Vrije Universiteit Brussel

Myths about Fundamental Indexing

De Moor, Lieven; Liu, Fang; Sercu, Piet

Published in:
Investment Analysts Journal

DOI:
10.1080/10293523.2018.1483050

Publication date:
2018

License:
Unspecified

Document Version:
Accepted author manuscript

Link to publication

Citation for published version (APA):
De Moor, L., Liu, F., & Sercu, P. (2018). Myths about Fundamental Indexing. *Investment Analysts Journal*, *47*(4), 304-326. https://doi.org/10.1080/10293523.2018.1483050

Copyright
No part of this publication may be reproduced or transmitted in any form, without the prior written permission of the author(s) or other rights holders to whom publication rights have been transferred, unless permitted by a license attached to the publication (a Creative Commons license or other), or unless exceptions to copyright law apply.

Take down policy
If you believe that this document infringes your copyright or other rights, please contact openaccess@vub.be, with details of the nature of the infringement. We will investigate the claim and if justified, we will take the appropriate steps.

Download date: 10. Apr. 2024

Full-text version

Myths about Fundamental Indexing[*]

Lieven De Moor,[†] Fang Liu,[‡] Piet Sercu[§]

Thursday 5th July, 2018

Abstract

'Fundamental Indexing' starts from the observation that in a value-weighted portfolio, any overpricing affects the stock's portfolio weight upward and its typical return downward, and vv; but on average the 'drag' on the portfolio's expected return caused by this negative interaction is avoided if weights are based, instead, on accounting-based instruments for true value. We find that the drag effect actually is statistically and economically unimportant. Our empirical work avoids regression-based alphas, which are flawed by demonstrable instabilities in the exposures.

JEL classification: G11, G12, G14
Key words: portfolio management, drag, pricing errors.

[*]A shorter version was published in the Investment Analyst Journal in 2018. This version still contains a discussion about FI's theoretical validity, a more extensive review of the empirical literature, the Multi-CAPM alpha analysis, and about 10 more footnotes. An earlier version was called "An anatomy of fundamental indexing". We thank Anup Basu, Gaël Imad Eddine, Aurobindo Ghosh, Hening Liu, Vestman Roine, Martina Vandebroek, an anonymous but inspired referee for the JFQA, and participants in presentations at the authors' home institutions and at the Indian Finance Conference (Bangalore), the BRForum (Antwerp), a Paris Finance Meeting, a Symposion on Banking and Finance (Shanghai), the Inquire Europe Conference (Munich) for useful comments. Any remaining errors are the authors' sole responsibility.

[†]Vrije Universiteit Brussel, Faculty of Economic and Social Sciences and Solvay Business School; lieven.de.moor@vub.ac.be; +32 498 616463.

[‡]Beijing Foreign Studies University; fang.liu@bfsu.edu.cn

[§]KU Leuven, Faculty of Economics and Business; piet.sercu@kuleuven.be; +32 16 326756.

Introduction

The familiar recommendation of portfolio theory in efficient markets is to mimic the market portfolio, *i.e.* to choose weights proportional to the firms' relative market capitalizations. Proponents of the Fundamental Indexing investment strategy (henceforth FI) disagree: while more money should be allocated to big stocks, weights should be based not on market capitalization but on instrumental-variable-like alternatives to true value, such as relative book value, payout, cash flow, or sales. The reason is that, in markets with noisy prices, a cap-based weight will over-weight overpriced stocks (the very stocks that tend to underperform) and will under-weight the underpriced stocks (the stocks that tend to overperform). This negative interaction between weighting errors and subsequent return drags down the returns. If the market-value weight is replaced by an instrument, the expectation is that this drag is avoided or at least reduced, so that typically the average return should improve.

FI is not 'fundamentally flawed', *pace* Perold (2007) or Kaplan (2008): as discussed in Section 1.1, the existence of drag follows straight from a model where, generalising Roll (1984), the market price deviates from the true value by an error orthogonal on that true value. The effect was independently derived by *e.g.* Brennan and Wang (2010), even though they do not refer to FI or its terminology. Perold (2007)'s critique is insightful but aims at a claim not made by FI, as we point out in Section 1.2, while Kaplan's critique addresses an assumption that is possibly just imprecise wording by Hsu (2006) and certainly inessential to FI.

Saying that drag is conceptually plausible does not imply that its economic relevance is non-trivial, though. Brennan and Wang's estimate would imply a potential drag of 1/3 to 1/2 percent for the average stock, and more for noisy ones. This is substantially below the *prima facie* extra return that FI seems to pay, suggesting that other risk factors are at play here. In addition, any market-wide component in mispricing would further reduce the effectiveness of re-weighting, as would residual correlation between the new errors and mispricing or autocorrelation in the market's errors. The empirical challenge is to estimate how much FI really adds once value and size effects have been controlled for, and to do so with proper recognition of demonstrably unstable size and value exposures.

In most empirical studies, the proposed alternative weighting schemes are effectively found to boost returns and Sharpe ratios.[1] Most authors also find that a substantial part of the gain is factor-related, but there is far less consensus as to whether there also is any net alpha left due to avoiding the drag effect.[2] One problem with return-based style analysis is that the evidence is indirect: alpha is what's left of the average return after eliminating the part that reflects exposure to all assumedly priced factors or styles, and one can always raise doubts about the choice of factors and the estimation of the portfolio's exposures.

This motivates our paper. In fact, we document that FI alphas suffer on both the above counts: they are quite model-sensitive, and look flawed because of instabilities in exposures. Our conceptual chief contribution is a way to estimate the drag effect that does not rely on regression, and our main empirical results suggest that drag is empirically meaningless, if it exists at all.[3] This does not mean FI-based investment is useless: it does offer an operationally easy efficient way of using the size and value effects without hurting diversification. What FI does not do, or at least not in a meaningful way, is to neutralize drag effects.

Our estimator works as follows. We first sort all stocks into 20 buckets (vigintiles), based on either size or value or on a double sort. We then assign capital to each bucket on the basis of its aggregate market cap, but within the buckets we weight the stocks equally. This mixed portfolio strategy, labeled 'VW/EW', is inspired by Treynor's (2005) suggestion to use equal weighting as a way to avoid the drag effect. Obviously, if equal weighting is applied to the whole portfolio, the

[1]See e.g. Arnott, Hsu and Moore (2005), Hemminki and Puttonen (2008), Stotz, Dhnert and Wanzenfried (2007), Neukirch (2008), Jun an d Malkiel (2008), Asness (2006), Walkhusl and Lobe (2010), Houwer and Plantinga (2009), Peltomki (2010), and Mihm and Locarek-Junge (2010, Ferreira and Krige (2011), Forbes and Basu (2011), Balatti, Brooks and Kappou (2017). Less satisfactory results are reported by e.g. Biltz, Van de Grient and Van Vliet (2010), Boudt, Darras, Nguyen and Peeters (2015), Zaremba and Miziolek (2017), Piljak and Swinckels (2017), Chen et al. (2007), Graham (2011).

[2]Claims of value/size biases include Fama and French, 2007; Dopfel, 2008; Subramaniam, Kulkarni, Kouzmenko and Melas, 2011; Perold, 2007. Positive alpha is reported by e.g. Stotz et al., 2007; Houwer and Plantinga, 2009; Peltomki, 2010; Mihm and Locarek-Junge, 2010; and Forbes and Basu, 2011; Balatti, Brooks and Kappou, 2017). No alpha was found by Jun and Malkiel (2008), Amenc, Goltz and Le Sourd (2008), and Walkhusl and Lobe (2010).

[3]We could, indeed, also question the underlying assumptions. While a Roll-style model is quite appealing, it is also stylized—for instance by assuming the error is independent of true value. So in reality the drag may not even exist at all rather than being just statistically and economically undetectable.

resulting portfolio is massively more exposed to the size factor, Fama-French's SMB,[4] than is a value-weighted index. But if the sort is size-based, for instance, under our rule equal weighting is applied only among stocks of very similar sizes (within the vigintiles), implying that the size bias does not arise. In that argument, we follow the logic behind FF's way of constructing both their factors and their test portfolios: even though exposure to SMB is not necessarily linearly related to size itself, the two should still be quite rank-correlated. As a result, buckets of firms with e.g. the same size should have very similar SMB exposures. So even if the buckets' size exposures change over time, those changes are largely shared by the firms within the bucket. Comparison of value- versus equal-weighted returns within each bucket therefore neutralizes the size factor in a non-parametric way. The same holds for Book-to-Market-sorted vigintiles, which control for Book to Market (BtM) effects, or double-sort-based buckets, which adjust for both effects, albeit in a less fine-grained way.

Inside a bucket of similar firms, the Treynor-FI logic still holds, though: (i) if one uses value weights there will be drag, while (ii) equal weighting ensures that the weights are not correlated with pricing errors, so that the drag effect is expected to be avoided. For these reasons the spread, per vigintile, between its equally-weighted and its value-weighted mean return provides an estimate of the drag effect, within its bucket, uncontaminated by size differences, or 'value' effects, or both. The weighted average of these twenty estimates of the drag effect, per size class, then gives us the aggregate drag effect present in the value-weighted index.

We find that the aggregate drag effect is economically and statistically insignificant. A first take-way, then, is that the extra return delivered by FI is not just partly style-related, as is widely accepted: it is essentially *all* about style. The second take-away is about the nature and importance of mispricing, an issue also relevant to academia: if, in a generalised Roll (1984) model, drag is trivial, then valuation errors must be either trivial or, perhaps more plausibly, quite persistent and/or largely reflecting market-wide mispricing. Lastly, we document how the positive alphas of FI estimated by return-based style regressions *à la* Carhart provide unreliable

[4] Fama and French's Small Minus Big factor is a portfolio where small-cap stocks are held long and big-cap stocks short. It picks up the size anomaly in average returns.

Table 1: **Notation**

$\epsilon_{j,t}$	the percentage pricing error: $V_{j,t} = v_{j,t}(1 + \epsilon_{j,t})$
$\epsilon^*_{j,t}$	the error in the alternative weight relative to the true weight: $w^*_{j,t} = w_{j,t}(1 + \epsilon^*_{j,t})$
$\epsilon_{m,t}$	the market-wide pricing error: $W_{j,t} = w_{j,t}\frac{1+\epsilon_{j,t}}{1+\epsilon_{m,t}}$
γ_j	the cross-correlation between $\epsilon_{j,t}$ and $\epsilon_{m,t-1}$
$R_{j,t}$	the noisy observed return, i.e. $R_{j,t} = r_{j,t}(1 + \epsilon_{j,t})/(1 + \epsilon_{j,t-1})$
$r_{j,t}$	the true return, i.e. $v_{j,t}/v_{j,t-1} - 1$
ρ_j	the autocorrelation in $\epsilon_{j,t}$
$V_{j,t}$	the observed market value of stock j at time t
$v_{j,t}$	the unobservable true market value of stock j at time t
$W_{j,t}$	the observed market weight of stock j at time t
$w_{j,t}$	the unobservable true value weight of stock j at time t
$w^*_{j,t}$	a proposed alternative weight of stock j at time t

measures for the drag-related gain of FI. We discuss these implications and their academic and practical relevance further in the paper.

The paper continues as follows. In Section 1 we present some background information: FI's three central propositions and prior empirics. Section 2 critically discusses regression-based measures of FI's drag-related gains. In Section 3 we introduce our direct measure of the drag effect and discuss the empirical results. Section 4 concludes.

1 Background material

Hsu (2006) provides a formal treatment, including some caveats not included in the rather informal treatment by Arnott *et al.* (2005). In this section we just review the key elements, using the notation set forth in Table 1.

1.1 What does FI promise?

The basic assumption is a generalisation of Roll's (1984) noisy-price equation: the observed price V deviates from the underlying true value v by a multiplicative noise term ϵ independent of v:[5]

$$V_{j,t} = v_{j,t}(1 + \epsilon_{j,t}), \text{ with } \mathrm{E}(\epsilon|v_{j,t}) = 0. \tag{1}$$

[5]The error may conditionally exhibit autocorrelation, but at this stage we are after the links between V and v.

On the basis of this, FI theory makes three main propositions. First, if prices are noisy and mispricing is unrelated to true value v or true return r, like in Roll's (1984) seminal model, the expectation of the noisy observed return R is biased upward by a Jensen's convexity effect (see also Brennan and Wang, 2010):

$$\begin{aligned} 1 + R_{j,t} &= (1 + r_{j,t}) \frac{1 + \epsilon_{j,t}}{1 + \epsilon_{j,t-1}} && (2) \\ &\approx (1 + r_{j,t})(1 + \epsilon_{j,t})(1 - \epsilon_{j,t-1} + \epsilon^2_{j,t-1}), \\ \Rightarrow \mathrm{E}(1 + R_{j,t}) &\approx \mathrm{E}(1 + r_{j,t}) \left[1 + \mathrm{var}(\epsilon)(1 - \rho_j)\right]. && (3) \end{aligned}$$

Second, mispricing also drives a wedge between observed value weights W and the true ones, w. If the errors are not diversified away at the market level, also the market-wide average mispricing enters into the picture:

$$W_{j,t-1} = w_{j,t-1} \frac{1 + \epsilon_{j,t-1}}{1 + \epsilon_{m,t-1}}. \tag{4}$$

Multiplying Equations (2) and (4), the firm-specific noise term in the denominator is replaced by the market-wide noise term. That is, value weighting means that the upward bias from $\mathrm{E}(1 + \epsilon_{j,t-1})^{-1}$ disappears, only to be replaced by a similar market-wide valuation term:

$$\begin{aligned} (1 + R_{j,t}) W_{j,t-1} &= (1 + r_{j,t})\, w_{j,t-1} \frac{1 + \epsilon_{j,t}}{1 + \epsilon_{m,t-1}}, && (5) \\ \Rightarrow \mathrm{E}[(1 + R_j) W_{j,t-1}] &= \mathrm{E}\left((1 + r_{j,t}) w_{j,t-1} \frac{1 + \epsilon_{j,t}}{1 + \epsilon_{m,t-1}} \right); \\ &\approx w_{j,t-1}\, \mathrm{E}(1 + r_{j,t}) \left[1 + \mathrm{var}(\epsilon_m)(1 - \gamma_j)\right]. && (6) \end{aligned}$$

The market's aggregate pricing error, being more diversified, has a smaller variance and covariance with the individual stock's subsequent return. So value-weighting weakens the upward Jensen's Inequality effect. If noise would be fully idiosyncratic, the market portfolio's value would be practically error-free, so that the entire Jensens' Inequality effect would be wiped out. This maximal loss from value weighting is FI's 'drag'. At the other extreme, if all stocks would be equally mispriced, all relative capitalisation weights would be unaffected, drag would not arise at all, and FI would have nothing to undo.

