  are fundamental values of  $A$  and  $B$  respectively, and  $P_{A,t}$  and  $P_{B,t}$  are observed prices at time  $t$ . Prices differ by  $\Delta$  from fundamentals at time 1. There is an even, 50–50 chance that a stock is over (under) valued, but we do not know which one. In this simplified illustration, we assume that prices exactly revert to fundamentals at time 2. The one-period return for the two stocks is therefore

$$R_A = \frac{-\Delta}{F_A + \Delta}, \quad R_B = \frac{\Delta}{F_B - \Delta}.$$

The capitalization-based weights at time 1 are

$$w_A^{\text{cap}} = \frac{F_A + \Delta}{F_A + F_B}, \quad w_B^{\text{cap}} = \frac{F_B - \Delta}{F_A + F_B}.$$

So the one-period return of the cap-weighted portfolio is

$$R^{\text{cap}} = \frac{F_A + \Delta}{F_A + F_B} \cdot \frac{-\Delta}{F_A + \Delta} + \frac{F_B - \Delta}{F_A + F_B} \cdot \frac{\Delta}{F_B - \Delta} = 0.$$

Alternatively, consider the portfolio weighted by fundamentals at time 1. Assuming the fundamental values are known, the fundamentals-based weights would be,<sup>4</sup>

$$w_A^{\text{fund}} = \frac{F_A}{F_A + F_B}, \quad w_B^{\text{fund}} = \frac{F_B}{F_A + F_B},$$

so the return will be

$$R^{\text{fund}} = \frac{\Delta^2}{P_{A,1}P_{B,1}} > 0. \tag{1}$$

Hence, weights based on fundamentals earn greater returns than cap weights. The reason is evident: capitalization puts too much weight on the stock with  $\Delta > 0$  (a positive deviation from fundamental value) and too little on the one with  $\Delta < 0$  (a negative deviation from fundamental value). The magnitude of excess returns depends on the extent to which prices deviate from fundamentals,  $|\Delta|$ .

This illustration highlights two key assumptions behind the excess returns. The first is that fundamental values can be, at least roughly, estimated. The second is that prices revert to fundamentals after one period. In more complicated models, as well as empirically, analogous conditions need to hold so that fundamental weights outperform cap weights. Fundamental weighting will not be profitable if actual prices stray far from fundamentals for long periods. (Indeed, prices far from fundamentals for too long would jeopardize “fundamentals” as the underlying rationale for prices). In the more complicated models in the next section prices do differ from fundamentals (analogous to the  $\Delta$ -factor), but are unbiased and do not drift far from fundamentals. In this case fundamental-based weights outperform cap-weights. In the historical data we also find overall superior performance of fundamental weighting though there are periods where fundamental

<sup>4</sup> Clearly, if fundamental values really were exactly known, then returns would be maximized by investing all resources in the under-valued stock, if the portfolio is long only.

weights do under perform. These correspond to periods where (arguably) prices strayed from fundamentals for significant periods (e.g. 1998 and 1999).

To exploit the excess returns of fundamental weighting we need estimates for the unobserved fundamental values,  $F_A$  and  $F_B$ . AHM (2005) proposed estimates based on accounting and economic factors for each stock. We propose estimates based on past prices as described in the next two sections.

### 3. Theoretical justifications

Let  $P_{i,t}$  denote the price of the  $i$ th stock at time  $t$ . We assume that the market price is a noisy signal of an underlying fundamental value  $F_{i,t}$ . Similar to Hsu (2004), we assume the price of a stock can be decomposed into a part due to its fundamentals and a part due to deviations from fundamentals;

$$P_{i,t} = F_{i,t}(1 + \epsilon_{i,t}), \quad (2)$$

where  $F_{i,t}$  is the unobserved price based on fundamentals at time  $t$ , and  $\epsilon_{i,t}$  is a white noise with mean zero and variance  $\sigma_\epsilon^2$ , capturing the over/under-valuation at time  $t$ .

With the assumed noisy market price we first derive in Theorem 1 the expected return for a cap-weighted relative to the fundamental-weighted portfolio. Similar to the above we show that the expected return of the fundamental-based portfolio exceeds the cap-weighted portfolio by an amount that depends on the variance of the noise,  $\epsilon$ . We then consider estimates for fundamental weights based on a smoothed average of cap weights. We show in Theorems 2 and 3 how returns of the fundamental-based portfolio depend on the evolution of the  $F_{i,t}$  process and on the size of the estimation window used to estimate fundamentals.

Let

$$R_{i,t+1} = \frac{P_{i,t+1} - P_{i,t}}{P_{i,t}},$$

be the return of  $i$ th stock in the holding period  $[t, t+1]$  based on observed price  $P_{i,t}$  and

$$w_{i,t}^{\text{cap}} = \frac{P_{i,t}S_i}{\sum_k P_{k,t}S_k}, \quad w_{i,t}^{\text{fund}} = \frac{F_{i,t}S_i}{\sum_k F_{k,t}S_k}$$

be the portfolio weight of the  $i$ th stock in the corresponding cap-weighted and fundamental-weighted portfolio. Here  $S_i$  is the number of outstanding shares of stock  $i$ . Then returns of the cap-weighted portfolio and the fundamental-weight based portfolio are

$$R_{t+1}^{\text{cap}} = \sum_i w_{i,t}^{\text{cap}} R_{i,t+1}, \quad R_{t+1}^{\text{fund}} = \sum_i w_{i,t}^{\text{fund}} R_{i,t+1},$$

respectively.

The relationship between the expected returns of the conventional cap-weighted portfolio and fundamental-weighted portfolio is summarized in Theorem 1.

**Theorem 1.** *If stock prices follow Eq. (2), then the relationship between the expected returns of the cap- and fundamental-weighted portfolio is, to a second order approximation,<sup>5</sup> given by*

<sup>5</sup> As indicated in the proof, we ignore terms with orders higher than  $\sigma_\epsilon^2$ .

$$E(R_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) = \sigma_\epsilon^2 E \left[ (1 + R_{t+1}^{\text{aux}}) \left( 1 - \sum_i w_{i,t}^{\text{fund}^2} \right) \right], \quad (3)$$

where  $R_{i,t+1}^f = \frac{F_{i,t+1} - F_{i,t}}{F_{i,t}}$  and  $R_{t+1}^{\text{aux}} = \sum_i w_{i,t}^{\text{fund}} R_{i,t+1}^f$ .

