Journal of
Risk and Financial Management

Article

Portfolio Strategies to Track and Outperform a Benchmark

Paskalis Glabadanidis

Finance Discipline, University of Adelaide Business School, Level 12, 10 Pulteney Street, Adelaide, SA 5005, Australia; paskalis.glabadanidis@adelaide.edu.au; Tel.: +61-(8)-8303-7283; Fax: +61-(8)-8223-4782

Received: 13 July 2020; Accepted: 29 July 2020; Published: 1 August 2020

Abstract: I investigate the question of how to construct a benchmark replicating portfolio consisting of a subset of the benchmark's components. I consider two approaches: a sequential stepwise regression and another method based on factor models of security returns' first and second moments. The first approach produces the standard hedge portfolio that has the maximum feasible correlation with the benchmark. The second approach produces weights that are proportional to a "signal-to-noise" ratio of factor beta to idiosyncratic volatility. Using a factor model of securities returns allows the use of a larger number of securities than the number of time periods used to estimate the parameters of the factor model. I also consider a second objective that maximizes expected returns subject to a target tracking error variance. The security selection criterion naturally extends to the product of the information ratio and the signal-to-noise ratio. The optimal tracking portfolio is either a one-fund or a two-fund portfolio rule consisting of the optimal hedging portfolio, the tangent portfolio or the global minimum variance portfolio, depending on what constraints are imposed on the objective function. I construct buy-and-hold replicating portfolios using the algorithms presented in the paper to track a widely followed stock index with very good results both in-sample *and* out-of-sample.

Keywords: optimal portfolio weights; benchmarking

JEL Classification: G11; G12

1. Introduction

A frequent question that arises in portfolio management is how on can construct a portfolio of securities that will best mimic the performance of a benchmark index. A passive investment strategy may indicate that the objective of the portfolio is to track the benchmark as closely as possible, while an active investment strategy will mandate that the portfolio outperforms the benchmark. The practitioner literature abounds with many approaches to this problem ranging from the standard stepwise regression through neural networks to genetic algorithms. Unfortunately, most of these applications are numerical in nature and do not yield much intuition regarding how to build a replicating portfolio that is compact and correlates highly with its benchmark.

Roll (1992) is an example of an early paper targeted at practitioners arguing against some common practices of fixing a target portfolio volatility while tracking a benchmark. He shows that unless the portfolio manager gets the volatility right ex ante, the replicating portfolio will do a poor job of tracking the index. In the same spirit, Jorion (2003) demonstrates how additional constraints such as value-at-risk may be necessary to align the incentives of portfolio managers and investors. Stutzer (2003) finds that in equilibrium the benchmarks may become priced risk factors when fund managers try to replicate or outperform the benchmarks. Rudolf et al. (1999) argue against using a mean squared error loss function and in favor of mean absolute deviations between the benchmark and the replicating portfolio returns. They show convincing evidence that this loss function results in more stable portfolio weights that are less sensitive to outliers.

Various statistical techniques have been applied toward the objective of benchmark index replication ranging from time series clustering (Focardi and Fabozzi 2004) to cointegration (Dunis and Ho 2005). Many studies have also investigated the question of actively managing a portfolio that replicates the performance of a benchmark index subject to limits on the tracking error (Burmeister et al. 2004; El-Hassan and Kofman 2003; Israelsen and Cogswell 2006).

Corielli and Marcellino (2006) were among the first to introduce factor models in the analysis of the benchmark replicating problem. Intuitively, factor models for the first two moments of securities returns provide "shrinkage" and reduce the estimation error involved with modeling the expected returns and variance-covariance matrix of returns. Stoyanov et al. (2008) provided an axiomatic approach for general loss functions that are similar conceptually to higher-order lower partial moments of the tracking error. This study included an empirical application with very good results, but unfortunately the authors were forced to use numerical methods to obtain a solution.

Chan et al. (1999) applied multi-factor models to securities' expected returns to find the minimum variance portfolio and maximum information ratio benchmark tracking portfolio. They reported that up to three factors were sufficient in describing the variance-covariance matrix of securities returns and found an adequate minimum variance portfolio. More factors appeared to be needed in order to find a portfolio that minimized the volatility of the tracking error. More recently, Glabadanidis and Zolotoy (2013) proposed several different objective functions for an active portfolio relative to a benchmark and provided optimal portfolio weights in closed-form under a mean-variance framework. Glabadanidis (2014) applied an approximate factor structure to the mean and variance of security returns, thereby providing more straightforward optimal portfolio weight solutions. A further application to an industry standard, mean-variance, value-added function of delegated portfolio management was solved explicitly in Glabadanidis (2020).

In the context of replicating hedge fund returns, Hasanhodzic and Lo (2007) apply linear multi-factor regressions of hedge fund returns on the returns of several asset classes encompassing a broad spectrum of risk exposures. They offer a good way of scaling the replicating portfolio weights to match the volatility of the target in sample. However, the focus of their work is purely on the performance of their replicating strategies relative to the target hedge fund returns and they offer little intuition on how the optimal weights depend on the returns' factor structure. In recent work, Amenc et al. (2010) went beyond the case of linear portfolio weights, only to find that this does not necessarily improve the replication power. However, more recent non-linear innovations using machine-learning like the autoencoder-based strategies seem quite promising. Examples of this approach include Ouyang et al. (2019) and Heaton et al. (2017), among others.

The contribution of this paper is three-fold. First, I extend the factor return framework to allow for multiple pre-specified factors driving the securities returns. I show that in the context of minimizing the tracking error variance, the replicating portfolio weights are proportional to the tangent portfolio weights scaled by the benchmark beta. This approach allows for a number of securities that can be much larger than the number of periods used to estimate the factor model parameters. Second, in a step-down approach I use another objective of picking securities that result in the highest possible deviation from the benchmark given the smallest possible tracking error variance from the previous results. Thirdly, I apply the theoretical results to a widely followed US stock index, the Dow Jones Industrial Average (DJIA). I use two time periods that consist of two years of daily data. The first year is used to estimate the necessary parameters, while the second year is reserved for an out-of-sample test of the replicating portfolio relative to the benchmark. For a narrow index, such as the DJIA, it is possible to replicate their return performance with up to one third of the component stocks.

This paper proceeds as follows. Section 2 presents the theoretical framework and the algorithms for finding subsets of securities that closely track benchmark returns. Section 3 provides an empirical illustration using a popular equity index as a benchmark. Section 4 offers a few concluding comments and suggestions for future research.