Equation (7) shows that if one had been able to observe the noise-free weights w, all of the original boost in expected return would have been preserved. The third FI proposition is that, if one works with alternative weights w^* whose errors ϵ^* are uncorrelated with the other components of return, then in terms of expectations this is as good as the true weights:

$$\begin{aligned}
\text{If } w^*_{j,t-1} &= w_{j,t-1}(1+\epsilon^*_{j,t-1}) \text{ with } \epsilon^*_{j,t-1} \perp (\epsilon_{j,t-1}, \epsilon_{j,t}, \mathrm{E}_{t-1}(r_{j,t})), \\
\text{then } \mathrm{E}[(1+R_{j,t})w^*_{j,t-1}] &= \mathrm{E}\left((1+r_{j,t})w_{j,t-1}\frac{1+\epsilon_{j,t}}{1+\epsilon_{m,t-1}}(1+\epsilon^*_{j,t-1})\right); \\
&\approx w_{j,t-1}\,\mathrm{E}(1+r_{j,t})\,[1+\mathrm{var}(\epsilon)(1-\rho_j)]. \qquad (7)
\end{aligned}$$

The assumption that FI's error ϵ^* is independent of the market's error ϵ is the one that eliminates drag; independence of expected true return, in contrast, makes FI style-eutral. The first one is more key to FI and is what this paper is about.

Note that nowhere in the above is it claimed that the FI weights w^*_j are closer to the true weights than are the market weights W_j, let alone that the W^*_j-s would be error-free. The objective is to avoid any systematic interaction between mispricing and weighting. FI is silent about whether or not a particular stock is under- or overpriced, whether in actual fact or in terms of likelihoods. FI is even silent on whether or not its own errors are smaller than the market's. Its only claim is that its own errors are bound to be less correlated with the market's mispricing, which weakens (and in the limit eliminates) the interaction with subsequent returns.

1.2 Is FI fundamentally flawed?

Perold (2007) and Kaplan (2008) claim that FI makes fundamental logical or mathematical mistakes. Even though in this paper we doubt the economic relevance of FI's drag, we think the theoretical objections are misdirected.

Part of the problem seems to be Hsu (2006)'s statement that (emphasis added)

> "[FI's weighting mistake] is a mean zero white noise uncorrelated with *other random variables*. This is to say that the selected portfolio weights may deviate significantly and across the board from the "true-value-weight" but these mistakes in assigning weights are not related to *other variables, such as market prices or firm capitalization*".

The first sentence is unnecessarily sweeping: as we saw, for FI to work it suffices that FI's mistakes be independent of the market's.[6] In Kaplan's maximalist view, though, Hsu's "other variables"

[6]The statement is also ambiguous: does "other random variables" cover anything, or does he mean just "variables

then also include accounting variables. In $V = V^*U = (V^*/F)\,F\,U$,[7] Kaplan continues, all three factors (V^*/F, F, and U) are therefore assumed to be mutually independent, and V^*/F independent of the product of the three. Kaplan's refutation consists of a proof that, unsurprisingly, all this leads to contradictions; therefore, he concludes, FI is wrong. But whether or not Hsu really intended V^*/F to independent of F (which some may find hard to believe), the objection is irrelevant because what FI needs is just that V^*/F be independent of the U component of V^*U. Then Kaplan's objection becomes moot in the sense that there is no internal contradiction. Whether FI actually achieves that hoped-for independence, and if so whether it does so without picking up size effects etc., is, of course, another issue.

Perold's objection, in his 2007 comment and his 2008 response to Arnott and Markowitz' (2008) reply, is that, if the investor's prior is diffuse (that is, if a price of V is equally likely to reflect under- as over-pricing relative to v), one cannot improve on the market price: the noisy-markets assumption $\mathrm{E}(\ln V|\ln v) = \ln v$ still implies that $\mathrm{E}(\ln v|\ln V) = \ln V$. The conclusion is correct under that prior, but one should realise that a uniform prior on $[-\infty, +\infty]$ for $\ln v$ is an assumption *he* adds to the noisy-markets model. Equally correctly, Arnott and Markovitz (2008) reply that, in a theoretical regression of $\ln v$ on $\ln V$, the slope should be below unity if the variance of $\ln v$ is finite, not infinite as it is with a diffuse prior; so for lower-cap stocks underpricing is (weakly) more likely, and *vice versa.* As a finite variance for v's prior distribution is conceivable, Perold's claim is not self-evidently better than Arnott and Markowitz'. More fundamentally though, the whole debate is besides the point. For FI to work the question is not whether one can beat the market valuations as estimates of true value, and whether FI achieves that. What matters is whether the mistakes in the fundamentals are uncorrelated with the market's mistakes, and whether that makes enough of a difference to make FI worthwhile.

Relatedly, both Kaplan and Perold state that FI even assumes that they do know the true

---

such as market prices or firm capitalisations", the version in the next sentence? The comma before "such as" should mean that he does not restrict "other variables" to prices and caps, but not everyone uses punctuation equally carefully all the time.

[7] We locally adopt Kaplan's notation: V is the observed value, V^* the true value, U the multiplicative error V/V^*, and F the fundamental. V^*/F corresponds to unity plus FI's weighting error. Kaplan's claim that V^*/F, F and U are all assumed to be independent of each other is made on page 34, following Eq (7).

values. It is true that FI's identification of the drag effect stems from a comparison of value-weighted returns with versus without noise, but there is no claim they know these true weights. FI's promotors explicitly go for proxies, and show that if the proxies are well-chosen (by avoiding mistakes correlated with the markets) then *on average* the drag is still avoided. In a way this is similar to statisticians discussing properties of estimators of $y = a+bx+e$: given one's assumptions about e, one can figure out whether, say, an instrumental-variable estimator offers any gain, on average, over an OLS one without actually knowing a and b.[8]

Other critics question the execution rather than the concept: does FI actually deliver what it claims to do? In the spirit of Kaplan (2008), if FI uses noisy accounting information and if that noise also causes part of the market's mistakes, the independence assumption would be violated and FI's advantage weakened. The effect of market-wide mispricing, pointed out by Hsu (2006), provides another reason why the avoidance of interaction may amount to less than $\text{var}(e_j)(1-\rho_j)$. Lastly, even if the weighting scheme introduced by FI actually would largely avoid correlation with the market's error, it may still be correlated with expected returns, for instance by biasing portfolios towards small and 'value'-type stocks. Many critics point out that this is exactly what FI does; if so, much of FI's extra return could then stem from exploiting size and value effects *etc.* rather than recovering drag.

Arnott, Liu and Markowitz (2007) riposte that much of the value and size effects may be reflections of noise. Whether that is actually the case is not immediately clear, though; judging by citations, most of finance academia regards size and value as plausible factors in their own right. Accordingly, our objective is to investigate whether pure drag avoidance, FI's key value proposition, adds anything when one filters out the effects of size etc. The empirical challenge is how to distinguish between the two when factor exposures are unstable.

[8] Ambiguous writing by Treynor (2005) may have contributed to the misunderstanding. Kaplan (2008, footnote 5) writes that

> "Treynor assumed that "MVI market-valuation indifferent investors spend the same number of dollars on the underpriced as they spend on the overpriced stocks" (p. 67)." Of course, they can do this only if they know the relative true values of all stocks, a condition that proponents of fundamental indexation (including Treynor in the same paper) deny is necessary for their systems to work. See Perold (2007)."

It is true that if one wants to invest exactly as much in undervalued stocks as in overvalued ones, one needs to know which is which. But Treynor (2005) is writing about equally weighted portfolios; so it looks plausible that his "equal amounts" may refer to the $1/N$ weights per stock, whether under- or overvalued, or to the mathematical expectation that, with random weighting errors, under- and overvalued stocks will be equally present. Whatever Treynor's true intent, FI does not need to know w_j; all is needed is that its own errors around w_j be uncorrelated with the market's.

1.3 Empirical findings in the literature

Most studies agree that FI-based portfolios do outperform value-weighted ones. Arnott, Hsu and Moore (2005), first, note that FI portfolios outperform the S&P 500 by 2% on average per year between 1962 and 2004 while the volatility is very close to that of a cap-weighted index. Hemminki and Puttonen (2008) reweight the components of the Dow Jones Euro Stoxx 50 and find a similar return gain. Stotz, Döhnert and Wanzenfried (2007) confirm these conclusions for the DJ Euro Stoxx 600 stocks. Neukirch (2008) compares an internationally diversified portfolio of FI-weighted ETFs to the MSCI cap-weighted world index and finds overperformance. Similarly promising extra returns have also been documented by e.g. Jun and Malkiel (2008), Asness (2006), Walkhäusl and Lobe (2010), Houwer and Plantinga (2009), Peltomäki (2010), and Mihm and Locarek-Junge (2010). Sometimes gains seem to be huge: Ferreira and Krige (2011) come forth with an extra return of 4.7% for South Africa, while Forbes and Basu (2011) conclude that in Australia the gain exceeded 5%. Balatti, Brooks and Kappou (2017), more recently, report that for U.K. stocks, a modified FI strategy based on profitability rather than FI's traditional book value, cash flow, and sales provides a 3% extra return.

Others report mixed results. Biltz, Van de Grient and Van Vliet (2010) obtain a downright impressive 10% extra return if re-weighting is done as early as March, but they hasten to add that this may be optimistic: if on March 1 the accounting information is not actually available yet, these calculations would definitely create a look-ahead bias. Actually, Biltz *et al.* note, re-weighting in September instead of in Spring, for instance, kills all extra returns, a result that flatly contradicts the FI logic. Boudt, Darras, Nguyen and Peeters (2015), find that, for S&P500 stocks 1985-2014, the FI-weighted portfolio has just a marginally better Sharpe ratio (0.75 versus 0.71), and features the worst maximal drawdown of all strategies considered in that paper. Zaremba and Miziołek (2017) note that in their international application "benefits are largely limited to emerging and frontier markets and almost disappear from the post-2007 period". In the bond market, results by Piljak and Swinckels (2017) are equally mixed: for investment-grade bonds the gain disappears when positions are hedged against exchange risk, while for non-investment-grade bonds the significance is marginal.

Some authors are downright skeptical. Chen *et al.* (2007) propose lagged cap weights as an alternative to accounting-based weights. They find that this trading rule also outperforms the cap-weighted index; still, the extra return is below that of an FI strategy.[9] Graham (2011) experiments with random weights, pointing out that such weights should eliminate drag as efficiently as equal or lagged weights. The extra return turns out to be small, though, except during the tech bubble. He also tests whether stocks that experienced extreme returns, being candidates for larger mispricing, offer higher averaged returns immediately after the jump. They do not.

But most critics take the line that FI just introduces size and value biases, so that the extra return is the reward for non-beta exposures (e.g. Fama and French, 2007; Dopfel, 2008; Subramaniam, Kulkarni, Kouzmenko and Melas, 2011; Perold, 2007; and others cited below). While Balatti *et al.* (2017) do not explicitly refer to the profitability anomaly or factor, [10] their positive results are consistent with the idea that that factor plays a major role in their profitability-based FI variant. Chen *et al.* (2007)'s weighting by lagged market values may pick up momentum/reversal effects, if any, while Graham's (2011) random weights are likely to bring in size effects, similar to equal weighting. In papers that address such factor effects, the proposed solution is to consider risk-adjusted returns à la Carthart. Again, the results are mixed. Some find that, after the standard regression-based correction for market, size and value exposure, there still is an alpha return left (Stotz *et al.*, 2007; Houwer and Plantinga, 2009; Peltomäki, 2010; Mihm and Locarek-Junge, 2010; and Forbes and Basu, 2011; Balatti, Brooks and Kappou, 2017), but others disagree, like Jun and Malkiel (2008), Amenc, Goltz and Le Sourd (2008), and Walkhäusl and Lobe (2010).

One possible explanation of the conflicting findings is that the factor exposures of FI-weighted portfolios are not constant and that changes in the betas may be correlated with factor returns, in

[9] Not directly related to the FI literature but relevant to our methodology is work by Plyakha, Uppal and Vilkov (2014) on value- versus equal weighting in factor portfolios. We return to that work later.

[10] Already in Ball and Brown (1968), net income (excluding extraordinary items) scaled by book value of assets, is found to predict the cross section of average returns. Novy-Marx (2013) proposes gross profit (revenue minus cost of goods sold) rather than net. Ball, Gerakos, Linnainmaa and Nikolaev (2015) find that operating profit (revenue minus COGS and costs other that R&D) outdoes gross profit and is not subsumed by the FF model. Like other 'anomalies' before, profitability, alongside investments, became a priced CAPM factor, see Hou, Xue and Zhang (2015)'s four factor 'q' model and Fama and French (2016) five-factor model. More recently, Ball, Kerakos, Linnainmaa and Nikolaev (2016) merge the profitability issue with the 'accrual anomaly' (Sloan, 1996; see Ball *et al.* (2016) for a review of subsequent work) to document that cash-based profitability (excluding accruals) survives both modified CAPMs.

which case the standard regression intercept does not provide a reliable estimate of the abnormal return. Walkhäusl and Lobe (2010), in fact, provide evidence that correlations between exposures and factor returns do exist: the regression coefficients for Treynor-Mazuy (1966) squared factor returns are significant. The challenge, then, is how to eliminate this nuisance factor in the diagnosis.[11]

2 Data, portfolio formation procedures, and style patterns

2.1 Data source and filtering

We obtain monthly market caps and returns for U.S. common stocks, adjusted for stock splits and dividend payments, from Thomson Reuters Datastream (TRD) for the period January 1990 till December 2014. The data set is corrected for the survivorship bias present in the TRD 'research list' as we add back the data in the 'dead stock' lists. The data are carefully filtered for errors following the procedures adopted in *e.g.* Ince and Porter (2006). The corresponding beginning-of-the-month fundamentals (*i.e.* book values, sales and free cash flows) are also from TRD.

The Pink Sheet part of the dataset includes the smallest firms, usually not present in CRSP-based datasets commonly used in this kind of research, but in any given month we do eliminate stocks with a market capitalization smaller than \$10 million, a monthly trading volume smaller than \$100,000 or a price smaller than \$1.[12] Each month we also exclude firms with negative or zero book value or sales.[13] After all this whittling the average cross-section contains about 3,300

[11]The Treynor-Mazuy (1966) squared factor return is able to detect, at best, linear relations between exposure and expected return. That is, it may document that variations in exposure are correlated with changes in expected factor returns (market timing), but it still misses changes in exposure uncorrelated with the market return and, therefore, does not allow any generally valid conclusions about whether there actually is a genuine alpha.