**Proof.** See Appendix.  $\square$

Since  $\sum w_{i,t}^{\text{fund}} = 1$  and  $w_{i,t}^{\text{fund}} > 0$ , we have  $\sum w_{i,t}^{\text{fund}^2} < 1$ . And since  $R_{t+1}^{\text{aux}} \geq -1$ , we have  $E \left[ (1 + R_{t+1}^{\text{aux}}) \left( 1 - \sum_i w_{i,t}^{\text{fund}^2} \right) \right] \geq 0$ . Hence, the cap-weighted portfolio earns an expected return that is  $\sigma_\epsilon^2 E \left[ (1 + R_{t+1}^{\text{aux}}) \left( 1 - \sum_i w_{i,t}^{\text{fund}^2} \right) \right]$  less than the fundamental-weighted portfolio. We see that this return drag increases with the size of the price inefficiency ( $\sigma_\epsilon^2$ ) in the equity market.

AHM (2005) propose estimates of fundamental weights based on accounting and economic characteristics of firms. Instead, we estimate the fundamental weights by a simple moving average of capitalization weights for the previous  $T$  periods

$$\hat{w}_{i,t}^{\text{fund}} = \frac{1}{T} \sum_{\tau=t-T+1}^t w_{i,\tau}^{\text{cap}}. \quad (4)$$

The return of the corresponding portfolio is denoted by  $\hat{R}_{t+1}^{\text{fund}} = \sum_i \hat{w}_{i,t}^{\text{fund}} R_{i,t+1}$ . Note that in the special case where  $T = 1$ ,  $\hat{w}_{i,t}^{\text{fund}} = w_{i,t}^{\text{cap}}$ .

The next theorem demonstrates the relationship between the cap-weighted portfolio and the one based on estimated fundamental weights.

**Theorem 2.** If prices follow Eq. (2) and the fundamental weights are estimated using (4), then to a second order approximation, we have

$$\begin{aligned} E(\hat{R}_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) &= \sigma_\epsilon^2 \left( 1 - \frac{1}{T} \right) \sum_{i=1}^N E \left[ (1 + R_{i,t+1}^f) (w_{i,t}^{\text{fund}} - w_{i,t}^{\text{fund}^2}) \right] \\ &\quad + (1 + \sigma_\epsilon^2) \sum_{i=1}^N E(R_{i,t+1}^f B_1(w_{i,t}^{\text{fund}}, T)) - \sigma_\epsilon^2 \sum_{i=1}^N E(R_{i,t+1}^f B_2(w_{i,t}^{\text{fund}}, T)) \\ &\quad + \sigma_\epsilon^2 \sum_{i=1}^N E(R_{i,t+1}^f B_3(w_{i,t}^{\text{fund}}, T)), \end{aligned} \quad (5)$$

where the difference between current fundamentals and smoothed fundamentals in the estimation window is represented by,

$$\begin{aligned} B_1(w_{i,t}^{\text{fund}}, T) &= \frac{1}{T} \sum_{\tau=t-T+1}^t w_{i,\tau}^{\text{fund}} - w_{i,t}^{\text{fund}}, \\ B_2(w_{i,t}^{\text{fund}}, T) &= \frac{1}{T} \sum_{\tau=t-T+1}^t w_{i,\tau}^{\text{fund}^2} - w_{i,t}^{\text{fund}^2}, \\ B_3(w_{i,t}^{\text{fund}}, T) &= \frac{1}{T} \sum_{\tau=t-T+1}^t w_{i,\tau}^{\text{fund}} \sum_{k=1}^N w_{k,\tau}^{\text{fund}^2} - w_{i,t}^{\text{fund}} \sum_{k=1}^N w_{k,t}^{\text{fund}^2}. \end{aligned}$$

**Proof.** See Appendix.  $\square$

The theorem shows that the excess return on the portfolio with estimated fundamental weights relative to the conventional cap-weighted portfolio depends on the window size  $T$  and how fundamental values change over time. This theorem does not specify the underlying process of the fundamental value. In the following, we compare  $E(\hat{R}_{t+1}^{\text{fund}})$  and  $E(R_{t+1}^{\text{cap}})$  under three alternative specifications regarding the  $F_{i,t}$  process: (i) the fundamental price remains constant over time; (ii) the fundamental price is a constant plus a noise; (iii) the fundamental price is a random walk.

**Theorem 3.** Under the conditions of Theorem 2, the following three results hold:

- (i) If the fundamental price is a constant,  $F_{i,\tau} = C_i$ ,  $\tau = t - T + 1, \dots, t, t + 1$ , then to a second order approximation, we have

$$E(R_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) = \sigma_\epsilon^2 \left( 1 - \sum_{i=1}^N w_{i,t}^{\text{fund}^2} \right), \quad (6)$$

$$E(\hat{R}_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) = \left( 1 - \frac{1}{T} \right) \sigma_\epsilon^2 \left( 1 - \sum_{i=1}^N w_{i,t}^{\text{fund}^2} \right). \quad (7)$$

- (ii) If the fundamental price is a constant plus noise,  $F_{i,\tau} = C_i(1 + \delta_{i,\tau})$ ,  $\tau = t - T + 1, \dots, t, t + 1$ , where  $\delta_{i,\tau}$  is a white noise with mean 0 and variance  $\sigma_\delta^2$  and independent to the pricing noise  $\varepsilon_{i,t}$  in (2), then to a second order approximation, we have

$$E(R_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) = \sigma_\epsilon^2 \left( 1 - \sum_{i=1}^N w_{i,t}^c{}^2 \right), \quad (8)$$

$$E(\hat{R}_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) = \left( 1 - \frac{1}{T} \right) (\sigma_\epsilon^2 + \sigma_\delta^2) \left( 1 - \sum_{i=1}^N w_{i,t}^c{}^2 \right), \quad (9)$$

where  $w_{i,t}^c = \frac{C_i S_i}{\sum_k C_k S_k}$  is the weight based on the constant part in fundamental price.

- (iii) If the fundamental price is a random walk,  $F_{i,t} = F_{i,t-1}(1 + \delta_{i,t})$ , where  $\delta_{i,t}$  is a white noise with mean 0 and variance  $\sigma_\delta^2$  and independent to the pricing noise  $\varepsilon_{i,t}$  in (2), then to a second order approximation, we have

$$E(R_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) = \sigma_\epsilon^2 \left( 1 - \sum_{i=1}^N w_{i,0}^2 \right), \quad (10)$$

$$E(\hat{R}_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) = \sigma_\epsilon^2 \left( 1 - \frac{1}{T} \right) \left( 1 - \sum_{i=1}^N w_{i,0}^2 \right), \quad (11)$$

where  $w_{i,0}$  is the fundamental weight of stock  $i$  at the starting point.