2. Theoretical Motivation

Let $R_{y,t}$ be the simple return on a benchmark index y in period t, $R_{j,t}$ be the simple return on basis security j in period t and $R_{f,t}$ be the risk-free rate of return in period t. A tracking portfolio p is composed of N basis securities and the risk-free asset. The simple rate of return of portfolio p in time period t is given by

$$R_{p,t} = \sum_{j=1}^{j=N} w_j R_{j,t} + \left(1 - \sum_{j=1}^{j=N} w_j\right) R_{f,t}. \tag{1}$$

The tracking error ϵ_t is defined as the difference between the simple returns of the tracking portfolio and the index benchmark:

$$\begin{aligned} \epsilon_t &= R_{p,t} - R_{y,t}, \\ &= \sum_{j=1}^{j=N} w_j R_{j,t} + \left(1 - \sum_{j=1}^{j=N} w_j\right) R_{f,t} - R_{y,t}, \\ &= \sum_{j=1}^{j=N} w_j \left(R_{j,t} - R_{f,t}\right) - \left(R_{y,t} - R_{f,t}\right), \\ &= \sum_{j=1}^{j=N} w_j r_{j,t} - r_{y,t}. \end{aligned} \tag{2}$$

where $r_{y,t}$ and $r_{j,t}$ are the simple excess returns of the benchmark index and the basis securities, respectively. Denoting by w the vector of tracking portfolio weights and using matrix notation we can express the tracking error more compactly as

$$\epsilon_t = w' r_t - r_{y,t}. \tag{3}$$

Most of the analysis in the literature revolves around optimizing an objective function of the sequence of tracking errors over the weights of the tracking portfolio.

2.1. Multiple Basis Securities and Fixed Portfolio Weights

2.1.1. General Mean-Variance Specification of Asset Returns

Problem 1. *Choose*

$$w^\star = argmin \;\; var\left(w'\tilde{r} - \tilde{r}_y\right). \tag{4}$$

Let the second moments of the basis and index assets be denoted by $\Sigma_{rr} = \mathrm{cov}(\tilde{r})$, $\sigma_{ry} = \mathrm{cov}(\tilde{r}, \tilde{r}_y)$, and $\sigma_y^2 = \mathrm{var}(\tilde{r}_y)$. Then we can make the following proposition.

Proposition 1. *The solution to (4) is given by:*

$$w^\star = \Sigma_{rr}^{-1} \sigma_{ry}. \tag{5}$$

Proof. See Appendix A. □

The intuition behind the tracking problem and its solution is straightforward. Note that $w^\star$ is the minimum variance hedging portfolio for the index return using the basis asset returns as instruments (Merton 1973; Ingersoll 1987). This portfolio has the highest feasible correlation with the instrument that we need to hedge. Moreover, as the following corollary illustrates, the optimal replicating portfolio weights obtain, in a multivariate linear regression of the index, excess return on the set of basis asset excess returns with an intercept.

Corollary 1. *The solution in (5) is given by* $\hat{\theta}_1$ *in the following multivariate linear regression:*

$$r_{y,t} = \theta_0 + r_t\theta_1 + u_t. \tag{6}$$

Proof. See Appendix A. □

Alternatively, (4) represents the solution to a projection of the returns of the basis assets on the returns of the target benchmark. This is essentially a spanning problem which is related to the mean-variance spanning literature (Huberman and Kandel 1987). The only difference is that in this problem we are asking the reverse question of what the best way is to span a single return series with a set of multiple basis return series. Namely, we are looking for the best way to span a single asset return (an index) with a set of multiple basis assets.

Corollary 2. *The solution in (5) is fully invested in the replicating securities and the replicating portfolio spans the index in a mean-variance sense, if and only if:*

$$\begin{aligned} \hat{\theta}_0 &= 0, \\ \hat{\theta}_1' 1_N &= 1. \end{aligned}$$

Proof. See Appendix A. □

The intuition behind this result is that if an investor's capital is fully (100%) invested in the basis assets and $\hat{\theta}_0 = 0$, then the mean-variance frontier remains unchanged after the addition of the benchmark asset to the set of basis assets. Furthermore, if exact spanning fails, we are able to determine whether the replicating portfolio outperforms the benchmark index, as the following corollary shows.

Corollary 3. *The replicating portfolio in (5) outperforms the index in-sample provided that* $\hat{\theta}_0 < 0$.

Proof. See Appendix A. □

If the intercept in the regression above is negative, then the excess return of the replicating portfolio exceeds the excess return of the benchmark. Conversely, if the intercept happens to be positive, we know that replicating portfolio's return lags the benchmark index return.

The above results suggest an algorithm for finding the best set of spanning basis assets. One can perform a stepwise linear regression of the excess return of the benchmark index on a set of candidate basis asset excess returns until a predetermined level for the tracking error variance is reached. This result provides validation for this common practice that is widely used by institutional investors who are managing index funds.

Decompose the variance-covariance matrix of basis asset returns and the covariance vector between the basis asset returns and the index return as follows:

$$\Sigma_{rr} = \Sigma_r \Phi_r \Sigma_r, \tag{7}$$

$$\sigma_{yr} = \Sigma_r \phi_{yr} \sigma_y, \tag{8}$$

where $\Sigma_r = \text{diag}[\sigma_i]$ is a diagonal matrix with the basis assets total return standard deviations along the diagonal, $\Phi_r = [\rho_{ij}]$ is the correlation matrix of the basis assets and $\phi_{yr} = [\rho_{yi}]$ is the vector of correlations between the basis asset returns and the index return. A few linear algebraic manipulations yield the optimal tracking portfolio weights as follows

$$w^\star = \Sigma_r^{-1}(\Phi_r^{-1}\phi_{yr})\sigma_y. \tag{9}$$

Alternatively, we can conveniently express each individual optimal portfolio weight as

$$w_i^\star = \left(\frac{\sigma_y}{\sigma_i}\right)\kappa_i \tag{10}$$

where $\kappa = [\kappa_i] = \Phi_r^{-1}\phi_{yr}$ is a vector containing the multivariate correlations between each security and the index. Note that the special case of two basis assets considered in the next subsection has the following values for κ:

$$\kappa_1 = \left(\frac{\rho_{1y} - \rho_{2y}\rho_{12}}{1-\rho_{12}^2}\right), \tag{11}$$

$$\kappa_2 = \left(\frac{\rho_{2y} - \rho_{1y}\rho_{12}}{1-\rho_{12}^2}\right). \tag{12}$$

The variance of the tracking error under the optimal portfolio weights, $v(w^\star)$, is given by

$$v(w^\star) = \sigma_y^2\left(1 - \phi_{yr}'\Phi_r^{-1}\phi_{yr}\right). \tag{13}$$

Basis securities that correlate more highly with the benchmark are more useful in reducing the variance of the tracking error. On the contrary, basis securities the are highly correlated with each other tend to increase the variance of the tracking error. Ideally, replicating the benchmark with low levels of the tracking error variance will require securities that are more highly correlated with the index and less so with each other. To see this more clearly, suppose we replace every correlation coefficient in Φ_r with the average correlation between any two basis securities $\bar{\rho}_{ij}$ and every correlation in ϕ_{yr} with the average correlation between the index and any basis security $\bar{\rho}_{yr}$. The correlation structure of the model then becomes:

$$\Phi_r = (1-\bar{\rho}_{ij})I_N + \bar{\rho}_{ij}1_N1_N', \tag{14}$$

$$\phi_{yr} = \bar{\rho}_{iy}1_N, \tag{15}$$

and the optimal tracking error variance becomes:

$$v(w^\star) = \sigma_y^2\left(\frac{(N-1)\bar{\rho}_{ij} + 1 - \bar{\rho}_{iy}^2 N}{(N-1)\bar{\rho}_{ij}+1}\right). \tag{16}$$

This quantity is clearly increasing in $\bar{\rho}_{ij}$ and decreasing in $\bar{\rho}_{iy}$.