[12]This is to eliminate tiny, illiquid and penny stocks which are reasonably more likely to contain data errors. Penny stocks are often fallen angels (Chan and Chen, 1991) which are highly speculative and illiquid. Tiny companies likewise have limited liquidity, can be subject to high price pressure or price manipulation, and often represent too little value to warrant attention.

When a stock falls below a floor on date t, both the return before and after that date are removed. That is, we do not remove the entire return series of a firm that does badly at one point, as this would have introduced a classical survivorship bias. Our way of filtering may still affect returns to some extent, but (i) it cannot have a serious effect on the *difference* between EW and VW returns, which is what matters to us; and (ii) it occurs only for very small firms, whose impact on total portfolio performance is too minute to matter.

[13]Negative and zero book values and sales are almost surely mistakes. Negative free cash flows, in contrast, make

stocks, over a period of 25 years. The asset-pricing factors and the T-bill rates, lastly, are from K. French's website.

2.2 FI portfolio formation procedures

We now discuss some details about the FI and lagged-weight (LW) portfolios. FI usually adopts weights that are synthetic (*i.e.* they are based on several size indicators at once), and smoothed via averaging over time. We adopt single-dimension weights, like either pure book-value weighting or pure sales-based weighting. One motivation for preferring multiple pure weighting schemes instead of a single synthetic one is that it allows us to better see where FI's extra returns are coming from. Second, it offers a robustness check: drag stems from cap-weighting, so if what we see is really just drag, not factor risk premiums, the extra return should be quite similar regardless of how the alternative (non-cap) weights are chosen. A third consideration is data availability. Both the market cap and all fundamental data need to be available if we want to work with one balanced sample for all weighting strategies. Working with such a balanced sample has the advantage that differences in performance across the weighting schemes cannot be driven by differential data availability. The drawback of using all criteria at once, however, is that the sample size depends on the least available fundamental, namely the free cash flow.

It is true that using FI indices based on individual criteria could alter the profitability of fundamental indexing. Our reply is that, first, it actually does not occur: our FI strategies are about as profitable as what others report. Second, the question is not whether FI boosts returns (it does), but how much is left after value and size effects have been stripped away. By FI's book, drag should be avoided no matter whether the alternative weights are syncretic or single-dimensioned.

Our three pure fundamental measures are book value, free cash flow, and sales. We do not weight on the basis of dividends because too many stock/year combinations then produce zero weights, which leads to a portfolio that is biased against small and distressed stocks. When weights are based on free cash flow, lastly, we work with its absolute value. Empirically, weighting by absolute free cash flow, $|CF|$, provides higher returns than the standard versions, namely CF

more sense and are much more prevalent.

itself or Max(CF, 0). Also *a priori* this makes sense: a hugely negative cash flow tends to occur only in big firms.

Our version of lagged-weight portfolios is likewise modified relative to the original. Chen *et al.* use the median value weight over the preceding 18 months. The attractive feature in this is a certain degree of smoothing, akin to averaging. One drawback, though, is that one has no clear idea as to how old the weights really are: a drop in median value over the last N months may reflect quite recent price movements, and if those would be mostly noise that partly persists until t, then such a median value would not be efficient at avoiding the drag effect. We accordingly use simple alternative lagging schemes, with lags of either one, three or six months (LW1, LW3, LW6). This provides an additional testable proposition: in the logic underlying FI, older weights should suffer less from drag (*i.e.* provide better returns) because the age of the weighting error weakens the link with current mispricing.

For the rebalancing frequency of the portfolios we use monthly revisions, the standard in this line of research. Monthly rebalancing of the portfolios is realistic for both institutional and private investors who need to weigh transaction costs against tracking error. We also keep track of stock delistings and IPO's on monthly basis. We do not correct for transaction costs.

The data base differs from the usual CRSP in terms of initial coverage and because of the presence of OTC stocks. In the main section, all results that could suggest drag effects are actually confined to smaller stocks and 'value' stocks. In light of this it is relevant to quickly document the size, value and momentum/reversal patterns present in this data set.

We start with the factor returns, both in the entire sample period and in the non-dotcom subsample, leaving out March 2000-October 2002 for reasons to be discussed later. Table 2 provides the means and standard deviations. In terms of average return the market factor did best (0.62 percent per month), followed by momentum (0.55) and, at a distance, the other three (about 0.20). Leaving out the dotcom crash improves the market return but lowers the HML spread. MOM has the highest volatility (4.91), followed by the market (4.36). Leaving out the dotcom crash months dampens all volatilities, but especially those of HML, MOM and STR. We next turn to 'extra' expected returns over and above value-weighting, reported per vigintile under

Table 2: **Factor returns Jan 1990 to Dec 2014, with and without the dotcom crash (Mar 2001 to Oct2002)**

	$r_m - r_f$	SMB	HML	MOM	ST-REV
Averages					
Full	0.625	0.181	0.223	0.552	0.208
Non-dotcom	0.892	0.221	-0.025	0.505	0.066
change	0.266	0.040	-0.248	-0.047	-0.142
StDevs					
Full	4.355	3.315	3.130	4.908	3.623
Non-dotcom	4.099	2.947	2.621	4.362	2.950
change	-0.256	-0.369	-0.509	-0.546	-0.673

Key The main sample is Jan1990-Dec2014. The dotcom period is Mar2001-Oct2002. We show returns on the market factor and the SMB, HML, MOM and STR factors, all from K. French's website.

various sorting schemes, for each of FI's investment rules.

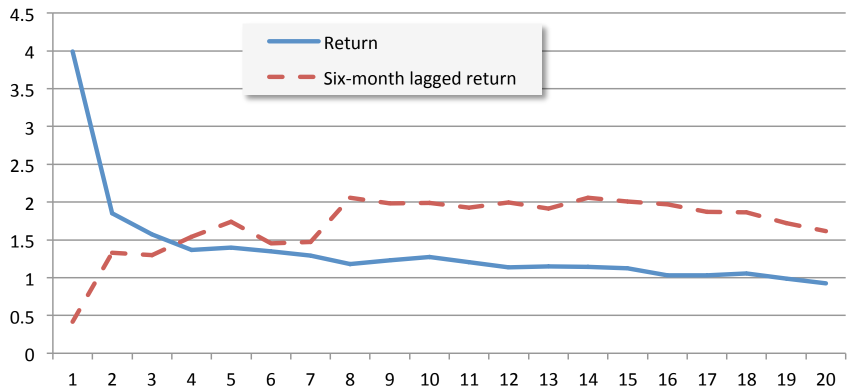
2.3 Size, value and momentum/reversal patterns in the TRD data set

In Figure 1 we visualise the extra-return spreads; Appendix Table 1 provides the numbers themselves, as well as the t-statistics. The full line in Figure 1 show the returns per vigintile. The sorts are, sequentially, into (i) twenty size vigintiles, (ii) twenty 'value' (i.e. book-to-market, BtM) vigintiles, and (iii) a double sort, where stocks are first grouped into size quintiles and, within each of these, into BtM quartiles.

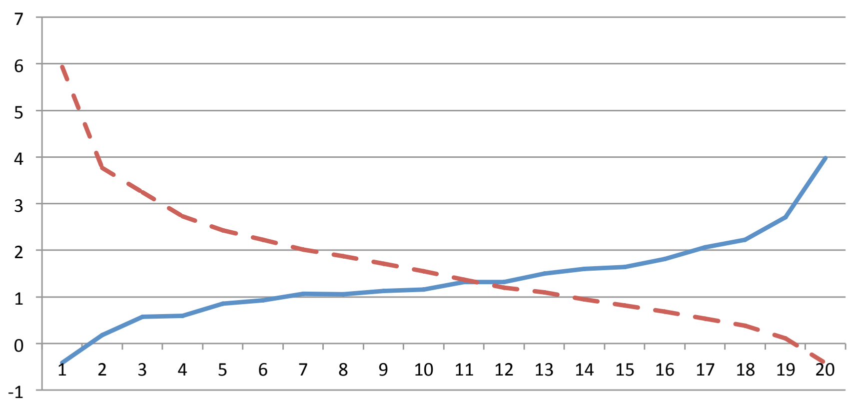
In Panel A, the full line reveals a strong size effect, with an average return spread of 3% per month between the first and last size vigintiles. (Much of that is reversal, as we show shortly.) Over two thirds of that is taken up by vigintile 1, and five sixths by vigintiles 1-3; that is, between buckets 4 and 20 the remaining size spread is just 0.5%.

It is equally interesting to see where these stocks came from. The broken line shows the average monthly return in the six months before, regardless of whether or not the stocks belonged to that vigintile at that time. Not surprisingly, mid- and large-cap stocks overperformed in the preceding half year, while small stocks tend to be the ones that did poorly. That underperformance is quite pronounced: class 1's return is about 1.5% lower than that in the modal buckets 8–16, for instance; and this is a monthly average over six months, implying an underperformance of about

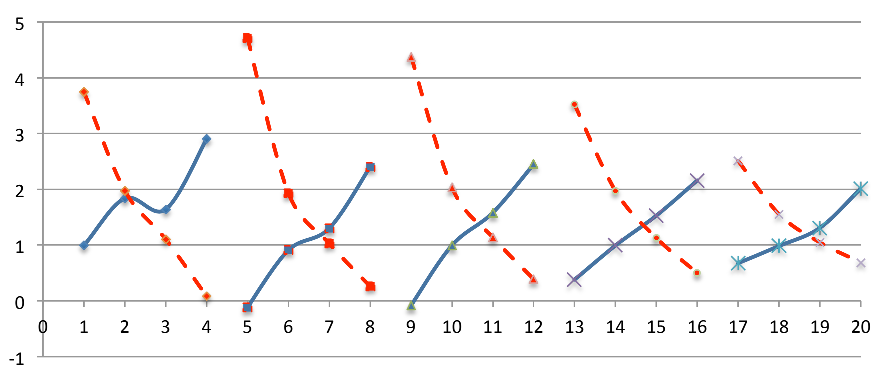
Figure 1: **Average monthly return, per vigintile, for a stock (i) when inside the bucket and (ii) over the six months before the sort**



Panel A: average return per size vigintile versus average return lagged six months



Panel B: average return per book-to-market vigintile versus average return lagged six months



Panel C: average return per size/BtM class versus average return lagged six months

Key At the start of each month t stocks are allocated to 20 buckets, by size (Panel A), BtM (Panel B), or first into size quintiles and then, per quintile, into four BtM quartiles (Panel C). "Return" is the average EW return per bucket. "Six-month lagged return" is a monthly average—not the total (accumulated) return—over the six months $t-2$ to $t-7$, on the stocks that, in month t, will be in vigintile v.

9.5% over the preceding half-year. To sum up, then, among the small stocks there are many firms that lost substantial market value and then rebounded, partially but relatively fast,[14] echoing the Fallen Angel pattern.[15]

The figure shows just point estimates. Appendix Table 2 also provides a portmanteau t-test of the difference between each bucket's current and past returns. Unsurprisingly, the extreme averages for the first vigintiles are always significant. Buckets 3 to 7, where in the Figure the two return patterns are similar, have insignificant differences; for buckets 8 to 20, where the graphical pattern was one of clear momentum rather than reversal, we again obtain statistically significant differences.

Panel B displays the results from a sort on book-to-market, with the higher-'value' stocks towards the right. The 'value'-premium pattern that emerges is stronger than the size effect: the full line suggests a spread of 4.5 percent per month between the extremes. The strongest effect is again found for the riskiest-looking stocks (classes 17-20), where BtM-weighting pays 2 to 4% more than market-cap-weighting; but also at the other end, in classes 1–3, the effect is pronounced. The spread for the more modal classes 4–17 is just 1.5%. Differences are significant for all buckets except 10 to 14, as Appendix Table 1 shows.

This strong 'value' pattern comes with a marked reversal pattern, as becomes clear when the return over the six preceding months (broken line) is brought into the picture. The top decile, with the lowest BtM ratio, has soared by 4–6 percent per month over the preceding half year and then underperforms the average stock. At the other end of the spectrum we see the opposite pattern. 'Value', in short, seems to have two reverting ends with, at one extreme, recent star performers that then fall back, and fallen angels that rebound at the other.

---

[14]With a faster correction, chances of capturing that correction within the horizon are better; and even if all corrections do happen inside the horizon, a sudden correction produces a higher mean monthly return than a gradual one. For instance, if a price goes from 50 to 100 in five equal percentage steps, the average per period return is 0.15. In contrast, if it jumps from 50 to 100 in period 1 and then pays four zero returns, the average return is 0.20.
Note that the reversal we document here relies on hindsight; we observe a group of small stocks (or, in the next paragraph, a high BtM stock) and then check whether they went down recently, on average.

[15]The fallen-angel effect is documented by e.g. Chan and Chen (1991), Chen and Zhang (1998), Lakonishok and Vermaelen (1990), Ikenberry, Lakonishok and Vermaelen (2002) or Peyer and Vermaelen (2009). Bearing in mind their own resumé, the story goes, professional portfolio managers are reluctant to invest in stocks that did quite badly recently, so prices are depressed and expected returns high.

Panel C, lastly, demonstrates that the above 'value' patterns are present in all size quintiles, but are more pronounced the smaller the firm. Again, the differences are significant in all buckets except where the full and dotted lines cross (the third BtM quartiles in each SMB quintile, see Table 2).

3 The failure of regression-based measures

Multifactor alphas estimate drag as a mean return purged of its risk-based components; as such, they depend on a factor model and an exposure estimation technology. To see why return-based style regressions may not provide reliable alphas in the case of FI portfolios, we first consider the pattern of weight adjustments that FI administers. We show that (i) the patterns are quite variable over time, which raises doubt about the stability of exposures to the size and book-value factors, and (ii) the dotcom crash is very different from the rest of the period. These findings (subsection 3.1) then motivate the robustness tests we append to the style regressions (subsections 3.2 to 3.5).