**Proof.** See Appendix.  $\square$

**Remark 1.** In each case the fundamental-weighted portfolio earns an expected return above that based on capitalization weights. The portfolio with *estimated* fundamental

weights outperforms the expected return for the cap-weighted portfolio and earns an expected return close to the one based on exact fundamental weights. The magnitude of the excess return is determined by the variance  $\sigma_\epsilon^2$ , and the estimation window size  $T$ .

**Remark 2.** If the fundamental price process is a constant plus noise (the second case), then the observed price of a stock can be decomposed into,

$$P_{i,\tau} = C_i(1 + \sigma_\delta)(1 + \sigma_\epsilon).$$

This means that when estimating fundamental weights with a moving average, we not only reduce the variance of the “mis-pricing” noise, but also the variance of the disturbance in the fundamental process.

## 4. Market returns with estimated fundamental weights

### 4.1. Estimated fundamental weights

We consider a monthly “market” index consisting of the 1000 largest cap stocks between January 1962 and December 2003. Data is obtained for each month from CRSP. The roster of stocks and portfolio weights are generated from the data on the last trading day of the prior month.

Let  $K_{i,t} = P_{i,t}S_{i,t}$  denote the capitalization of stock  $i$  at time  $t$ , where  $P_{i,t}$  is the price of stock at time  $t$  and  $S_{i,t}$  is the outstanding shares of stock  $i$ . Let  $G_t$  is the total capitalization of the largest 1000 stocks at time  $t$ . Then for any stock (regardless of whether it is among the largest 1000 stocks) in the CRSP data set at time  $t$ , we define  $w_{i,t} = K_{i,t}/G_t$ . This represents the capitalization of stock  $i$  relative to the total capitalization of the top 1000 stocks at time  $t$ . When stock  $i$  is indeed among the largest 1000 stocks, then  $w_{i,t}$  is the conventional capitalization weight. A portfolio of the largest 1000 stocks, with weight  $w_{i,t}$ , determines the cap-weighted index (CWI).

To estimate the fundamental weights we use the median of the relative cap-weights within a fixed window of the immediate past. The estimation window,  $Z$ , consists of the previous  $12 \times n$  months. The estimated weight at time  $t$  is then given by<sup>6</sup>

$$\hat{w}_{i,t} = \text{median}\{w_{i,t}, w_{i,t-1}, \dots, w_{i,t-Z}\}.$$

We use the median instead of the mean for robustness purposes. The resulting portfolio index is referred to as the smoothed cap-weighted index (SCWI). The portfolio is held for 12 months,  $t + 1$  until  $t + 12$ , and rebalanced every 12 months. In order to account for possible January effects, we consider two alternative rebalancing cycles, one that starts with January and the other with June.

New firms, or mergers and spin-offs of existing firms that result in new names, will not have a full set of observations in the  $12 \times n$  months estimation window. In this case we use the median of available observations. The associated weights are labeled, SW. We found that alternate ways of handling new firms produced similar returns (results available on

<sup>6</sup> Note that  $w_{i,t}$  is defined for all stocks in CRSP not just those in the top 1000. The stocks in the SCWI are determined by the 1000 largest cap stocks at time  $t$ , even if some were not in the top 1000 inside the look-back window.

request). This is not surprising because the proportion of young firms in the top 1000 is relatively small and cap-weighted returns are most strongly influenced by the largest firms which tend to be long-lived.

When companies merge, acquire other firms (without changing names), or undergo other significant corporate events, the historical weights will not be representative of current size or value. To deal with such cases we consider a modified estimate for fundamentals, denoted by SW2 with associated index returns denoted by SCWI2. The modified scheme identifies firms whose capitalization changes more by than 20% in a month. Such

Table 1  
Ten largest companies

| Company name                                | SW (%)  | SW2 (%) | CW (%)  |
|---------------------------------------------|---------|---------|---------|
| <i>Top 10 SW stocks at the end of 2003</i>  |         |         |         |
| GENERAL ELECTRIC CO                         | 3.09    | 3.25    | 2.42    |
| MICROSOFT CORP                              | 2.32    | 2.44    | 2.77    |
| EXXON MOBIL CORP                            | 2.16    | 2.27    | 2.35    |
| WAL MART STORES INC                         | 1.70    | 1.79    | 2.23    |
| COCA COLA CO                                | 1.63    | 1.71    | 1.09    |
| INTEL CORP                                  | 1.60    | 1.68    | 1.03    |
| MERCK & CO INC                              | 1.49    | 1.57    | 1.27    |
| INTERNATIONAL BUSINESS MACHS COR            | 1.35    | 1.42    | 1.31    |
| AMERICAN TELEPHONE & TELEG CO               | 1.31    | 0.00    | 0.20    |
| PHILIP MORRIS COS INC                       | 1.30    | 1.37    | 0.84    |
|                                             | SW2 (%) | SW (%)  | CW (%)  |
| <i>Top 10 SW2 stocks at the end of 2003</i> |         |         |         |
| GENERAL ELECTRIC CO                         | 3.25    | 3.09    | 2.42    |
| MICROSOFT CORP                              | 2.44    | 2.32    | 2.77    |
| EXXON MOBIL CORP                            | 2.27    | 2.16    | 2.35    |
| WAL MART STORES INC                         | 1.79    | 1.70    | 2.23    |
| COCA COLA CO                                | 1.71    | 1.63    | 1.09    |
| INTEL CORP                                  | 1.68    | 1.60    | 1.03    |
| MERCK & CO INC                              | 1.57    | 1.49    | 1.27    |
| INTERNATIONAL BUSINESS MACHS COR            | 1.42    | 1.35    | 1.31    |
| PHILIP MORRIS COS INC                       | 1.37    | 1.30    | 0.84    |
| PROCTER & GAMBLE CO                         | 1.23    | 1.17    | 1.12    |
|                                             | CW (%)  | SW (%)  | SW2 (%) |
| <i>Top 10 CW stocks at the end of 2003</i>  |         |         |         |
| MICROSOFT CORP                              | 2.77    | 2.32    | 2.44    |
| GENERAL ELECTRIC CO                         | 2.42    | 3.09    | 3.25    |
| EXXON MOBIL CORP                            | 2.35    | 2.16    | 2.27    |
| WAL MART STORES INC                         | 2.23    | 1.70    | 1.79    |
| PFIZER CHAS & CO INC                        | 1.88    | 1.13    | 1.19    |
| COMMERCIAL CREDIT CO                        | 1.78    | 0.71    | 0.74    |
| JOHNSON & JOHNSON                           | 1.60    | 1.16    | 1.22    |
| AMERICAN INTERNATIONAL GROUP INC            | 1.51    | 1.01    | 1.06    |
| INTERNATIONAL BUSINESS MACHS COR            | 1.31    | 1.35    | 1.42    |
| MERCK & CO INC                              | 1.27    | 1.49    | 1.57    |

SW and SW2 are the corresponding portfolio weights in SCWI and SCWI2; CW are the portfolio weights in CWI.