2.1.2. Market Model Specification of Returns

For simplicity we consider the case where the asset and index returns are driven by the market model of Sharpe (1963). This may be useful in cases where the number of basis assets is very large. Consider the following standard return variance decomposition

$$\Sigma_{rr} = \beta_r\beta_r'\sigma_m^2 + D, \tag{17}$$

$$\sigma_{ry} = \beta_r\beta_y\sigma_m^2, \tag{18}$$

$$\sigma_y^2 = \beta_y^2\sigma_m^2 + \sigma_{\epsilon_y}^2, \tag{19}$$

where β_r is the vector of market betas of the basis assets, β_y is the market beta of the index, σ_m^2 is the market return variance and D is a diagonal matrix consisting of the idiosyncratic return variances along the diagonal.

Problem 2. *Choose*

$$w^{\star} = argmin \;\; var\left(w'\tilde{r} - \tilde{r}_y\right) \tag{20}$$

where the index and basis asset returns are driven by (17)–(19).

Given the specific structures of the variance-covariance matrix of asset returns and the covariance of asset returns with the benchmark index, the following results can be obtained.

Proposition 2. *The optimal index-tracking portfolio weights in this case are given by*

$$w^{\star} = \frac{D^{-1}\beta_r\beta_y}{\frac{1}{\sigma_m^2} + (\beta_r' D^{-1}\beta_r)}. \tag{21}$$

The variance of the tracking error under the optimal portfolio strategy, $v(w^{\star})$, is as follows:

$$v(w^{\star}) = \sigma_{\epsilon_y}^2 + \frac{\beta_y^2}{\frac{1}{\sigma_m^2} + (\beta_r' D^{-1}\beta_r)}. \tag{22}$$

Proof. See Appendix A. □

Note that in this case, there is no need to solve for all replicating portfolio weights jointly as the solution to Problem 1 in Proposition 1. When we impose the market model of returns we can obtain that optimal weights are:

$$w_i^{\star} = \left(\frac{\beta_i}{\sigma_{\epsilon_i}^2}\right) \frac{\beta_y}{\left[\left(\frac{1}{\sigma_m^2}\right) + \Sigma_j\left(\frac{\beta_j^2}{\sigma_{\epsilon_j}^2}\right)\right]}. \tag{23}$$

Basis assets with high betas are expected to have high expected returns. Similarly, basis assets with low idiosyncratic volatility are more valuable in terms of reducing the variance of the tracking error of the replicating portfolio. Note that each security's weight is proportional to a "signal-to-noise" ratio of market beta to idiosyncratic variance, while the coefficient of proportionality (the second fraction in the above equation) is common to all securities. Moreover, as a result of the fact that the market model implies an approximation in the variance-covariance matrix decomposition ((17) and (18)), we have the following corollary.

Corollary 4. *The beta of the replicating portfolio is always less than the beta of the index benchmark.*

Proof. See Appendix A. □

Note that the shortfall between β_p and β_y is greatest when σ_m^2, the market return volatility, is low. Conversely, in times of high market return volatility, higher σ_m^2 leads to a replicating portfolio beta that is much closer to the benchmark index beta. Finally, we can demonstrate that the optimal replicating portfolio weights are proportional to the tangent portfolio weights that result from the set of basis assets.

Corollary 5. *The optimal replicating portfolio weights in (21) are proportional to the tangent portfolio weights where the coefficient of proportionality is increasing in β_y and decreasing in the factor's reward-to-risk ratio.*

Proof. See Appendix A. □

The expression for the optimal tracking error clearly suggests an algorithm for picking the stocks that enter the tracking portfolio: rank all candidate securities by their ratios of standardized systematic risk (i.e., market betas β_i) to standard deviation of idiosyncratic return risk (σ_{ϵ_i}) and pick the top ones. This will result in the highest possible value for $(\beta_r' D^{-1}\beta_r)$ in the denominator of (22).

A more general comment is in order. The variance decomposition is valid for any mean-variance return specification. The content of the market model is in the diagonal structure of D. However, even if D is not diagonal (and we are no longer in the market model world) the above formulae are still valid, even if the general algorithm needs to be modified to maximize the quantity $(\beta_r' D^{-1} \beta_r)$. Furthermore, these results are not restricted to single-factor models. Multiple-factor models, along the lines of the APT, can be similarly applied to this setting.

Consider the following K-factor model driving the second moments of the excess returns of the index and the basis assets:

$$\Sigma_{rr} = B_r V_f B_r' + D, \tag{24}$$

$$\sigma_{ry} = B_r V_f b_y, \tag{25}$$

$$\sigma_y^2 = b_y' V_f b_y + \sigma_{\epsilon_y}^2. \tag{26}$$

where V_f is a $K \times K$ variance-covariance matrix of factor returns, B_r is an $N \times K$ matrix of the basis assets' factor loadings and b_y is a K vector of the factor loadings of the index.

Problem 3. *Choose*

$$w^\star = argmin \;\; var\left(w'\tilde{r} - \tilde{r}_y\right) \tag{27}$$

where the index and basis asset returns are driven by (24)–(26).

The solution to this problem is stated in the following proposition.

Proposition 3. *The optimal tracking portfolio weights for this model are*

$$w^\star = D^{-1} B_r \left[V_f^{-1} + \left(B_r' D^{-1} B_r \right) \right]^{-1} b_y. \tag{28}$$

The variance of the tracking error under this portfolio strategy is given by:

$$v(w^\star) = b_y' \left[V_f^{-1} + \left(B_r' D^{-1} B_r \right) \right]^{-1} b_y + \sigma_{\epsilon_y}^2. \tag{29}$$

Proof. See Appendix A. □

This is the multi-factor generalization to the tracking error variance under the optimal tracking portfolio strategy. It is intuitively similar in spirit to the expression for the variance of the tracking error in the single-factor case in (22) above. Assuming that the index factor exposures are positive, securities with high ratios of factor betas to idiosyncratic excess return variance will command higher weights and be more useful in mimicking the excess returns of the index.

Coming up with an analytical way of identifying a tracking portfolio that minimizes $v(w^\star)$ is not easy, especially in light of the fact that the index may have factor loadings b_y of differing signs. Therefore, a prudent way to identify the best tracking portfolio might be given by the following.