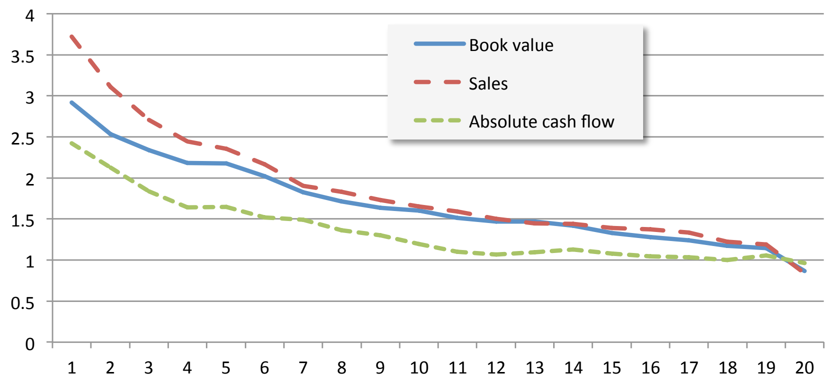
3.1 The behavior of FI's cross-sectional weight adjustments

In this section we consider cross-sections of the FI weight adjustments. One interesting feature of cross-sections is that, at any point in time, the aggregate market noise $\epsilon_{m,t-1}$ is a constant, so that noise in individual prices is the only differentiating factor. We first look at the average patterns across the size or BtM spectrum, and then document the strong longitudinal variability of the cross-sectional effects.

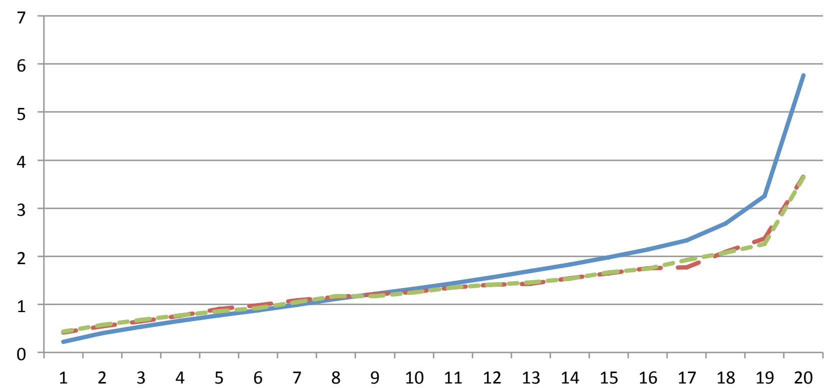
The information is processed as follows. We first sort stocks into 20 size buckets, and then compute the ratio of FI-administered weights (BV-, SL- or CF-based) over market-cap-based weights. This hows what size classes are overweighted by orthodox norms, and to what extent. Sorting the stocks by BtM, we proceed similarly to see to in what 'value' classes FI portfolios are overweighted, and so on. The point estimates are again visualised graphically, with the numbers provided in Appendix Table 2.

It is known that FI is more exposed to size than is the market on average; Panel A of Figure

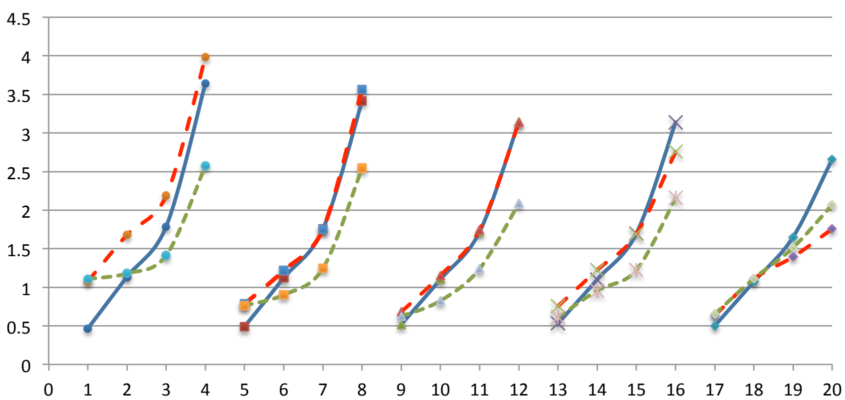
Figure 2: **Vigintiles' FI-assigned weight divided by its market weight**



Panel A: FI's weights over market weights, per size vigintile; small stocks to the left



Panel B: FI's weights over market weights, per BtM vigintile; value stocks to the right



Panel C: FI's weights over market weights, double-sorted (five size classes, four value classes)

Key This figure displays the mean of vigintile's FI-assigned weight (measured by book value, sales or absolute cash flow) divided by vigintile's market-value weight. The sorting is first by size (Panel A), then by BtM (Panel B) and, in Panel C, first by size (five quintiles) and then by BtM (four quartiles per size group).

2, for instance, shows that the lowest-cap vigintiles tend to receive weights that are 2-3 times larger than their market-value weights, with a peak of 3.75 times for the sales variant applied to the smallest stocks. This far exceeds anything justifiable on Bayesian error-correction grounds. Panel B documents an even more extreme overweighting of 'value' stocks. The most speculative shares, in class 20, receive weights that are on average six times their market value if FI applies book-value weighting, and still 3.5 times by the other standards.

Appendix Table 1 adds that these differences are overwhelmingly significant. The Table shows p-values for each of the averages of the ratios w^*/W under each of the three FI-inspired investment rules, reported by size, value or double sorts. Only five of the 180 t-tests produce a two-sided p-value exceeding 0.0005. None of this is a surprise, qualitatively: also earlier work has documented 'value' exposure and size bias. What is less well-known is that these biases are strongly variable over time.

Start with size. We summarize each cross-section of weight adjustments, $\ln(w_j^*/W_j)$, by its regression relation to size, $\ln V$. For a proper interpretation we first consider the numerator in that slope, the cross-sectional covariance. For ease of reading we temporarily drop time subscripts. Log percentages are denoted by primes, like in $\epsilon' := \ln(1+\epsilon)$, and cross-sectional (co)variances, unconditional on any information about the firm, are denoted by $\overline{\text{cov}}$ and $\overline{\text{var}}$. Below, we first write out the weight adjustments, and then look at the covariance with size, bearing in mind that within a given cross-section the market-wide mispricing, if any, is a constant. Then

$$\begin{aligned} \ln(w^*/W) &= [\ln w + \epsilon^{*'}] - [\ln w + \epsilon' - \epsilon_m'], \\ &= \epsilon^{*'} - \epsilon' + \epsilon_m'; \\ \Rightarrow \overline{\text{cov}}(w^*/W, \ln V|\epsilon_m) &= \overline{\text{cov}}(\epsilon^{*'} - \epsilon', \ln v + \epsilon'), \\ &= \overline{\text{cov}}(\epsilon^{*'}, \ln v) - \overline{\text{var}}(\epsilon'). \end{aligned} \tag{8}$$

The first term on the right is a measure of size style change, the second is a noise variance unconditional on j, i.e. mixing all ϵ_j's together. So variability over time in the covariance could come from either source. Brennan and Wang (2010) estimate the standard deviation of noise to be typically 0.06. But the cross-sectional variance of ln$V$ is at least 3. So for a style-neutral update with $\overline{\text{cov}}(\epsilon^{*'}, \ln v) = 0$, the equation

$$\ln(w_j^*/W_j) = a + b\,(\ln V_j - \overline{\ln V}) + \epsilon_j, \tag{9}$$

would have a slope no worse than $-0.06^2/3 = -0.0012$. For the very noisiest stocks, Brennan and Wang's estimate is 0.40, implying a bound of $-0.4^2/3 = -0.053$ for a cross-sectional slope. If actual slopes are more negative, the weight adjustments must also contain a style shift component.

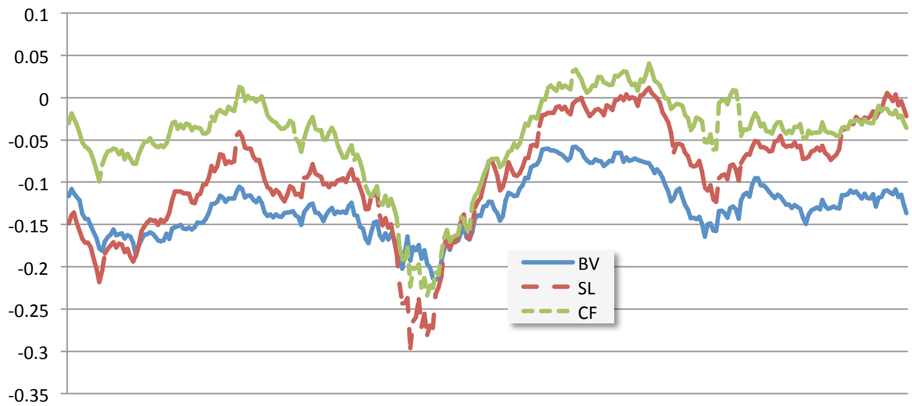
The Brennan-Wang estimates provide not just a standard for our estimates, but also suggest that the slope b should be quite different across noise classes. Since smaller stocks tend to be noisier, we could accordingly allow separate slopes conditional on the size vigintile. We adopt a simpler way to let the above slope b vary with size: we run a quadratic model, which is equivalent to letting the slope be linear in size:

$$\begin{aligned} \text{If } b(V_j) &= b_0 + b_1\,(\ln V_j - \overline{\ln V}), && (10)\\ \text{then } \ln\frac{w_j^*}{W_j} &= a + [b_0 + b_1\,(\ln V_j - \overline{\ln V})]\,(\ln V_j - \overline{\ln V}) + \epsilon_j'. && (11) \end{aligned}$$

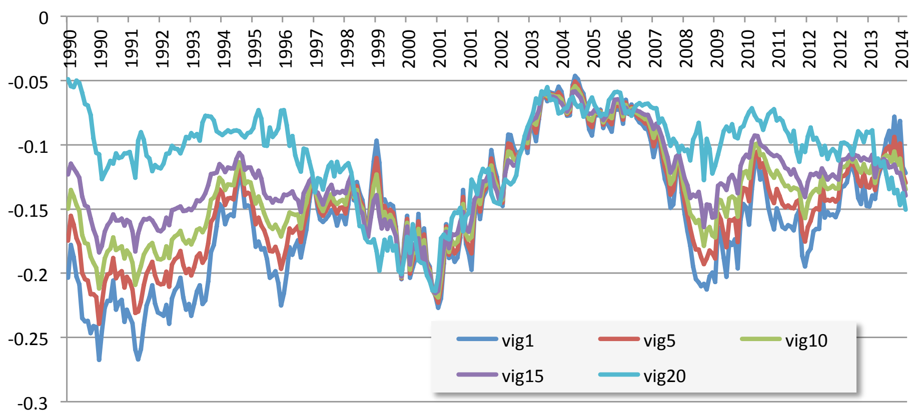
First consider the results for the linear equation. Panel A of Figure 3 shows time-series plots of the month-by-month cross-sectional slope coefficients for each of the three FI weighting schemes. The slopes are, first, mostly negative, confirming that FI tends to shrink the weights towards the mean. Only BV produces a consistently negative slope in every cross-section, though. Next, consider the magnitude of the slopes. A typical value for b is -0.10 to -0.15. This is far more pronounced than what we would expect, on average, if there would be no style shift—but the existence of size bias in FI portfolios is hardly news. Third, and more pertinent for current purposes, the slope is very variable over time. Given the prediction of $b = -0.0012$ under the assumption of no style shift, we therefore infer that most of the variability must come from variability in that style shift. Fourth, the dot-com period is unusual. For much of the 1990s the size correction in the log weight adjustments fluctuates around –0.15; it then becomes even more pronounced during the bubble year, only to shrink to about –0.05 (BV) or even zero (SL and CF) during the 2003–2007 lull. After that, the bias resumes its original pattern.

The above does not at all differentiate between big and small stocks, but the quadratic regression, (11), does. We apply this to the BV weighting scheme. Panel B of Figure 3 shows

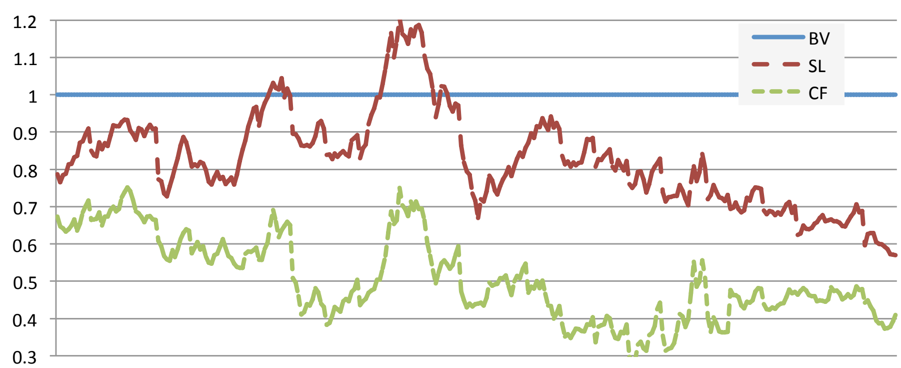
Figure 3: **Plot of slopes b over time from cross-sectional regressions of w_j^*/W_j on either $\ln(V)$ or $\ln(BtM)$, for BV, SL and CF weighting schemes**



Panel A: linear regression on ln(*V*), for each of the three FI weighting schemes



Panel B: quadratic regression on ln(*V*), for book-value FI weights



Panel C: linear regression on ln(*BtM*), for each of the three FI weighting schemes

Key Every month, at stock level, logs of FI weights over market weights are regressed cross-sectionally on log market values (linear in the top graph, quadratic in the middle one), or book-to-market (bottom). Each plot shows, for each of the three FI weighting schemes, the time series of slopes.

time-series plots of the month-by-month slope coefficients $b(V_j) = b_0 + b_1\,(\ln V_j - \overline{\ln V})$ for V_j set to the month's average company market value for the 1st, 5th, 10th, 15th and 20th vigintile, respectively. We see that, initially, the adjustments were more markedly size-related for the smaller stocks (whose bs start at minus 0.20-0.25, compared to the 0.10 downsizing typical for blue chips). But during the dotcom episode size discrimination all but disappears: all bs converge to –0.20 (including this time also large stocks) and then they all shrink, in line with each other, to (minus) 0.05 during the intercrises lull. The initial pattern then reappears as of the onset of the financial crisis.

Panel C, lastly, presents the slopes of linear regressions of the three weight adjustments on BtM. There is one instance where the regression produces a trivial result, namely BV: when $w_j^* = BV_j/BV_m$ (with subscript m referring to the market aggregate) and $W_j = V_j/V_m$, we get $\ln(w_j^*/W_j) = \ln(BtM_j) - \ln(BtM_m)$. The second item, $\ln(BtM_m)$, is cross-sectionally a constant, so when we regress the BV-based log weights ratio on $\ln(BtM)$ we get a unit slope.[16] For the other FI instruments, SL and CF, we see that, in addition, the log corrections are rather unstable. There is, notably, a clear downward trend, with substantial variability especially around—again—the dotcom bubble and bust.