In the top table stocks are sorted by largest SW. In the middle table stocks are sorted by largest SW2. In the bottom table stocks are sorted by largest CW.

a large change is taken to signal a merger, spin-off, or acquisition and in this case we treat the firm as if it were *new*.

Weights are compared in Table 1. It shows the largest 10 stocks at the end of 2003 according to the alternative measures of size. The table shows that stocks and weights in the top 10 can differ. As expected, SCWI tends to put greater weight on old economy stocks like GE (3.09% versus 2.42% in CWI), while CWI favors new economy stocks like Microsoft (2.77% versus 2.32% in SCWI).

The table also shows that corporate events matter. For example, AT&T is in the top 10 according to SW but (correctly) near the bottom of the 1000 largest current-cap firms. The large weight occurs because the firm named “AT&T” in the estimation window is very different than the one existing at the end of 2003 (due to spin-offs from the parent AT&T between 1993 and 2003). Without any indication in the database and without a change in company name, it is difficult to detect such data anomalies. As a result the smoothed weight without correction becomes a poor estimate of AT&T’s current fundamental value. The refined weighting of SW2 that defaults to a new firm when there has been a large change in current cap correctly filters AT&T out of the top 10 and gives it less weight in the market index.

#### 4.2. Portfolio performance

Table 2 shows the attributes of the SCWI and SCWI2 indices reconstituted on Dec 31 and Jun 30 and the benchmark CWI for the 42 years from Jan 1, 1962 to Dec 31, 2003.

The table shows that: (1) the indexes rebalanced on January and June have similar performance; (2) performance improves with longer estimation windows; (3) the SCWIs tend to have significantly lower return volatility; (4) most importantly, SCWI and SCWI2 outperform the benchmark CWI by about 1.00% per annum.

Table 3 presents a year-by-year comparison. It shows periods when SCWI performed better and worse than the usual cap weights. During the boom of the 1990s, for example,

Table 2  
Return attributes of SCWI SCWI2 and CWI

|            | December |         |        | June     |         |        |
|------------|----------|---------|--------|----------|---------|--------|
|            | Mean (%) | Std (%) | Beta   | Mean (%) | Std (%) | Beta   |
| CWI        | 10.84    | 15.33   |        | 10.92    | 15.42   |        |
| SCWI 12m   | 10.81    | 15.51   | 1.0067 | 10.84    | 15.46   | 0.9999 |
| SCWI 24m   | 10.81    | 15.43   | 0.9979 | 11.10    | 15.38   | 0.9904 |
| SCWI 60m   | 11.41    | 14.94   | 0.9616 | 11.57    | 14.96   | 0.9562 |
| SCWI 120m  | 11.85    | 14.57   | 0.9317 | 11.91    | 14.62   | 0.9296 |
| SCWI2 12m  | 10.82    | 15.50   | 1.0060 | 10.85    | 15.44   | 0.9990 |
| SCWI2 24m  | 10.83    | 15.40   | 0.9962 | 11.12    | 15.36   | 0.9893 |
| SCWI2 60m  | 11.42    | 14.90   | 0.9587 | 11.56    | 14.94   | 0.9548 |
| SCWI2 120m | 11.86    | 14.55   | 0.9302 | 11.92    | 14.61   | 0.9289 |

SCWI is the smoothed cap-weighted index, SCWI2 is merge and acquisition events treated smoothed cap-weighted index and CWI is the current cap-weighted index as described in Section 4.1. December and June correspond to alternative rebalancing dates. The 12m, ..., 120m correspond to alternative estimation windows. Mean and standard deviation are annualized. The Beta comes from regressing SCWI, SCWI2 on CWI for each month, 1962–2003.