2.2. The Case of Two Basis Assets

Consider the problem of replicating a portfolio with return $\tilde{r}_{y,t}$ using a linear combination of two stocks with returns $\tilde{r}_{1,t}$ and $\tilde{r}_{2,t}$ with portfolio weights w_1 and w_2. Denote the second moments of the excess returns as σ_1^2, σ_2^2 and σ_y^2; and the excess return correlations as ρ_{12}, ρ_{1y} and ρ_{2y}. The optimal (unconstrained) y-tracking portfolio weights can be shown to be equal to:

$$w_1^\star = \left(\frac{\sigma_y}{\sigma_1}\right)\left(\frac{\rho_{1y}-\rho_{2y}\rho_{12}}{1-\rho_{12}^2}\right), \tag{30}$$

$$w_2^\star = \left(\frac{\sigma_y}{\sigma_2}\right)\left(\frac{\rho_{2y}-\rho_{1y}\rho_{12}}{1-\rho_{12}^2}\right). \tag{31}$$

Both $w_1^\star$ and $w_2^\star$ above are the OLS regression estimates that would be obtained in a multivariate regression of the excess return of the benchmark index on the excess returns on both securities. We can also interpret the weights as hedge ratios. If we are trying to hedge the values of the benchmark index, we need hold $w_1^\star$ of the first security and $w_2^\star$ of the second. Note that both portfolio weights are given by the ratio of the standard deviation of the index return to the standard deviation of each security's return, multiplied by the multivariate correlation between the excess return of the index and the excess return of the security.

The variance of the tracking error under the optimal portfolio weights above is given as follows:

$$v(w_1^\star, w_2^\star) = \sigma_y^2\left(\frac{1-\rho_{1y}^2-\rho_{2y}^2+2\rho_{1y}\rho_{2y}\rho_{12}}{1-\rho_{12}^2}\right). \tag{32}$$

Consider a hypothetical scenario with two basis securities (not necessarily components of the benchmark index) that have $\rho_{12} = 0$ and $\rho_{1y} = \rho_{2y} = 0.7$. In this case, $v(w_1^\star, w_2^\star) = 0.02\sigma_y^2$ and we can almost perfectly replicate the index by holding the portfolio $w_1^\star = \rho_{1y}\sigma_y/\sigma_1$ and $w_2^\star = \rho_{2y}\sigma_y/\sigma_2$ with any excess funds available to be invested in the risk-free asset. We cannot drive the tracking error variance all the way down to zero though, since the correlation matrix needs to be positive definite.

To obtain further intuition from the OLS formula for the optimal tracking portfolio weights, let us consider the case in which the benchmark index and the two basis asset returns are driven by the market model:

$$\tilde{r}_y = \beta_y\tilde{r}_m + \tilde{\epsilon}_y, \tag{33}$$

$$\tilde{r}_1 = \beta_1\tilde{r}_m + \tilde{\epsilon}_1, \tag{34}$$

$$\tilde{r}_2 = \beta_2\tilde{r}_m + \tilde{\epsilon}_2. \tag{35}$$

The optimal index-tracking portfolio weights for the market-model-driven returns are

$$w_1^\star = \left(\frac{\beta_1}{\sigma_{\epsilon_1}^2}\right)\frac{\beta_y}{\left[\left(\frac{1}{\sigma_m^2}\right)+\left(\frac{\beta_1^2}{\sigma_{\epsilon_1}^2}+\frac{\beta_2^2}{\sigma_{\epsilon_2}^2}\right)\right]}, \tag{36}$$

$$w_2^\star = \left(\frac{\beta_2}{\sigma_{\epsilon_2}^2}\right)\frac{\beta_y}{\left[\left(\frac{1}{\sigma_m^2}\right)+\left(\frac{\beta_1^2}{\sigma_{\epsilon_1}^2}+\frac{\beta_2^2}{\sigma_{\epsilon_2}^2}\right)\right]}. \tag{37}$$

The market model shows that the optimal replicating portfolio essentially represents a scaled position in the (un-normalized) tangent portfolio with a scaling factor given by the beta of the benchmark index. The optimal tracking error variance simplifies to:

$$v(w_1^\star, w_2^\star) = \sigma_{\epsilon_y}^2 + \frac{\beta_y^2}{\left(\frac{1}{\sigma_m^2}\right)+\left(\frac{\beta_1^2}{\sigma_{\epsilon_1}^2}+\frac{\beta_2^2}{\sigma_{\epsilon_2}^2}\right)}, \tag{38}$$

where the motivation for Algorithm 2 is made clear by the appearance of the squared signal-to-noise ratios in the denominator.

2.3. Index Beating Strategies

If we do not put any structure on the first two moments of the security and index returns, then we are left with trying to select the best possible set of replicating securities which also maximizes the following quantity:

$$\Delta = \mu_r' \Sigma_{rr}^{-1} \sigma_{ry} - \mu_y. \tag{39}$$

Finding the best set of replicating securities that maximizes Δ is a feasible combinatorial problem. However, given the difficulties associated with predicting μ_r, a viable alternative would be to use a factor model for the expected return vector. One such model for the expected excess returns of the basis securities and the index could be the following:

$$\mu_r = \alpha_r + \beta_r \mu_m, \tag{40}$$
$$\mu_y = \alpha_y + \beta_y \mu_m. \tag{41}$$

Under this specific model for the first moments, the outperformance of the replicating portfolio relative to the index is given by

$$\Delta = \beta_y \left[\frac{(\alpha_r' D^{-1} \beta_r) - \left(\frac{\mu_m}{\sigma_m^2}\right)}{\left(\frac{1}{\sigma_m^2}\right) + (\beta_r' D^{-1} \beta_r)} \right] - \alpha_y, \tag{42}$$

assuming that we use the optimal portfolio weights from the previous subsection.

In the following discussion, I will assume, without loss of generality, that the index is a broadly diversified equity portfolio with $\alpha_y = 0$ and $\beta_y = 1$. This simplifies the value of the index outperformance Δ as follows:

$$\Delta = \frac{(\alpha_r' D^{-1} \beta_r) - \left(\frac{\mu_m}{\sigma_m^2}\right)}{\left(\frac{1}{\sigma_m^2}\right) + (\beta_r' D^{-1} \beta_r)}. \tag{43}$$

The optimal replicating strategy is already designed to deliver the smallest possible tracking error variance for a given set of basis securities. If, in addition, we want to maximize Δ, then we have to pick the set of replicating securities so that the quantity $(\alpha_r' D^{-1} \beta_r)$ is as large as possible. This suggests a straightforward way of ranking each candidate security i by the value of $\alpha_i \beta_i / \sigma_{\epsilon_i}^2$. The strategy that maximizes Δ is the one that keeps adding securities with positive values of $\alpha_i \beta_i / \sigma_{\epsilon_i}^2$ to the tracking portfolio until

$$\alpha_r' D^{-1} \beta_r > \frac{\mu_m}{\sigma_m^2}. \tag{44}$$

The denominator in (43) is strictly positive and we are guaranteed that $\Delta > 0$ in expectation. The analysis suggests the following algorithm for selecting securities that, combined in a portfolio, will outperform the benchmark in expectation. First, estimate the market model for all candidate securities for the replicating portfolio. Next, pick the security with the highest value of Δ. Keep adding securities to the replicating portfolio while $\Delta > 0$. Finally, terminate this process when the desired number of securities in the portfolio is reached.