The main takeaways, then, from this section are that (i) it is implausible that, in style regressions and the like, exposures would be constant, and (ii) the dotcom crash is especially atypical in that respect.

3.2 Gross returns prior to any risk or style correction

We now proceed to a standard performance analysis, starting with gross returns without any correction for style shifts. Table 3 shows the *prima facie* returns from FI, in percent per month, when portfolio weights are based on book value, sales and the absolute value of free cash flow, respectively. We also show the returns from applying EW and lagged weights (LW1, LW3 and LW6).

[16] This does not yet mean that the style is constant: the adjustment is nonlinear, being proportional to the market weight and its interaction with BtM's deviation from the average: $w_j^* - W_j = W_j\left(\frac{BtM_j - BtM_m}{BtM_m}\right)$.

Table 3: **Gross performance of FI, EW and LW indices v value-weighted index (VW)**

	Mean	Stand.	Sharpe	extra return		annual returns	
	(%/mo)	dev.	ratio	(%/mo)	t-stat	(%/yr)	extra
VW	0.913	4.56	0.145			10.14	
BV	1.189	4.91	0.191	0.276	3.87	13.59	3.45
SL	1.161	4.83	0.188	0.249	2.61	13.28	3.14
CF	1.135	4.87	0.181	0.222	2.55	12.90	2.76
EW	1.319	5.82	0.183	0.406	2.42	14.69	4.55
LW1	0.911	4.64	0.142	0.021	1.12	10.03	-0.11
LW2	0.888	4.74	0.133	0.016	0.45	9.60	-0.54
LW3	0.877	4.88	0.127	0.009	0.17	9.28	-0.86

Key The table shows (i) mean monthly returns, (ii) standard deviations, (iii) Sharpe ratios and (iv) excess returns relative to the value-weighted index ("extra" returns) for various strategies. The fundamental indexing strategies have weights based on book value (BV), sales (SL) and the (absolute value of) the free cash flow (CF). Also shown are results for portfolios where weights are equal (EW), or set equal to the value weight as observed 1, 3 or 6 months ago (LW1-6). The Sharpe ratios are not annualized. The last two columns show the geometric *p.a.* mean returns from following the strategy for 25 years, and the differences of the latter relative to the VW figure ("extra").

The returns do look promising for EW and FI, but far less so for LW. Monthly returns increase on average by 0.28% when weighting is based on book values, and by 0.25% and 0.22% when weighting is by sales and the absolute value of free cash flow, respectively. In terms of geometric average returns *per annum*, this means 2.8–3.5% extra.[17] In our experiment, this comes with a small increase in standard deviation. Unlike the rise in return, an increased portfolio risk is not a common finding in empirical work of this kind, but our Sharpe ratios are still up; that is, by de-levering, one could have maintained volatility and still come out ahead in terms of return.

EW pays an even more generous extra return, nigh 1.2 percent per month, with a predictably larger standard deviations (5.8% per month, against 4.5-5% for the other strategies). While its Sharpe ratio is not as good as the BV-based one, it still beats the other FI variants and VW. Lagged returns, finally, barely (and insignificantly) outclass VW in terms of mean return and Sharpe ratio. Actually, against the FI prediction, the extra return falls with the lag length. Whether this reflects an autocorrelation pattern in the true returns, or changing style shifts, or

[17]The annual return is calculated as $(\Pi_{t=1}^{N}(1+R_t))^{12/N}$, with N the number of months, 360. The remainder of the analysis is done in terms of arithmetc average monthly returns. These mean monthly returns could be poor indicators of long-term cumulative returns if there are serious autocorrelation patterns that differ across strategies. But the geometric average annual returns do not bear this out. Across the seven strategies, there is a 0.997 correlation between simple monthly averages and geometric annual averages. Thus, it looks correct to study just the mean monthly returns.

just a coincidence is far from clear: the drop is not significant. Still, our results for LW suggest that drag may be statistically and economically insignificant and, by implication, that FI's extra return probably has another source than drag avoided.

We now proceed to a standard multi-factor alpha analysis.

3.3 Standard multi-factor performance analysis

For the FI portfolio strategies we compute returns in excess of the value-weighted index (henceforth referred to as 'extra returns'), we regress them on the Fama-French factors augmented with the momentum factor, following Carhart (1997), and we produce the standard Newey-West t-test for a zero value of the intercept, alpha. Unlike in Balatti *et al.* (2017), our FI strategies are not explicitly profitability-oriented, so we do not include profitability and investment factors.

The results are summarized in Panel A of Table 4. As the left-hand variable is a return in excess of the value-weighted index (VW), the coefficients estimate the exposures to MKT, SMB, HML and momentum as differentials relative to those of the value-weighted index.[18] These numbers do reveal some style shift, as usual. Specifically, relative to the value-weighted index, FI imparts a statistically significant but small boost to the portfolio's sensitivity to the market (MKT), somewhat more to its exposure to SMB, and substantially to its sensitivity to HML. More unexpectedly perhaps, there is negative exposure to momentum, a phenomenon we return to below.

Among FI's competitors, EW is especially exposed to SMB, as one would expect, but hardly more to MKT and HML than is the value-weighted portfolio. Using lagged weights affects style far less than adopting FI, again as expected, except for some underexposure to SMB (LW3 and LW6) and a tiny overexposure to MKT.

So like others before us we conclude that FI does induce style shifts. But the alphas tell us that these increased exposures still do not seem to explain away much of the extra return we have noted before. In fact, all FI strategies provide a positive risk-adjusted return of 0.20-0.25

[18]The exposures to SMB, HML and momentum, for our value-weighted portfolios are not exactly zero because our sample does not coincide with the CRSP market portfolio.

Table 4: **Multi-factor performance analysis of FI, EW and LW portfolio strategies (based on extra returns)**

Panel A: FF factors plus momentum

	α (%/month)		MKT (β)		SMB (γ)		HML (δ)		Momentum (ζ)	
	coeff	tstat	coeff	tstat	coeff	tstat	coeff	tstat	coeff	tstat
BV	0.253	6.129	0.030	2.964	0.075	5.778	0.255	18.193	-0.120	-13.892
SL	0.241	4.095	0.000	-0.028	0.064	3.429	0.344	17.196	-0.145	-11.796
CF	0.207	3.889	0.033	2.561	-0.031	-1.862	0.303	16.728	-0.122	-10.972
EW	0.351	3.931	-0.028	-1.268	0.760	26.930	0.106	3.502	-0.161	-8.630
LW1	0.025	1.315	0.012	2.589	-0.011	-1.798	0.002	0.351	-0.015	-3.883
LW3	0.044	1.428	0.022	2.958	-0.028	-2.908	-0.006	-0.609	-0.058	-9.028
LW6	0.090	2.330	0.023	2.473	-0.061	-4.992	-0.013	-0.998	-0.126	-15.427

Panel B: FF factors plus short-term reversal

	α (%/month)		MKT (β)		SMB (γ)		HML (δ)		STR (η)	
BV	0.153	3.135	0.047	3.889	0.060	3.848	0.284	17.143	0.092	6.606
SL	0.120	1.782	0.024	1.449	0.046	2.119	0.381	16.593	0.098	5.053
CF	0.105	1.784	0.050	3.401	-0.047	-2.489	0.333	16.575	0.098	5.814
EW	0.216	2.284	-0.006	-0.256	0.740	24.351	0.145	4.493	0.130	4.797
LW1	0.010	0.704	0.000	-0.018	-0.014	-3.085	0.001	0.307	0.063	15.861
LW3	-0.006	-0.225	0.015	2.122	-0.036	-4.076	0.002	0.253	0.101	12.857
LW6	-0.019	-0.411	0.033	2.847	-0.077	-5.214	0.009	0.567	0.114	8.648

Panel C: FF factors plus a dot.com dummy

	α (%/month)		MKT (β)		SMB (γ)		HML (δ)		dot.com (θ)	
	coeff	tstat	coeff	tstat	coeff	tstat	coeff	tstat	coeff	tstat
BV	0.069	1.293	0.080	6.626	0.058	3.563	0.275	15.765	0.799	4.801
SL	0.031	0.425	0.059	3.580	0.043	1.953	0.371	15.566	0.849	3.740
CF	0.012	0.183	0.085	5.895	-0.050	-2.585	0.322	15.400	0.890	4.466
EW	0.167	1.607	0.033	1.426	0.741	23.518	0.146	4.288	0.481	1.486
LW1	-0.015	-0.758	0.019	4.419	-0.014	-2.318	0.002	0.254	0.245	4.056
LW3	-0.047	-1.315	0.046	5.643	-0.036	-3.359	0.003	0.225	0.385	3.485
LW6	-0.053	-0.979	0.067	5.461	-0.076	-4.591	0.014	0.772	0.317	1.884

Key Panel A of the table shows Carhart-style (1997) performance analysis regressions, using percentage monthly returns. The left-hand-side variable is the return in excess of the value-weighted index for fundamental indexing strategies with weights based on book value (BV), sales (SL) and the (absolute value of) the free cash flow CF. Also shown are results for portfolios where weights are equal (the 1/N rule), or set equal to the value weight as observed 1, 3 or 6 months ago. In Panels B and C, momentum is replaced by short-term reversal and a dot.com crash dummy (March 2000 till October 2002, respectively). T-stats are Newey-West-corrected, and all factors are from K. French's website.

percent per month, which is large in both statistical and economic terms; EW, with almost 0.40 percent per month, scores even better; and also LW6 seems to add value, once its lower risk is accounted for, even though economically it does not come near any of the FI strategies. But the direct estimates of drag in Section 4 are quite different from these alphas, as we shall see; relatedly perhaps, the alphas are not very robust, as we first show.

3.4 Sensitivity to factor choice

The conclusion may be less straightforward than it seems from the above. The strongly negative extra exposure to momentum, for one, is somewhat puzzling and may really mean a positive exposure to reversal.

The potential for a link between momentum/reversal and the weighting scheme is logically clearest for the LW strategies. The difference between the LW portfolio strategy and the value-weighted index (VW) is driven by the change of weights over the preceding L months; so weighting based on lagged weights (LW) means betting on reversal, against momentum.[19] For the other FI strategies there is a link with reversal too, albeit weaker: the FI weights are insensitive to recent changes in market value, so the difference with the value-weighted (VW) index is driven by two elements: (i) the gap between the FI weights at time t and the value weights at time $t-L$, and (ii) the changes in the value weights between $t - L$ and t. This last item is the same bet on reversal like for LW, but it may be partly obscured by the first.

The finding of strong reversal prompted us to replace, in the performance analysis regressions, the momentum factor by French's short-term reversal factor, STR. Panel B of Table 4 shows the new results. We find clear evidence of positive extra exposure to STR in all strategies, and most clearly so in the LW portfolios. In addition, each alpha drops substantially (typically to half its original level), and all intercepts except BV's are now insignificant. So it seems that the performance analysis results are not robust with respect to the choice of the factors, even if the alternatives are as closely related as momentum and reversal. Reversal dominates: FI, as we saw, substantially overweights small and high-BtM stocks, which in turn exhibit strong reversal patterns.

3.5 Unstable exposures

[19] For example, if a stock has underperformed the market, then the lagged weight it gets is higher than the current weight. If giving a positive differential weight to past losers seems to pay, as it does in our sample, we can conclude there was reversal.

Table 5: **Results for Carhart regressors and their interactions with the market factor: FI extra returns**

	alpha	MKT	interactions of MKT with				SMB	HML	MOM
			MKT	HML	SMB	MOM			
BV	0.269	0.025	0.002	-0.005	-0.016	-0.003	0.086	0.247	-0.124
t-test	4.778	1.972	1.331	-2.067	-4.885	-1.824	5.779	13.782	-12.369
SL	0.225	-0.004	0.003	-0.001	-0.020	-0.002	0.089	0.370	-0.155
t-test	2.724	-0.202	1.265	-0.203	-4.170	-0.855	4.118	14.094	-10.588
CF	0.213	0.029	0.003	0.000	-0.019	-0.003	-0.015	0.315	-0.124
t-test	2.884	1.707	1.223	0.106	-4.480	-1.541	-0.781	13.425	-9.450

Key For each of the FI portfolio strategies (extra returns) we run a Carhart regression augmented by interactions with the market factor. The table shows the estimated coefficients and their t-statistics.

There may be other problems with the standard regressions. Familiarly, exposures are assumed to be constant. Variability over time is especially problematic if it is correlated with the factor's price of risk: then the cross-product gets mistaken for genuine alpha.

In our analysis of the weight adjustments (Section 3.1) we do find that FI's size mix varies substantially over time, with the end of the dotcom period being especially atypical. For that period SMB, MOM and STR returns are unusual, and the episode was identified as special by Graham (2012). When we accordingly add to the FF model a dotcom crash dummy (March 2000-October 2002)[20] as a crude way to separate this unusual period, see Panel C of Table 4, all alphas again become insignificant even without STR in the regression.

A more direct illustration of changing exposures is provided in Table 5, which shows extra exposures and their t-tests for the parameters in a Carhart regression augmented with interactions between the market and each of the four factors, *à la* Treynor-Mazuy (1966).[21]

[20]The dot-com bubble burst on Friday, March 10, 2000. The technology heavy NASDAQ Composite Index had just peaked at 5,048.62 (intra-day peak 5,132.52), more than double its value just a year before. USD 5 trillion in market cap disappeared from March 2000 to October 2002.

[21]The original idea was that fund managers could increase beta when expected market returns are high, and a standard linear test would then mistake the covariance between beta and expected market return for alpha. Later motivations stress that by buying index calls one could again fool a standard Jensen regression; and also here, including the squared return would act as a signal that something's wrong with the standard test. The test is a low-power version, as it looks for correlation between beta and the realised market return rather than the expected return (or any other driver of changes in beta, for that matter), implying an error-in-the-regressor problem; so if one does find something, beta is definitely not constant.