Table 3  
Yearly returns of SCWI, SCWI2, CWI and S&P 500 index

| Year       | SCWI          |           | SCWI2         |           | CWI      |         | SPRTN   |
|------------|---------------|-----------|---------------|-----------|----------|---------|---------|
|            | December/120m | June/120m | December/120m | June/120m | December | June    |         |
| \$1 growth | 90.87         | 93.08     | 91.62         | 93.30     | 56.89    | 58.48   | 15.54   |
| Average    | 11.85%        | 11.91%    | 11.86%        | 11.92%    | 10.84%   | 10.92%  | 7.70%   |
| 1962       | −4.83%        | −4.44%    | −4.82%        | −4.43%    | −8.19%   | −7.96%  | −10.72% |
| 1963       | 20.17%        | 20.50%    | 20.17%        | 20.5%     | 19.58%   | 19.79%  | 17.88%  |
| 1964       | 15.64%        | 15.56%    | 15.63%        | 15.55%    | 15.26%   | 15.09%  | 12.35%  |
| 1965       | 12.68%        | 12.56%    | 12.67%        | 12.56%    | 12.25%   | 12.83%  | 9.05%   |
| 1966       | −9.70%        | −9.73%    | −9.7%         | −9.73%    | −8.46%   | −9.23%  | −13.36% |
| 1967       | 21.98%        | 21.93%    | 21.98%        | 21.93%    | 23.04%   | 23.85%  | 19.16%  |
| 1968       | 14.85%        | 14.77%    | 14.84%        | 14.77%    | 13.03%   | 12.89%  | 8.16%   |
| 1969       | −10.58%       | −10.14%   | −10.58%       | −10.14%   | −10.39%  | −9.08%  | −11.22% |
| 1970       | 8.17%         | 8.86%     | 8.18%         | 8.87%     | 3.68%    | 3.84%   | 1.98%   |
| 1971       | 14.12%        | 13.91%    | 14.12%        | 13.91%    | 15.51%   | 15.17%  | 11.14%  |
| 1972       | 16.07%        | 15.64%    | 16.07%        | 15.64%    | 17.11%   | 16.05%  | 14.79%  |
| 1973       | −16.51%       | −17.47%   | −16.51%       | −17.49%   | −18.41%  | −20.07% | −17.95% |
| 1974       | −26.78%       | −26.51%   | −26.8%        | −26.52%   | −27.87%  | −27.87% | −32.19% |
| 1975       | 37.68%        | 37.80%    | 37.67%        | 37.79%    | 32.22%   | 34.78%  | 29.15%  |
| 1976       | 25.32%        | 25.27%    | 25.37%        | 25.31%    | 24.61%   | 24.39%  | 18.51%  |
| 1977       | −3.99%        | −2.77%    | −3.97%        | −2.73%    | −3.21%   | −3.87%  | −11.76% |
| 1978       | 11.27%        | 11.39%    | 11.18%        | 11.29%    | 11.59%   | 11.56%  | 2.32%   |
| 1979       | 17.07%        | 17.27%    | 16.97%        | 17.2%     | 17.91%   | 19.34%  | 12.50%  |
| 1980       | 23.84%        | 23.98%    | 23.79%        | 23.92%    | 28.67%   | 29.75%  | 24.67%  |
| 1981       | 1.22%         | 1.16%     | 1.23%         | 1.17%     | −2.06%   | −0.14%  | −9.44%  |
| 1982       | 21.37%        | 22.06%    | 21.37%        | 22.16%    | 17.95%   | 18.92%  | 15.45%  |
| 1983       | 22.80%        | 22.04%    | 22.66%        | 22.03%    | 23.78%   | 20.14%  | 16.49%  |
| 1984       | 8.64%         | 9.18%     | 8.64%         | 9.04%     | 5.57%    | 5.65%   | 2.26%   |
| 1985       | 29.27%        | 28.35%    | 29.12%        | 28.17%    | 28.35%   | 27.73%  | 24.25%  |
| 1986       | 17.07%        | 16.60%    | 17.46%        | 17.01%    | 18.15%   | 16.53%  | 15.18%  |
| 1987       | 10.63%        | 10.39%    | 10.42%        | 10.26%    | 9.11%    | 8.44%   | 6.65%   |
| 1988       | 18.09%        | 18.10%    | 18.31%        | 18.36%    | 16.46%   | 17.14%  | 12.23%  |
| 1989       | 26.14%        | 25.92%    | 25.37%        | 25.57%    | 26.44%   | 26.46%  | 25.04%  |
| 1990       | −3.77%        | −3.64%    | −3.64%        | −3.51%    | −2.98%   | −3.21%  | −5.25%  |
| 1991       | 25.65%        | 26.27%    | 25.46%        | 26.13%    | 27.52%   | 29.79%  | 24.69%  |
| 1992       | 9.03%         | 9.49%     | 9.13%         | 9.61%     | 7.28%    | 8.82%   | 4.63%   |
| 1993       | 13.72%        | 14.00%    | 13.63%        | 13.86%    | 10.28%   | 9.89%   | 7.00%   |
| 1994       | 2.26%         | 2.37%     | 2.12%         | 2.3%      | 1.04%    | 1.13%   | −1.03%  |
| 1995       | 30.98%        | 30.86%    | 30.96%        | 30.84%    | 31.82%   | 31.22%  | 29.83%  |
| 1996       | 20.03%        | 19.95%    | 20.39%        | 20.19%    | 20.95%   | 21.35%  | 19.12%  |
| 1997       | 28.74%        | 28.77%    | 28.89%        | 28.35%    | 29.21%   | 28.99%  | 28.47%  |
| 1998       | 21.14%        | 21.19%    | 21.23%        | 21.26%    | 24.06%   | 25.06%  | 26.10%  |
| 1999       | 12.37%        | 13.34%    | 12.52%        | 13.58%    | 16.97%   | 19.73%  | 18.75%  |
| 2000       | 4.32%         | 3.19%     | 5.88%         | 4.61%     | −10.15%  | −11.01% | −9.32%  |
| 2001       | −2.25%        | −1.68%    | −2.7%         | −2.73%    | −9.80%   | −9.72%  | −12.07% |
| 2002       | −16.27%       | −16.34%   | −16.1%        | −15.77%   | −20.21%  | −23.82% | −24.35% |
| 2003       | 29.88%        | 30.29%    | 29.65%        | 29.76%    | 27.60%   | 28.31%  | 24.22%  |

SCWI is the smoothed capitalization-weighted index; SCWI2 is the events treated smoothed capitalization-weighted index; and CWI is the current capitalization-weighted index as described in Section 4.1. SPRTN is the return of the S&P 500 index; December and June are alternative rebalancing months. The look-back window is 120 months; “\$1 growth” is the value of \$1 invested in 1962; “average” is the annualized mean monthly return.

Table 4  
Bear and bull market return characteristics

|             |               | Return (%) | Volatility (%) |
|-------------|---------------|------------|----------------|
| <i>Bear</i> |               |            |                |
| SCWI        | December/120m | −17.10     | 18.07          |
|             | June/120m     | −17.05     | 18.08          |
| CWI         | December      | −22.25     | 19.89          |
|             | June          | −22.40     | 20.13          |
| SPRTN       |               | −24.02     | 18.68          |
| <i>Bull</i> |               |            |                |
| SCWI        | December/120m | 19.44      | 12.61          |
|             | June/120m     | 19.50      | 12.69          |
| CWI         | December      | 19.44      | 12.82          |
|             | June          | 19.58      | 12.84          |
| SPRTN       |               | 15.95      | 13.06          |

SCWI is the smoothed capitalization-weighted index and CWI is the current capitalization-weighted index as described in Section 4.1. SPRTN is the return of S&P 500 index; December and June are alternative rebalancing months; the look-back window is 120 months; return and volatility values are annualized. A bull market is defined simplistically (and ex post) by a 20% rally from the previous low and a bear market by 20% decline from the previous high.

SCWI lagged CWI. In 1997 CWI (rebalanced on December 31) had returns of 29.21% while the return for the SCWI lagged at 28.74%. In contrast, in 2000, the CWI (based on what turned out to be inflated weights of over-valued firms) lost more than 10%, while the SCWI was up 3%.

Table 4 shows portfolio performances in bear and bull markets. A bull market is defined by a 20% rally from the previous low and a bear market by 20% decline from the previous high. The table shows that most of the improved performance of the SCWI occurred in bear markets.