A similar approach can be taken within a multi-factor model for the security and index returns. Suppose that the second moments are again driven by (24)–(26) while the expected excess returns are given by:

$$\mu_r = \alpha_r + B_r \mu_F, \tag{45}$$
$$\mu_y = \alpha_y + b_y' \mu_F, \tag{46}$$

where μ_F is a vector of the expected factor returns. The expected return differential between the optimal tracking portfolio and the index is

$$\Delta_{MF} = \left[(\alpha_r' D^{-1} B) - \mu_F' V_F^{-1}\right] \cdot \left[V_F^{-1} + (B_r' D^{-1} B_r)\right]^{-1} b_y - \alpha_y \tag{47}$$

which is a natural extension of the Δ obtained under the single-factor model in (43). One possible way to sequentially pick candidate securities for a portfolio that minimizes the tracking error variance while attempting the beat the index would be the following. First, estimate the respective multi-factor model for all candidate securities for the replicating portfolio. Next, pick the security with the highest value of Δ_{MF}. Keep adding securities to the replicating portfolio while $\Delta_{MF} > 0$. Finally, terminate this process when the desired number of securities in the portfolio is reached.

2.4. Further Results

In this section, I consider several extensions that are of further interest to the theorist and the practitioner. I present the theoretical results and discuss the intuition behind them, along with their implications. First, I develop the optimal replicating portfolio that is fully invested in the risky assets without any position on the risk-free asset. Second, I present the results for the optimal portfolio in the presence of a constraint on the total risk (i.e., return standard deviation) of the replicating portfolio. Third, in the context of factor return models I derive the optimal portfolio weights when there is a constraint on the factor loadings of the replicating portfolio. Finally, I make some suggestions for solving this problem in the presence of multiple linear and/or quadratic constraints on the replicating portfolio weights.

2.4.1. Fully Invested Replicating Portfolio Weights

Let us go back to the simple replicating portfolio return, except now we will require that the portfolio is fully invested in the risky benchmark assets:

$$R_{p,t} = \sum_{j=1}^{j=N} w_j R_{j,t}. \tag{48}$$

The tracking error ϵ_t is defined as the difference between the simple returns of the tracking portfolio and the index benchmark:

$$\begin{aligned} \epsilon_t &= R_{p,t} - R_{y,t}, \\ &= \sum_{j=1}^{j=N} w_j R_{j,t} - R_{y,t}, \end{aligned} \tag{49}$$

where $R_{y,t}$ and $R_{j,t}$ are the simple total returns of the benchmark index and the basis securities, respectively. Denoting by w the vector of tracking portfolio weights and using matrix notation, we can express the tracking error more compactly as

$$\epsilon_t = w' R_t - R_{y,t}. \tag{50}$$

Problem 4. *Choose*

$$\begin{aligned} w^\star &= argmin \;\; var\left(w'\tilde{R} - \tilde{R}_y\right) \\ s.t. &\quad w' 1_N = 1. \end{aligned} \tag{51}$$

Proposition 4. *The solution to (51) is given by:*

$$w^{\star} = \Sigma_{RR}^{-1}\sigma_{Ry} + \left(1 - 1_N'\Sigma_{RR}^{-1}\sigma_{Ry}\right)\left(\frac{\Sigma_{RR}^{-1}1_N}{1_N'\Sigma_{RR}^{-1}1_N}\right). \tag{52}$$

Proof. See Appendix A. □

The intuition behind the result in Proposition 4 is quite straightforward. The optimal replicating portfolio weights follow the same strategy as in the unconstrained case (with Σ_{RR} and σ_{Ry} in place of Σ_{rr} and σ_{ry}, respectively), and any remaining funds are invested in the minimum variance portfolio generated by the returns of the benchmark assets.

2.4.2. Replicating Portfolio Weights with Constraints on Total Risk

Another useful alternative to consider here is to find the best replicating portfolio under a constraint on the total risk σ_0^2 of the portfolio, which may or may not equal the total risk of the underlying benchmark σ_y^2.

Problem 5. *Choose*

$$\begin{aligned} w^{\star} \quad &= \quad \mathit{argmin} \;\; \mathit{var}\left(w'\tilde{r} - \tilde{r}_y\right) \\ \mathit{s.t.} \quad &\quad\;\; w'\Sigma_{rr}w = \sigma_0^2. \end{aligned} \tag{53}$$

Proposition 5. *The solution to (53) is given by:*

$$w^{\star} = \left(\frac{\sigma_0}{\sqrt{\sigma_{ry}'\Sigma_{rr}^{-1}\sigma_{ry}}}\right)\Sigma_{rr}^{-1}\sigma_{ry}. \tag{54}$$

Proof. See Appendix A. □

The intuition behind the result in Proposition 5 is easy to follow. The optimal replicating portfolio weights follow the same strategy as in the unconstrained case presented in Proposition 1 with a scaling factor that brings the return variance of the replicating portfolio down or up to σ_0^2. Note that $\sigma_{ry}'\Sigma_{rr}^{-1}\sigma_{ry}$ is the total risk of the unconstrained maximum correlation portfolio. Hence, if the target risk control exceeds that value, we need to invest more than 100% of our funds in the unconstrained portfolio in order to meet the risk target, and vice versa; if the risk target is below the total risk of the unconstrained portfolio, we need to invest less than 100% in the unconstrained portfolio if we are to meet the risk target.

2.4.3. Replicating Portfolio Weights with a Linear and a Quadratic Constraint

Consider combining the two constraints in the previous subsections. Specifically, investors may require that the replicating portfolio is fully invested in the risky assets *and* that there is a limit of the replicating portfolio's return standard deviation. Such joint constraints have been implemented empirically in the context of cloning hedge fund returns by Hasanhodzic and Lo (2007). Next, I state the problem more formally and offer an analytical solution:

Problem 6. *Choose*

$$
\begin{aligned}
w^\star &= \text{argmin } \text{var}\left(w'\tilde{R} - \tilde{R}_y\right) \\
\text{s.t.} \quad & w'1_N = 1, \\
& w'\Sigma_{RR} w = \sigma_0^2.
\end{aligned} \tag{55}
$$

Proposition 6. *The solution to (55) is given by:*

$$
w^\star = \sqrt{\frac{\sigma_0^2 - \sigma_{mv}^2}{Q}} \Sigma_{RR}^{-1}\sigma_{Ry} + \left[1 - \sqrt{\frac{\sigma_0^2 - \sigma_{mv}^2}{Q}} (1_N' \Sigma_{RR}^{-1}\sigma_{Ry})\right] \left(\frac{\Sigma_{RR}^{-1} 1_N}{1_N' \Sigma_{RR}^{-1} 1_N}\right), \tag{56}
$$

where

$$
\begin{aligned}
Q &= (\sigma_{Ry}' \Sigma_{RR}^{-1} \sigma_{Ry}) - \frac{(1_N' \Sigma_{RR}^{-1}\sigma_{Ry})}{(1_N' \Sigma_{RR}^{-1} 1_N)}, \\
\sigma_{mv}^2 &= \frac{1}{1_N' \Sigma_{RR}^{-1} 1_N}.
\end{aligned}
$$

Proof. See Appendix A. □

The intuition behind the result in Proposition 6 is as follows. The optimal replicating portfolio is still split between the minimum variance portfolio and the maximum correlation portfolio from previous sections. Furthermore, the investment in the unconstrained portfolio is scaled as necessary to meet the risk target as in the previous proposition, while any remaining funds are invested in the global minimum variance portfolio.