The regressee is still the extra return for each FI portfolio strategy (*i.e.* in excess of the value-weighted index return), and a significant t-statistic signals that changes in the differential beta are correlated with unusual returns in the factor. For all three FI strategies, β is manifestly down when small stocks do well. For BV strategy, β is down also when 'value' stocks pay off more, and maybe also when momentum is high ($p = 0.068$, two-sided). These results suffice to conclude that standard return-based style regressions do not provide reliable alphas as indirect measures of the drag effect. Once the presence of nonlinearity is established, Treynor-Mazuy alphas have unclear validity as estimates of abnormal return because, when betas are changing over time, one's chances of picking up all of that variability via quadratic terms are slim. For that reason we did not even bother to add the full set of Treynor-Mazuy second-order regressors.

We sum up the results of this section as follows: FI's alphas are quite sensitive to the factors chosen, they change over time, and they seem unreliably estimated as exposures to risks are changing over time in ways that are partly related to other risk factors. Hence our interest in a direct estimate of drag instead of an α.

4 A direct estimate of the drag effect

4.1 Proposed test procedure: mixed portfolio strategies

If FI is just about avoiding the drag effect, any set of weights whose errors are uncorrelated with the market's will do. Treynor (2005) already notes that equal weighting (EW) should dilute away essentially all mispricing problems, thus efficiently avoiding any interaction between weighting errors and abnormal returns. FI proponents reply that EW is infeasible in practice and creates a pronounced style shift, two problems that are substantially reduced when FI is adopted instead of EW. Chen *et al.* (2007) suggest yet another alternative: use lagged weights (LW). This is simpler than FI, and probably largely preserves the style chosen by conventional value weighting (VW) because weights tend not to change drastically over a few months. The snag in employing a lagged value weightis, as we saw, that the return spread relative to value weighting is driven by the recent performance of the stock, a procedure that brings in momentum/reversal effects. In general, then, the problem is how to make sure the alternative weights do not introduce an unexpected style shift (like reversal, in LW). Our proposed solution starts from Treynor's (2005) EW idea. To avoid the concomitant style shift, we mix EW with VW. Suppose we want to control for size. To achieve that, at beginning of every month we sort all stocks into 20 equally-populated size buckets (vigintiles). We assign capital to each size bucket as a whole on the basis of the bucket's aggregate market cap, like in VW. Within the bucket, however, we weight equally, Treynor-style. In other words, each company's own value weight is replaced by the average market-value weight of all stocks in the company's size vigintile, ensuring that the alternative weights are comparable in magnitude to VW ones. The strategy is referred to as the VW/EW mixture — value weighted across buckets, and equally weighted inside — applied, here, to size buckets. Formally, every period, every firm j is associated with an index set $J(j, t-1)$ containing $N_{J(j,t-1)}$ subscript-IDs of firms k that are all very similar to j w.r.t. size. We denote the average true weight for the firms in $J(j, t-1)$ by $\overline{w}_{j,t-1}$, the average pricing error for those firms by $\overline{\epsilon}_{j,t-1}$, and the mean cross-product of de-meaned true weights and pricing mistakes by $\widehat{\text{cov}}_{J(j,t-1)}(w, \epsilon)$. Then

$$
\begin{aligned}
W_{j,t-1}^{\text{VW/EW}} &:= \sum_{\forall k \in J(j)} \frac{W_{k,t-1}}{N_{J(j)}}, \\
&= \overline{w}_{j,t-1}(1 + \overline{\epsilon}_{j,t-1}) + \widehat{\text{cov}}_{J(j,t-1)}(w, \epsilon), \\
&= \approx \overline{w}_{j,t-1}.
\end{aligned} \tag{12}
$$

In the last line we relied on the fact that, in our calculations, the average size of each bucket is about 160 firms. Under FI assumptions, the mean mispricing $\overline{e}$ must therefore be quite small, and so must be the average crossterm. It follows that the weight adjustment versus the original market weights, $W_{j,t-1} - W_{j,t-1}^{\text{VW/EW}}$, loads quite heavily on the mispricing error:

$$
\begin{aligned}
W_{j,t-1} - W_{j,t-1}^{\text{VW/EW}} &\approx w_{j,t-1}(1 + e_{j,y-1}) - \overline{w}_{j,t-1}, \\
&= (w_{j,t-1} - \overline{w}_{j,t-1}) + w_{j,t} e_{j,t-1}.
\end{aligned} \tag{13}
$$

Again, since all firms are quite similar, the deviations in the true weights from their local average are small so that any covariance between true weight gaps and expected returns must be small. That is, there is not much room left for 'value' or especially size effects. The second term on the right however, contains all of the error, so that the drag effect should still be fully present up. In short, the difference between the mean returns from the standard VW and the VW/EW portfolios contains mostly drag.

In all this, the VW/EW strategy differs systematically from FI. FI is not taking care to keep the adjustments in the weights minimal. Moreover, as we saw, the adjustments are systematically related to size and 'value', and therefore induce a style effect, although in an unstable way over time. To sum up, then, the comparison VW versus VW/EW represents essentially the drag effect, and VW/EW versus FI estimates mostly the style effect in FI. Since we do not use regression, we need no assumption that the style exposures in FI's return are constant in any way.

In the above, the underlying vigintiles were size-based. We can then repeat the same analysis starting from BtM-sorted vigintiles, a procedure that expressedly neutralises the value factor; and finally we work with a double sort (five size classes, four 'value' classes). Each of the procedures may leave some priced risk in the mixed portfolios; the double-sort version, for instance, still does not neutralise reversal. So if we nevertheless find that, after controlling for just one or two types of risk, the extra return becomes unimportant, our conclusion that drag is insignificant is conservative.

In addition, we can apply variants of the above design to study how much return FI obtains from its risk-loaded weights, and get this answer without resorting to regression. In step 1 we again put stocks into buckets (per size vigintile, for instance) and again assign money to each bucket proportionally to the bucket stocks' aggregate market cap,[22] but within each of these size buckets we now assign, in step 2, weights based on fundamental variables or lagged market cap. These mixed strategies are again referred to as $\text{Weighting}_{across}/\text{Weighting}_{inside}$ with, for instance, VW/BV indicating value weighting across vigintiles and book-value based weighting within the vigintiles. In terms of drag avoidance, bucket per bucket, the FI logic holds: fundamental weights should avoid most or all of the correlation with pricing errors. But so do EW vigintile portfolios, and they do so without picking up much of a style shift. So if FI-weighted vigintile portfolios do generate higher returns, we have a direct estimate of how much return is a reflection of style shifts within buckets. Shifts in exposure to HML, for instance, are especially likely in BV-weighted

[22] We always weight by market cap across buckets. Weighting in proportion to BtM across cells would not make sense: we do know that BtM matters. What FI claims to beat is VW, and we want to know why that happens.

portfolios, even when overall bucket weights are set by market value: since within each size bucket the market values are rather similar, weighting by book value is close to weighting by BtM, creating HML exposure that should show up as an extra return above the bucket's EW return. Lagged weighting, similarly, is likely to load on reversal, especially among small stocks.

To sum up: the key figure is the extra return from VW/EW: it shows the extra return when value weighting within cells is replaced by equal weighting, and it provides an estimate of drag with little or no style effects. The other strategies, corresponding to alternative weights within cells, just tell us whether chasing a style within a cell still matters; it provides evidence of the local importance of priced risk, not of drag.

4.2 FI's performance: empirics

Recall how in Table 3 FI was found to offer extra returns of 3–4 % *p.a.* relative to a VW strategy; EW did even better; and LW performed poorly. That information, with the associated t-statistics, is repeated in the first panel of Table 6, in the columns under the heading 'pure strategies'. Recall also that under VW, weighting inside and across vigintiles is based on the same criterion. In the three additional panels, we show results when a FI weighting scheme applies within each of the 20 cells; across cells we still weight by market value. The three panels differ in terms of how these cells are composed: either 20 size vigintiles, or 20 'value' vigintiles, of five size quintiles each subdivided into four 'value' quartiles.

The VW/EW line in Table 6 tells us that, regardless of how we do the sorting, equal weighting within vigintiles always reduces the raw 0.40% extra return to insignificance, and often even to negative levels. The most flattering estimate, after stripping out size-related elements, is 5 bp/month, surrounded by statistical question marks. Working with just BtM-based vigintiles, to control for BtM instead of size, further lowers the extra return to minus 11 bp/month; the double-sort-based estimate is essentially zero.

We conclude, first, that in Treynor's EW strategy drag is economically and statistically small, however convincing the *a priori* case may sound. Second, when we neutralise just the BtM aspect, we get the strongest drop in value, followed by the double-sort neutralisation, and then

Table 6: **Performance of mixed portfolio strategies (VW/.)**

[A] pure strategies (table 3)			VW/[.] mixtures, with vigintiles set up to neutralise ...						
				[B1] size		[B2] 'value'		[B3] both	
style	$\Delta\bar{r}$	t-stat	style	$\Delta\bar{r}$	t-stat	$\Delta\bar{r}$	t-stat	$\Delta\bar{r}$	t-stat
EW/EW	0.406	2.423***	VW/EW	0.050	1.247	-0.115	-0.650	-0.004	-0.048
BV/BV	0.276	3.873***	VW/BV	0.211	3.236***	0.012	1.903*	0.039	1.907*
SL/SL	0.249	2.612***	VW/SL	0.168	1.889	0.038	0.628	0.063	0.915
CF/CF	0.222	2.550***	VW/CF	0.198	2.333***	0.041	0.866	0.070	1.209
LW1/LW1	0.021	1.120	VW/LW1	-0.002	-0.090	0.036	2.400***	0.037	2.347***
LW3/LW3	0.016	0.448	VW/LW3	-0.036	-1.040	-0.010	-0.441	-0.013	-0.561
LW6/LW6	0.009	0.167	VW/LW6	-0.047	-0.895	-0.027	-0.727	-0.029	-0.774

Key The table shows mean extra monthly returns, $\Delta\bar{r}$, *i.e.* average returns in excess of the value-weighted index, for the pure strategies (first panel, a repeat of the fifth and sixth column of Table 3) and three mixed portfolio strategies (panels 2-4). For the mixed portfolio strategies, every month, all stocks are assigned to one of the 20 size vigintiles, or 20 'value' vigintiles, or five size quintiles each subdivided into four 'value' quartiles. Capital is allocated to each vigintile on the basis of market weight, but within vigintiles we use either equal weighting (EW), lagged value weights as observed 1, 3 or 6 months ago (LW1-6), or weights based on book value (BV) or free cash flow (CF), or sales (SL). Only the VW/EW numbers test for drag; the other lines merely assess the within-bucket style effects.

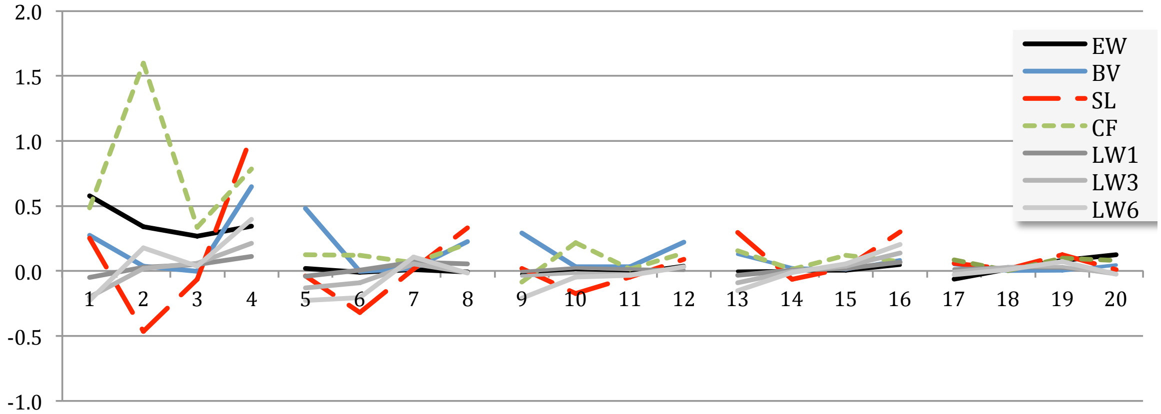
the correction for size effects.

Before turning to the results per vigintile we can still extract some useful insights from the other lines. The effect of stripping out the size factor is statistically and economically minute for the three LW strategies; no surprise there. It is true that in two cases the VW/LW1 strategy has a statistically unambiguous positive effect. But it is economically small, and it is not about drag: any drag should have been visible in the EW line already, and should have been stronger in the LW3 or LW6 line than in L1; we see none of that. The likely explanation is reversal, as documented before. Still, the VW/LW strategies come up with, economically, remarkably low extra returns over the VW/VW alternative.

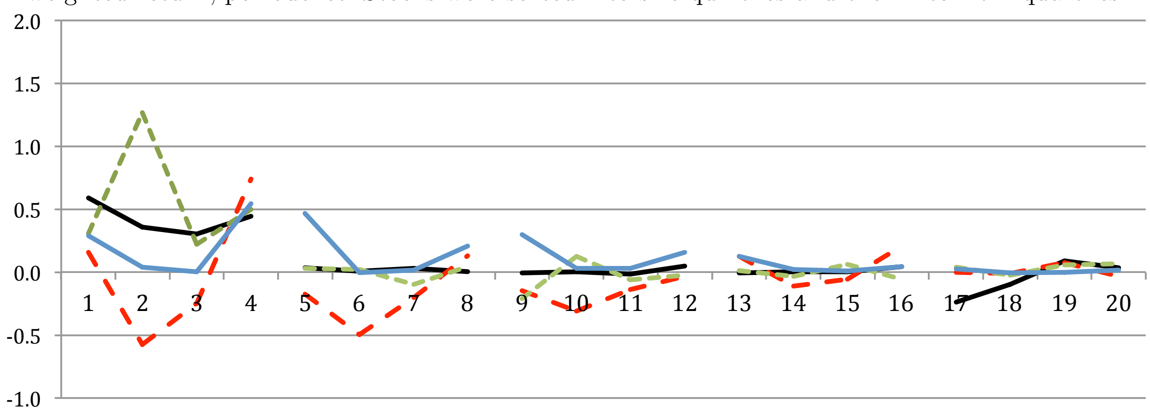
Among the three FI investment rules, only sales-based investing is partly size-loaded: one quarter of the return disappears after controlling for size, and what is left is statistically insignificant. The other FI strategies are mostly 'value'-loaded: size explains very little for BV- or CF-based portfolios, while 'value' does pick up far more and takes away the *prima facie* superior returns.

It must be acknowledged that underneath these unpromising results at the portfolio level there

Figure 4: **Performance of EW, FI and LW portfolio strategies per size×'value' cell, relative to value weighting within the vigintile.**



Panel A: Average extra returns for EW-, FI- and LW-weighted portfolios, in excess of value-weighted return, per bucket. Stocks were sorted into size quintiles and then into BtM quartiles.