## 5. Summary

When prices vary randomly around fundamental values there will be a gap between cap-weighted and fundamental-weighted indices. If, in addition, prices are eventually determined by fundamentals, then a portfolio based on good estimates of fundamental weights will outperform a standard cap-weighted index. Under these assumptions we have presented expressions for expected returns. Empirical results confirm that returns of the estimated fundamentals-based portfolio using smoothed cap weights exceed standard indices by an amount comparable to prior estimates that used accounting data to determine size.

## Acknowledgements

We are grateful for very helpful suggestions and comments from Rob Arnott, Jason Hsu, Pat Rocco and two anonymous referees.

## Appendix

The holding period return for a single stock is

$$R_{i,t+1} = \frac{P_{i,t+1} - P_{i,t}}{P_{i,t}} = \frac{P_{i,t+1}}{P_{i,t}} - 1 = \frac{F_{i,t+1}(1 + \epsilon_{i,t+1})}{F_{i,t}(1 + \epsilon_{i,t})} - 1,$$

which can be approximated using 2nd order Taylor expansion by

$$R_{i,t+1} = \frac{F_{i,t+1}}{F_{i,t}} (1 + \epsilon_{i,t+1} - \epsilon_{i,t} + \epsilon_{i,t}^2 - \epsilon_{i,t}\epsilon_{i,t+1}) - 1.$$

Substituting the fundamental return  $R_{i,t}^f = \frac{F_{i,t+1} - F_{i,t}}{F_{i,t}}$  in the above expression, we have

$$R_{i,t+1} = (R_{i,t+1}^f + 1)(1 + \epsilon_{i,t+1} - \epsilon_{i,t} + \epsilon_{i,t}^2 - \epsilon_{i,t}\epsilon_{i,t+1}) - 1.$$

By the definition of cap-weighting, the weight for stock  $i$  in a cap-weighted portfolio is

$$w_{i,t}^{\text{cap}} = \frac{P_{i,t}S_i}{\sum_k P_{k,t}S_k} = \frac{F_{i,t}(1 + \epsilon_{i,t})S_i}{\sum_k F_{k,t}(1 + \epsilon_{k,t})S_k} = \frac{F_{i,t}S_i(1 + \epsilon_{i,t})}{(\sum_k F_{k,t}S_k) \left(1 + \frac{\sum_k F_{k,t}S_k \epsilon_{k,t}}{\sum_k F_{k,t}S_k}\right)}.$$

By definition of the fundamental weights  $w_{i,t}^{\text{fund}} = \frac{F_{i,t}S_i}{\sum_k F_{k,t}S_k}$ , we have

$$w_{i,t}^{\text{cap}} = w_{i,t}^{\text{fund}} \frac{(1 + \epsilon_{i,t})}{1 + \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t}},$$

which can be approximated by

$$w_{i,t}^{\text{cap}} = w_{i,t}^{\text{fund}} \left( 1 + \epsilon_{i,t} - \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} + \left( \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} \right)^2 - \epsilon_{i,t} \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} \right).$$

**Proof of Theorem 1.** Recall that  $E(\epsilon_{i,t}) = 0$ ,  $V(\epsilon_{i,t}) = \sigma_{i,t}^2$ , and  $E(\epsilon_{i,t}\epsilon_{i,s}) = 0$ ,  $t \neq s$ ,  $E(\epsilon_{i,t}\epsilon_{j,t}) = 0$ ,  $i \neq j$ . Using the notations given in Section 3, cap-weighted portfolio return is  $R_{t+1}^{\text{cap}} = \sum_i w_{i,t}^{\text{cap}} R_{i,t+1}$  and the fundamental-weighted portfolio return is  $R_{t+1}^{\text{fund}} = \sum_i w_{i,t}^{\text{fund}} R_{i,t+1}$ . The auxiliary portfolio based on fundamental weights and calculated with fundamental prices is  $R_{t+1}^{\text{aux}} = \sum_i w_{i,t}^{\text{fund}} R_{i,t+1}^f$ .

Then the expected return for fundamental-weighted portfolio is

$$\begin{aligned} E(R_{t+1}^{\text{fund}}) &= \sum_{i=1}^N E(w_{i,t}^{\text{fund}} R_{i,t+1}) \\ &= \sum_{i=1}^N E\left(w_{i,t}^{\text{fund}} ((R_{i,t+1}^f + 1)(1 + \epsilon_{i,t+1} - \epsilon_{i,t} + \epsilon_{i,t}^2 - \epsilon_{i,t}\epsilon_{i,t+1}) - 1)\right) \\ &= E(R_{t+1}^{\text{aux}}) + \sigma_{\epsilon}^2 E(R_{t+1}^{\text{aux}}) + \sigma_{\epsilon}^2 \sum_{i=1}^N E(w_{i,t}^{\text{fund}}) \\ &= E(R_{t+1}^{\text{aux}}) + \sigma_{\epsilon}^2 E(R_{t+1}^{\text{aux}}) + \sigma_{\epsilon}^2. \end{aligned}$$

Notice that we drop terms with orders greater than  $\sigma^2$ .

On the other hand, the expected return for cap-weighted portfolio is

$$\begin{aligned}
 E(R_{t+1}^{\text{cap}}) &= E\left(\sum_{i=1}^N w_{i,t}^{\text{cap}} R_{i,t+1}\right) \\
 &= \sum_{i=1}^N E\left(w_{i,t}^{\text{fund}} \left(1 + \epsilon_{i,t} - \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} + \left(\sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t}\right)^2 - \epsilon_{i,t} \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t}\right) R_{i,t+1}\right) \\
 &= \sum_{i=1}^N E\left[w_{i,t}^{\text{fund}} R_{i,t+1}^f \left(1 + \epsilon_{i,t} - \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} + \left(\sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t}\right)^2 - \epsilon_{i,t} \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t}\right)\right. \\
 &\quad \times (1 - \epsilon_{i,t} + \epsilon_{i,t}^2 + \epsilon_{i,t+1} - \epsilon_{i,t} \epsilon_{i,t+1}) \\
 &\quad \left. - w_{i,t}^{\text{fund}} \left(1 + \epsilon_{i,t} - \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} + \left(\sum_k w_{k,t}^{\text{fund}^2} \epsilon_{k,t}\right) - \epsilon_{i,t} \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t}\right)\right. \\
 &\quad \left. \times (\epsilon_{i,t} - \epsilon_{i,t}^2 - \epsilon_{i,t+1} + \epsilon_{i,t} \epsilon_{i,t+1})\right] \\
 &= E(R_{t+1}^{\text{aux}}) + \sigma_\epsilon^2 E\left(R_{t+1}^{\text{aux}} \sum_{i=1}^N w_{i,t}^{\text{fund}^2}\right) + \sigma_\epsilon^2 \sum_{i=1}^N E(w_{i,t}^{\text{fund}^2}).
 \end{aligned}$$