2.4.4. Replicating Portfolio Weights with Constraints on Factor Loadings

In the context of factor models for the first two moments of asset returns, investors may insist that the replicating portfolio has certain exposures to the factors. This presents a natural extension to the single-factor results presented previously, leading to the following:

Problem 7. *Choose*

$$
\begin{aligned}
w^\star &= \text{argmin } \text{var}\left(w'\tilde{r} - \tilde{r}_y\right) \\
\text{s.t.} \quad & w'\beta_r = \beta_0, \\
& \sigma_{ry} = \beta_r \beta_y \sigma_m^2, \\
& \Sigma_{rr} = \beta_r \beta_r' \sigma_m^2 + D.
\end{aligned} \tag{57}
$$

Proposition 7. *The solution to (57) is given by:*

$$
w^\star = \frac{D^{-1}\beta_r}{\beta_r' D^{-1} \beta_r} \beta_0. \tag{58}
$$

Proof. See Appendix A. □

Note the similarity between the above portfolio rule and the result in Proposition 2 before. The result in Proposition in 7A is a cleaner version of the previous portfolio rule which scales the beta of the portfolio to the target β_0. The inverse of the market's total risk, σ_m^2, disappears from the denominator for technical reasons (please see proof in the Appendix A), and we are left with a simple beta scaling factor which meets the required factor loading. Note that β_0 need not necessarily be equal to β_y, the benchmark's factor exposure to the market factor.

A natural extension to multi-factor models of asset returns leads to the following:

Problem 8. *Choose*

$$\begin{aligned} w^\star &= \quad argmin \;\; var\left(w'\tilde{r} - \tilde{r}_y\right) \\ s.t. & \quad B_r' w = b_0, \\ & \quad \sigma_{ry} = B_r V_f b_y, \\ & \quad \Sigma_{rr} = B_r V_f B_r' + D. \end{aligned} \tag{59}$$

Proposition 8. *The solution to (59) is given by:*

$$w^\star = D^{-1} B_r (B_r' D^{-1} B_r)^{-1} b_0. \tag{60}$$

Proof. See Appendix A. □

Again, this is quite similar and simpler than the result in Proposition 3 before with the inverse of the factor variance-covariance matrix conveniently disappearing from the expression. The rule involves a scaling factor combination of the target factor loadings in the vector b_0 with the interpretation of factor signal-to-noise ratios contained in the $N \times K$ matrix $D^{-1}B_r$.

3. An Empirical Example: Replicating DJIA

In this section, I apply the theoretical results to the Dow Jones Industrial Average (DJIA). I use a two-year window of daily stock return data. The test time period uses 2018 data for the estimation and 2019 data for the out-of-sample performance. During the estimation period, I use daily stock returns for various subsets of the index components in order to estimate all the parameters that are needed to compute the optimal portfolio weights of the benchmark replicating portfolios. Throughout the rest of this section I construct buy-and-hold replicating portfolios. In the test period, I assume that the estimates are unbiased predictors of the true parameter values and track the returns of the benchmark portfolio relative to the index.

The benchmark index return data consist of total return series for DJIA. The data were obtained directly from Dow Jones indexes. The individual stock return data were obtained from the Center for Research in Securities Prices. Historical daily factor returns and risk-free rates were obtained from Ken French's Data Library online.

Empirical results from the replications of DJIA usingAlgorithm 1 are reported in Table 1. Straight replication-only Algorithms 1 through 3 are the same as in the previous section. Algorithm 3A uses the three-factor Fama and French (1992) model while Algorithm 3B uses the four-factor Carhart (1997) model. Algorithm 4 uses the single-factor market model, Algorithm 5A uses the three-factor Fama-French model and Algorithm 5B uses the four-factor Carhart model.

Table 1. Optimal tracking portfolios for DJIA index: Algorithm 1.

Number of Stocks	In-Sample ρ_{py}	$\sqrt{v^\star}$	RAR	α_p	β_p	Out-of-Sample ρ_{py}	$\sqrt{v^\star}$	RAR	α_p	β_p
1	0.8241	0.0086	−0.1273	−0.0006	1.1439	0.5929	0.0129	−0.3359	−0.0015	1.2640
2	0.8983	0.0056	−0.0073	−0.0001	1.0545	0.6499	0.0096	−0.1864	−0.0008	1.1033
3	0.9276	0.0044	0.0000	−0.0001	1.0144	0.7483	0.0068	−0.0824	−0.0004	1.0364
4	0.9457	0.0037	−0.0164	−0.0001	0.9959	0.8248	0.0054	−0.0051	−0.0001	1.0631
5	0.9672	0.0031	0.0166	−0.0000	1.0733	0.8840	0.0043	−0.0589	−0.0003	1.0823
6	0.9749	0.0026	0.0259	0.0000	1.0377	0.9012	0.0038	−0.0391	−0.0002	1.0487
7	0.9779	0.0023	0.0354	0.0001	0.9964	0.9143	0.0033	−0.0365	−0.0002	1.0018
8	0.9852	0.0020	0.0509	0.0001	1.0380	0.9490	0.0025	0.0330	0.0000	1.0269
9	0.9870	0.0017	0.0628	0.0002	1.0092	0.9543	0.0023	0.0463	0.0001	0.9975
10	0.9892	0.0016	0.0639	0.0002	1.0031	0.9648	0.0020	0.0243	0.0000	0.9979
11	0.9912	0.0014	0.0558	0.0001	1.0061	0.9745	0.0017	0.0505	0.0001	1.0202
12	0.9925	0.0013	0.0525	0.0001	1.0025	0.9781	0.0017	0.0551	0.0001	1.0338
13	0.9934	0.0012	0.0500	0.0001	0.9972	0.9804	0.0015	0.0576	0.0001	1.0193
14	0.9944	0.0011	0.0394	0.0001	0.9993	0.9833	0.0014	0.0519	0.0001	1.0257
15	0.9954	0.0010	0.0377	0.0001	1.0068	0.9857	0.0013	0.0658	0.0001	1.0271
16	0.9962	0.0009	0.0292	0.0000	1.0086	0.9889	0.0012	0.0571	0.0001	1.0292
17	0.9968	0.0008	0.0323	0.0001	1.0073	0.9896	0.0012	0.0574	0.0001	1.0358
18	0.9973	0.0008	0.0332	0.0001	1.0035	0.9900	0.0011	0.0591	0.0001	1.0306
19	0.9976	0.0007	0.0360	0.0001	1.0007	0.9918	0.0010	0.0595	0.0001	1.0228
20	0.9981	0.0006	0.0341	0.0001	0.9985	0.9925	0.0009	0.0614	0.0001	1.0178
21	0.9984	0.0006	0.0309	0.0000	0.9978	0.9953	0.0007	0.0590	0.0001	1.0078
22	0.9986	0.0005	0.0408	0.0001	0.9928	0.9960	0.0007	0.0565	0.0001	1.0010
23	0.9988	0.0005	0.0423	0.0001	0.9942	0.9968	0.0006	0.0551	0.0001	1.0040
24	0.9990	0.0004	0.0431	0.0001	0.9963	0.9973	0.0006	0.0533	0.0001	1.0063
25	0.9991	0.0004	0.0425	0.0001	0.9917	0.9975	0.0005	0.0511	0.0001	1.0049
26	0.9992	0.0004	0.0411	0.0001	0.9923	0.9980	0.0005	0.0464	0.0001	1.0075
27	0.9992	0.0004	0.0396	0.0001	0.9889	0.9980	0.0005	0.0492	0.0001	1.0042
28	0.9992	0.0004	0.0397	0.0001	0.9896	0.9984	0.0004	0.0450	0.0001	1.0071
29	0.9993	0.0004	0.0388	0.0001	0.9886	0.9986	0.0004	0.0414	0.0001	1.0061
30	0.9997	0.0004	0.0381	0.0001	0.9893	0.9989	0.0004	0.0353	0.0001	1.0031