Panel B: FF alphas for extra returns for EW- and FI-weighted portfolios, in excess of value-weighted return, per bucket. Stocks were sorted into size quintiles and then into BtM quartiles.

Key Panel A of this graph is the graphical representation of the vigintile-level results underlying the aggregate extra-return numbers $\Delta\bar{r}$ in the last panel of Table 6. Every month, stocks are sorted into size quintiles and, within each of these, into 'value' quartiles. The top figure shows, for each size-'value' bucket, its mean EW, FI and LW monthly return in excess of the bucket's value-weighted portfolio (i.e. the bucket's extra return $\Delta\bar{r}$), with FI = Sales (SL), Cash Flow (CF), Book Value (BV), and Lagged Weights (LW1, LW3, LW6). The extra return based on EW is an estimate of drag avoided; those based on FI are an indication of the local importance of style-chasing. In Panel B we regress the same extra returns on the FF factors and show the alphas.

are some significant figures at the individual bucket level. Panel A of Figure 4 offers the graphical representation of the vigintile-level results underlying the aggregate extra-return numbers $\Delta\bar{r}$ in the last panel of Table 6. Significant numbers occur at the small-stock end, and the numbers we are talking about for VW/EW, our estimate of drag avoided, are of the order of 0.30-0.50% per month. The likely explanation is not necessarily reversal. It is true that we do not control for momentum or reversal, in this test; but if reversal were very big in that group of stocks, it

Table 7: **Results for FF alphas of estimated drag per size, BtM, or double-sort bucket**

A. Size buckets			B. BtM buckets			C. Size-BtM buckets		
id	α	(t)	id	α	(t)	id	α	(t)
1	0.427	19.304	1	-1.216	-4.867	1.1	0.590	8.855
2	0.039	3.513	2	-0.879	-5.077	1.2	0.355	5.534
3	-0.001	-0.179	3	-0.366	-2.126	1.3	0.301	7.471
4	0.004	0.508	4	-0.420	-2.752	1.4	0.443	12.449
5	0.002	0.322	5	-0.081	-0.513	2.1	0.033	1.038
6	-0.007	-1.188	6	-0.087	-0.552	2.2	0.009	0.292
7	0.003	0.613	7	0.031	0.211	2.3	0.029	1.241
8	0.008	1.456	8	-0.069	-0.521	2.4	0.003	0.142
9	-0.005	-0.963	9	-0.051	-0.383	3.1	-0.008	-0.221
10	0.003	0.490	10	0.100	0.711	3.2	0.001	0.042
11	-0.004	-0.665	11	0.043	0.312	3.3	-0.018	-0.925
12	0.004	0.715	12	0.009	0.064	3.4	0.048	2.233
13	0.006	1.221	13	0.315	1.954	4.1	-0.009	-0.305
14	0.004	0.783	14	0.432	2.660	4.2	0.004	0.217
15	0.003	0.445	15	0.122	0.676	4.3	0.003	0.143
16	0.006	1.012	16	0.258	1.405	4.4	0.044	1.437
17	-0.004	-0.581	17	0.237	1.163	5.1	-0.240	-2.207
18	-0.002	-0.171	18	-0.022	-0.109	5.2	-0.097	-1.364
19	-0.021	-1.862	19	0.622	2.599	5.3	0.090	1.112
20	-0.008	-0.130	20	1.438	6.577	5.4	0.035	0.231

Key Every month, stocks are sorted into either 20 size vigintiles (Panel A), or 20 BtM vigintiles (B) or first into quintiles and, within each of these, into BtM quartiles (C). For each bucket, the EW monthly return in excess of the bucket's value-weighted portfolio (*i.e.* the bucket's extra return Δr_v) is regressed on the FF factors. The table provides the three-factor FF alphas (pictured in Panel B of Figure 4) and their t-statistics.

should have shown up in the VW/LW returns too, which is not really the case. Regardless of what the source is, the smallest-stock quintile represents too little money to be important in a large, diversified portfolio: in the portfolio-level numbers we just discussed, for instance, there is no trace of an extra performance.

Subject to the caveat of unstable factor exposures we lastly check whether there are any remaining systematic risks in the mixed portfolio returns and, if so, whether the extra returns are affected. Four-factor regressions reveal traces of exposure (not shown), but the returns we just saw hardly differ from the alphas that remain after risk—or some average version of risk—has been stripped out, as the comparison between the alphas in Panel B and the average extra returns in Panel A shows. As obvious also from the t-tests in Table 7, abnormal extra returns are confined to the smallest stocks (and to the extreme BtM stocks, if one does not also control for size), which do not really matter in a portfolio that does not chase styles.

5 Conclusion

Building portfolios based on weights from fundamental corporate data instead of weights equal to market cap should increase portfolio returns, and our first-pass results do look promising, like in many earlier studies. FI claims that these extra returns show how much of the drag effect is avoided when weights are less correlated with pricing errors. But also style shifts or market timing can be at work, which would mean that not all of the extra returns stem from genuine selectivity.

A quick look at the weights reveals, unsurprisingly, massive upward corrections for small and 'value' stocks. However, the correction patterns across the weight spectrum are also quite variable over time: sometimes small stocks' overweights are at least quadratically related to size, while at other times the link is more linear, albeit with very varying slopes over time. This fits in with the instability of factor exposures, and their correlation with factor returns, documented via Treynor-Mazuy (1966)-style non-linear regressions. The implication is that alphas estimated from return-based style regressions are unreliable measures for FI's drag-related gain, even if one takes the asset pricing models for granted. In addition, we find that the alphas are not robust to minor modifications in the multifactor model.

Our methodological contribution is to provide an alternative, and direct, way to estimate the drag-related gain from FI: we introduce mixed portfolio strategies, focusing primarily on 'VW/EW', i.e. value weighting across style vigintiles and equal weighting within vigintiles, with vigintiles being the result of a sort on size, or on 'value', or on both. This strategy should still avoid the drag effect and largely sidestep the style effect(s) targeted in the sorting. The smallest stocks, we find, are still exposed to reversal, but may also offer some alpha over and above that; as a three-way sort would leave us with buckets that contain too few stocks, we cannot really tell. That said, micro firms hardly matter in a large, diversified portfolio, as our overall-portfolio results demonstrate. For the all-stock VW/EW portfolio we find that the extra return from this mixed portfolio strategy relative to the value-weighted index is economically and statistically insignificant.[23]

Mixing value weighting (across vigintiles) with fundamental weights (within vigintiles), instead of with equal weights, pays extra returns for far more stocks than just the small or high-BtM ones, and shows that 'value' effects are the stronger component of FI's extra return, compared to size bias. The main message is that (i) the drag-related gain from FI can, and should, be estimated via other means than alphas from return-based style regressions; and (ii) if one does so; it turns out to be economically negligible. For the portfolio manager, this also implies that a genuine style-based strategy would be a more transparent and more efficient alternative for FI, with 'value' being the most profitable style.

How important is the 'value' factor in FI? The four-factor regressions in Table 4 may not provide reliable measures of abnormal return, but still offer indications of average exposures. On average, FIs exposure to 'value' is roughly 0.3.[24] Size bias is much less important, on average: the highest excess exposure estimate is just 0.075, and while size exposure is always positive for BV- and SL-based weights it turns negative for CF weighting. Taking into account that the premium for SMB exposure is comparable, in our sample, to HML's, the return contribution of size bias in FI is minor compared to that of the 'value' bias. Momentum bias, or more accurately reversal bias, is quite important for small and high-BtM stocks but those small stocks do not mean much in a large, diversified portfolio. Table 6 brings the same message: for the three FI-based strategies, correcting for size bis does not make much of a dent in FI's 0.22-0.28% raw extra return; correcting for BtM bias, in contrast, wipes out essentially all extra returns.

If FI's return wholly or mostly is factor-driven, should that be an alternative argument for applying FI's recommendations? The results in Table 6 show that, via mixed strategies, one can

[23]This consideration also implies that our findings cannot be due to our reliance on Thomson Reuters Datastream. Our dataset does differ from the usual CRSP or Compustat files, used in this kind of research, with respect to the presence of more tiny stocks. This does load the dice in favor of FI in the sense that these smallest stocks, being traded in illiquid markets and lacking analyst attention, are prime candidates for pricing errors; in that sense, our conclusion that the drag effect is trivial is conservative. Adding such an idiosyncratic group of stocks does create a risk that the overall picture may be obscured by noise. But the value weight of those smallest vigintiles is absolutely tiny, so the aggregate estimate is essentially unaffected.

[24]By construction, the VW market portfolio has a 'value' exposure of 0 and the HML fund has 'value' exposure 1.

also capture much of FIs return without massively deviating from the market weights. Most of the pure FI return can be captured by (i) sorting stocks by just size, (ii) setting weights across vigintiles on the basis of market values, and (iii) setting intra-vigintile weights on the basis of accounting data; see rows 2-4 in column [B1]. Using small-cap funds that are value-weighted rather than book-value-weighted is less efficient: they just pick up a (weak) size effect and a reversal effect. But the most direct and open way to get SMB or especially HML exposure is, of course, to construct short-long portfolios of that type.

Our main finding, we think, is that drag avoidance does not really work in any meaningful way. To academia, an implied message, relating to the level and nature of pricing errors, is that pricing errors are not simultaneously large, idiosyncratic and short-lived, except perhaps for the smallest 10-15% of stocks (mostly OTC). That does leave many candidate explanations: pricing errors, if any, are either small, or highly persistent, or have a big market-wide component.

References

Amenc, N., F. Goltz, and V. Le Sourd (2008). "The Performance of Characteristics-based Indices". *European Financial Management* 15(2), 241-278.

Arnott, R. D., J. Hsu and P. Moore (2005). "Fundamental Indexation". *Financial Analysts Journal* 61(2), 83-99.

Arnott, R. D., J. Hsu, J. Liu and H. M. Markowitz (2007). "Does Noise Create the Size and Value Effects?". Working paper, http://ssrn.com/abstract=936272.

Arnott, R. D., and H. M. Markowitz (2008). "Fundamentally Flawed Indexing: Comment". *Financial Analysts Journal* 64(2), 12-14.

Asness, C. (2006). "The Value of Fundamental Indexing". *Institutional Investor*, October 19, 67-71.

Asparouhova, E., H. Bessembinder, and I. Kalcheva (2013). "Noisy prices and inference regarding returns". *Journal of Finance* 68, 665-714.

Ball, R., and Brown, P. (1968).. "An empirical evaluation of accounting income numbers". *Journal of Accounting Research*, 6(1968), pp. 159-178

Ball, R., Gerakos J., Linnainmaa J. T., and Nikolaev V., (2015).. "Deflating profitability". *Journal of Financial Economics* 117, 225-248

Ball, R., Gerakos J., Linnainmaa J. T., and Nikolaev V., (2016).. "Accruals, cash flows, and operating profitability in the cross section of stock returns". *Journal of Financial Economics* 121 (2016) 2845

Balatti, M., Brooks, C., and Kappou, K. (2017),. "Fundamental indexation revisited: New evidence on alpha". *International Review of Financial Analysis* 51, 115

Blitz, D., B. Van der Grient, and P. Van Vliet (2010). "Fundamental Indexation: Rebalancing Assumptions and Performance". *Journal of Index Investing,* 1(2), 82-88.

Bhattacharya U., and N. Galpin (2005). "Is stock picking declining around the world?". AFA Chicago Meetings Paper, http://ssrn.com/abstract=849627.

Biltz D., and L. A. P. Swinkels L.A.P. (2008). "Fundamental Indexation: An active value strategy in disguise". *Journal of Asset Management* 9(4), 264-269.

Blume, E. Marshall, and R. F. Stambaugh (1983). "Biases in computed returns : An application to the size effect". *Journal of Financial Economics* 12(3), 387-404.

Bogle J. C., and B. G. Malkiel (2006). "Run on a paradigm". *Wall Street Journal*, June 27.

Boudt, K., J. Darras, G.H. Nguyen and B. Peeters, 2015. Smart beta equity investing through calm and storm. Forthcoming in E. Jurczendko,*Practical Guide to Risk-Based Investing*, 1st Edition, Elsevier, Amsterdam, ISBN 9781785480089, Nov 2015

Brennan, and Wang (2010). "The Mispricing Return Premium". *Review of Financial Studies* 23(9), 3437-3468.

Brown, S., and J. Warner.. "Using Daily Stock Returns: The Case of Event Studies.". Journal of Financial Economics, 14 (1985), 1-31.

Carhart, M. (1997). "On Persistence in Mutual Fund Performance". *Journal of Finance* 52(1), 57-82.

Chan, K. C., and N. Chen (1991). "Structural and return characteristics of small and large firms" *Journal of Finance* 46, 1467-1484.

C. Chen, R. Chen, and G.W. Bassett (2007). "Fundamental Indexation via Smoothed Cap Weights" *Journal of Banking and Finance* 31(11), 3486-3506.

Chen, N., and F. Zhang (1998). "Risk and Return of Value Stocks". *The Journal of Business* 71(4), 501-535.

De Moor, L. and P. Sercu, 2013. The Smallest Firm Effect: an International Study. *Journal of International Money and Finance*, 2013, vol. 32(1), 129-155

De Moor, L. and P. Sercu, 2015. The Smallest Stocks are Not Just Smaller: Global Evidence. *The European Journal of Finance*, 21(1), January 2015, 51-70

Dopfel, F.E., 2008, The arithmetic of Fundamental Indexing. *The Journal of Investing*, 17(2), 114-118

Estrada, J. 2008, Fundamental indexation and international diversification, Journal of Portfolio Management, (Spring), Vol.34, No.3, 93-109

Fama and French, 2007. Nouveau indexes, noise and the nonsense of active management. *Journal of Indexes*, 8

Fama, E.F., and French, K. R. (2016).. "Dissecting Anomalies with a Five-Factor Model". *Revue of Financial Studies* 29(1), 69-103.

Ferreira, R. and J. D. Krige, 2011. The application of fundamental indexing to the South African equity market for the period 1996 to 2009, *Investment Analysts Journal* 73, 1-12.