Compare it with the formula of fundamental-weighted portfolio returns, we have

$$E(R_{t+1}^{\text{fund}}) - E(R_{t+1}^{\text{cap}}) = \sigma_\epsilon^2 E\left[(1 + R_{t+1}^{\text{aux}}) \left(1 - \sum_i w_{i,t}^{\text{fund}^2}\right)\right]. \quad \square$$

**Proof of Theorem 2.** We take the average of the past capitalization weights as the estimated fundamental weights

$$\hat{w}_{i,t}^{\text{fund}} = \frac{1}{T} \sum_{\tau=t-T+1}^t w_{i,\tau}^{\text{cap}} = w_{i,t}^{\text{cap}} + \frac{1}{T} \sum_{\tau=t-T+1}^t (w_{i,\tau}^{\text{cap}} - w_{i,t}^{\text{cap}}).$$

Then the expected return for the fundamental estimation-based portfolio becomes

$$\begin{aligned}
 E(\hat{R}_{t+1}^{\text{fund}}) &= E\left(\sum_{i=1}^N \hat{w}_{i,t}^{\text{fund}} R_{i,t+1}\right) \\
 &= \sum_{i=1}^N E\left(\left(w_{i,t}^{\text{cap}} + \frac{1}{T} \sum_{\tau=t-T+1}^t (w_{i,\tau}^{\text{cap}} - w_{i,t}^{\text{cap}})\right) R_{i,t+1}\right) \\
 &= E(R_{t+1}^{\text{cap}}) + \sum_{i=1}^N E\left(\frac{1}{T} \sum_{\tau=t-T+1}^t (w_{i,\tau}^{\text{cap}} - w_{i,t}^{\text{cap}}) R_{i,t+1}\right).
 \end{aligned}$$

# Plugging

$$w_{i,\tau}^{\text{cap}} = w_{i,\tau}^{\text{fund}} \left( 1 + \epsilon_{i,\tau} - \sum_k w_{k,\tau}^{\text{fund}} \epsilon_{k,\tau} + \left( \sum_k w_{k,\tau}^{\text{fund}} \epsilon_{k,\tau} \right)^2 - \epsilon_{i,\tau} \sum_k w_{k,\tau}^{\text{fund}} \epsilon_{k,\tau} \right),$$

$$w_{i,t}^{\text{cap}} = w_{i,t}^{\text{fund}} \left( 1 + \epsilon_{i,t} - \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} + \left( \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} \right)^2 - \epsilon_{i,t} \sum_k w_{k,t}^{\text{fund}} \epsilon_{k,t} \right),$$

$$R_{i,t+1} = ((R_{i,t+1}^f + 1)(1 + \epsilon_{i,t+1} - \epsilon_{i,t} + \epsilon_{i,t}^2 - \epsilon_{i,t} \epsilon_{i,t+1}) - 1)$$

into the above equation, taking expectations and dropping terms with higher orders of  $\sigma_\epsilon^2$ , we have

$$\begin{aligned} E(\hat{R}_{t+1}^{\text{fund}}) &= E(R_{t+1}^{\text{cap}}) + \sigma_\epsilon^2 \left( 1 - \frac{1}{T} \right) \sum_{i=1}^N E[(1 + R_{i,t+1}^f)(w_{i,t}^{\text{fund}} - w_{i,t}^{\text{fund}^2})] \\ &\quad + (1 + \sigma_\epsilon^2) \sum_{i=1}^N E \left( R_{i,t+1}^f \frac{1}{T} \sum_{\tau=t-T+1}^t (w_{i,\tau}^{\text{fund}} - w_{i,t}^{\text{fund}}) \right) \\ &\quad - \sigma_\epsilon^2 \sum_{i=1}^N E \left( R_{i,t+1}^f \frac{1}{T} \sum_{\tau=t-T+1}^t (w_{i,\tau}^{\text{fund}^2} - w_{i,t}^{\text{fund}^2}) \right) \\ &\quad + \sigma_\epsilon^2 \sum_{i=1}^N E \left( R_{i,t+1}^f \frac{1}{T} \sum_{\tau=t-T+1}^t \left( w_{i,\tau}^{\text{fund}} \sum_{k=1}^N w_{k,\tau}^{\text{fund}^2} - w_{i,t}^{\text{fund}} \sum_{k=1}^N w_{k,t}^{\text{fund}^2} \right) \right). \quad \square \end{aligned}$$

**Proof of Theorem 3.** Based on Theorems 1 and 2, we only need to compute returns and weights under different scenarios.

For the first scenario, fundamental price of each stock  $F_{i,t}$  is a constant over time. Fundamental weight is also a constant, hence, the bias function of fundamental weights  $B_i(w_{i,t}^{\text{fund}}, T)$  equals 0 for all  $i$ . Since the fundamental return in a holding period for a single stock  $R_{i,t+1}^f = 0$ , then the return on the auxiliary portfolio  $R_{t+1}^{\text{aux}} = 0$ . Plugging these results into the formulas in Theorem 1 and we have Eqs. (6) and (7).