This table reports the in-sample and out-of-sample correlation of the tracking portfolio return with the index return ρ_{py}, daily tracking error $\sqrt{v^\star}$ and cumulative out-of-sample return differential between the tracking portfolio and the index RAR, and the intercept α_p and the slope β_p of a simple linear regression of the portfolio excess return on the benchmark index excess return. The estimation period covers 2 January 2018 until 31 December 2018. The out-of-sample test period starts 2 January 2019 and ends on 31 December 2019. RAR is the cumulative realized active return given by $\Pi_{t=1}^{t=T}(1+r_{p,t}) - \Pi_{t=1}^{t=T}(1+r_{y,t})$.

Table 1 reports the results for the replicating strategies with 2011 as the out-of-sample test period. In this instance, the stepwise regression model (Algorithm 1) does a better job than most of the other algorithms proposed in the paper. One possible reason for this superiority in this case could be the fact that all the replicating portfolios based on Algorithm 1 have the highest market betas compared with all the other replicating portfolios based on the remaining four algorithms.

Broadly speaking, the empirical results indicate that replication strategies that use factor models to track a benchmark perform just as well and sometimes better out-of-sample than the standard stepwise regression model (Algorithm 1). One reason behind this could potentially be the difficulty in estimating a larger set of stock return covariances with the same amount of historical data. The annualized standard deviation of the replication error ranges between 0.64% and 1.60% when more than half of the index components have been included in the replicating portfolio. The correlation between the replicating portfolio and the benchmark quickly increases and reaches almost 1 as we increase the number of securities in the replicating portfolio. Similarly, the portfolio beta with the benchmark approaches 1 while the abnormal return hovers around zero, though does remain positive at about 1 basis point per day. This performance is consistent both in-sample and out-of-sample.

Next, I consider using a single-factor model as the model of securities' expected returns, variances and covariances. The empirical performance of Algorithm 2 is presented in Table 2. Comparing the performance of Algorithm 2 to Algorithm 1, we note a few minor differences. First, the tracking error is a little higher with Algorithm 2, though less so out-of-sample than in-sample. The relative outperformance is greater now while the abnormal return virtually disappears especially out-of-sample. The replicating portfolio beta stabilizes eventually, though at a value slightly above 1 once all 30 stocks are included in the portfolio. Overall, Algorithm 2 performed well to the task that it was designed to perform, namely, to get as close as possible to perfect correlation with the benchmark while using a single-factor model of stock returns' first and second moments.

Table 2. Optimal tracking portfolios for DJIA index: Algorithm 2.

Number of Stocks	In-Sample ρ_{py}	$\sqrt{v^\star}$	RAR	α_p	β_p	Out-of-Sample ρ_{py}	$\sqrt{v^\star}$	RAR	α_p	β_p
1	0.8239	0.0061	−0.1227	−0.0005	1.1427	0.5924	0.0129	−0.3359	−0.0016	1.2625
2	0.9063	0.0048	0.0032	−0.0001	1.2329	0.8169	0.0066	0.0558	0.0000	1.2272
3	0.9445	0.0041	0.0498	0.0001	1.2877	0.8592	0.0056	−0.0457	−0.0003	1.2135
4	0.9467	0.0037	0.0684	0.0002	1.2618	0.8795	0.0053	−0.0716	−0.0005	1.2488
5	0.9422	0.0034	0.0856	0.0002	1.2398	0.9031	0.0044	0.0092	−0.0001	1.1920
6	0.9490	0.0031	0.0599	0.0001	1.2783	0.9263	0.0041	-0.0027	−0.0002	1.2266
7	0.9594	0.0029	0.0374	0.0001	1.2324	0.9435	0.0036	0.0313	−0.0001	1.2182
8	0.9620	0.0028	0.0417	0.0001	1.2036	0.9514	0.0033	0.0458	−0.0000	1.2061
9	0.9636	0.0027	0.0139	−0.0000	1.1994	0.9549	0.0033	0.0506	−0.0000	1.2224
10	0.9682	0.0026	0.0142	−0.0000	1.1725	0.9624	0.0029	0.0628	0.0000	1.1909
11	0.9755	0.0025	0.0258	0.0000	1.1519	0.9706	0.0025	0.0588	0.0000	1.1692
12	0.9785	0.0024	0.0225	0.0000	1.1383	0.9743	0.0024	0.0605	0.0000	1.1767
13	0.9792	0.0024	0.0202	−0.0000	1.1457	0.9761	0.0024	0.0952	0.0001	1.1935
14	0.9826	0.0023	0.0128	−0.0000	1.1209	0.9803	0.0021	0.0834	0.0001	1.1538
15	0.9851	0.0023	0.0141	−0.0000	1.1055	0.9814	0.0020	0.0609	0.0000	1.1451
16	0.9863	0.0022	0.0164	−0.0000	1.0887	0.9818	0.0019	0.0624	0.0001	1.1299
17	0.9856	0.0022	0.0190	−0.0000	1.0947	0.9808	0.0020	0.0656	0.0001	1.1354
18	0.9863	0.0022	0.0214	0.0000	1.0854	0.9811	0.0019	0.0566	0.0000	1.1251
19	0.9873	0.0021	0.0172	−0.0000	1.0812	0.9820	0.0018	0.0551	0.0000	1.1222
20	0.9878	0.0021	0.0231	0.0000	1.0644	0.9826	0.0017	0.0433	0.0000	1.1037
21	0.9883	0.0021	0.0389	0.0001	1.0469	0.9832	0.0016	0.0387	0.0000	1.0812
22	0.9888	0.0021	0.0431	0.0001	1.0455	0.9837	0.0016	0.0402	0.0000	1.0833
23	0.9899	0.0020	0.0426	0.0001	1.0352	0.9848	0.0015	0.0389	0.0000	1.0666
24	0.9902	0.0020	0.0414	0.0001	1.0312	0.9854	0.0015	0.0315	−0.0000	1.0667
25	0.9910	0.0020	0.0415	0.0001	1.0228	0.9863	0.0014	0.0324	0.0000	1.0554
26	0.9920	0.0020	0.0433	0.0001	1.0134	0.9868	0.0013	0.0331	0.0000	1.0450
27	0.9923	0.0020	0.0433	0.0001	1.0012	0.9868	0.0013	0.0367	0.0000	1.0316
28	0.9925	0.0020	0.0461	0.0001	0.9924	0.9871	0.0012	0.0386	0.0000	1.0221
29	0.9925	0.0020	0.0474	0.0001	0.9848	0.9871	0.0012	0.0375	0.0000	1.0152
30	0.9925	0.0020	0.0475	0.0001	0.9851	0.9871	0.0012	0.0377	0.0000	1.0156