Forbes, B. M. and A. K. Basu, 2011. Does Fundamental Indexation Lead to Better Risk Adjusted Returns? New Evidence from Australian Securities Exchange. WP, Queensland UT

Foerster S., Tsagarelis J., Wang W., (2017).. "Are Cash Flows Better Stock Return Predictors Than Profits?". *Financial Analysts Journal*, 73(1), 73-99

Graham, J., 2012. Comment on the Theoretical and Empirical Evidence of Fundamental Indexing. *Journal of Investment Management*, 10(1), 90-90

Hemminki, J and V. Puttonen, 2008, Fundamental indexation in Europe, *Journal of Asset Management*, Vol.8, No.6 (February), 401-405

Henriksson R.D. and Merton R.C. (1981).. "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills". *The Journal of Business*, 54(4), 513-533

Hou K., Xhen X., and Zhang L. (2015).. "Digesting Anomalies: An Investment Approach". *Revue of Financial Studies* 28(3), 650-705.

Houwer, Roel and Plantinga, Auke, Fundamental Indexing: An Analysis of the Returns, Risks and Costs of Applying the Strategy (February 15, 2009). Available at SSRN: https://ssrn.com/abstract=1343879 or http://dx.doi.org/10.2139/ssrn.1343879

Hsu, J.C. (2006), Cap-weighted portfolios are sub-optimal portfolios, *Journal of investment*, 4(3); 1-10.

Hsu, J.C. (2008), Why Fundamental Indexation Might—or Might Not—Work: A Comment, *Financial Analysts Journal,* March/April 2008, 64(2): 17-18.

Ikenberry, D., J. Lakonishok and T. Vermaelen, 2002. Stock Repurchases in Canada: Performance and Strategic Trading. *The Journal of Finance* 55(5), 2373-2397

Ince, O., S., Porter, R., B., 2006. Individual equity return data from Thomson Datastream: handle with care!. *Journal of Financial Research*, Vol. XXIX, No. 4, 463-479.

Jun, D., B. G. Malkiel, 2008. New Paradigms in Stock Market Indexing. *European Financial Management* 14(1), 118-126

Kaplan P.D. (2008), Why Fundamental Indexation Might—or Might Not—Work, *Financial Analysts Journal*, 64(1); 32-39.

Lakonishok, J. and T. Vermaelen, 1990. Anomalous Price Behavior Around Repurchase Tender Offers. *The Journal of Finance* 45(2), 455 - 477

McQuarrie, E. F., 2008. Fundamentally Indexed or Fundamentally Misconceived: Locating the Source of RAFI Outperformance. *The Journal of Investing*, 17(4), 29-37

Mihm, M. and H. Locarek-Junge, 2010. Empirical Examination of Fundamental Indexation in the German Market. In *Advances in data analysis, data handling, and business intelligence. Studies in Classification, Data Analysis, and Knowledge Organization*, Springer, Berlin-Heidelberg, 2010, Part 7, 471-479.

Miziolek, Thomas and Adam Zaremba. Romanian Journal of Economic Forecasting vol. 20 (1), 2017, p23-p37

Neukirch, T., 2008. Alternative Indexing with the MSCI World Index. http://ssrn.com/abstract=1106109 or http://dx.doi.org/10.2139/ssrn.1106109.

Novy-Narx, R., (2013).. "The other side of value: The gross profitability premium". *Journal of Financial Economics* 108, 128

Perold A. F., 2007, Fundamentally Flawed Indexing, *Financial Analysts Journal*, 63(6), 31-37.

Peltomäki, J., 2010. Nonlinear Exposures of Fundamental Index Returns. *The Journal of Wealth Management* 13(3), 96-106

Peyer, U. and T. Vermaelen, 2009. The Nature and Persistence of Buyback Anomalies. *Review of Financial Studies* 22(4), 1693-1745

Piljak, V. and Swinkels, L., 2017. Fundamental indexation for developed, emerging, and frontier government bond markets. *Journal of Asset Management* 18(5), 405-420

Plyakha, Y., R. Uppal and G. Vilkov, 2014. Equal or Value Weighting? Implications for Asset-Pricing Tests. Working paper

Roll R., 1984, A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market, *Journal of Finance*, 39 (4), 1127-1139.

Siegel , J., 2006. The Noisy Market Hypothesis, *Wall Street Journal*, June 14

Stotz O., D́ohnert K., and Wanzenried G., 2007. Do fundamental indexes produce higher risk-adjusted returns than market cap indexes? Evidence for European stock markets. *Financial Markets and Portfolio Management*, 24(3), 219-243

Subramanian, M, P. Kulkarni, R. Kouzmenko, D. Melas, 2011. Capturing the Value Premium. MSCI Barra Research Paper No. 2011-12, New York, NY. DOI: http://dx.doi.org/10.2139/ssrn.1915261

Tabner, I.T. (2009), Benchmark concentration: capitalization weights versus equal weights in the FTSE 100 Index, *Multinational finance journal*, 13(1/2); 161-180.

Treynor, J. and K. Mazuy, 1966. Can mutual funds outguess the market? *Harvard Business Review*, 44(4), 131136

Treynor, J., 2005. Why Market-Valuation-Independent Indexing Works. *Financial Analysts Journal* 61(5), 65-69

Treynor, J., 2008. Fundamentally Flawed Indexing: Comments. *Financial Analysts Journal*, 64(2): 14.

Walkshäusl, C. and S Lobe, 2010. Fundamental indexing around the world. *Review of Financial Economics*, 19(3), 117-127.

Zaremba, A. and and T. Miziołek (2017). Nothing Lasts Forever (and Everywhere): Fundamental Indexation at the Global Level. *The Journal of Index Investing* 8(3), 6-20.

Appendix Tables

Table 1: **(Appendix) Average return v. average monthly return over preceding six months, per bucket**

Sort on size			Sort on BtM			Double sort		
avg(r_t)	avg($r_{(t-6,t-1)}$)	p-val	avg(r_t)	avg($r_{(t-6,t-1)}$)	p-val	avg(r_t)	avg($r_{(t-6,t-1)}$)	p-val
3.989	0.417	0.000	-0.413	5.935	0.000	0.994	3.743	0.000
1.854	1.328	0.110	0.184	3.769	0.000	1.837	1.976	0.808
1.574	1.299	0.334	0.576	3.248	0.000	1.639	1.106	0.047
1.364	1.539	0.596	0.592	2.727	0.000	2.901	0.083	0.000
1.397	1.740	0.341	0.854	2.422	0.000	-0.114	4.715	0.000
1.349	1.452	0.798	0.925	2.230	0.000	0.908	1.933	0.004
1.294	1.471	0.672	1.072	2.017	0.003	1.300	1.028	0.320
1.181	2.059	0.041	1.060	1.868	0.010	2.402	0.260	0.000
1.228	1.984	0.057	1.127	1.708	0.069	-0.080	4.378	0.000
1.271	1.986	0.053	1.156	1.555	0.168	0.996	2.032	0.003
1.208	1.924	0.054	1.321	1.369	0.946	1.581	1.137	0.165
1.137	1.997	0.016	1.316	1.203	0.667	2.452	0.389	0.000
1.150	1.915	0.032	1.501	1.094	0.125	0.376	3.525	0.000
1.145	2.057	0.010	1.597	0.951	0.024	0.993	1.970	0.003
1.124	2.008	0.010	1.643	0.815	0.003	1.530	1.127	0.162
1.033	1.972	0.003	1.810	0.683	0.000	2.159	0.501	0.000
1.033	1.871	0.007	2.063	0.532	0.000	0.676	2.513	0.000
1.053	1.862	0.007	2.224	0.382	0.000	0.984	1.553	0.031
0.986	1.720	0.010	2.709	0.110	0.000	1.310	1.047	0.282
0.923	1.613	0.004	3.984	-0.421	0.000	2.014	0.688	0.000

Key Every month, all firms are sorted into twenty buckets on the basis of size (Panel A) or Book to Market (Panel B) or first size (quintiles) and then value (quartiles) (Panel C). Each panel shows the average monthly returns of the firms (i) when in the bucket and (ii) during the six months preceding their bucket membership, alongside a two-sided p on the average difference.

Table 2: **(Appendix) FI-proposed weights versus market weights, per bucket**

weights per vigintile, percent				FI weight/mkt weight			two-sided p value		
mkt W	w^* BV	w^* SL	w^* CF	w^* BV	w^* SL	w^* CF	w^* BV	w^* SL	w^* CF
A: Size sorted									
0.035	0.101	0.128	0.084	2.920	3.722	2.421	0.000	0.000	0.000
0.056	0.142	0.175	0.119	2.536	3.110	2.129	0.000	0.000	0.000
0.080	0.188	0.218	0.148	2.339	2.706	1.837	0.000	0.000	0.000
0.108	0.235	0.263	0.177	2.181	2.445	1.641	0.000	0.000	0.000
0.142	0.309	0.335	0.234	2.179	2.358	1.650	0.000	0.000	0.000
0.186	0.375	0.402	0.282	2.019	2.164	1.520	0.000	0.000	0.000
0.239	0.436	0.456	0.357	1.825	1.905	1.491	0.000	0.000	0.000
0.305	0.523	0.560	0.417	1.713	1.834	1.363	0.000	0.000	0.000
0.391	0.641	0.678	0.510	1.638	1.732	1.303	0.000	0.000	0.000
0.502	0.805	0.831	0.599	1.605	1.656	1.195	0.000	0.000	0.000
0.641	0.971	1.020	0.704	1.515	1.592	1.099	0.000	0.000	0.000
0.825	1.213	1.240	0.882	1.469	1.502	1.068	0.000	0.000	0.000
1.071	1.575	1.550	1.176	1.470	1.447	1.098	0.000	0.000	0.000
1.417	2.009	2.042	1.598	1.418	1.441	1.128	0.000	0.000	0.000
1.927	2.557	2.681	2.081	1.327	1.392	1.080	0.000	0.000	0.000
2.689	3.439	3.692	2.808	1.279	1.373	1.044	0.000	0.000	0.035
3.939	4.897	5.266	4.070	1.243	1.337	1.033	0.000	0.000	0.008
6.397	7.526	7.840	6.403	1.176	1.225	1.001	0.000	0.000	0.918
12.76	14.62	15.16	13.47	1.146	1.188	1.055	0.000	0.000	0.000
66.29	57.44	55.46	63.89	0.866	0.837	0.964	0.000	0.000	0.000
B: Book-to-Market sorted									
9.659	2.126	4.005	4.192	0.220	0.415	0.434	0.000	0.000	0.000
12.07	4.869	6.535	6.913	0.404	0.542	0.573	0.000	0.000	0.000
10.54	5.671	6.838	7.170	0.538	0.649	0.681	0.000	0.000	0.000
9.217	6.086	7.020	7.098	0.660	0.762	0.770	0.000	0.000	0.000
8.227	6.330	7.437	7.066	0.769	0.904	0.859	0.000	0.000	0.000
7.404	6.500	7.300	6.867	0.878	0.986	0.927	0.000	0.227	0.000
6.320	6.249	6.847	6.621	0.989	1.083	1.048	0.227	0.000	0.001
5.331	5.923	6.157	6.223	1.111	1.155	1.167	0.000	0.000	0.000
4.866	5.950	5.891	5.714	1.223	1.211	1.174	0.000	0.000	0.000
4.516	5.992	5.700	5.648	1.327	1.262	1.251	0.000	0.000	0.000
3.899	5.605	5.262	5.282	1.438	1.350	1.355	0.000	0.000	0.000
3.313	5.183	4.677	4.662	1.564	1.412	1.407	0.000	0.000	0.000
2.914	4.927	4.173	4.234	1.690	1.432	1.453	0.000	0.000	0.000
2.526	4.613	3.892	3.869	1.826	1.541	1.532	0.000	0.000	0.000
2.233	4.433	3.674	3.726	1.985	1.645	1.668	0.000	0.000	0.000
2.047	4.377	3.590	3.561	2.138	1.753	1.739	0.000	0.000	0.000
1.947	4.550	3.452	3.751	2.337	1.773	1.926	0.000	0.000	0.000
1.322	3.552	2.767	2.746	2.687	2.093	2.076	0.000	0.000	0.000
0.990	3.223	2.344	2.229	3.257	2.369	2.252	0.000	0.000	0.000
0.667	3.842	2.439	2.430	5.764	3.659	3.645	0.000	0.000	0.000
C: Double sort									
0.041	0.019	0.044	0.045	0.465	1.077	1.107	0.000	0.000	0.000
0.040	0.046	0.068	0.048	1.136	1.680	1.179	0.000	0.000	0.000
0.064	0.114	0.140	0.090	1.788	2.186	1.414	0.000	0.000	0.000
0.134	0.487	0.532	0.344	3.642	3.982	2.575	0.000	0.000	0.000
0.169	0.082	0.132	0.129	0.488	0.785	0.764	0.000	0.000	0.000
0.186	0.209	0.226	0.168	1.122	1.218	0.902	0.000	0.000	0.000
0.249	0.437	0.437	0.311	1.754	1.753	1.245	0.000	0.000	0.000
0.268	0.916	0.956	0.682	3.415	3.566	2.545	0.000	0.000	0.000
0.566	0.295	0.386	0.353	0.520	0.681	0.623	0.000	0.000	0.000
0.661	0.730	0.763	0.546	1.105	1.155	0.826	0.000	0.000	0.000
0.670	1.152	1.170	0.831	1.720	1.747	1.241	0.000	0.000	0.000
0.462	1.453	1.450	0.965	3.143	3.137	2.089	0.000	0.000	0.000
2.092	1.111	1.578	1.278	0.531	0.754	0.611	0.000	0.000	0.000
2.284	2.516	2.798	2.189	1.102	1.225	0.958	0.000	0.000	0.116
1.810	3.073	3.060	2.217	1.698	1.691	1.225	0.000	0.000	0.000
0.917	2.880	2.530	1.978	3.140	2.758	2.157	0.000	0.000	0.000
46.84	23.58	29.69	30.63	0.503	0.634	0.654	0.000	0.000	0.000
25.27	27.11	28.04	28.12	1.073	1.110	1.113	0.000	0.000	0.000
12.09	19.98	16.87	18.32	1.653	1.395	1.515	0.000	0.000	0.000
5.192	13.81	9.124	10.75	2.660	1.758	2.070	0.000	0.000	0.000

Key Every month, all firms are sorted into twenty buckets on the basis of size (Panel A) or Book to Market (Panel B) or first size (quintiles) and then value (quartiles) (Panel C). We first show what weights the buckets' stocks get, on average, when investment is by market cap, BV, SL or CF; then each FI-based weights divided by its market weight; and lastly the p-values of a two-sided t-test for the Null that the vigintile's average ratio has an expectation of unity.