In the second scenario, fundamental weights and fundamental holding period returns are functions of the noise term  $\delta_{i,t}$ . We have

$$w_{i,t}^{\text{fund}} = \frac{F_{i,t} S_i}{\sum_k F_{k,t} S_k} = \frac{C_i S_i (1 + \delta_{i,t})}{\sum_k C_k S_k (1 + \delta_{k,t})}.$$

Let  $w_{i,t}^c = \frac{C_i S_i}{\sum_k C_k S_k}$ . Then the fundamental weight becomes

$$w_{i,t}^{\text{fund}} = w_{i,t}^c (1 + \delta_{i,t}) \left( 1 - \sum_k w_{k,t}^c \delta_{k,t} + \left( \sum_k w_{k,t}^c \delta_{k,t} \right)^2 \right),$$

$$R_{i,t+1}^f = \frac{F_{i,t+1}}{F_{i,t}} - 1 \approx \frac{C_i (1 + \delta_{i,t+1})}{C_i (1 + \delta_{i,t})} - 1 = (1 + \delta_{i,t+1})(1 - \delta_{i,t} + \delta_{i,t}^2) - 1.$$

The auxiliary portfolio return has the form

$$\begin{aligned} R_{t+1}^{\text{aux}} &= \sum_{i=1}^N w_{i,t}^{\text{fund}} R_{i,t+1}^f \\ &= \sum_{i=1}^N w_{i,t}^c (1 + \delta_{i,t}) \left( 1 - \sum_k w_{k,t}^c \delta_{k,t} + \left( \sum_k w_{k,t}^c \delta_{k,t} \right)^2 \right) ((1 + \delta_{i,t+1})(1 - \delta_{i,t} + \delta_{i,t}^2) - 1). \end{aligned}$$

Based on the above expressions, the expected B-functions in Theorem 2 become

$$\begin{aligned} E(R_{i,t+1}^f B_1(w_{i,t}^{\text{fund}}, T)) &\approx \frac{1}{T} \sum_{\tau=t-T+1}^{t-1} (w_{i,\tau}^c - w_{i,t}^c)^2 \sigma_\delta^2, \\ E(R_{i,t+1}^f B_2(w_{i,t}^{\text{fund}}, T)) &\approx \frac{1}{T} \sum_{\tau=t-T+1}^{t-1} (w_{i,\tau}^c{}^2 - 2w_{i,t}^c{}^3 + w_{i,t}^c{}^2) \sigma_\delta^2, \\ E(R_{i,t+1}^f B_3(w_{i,t}^{\text{fund}}, T)) &\approx \frac{1}{T} \sum_{\tau=t-T+1}^{t-1} (w_{i,\tau}^c{}^3 - 3w_{i,t}^c{}^4 + w_{i,t}^c{}^3) \sigma_\delta^2. \end{aligned}$$

Dropping the terms with degrees higher than  $\sigma_\epsilon^2$  and  $\sigma_\delta^2$ , we have Eqs. (8) and (9).

In the third scenario, fundamental weights are

$$w_{i,t}^{\text{fund}} = \frac{F_{i,t} S_i}{\sum_k F_{k,t} S_k} = \frac{F_{i,0} S_i (1 + \delta_{i,t})(1 + \delta_{i,t-1}) \cdots (1 + \delta_{i,1})}{\sum_k F_{k,0} S_k (1 + \delta_{k,t})(1 + \delta_{k,t-1}) \cdots (1 + \delta_{k,1})}.$$

Dropping the high order terms of  $\delta_{i,t}$ , we can approximate  $(1 + \delta_{i,t})(1 + \delta_{i,t-1}) \cdots (1 + \delta_{i,1})$  by  $1 + \sum_{\tau=1}^t \delta_{i,\tau}$ . Hence the fundamental weights is approximated by

$$\begin{aligned} w_{i,t}^{\text{fund}} &\approx \frac{F_{i,0} S_i (1 + \sum_{\tau=1}^t \delta_{i,\tau})}{\sum_k F_{k,0} S_k (1 + \sum_{\tau=1}^t \delta_{k,\tau})} \\ &= w_{i,0} \left( 1 + \sum_{\tau=1}^t \delta_{i,\tau} \right) \left( 1 - \sum_k w_{k,0} \sum_{\tau=1}^t \delta_{k,\tau} + \left( \sum_k w_{k,0} \sum_{\tau=1}^t \delta_{k,\tau} \right)^2 \right). \end{aligned}$$

The fundamental return is:

$$R_{i,t+1}^f = \frac{F_{i,t+1}}{F_{i,t}} - 1 = \frac{F_{i,t}(1 + \delta_{i,t+1})}{F_{i,t}} - 1 = \delta_{i,t+1}.$$

Then,

$$\begin{aligned} (R_{i,t+1}^f) &= 0, \\ E(R_{i,t+1}^f B_1(w_{i,t}^{\text{fund}}, T)) &= 0, \\ E(R_{i,t+1}^f B_2(w_{i,t}^{\text{fund}}, T)) &= 0, \\ E(R_{i,t+1}^f B_3(w_{i,t}^{\text{fund}}, T)) &= 0. \end{aligned}$$

Substituting the fundamental weights and fundamental single stock return into equations in Theorems 1 and 2, we have Eqs. (10) and (11).  $\square$

## References

- Arnott, R.D., Hsu, J., Markowitz, H., 2005. Does noise create the Fama-French value and size factors? Research Affiliates Working Paper #WP275402.
- Arnott, Robert D., Hsu, Jason C., Moore, Philip, 2005. Fundamental indexation. *Financial Analysts Journal* 61 (2), 83–99.
- Campbell, J.Y., Shiller, R., 1988. Stock prices, earnings and expected dividends. *Journal of Finance* 43, 661–676.
- Campbell, J.Y., Vuolteenaho, T., 2004. Inflation illusion and stock prices. *American Economic Review, Papers and Proceedings* 94, 19–23.
- Chopra, Navin, Lakonishok, Josef, Ritter, Jay R., 1992. Measuring abnormal performance. *Journal of Financial Economics* 31, 235–268.
- De Bondt, Werner F.M., Thaler, Richard, 1985. Does the stock market overreact? *The Journal of Finance* XL (3), 793–805.
- De Bondt, Werner F.M., Thaler, Richard, 1987. Further evidence on investor overreaction and stock market seasonality. *The Journal of Finance* XLII (3), 557–581.
- diBartolomeo, Dan, 2000. The enhanced index fund as an alternative to indexed equity management. Working paper, Northfield Information Services.
- Hsu, Jason C., 2004. Cap-weighted portfolios are sub-optimal portfolios. Working paper, Research Affiliates, LLC.
- Hsu, Jason C., Campollo, Carmen, 2005. An examination of fundamental indexation. Working paper, Research Affiliates, LLC and FTSE Americas.
- Kim, Myung.J., Nelson, Charles R., Startz, Richard, 1991. Mean reversion in stock prices? A reappraisal of the empirical evidence. *Review of Economic Studies* 58, 515–528.
- Lo, Andrew W., MacKinlay, A. Craig, 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. *The Review of Financial Studies* 1 (1), 41–66.
- Riepe, Mark W., Werner, Matthew D., 1998. Are enhanced index mutual funds worthy of their name? *Journal of Investing* 7 (2, Summer), 6–15.
- Shiller, Robert J., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71 (3), 421–436.