This table reports the in-sample and out-of-sample correlation of the tracking portfolio return with the index return ρ_{py}, daily tracking error $\sqrt{v^\star}$ and cumulative out-of-sample return differential between the tracking portfolio and the index RAR, and the intercept α_p and the slope β_p of a simple linear regression of the portfolio excess return on the benchmark index excess return. The estimation period covers 2 January 2018 until 31 December 2018. The out-of-sample test period starts 2 January 2019 and ends on 31 December 2019. RAR is the cumulative realized active return given by $\Pi_{t=1}^{t=T}(1+r_{p,t}) - \Pi_{t=1}^{t=T}(1+r_{y,t})$.

In order to consider exposures to the size and value effect, I turn to the Fama and French three-factor model (Algorithm 3A) as a descriptor of stock returns. Table 3 reports the findings regarding the performance of the strict replication Algorithm 3A. One new feature that appears for this Algorithm is how smaller portfolios tend to disappoint out-of-sample both in terms of having lower correlations with the benchmark but also in terms of delivering lower absolute returns relative to the benchmark return. However, it is worthwhile noting that once most of the component stocks were included in the replicating portfolio, its performance improved considerably. Note also that this performance improvement did not happen at the expense of taking on an inordinate amount of systematic risk.

Table 3. Optimal tracking portfolios for DJIA index: Algorithm 3A.

Number of Stocks	In-Sample ρ_{py}	$\sqrt{v^\star}$	RAR	α_p	β_p	Out-of-Sample ρ_{py}	$\sqrt{v^\star}$	RAR	α_p	β_p
1	0.8241	0.0073	−0.125	−0.0006	1.1436	0.5928	0.0129	−0.3359	−0.0015	1.2635
2	0.9062	0.0048	0.0032	−0.0001	1.2329	0.8169	0.0066	0.0558	0.0000	1.2272
3	0.9446	0.0041	0.0497	0.0001	1.2877	0.8591	0.0056	−0.0461	−0.0003	1.2135
4	0.9468	0.0037	0.0680	0.0002	1.2617	0.8795	0.0053	−0.0724	−0.0005	1.2488
5	0.9424	0.0033	0.0851	0.0002	1.2397	0.9032	0.0044	0.0085	−0.0001	1.1920
6	0.9490	0.0031	0.0598	0.0001	1.2781	0.9262	0.0041	−0.0031	−0.0002	1.2265
7	0.9594	0.0029	0.0369	0.0001	1.2320	0.9434	0.0036	0.0307	−0.0001	1.2183
8	0.9620	0.0028	0.0413	0.0001	1.2035	0.9513	0.0033	0.0452	−0.0000	1.2063
9	0.9636	0.0027	0.0137	−0.0000	1.1992	0.9548	0.0033	0.0507	−0.0000	1.2224
10	0.9682	0.0026	0.0142	−0.0000	1.1722	0.9623	0.0029	0.0630	0.0000	1.1909
11	0.9755	0.0025	0.0256	0.0000	1.1516	0.9706	0.0025	0.0588	0.0000	1.1692
12	0.9786	0.0024	0.0223	0.0000	1.1381	0.9742	0.0024	0.0604	0.0000	1.1767
13	0.9792	0.0024	0.0198	−0.0000	1.1455	0.9761	0.0024	0.0949	0.0001	1.1937
14	0.9826	0.0023	0.0128	−0.0000	1.1206	0.9803	0.0021	0.0838	0.0001	1.1538
15	0.9851	0.0023	0.0140	−0.0000	1.1055	0.9814	0.0020	0.0612	0.0000	1.1452
16	0.9845	0.0022	0.0165	−0.0000	1.1112	0.9799	0.0021	0.0642	0.0001	1.1505
17	0.9856	0.0022	0.0187	−0.0000	1.0949	0.9808	0.0020	0.0657	0.0001	1.1357
18	0.9863	0.0022	0.0211	0.0000	1.0857	0.9811	0.0019	0.0569	0.0000	1.1255
19	0.9872	0.0021	0.0172	−0.0000	1.0819	0.9820	0.0018	0.0559	0.0000	1.1229
20	0.9878	0.0021	0.0228	0.0000	1.0645	0.9826	0.0017	0.0433	0.0000	1.1039
21	0.9883	0.0021	0.0385	0.0001	1.0471	0.9832	0.0016	0.0387	0.0000	1.0813
22	0.9895	0.0021	0.0379	0.0001	1.0364	0.9843	0.0015	0.0372	0.0000	1.0643
23	0.9900	0.0020	0.0417	0.0001	1.0354	0.9848	0.0015	0.0386	0.0000	1.0669
24	0.9903	0.0020	0.0405	0.0001	1.0316	0.9854	0.0015	0.0313	−0.0000	1.0671
25	0.9911	0.0020	0.0406	0.0001	1.0239	0.9862	0.0014	0.0324	0.0000	1.0564
26	0.9920	0.0020	0.0426	0.0001	1.0137	0.9867	0.0013	0.0331	0.0000	1.0453
27	0.9923	0.0020	0.0427	0.0001	1.0011	0.9867	0.0013	0.0368	0.0000	1.0314
28	0.9925	0.0020	0.0456	0.0001	0.9919	0.9870	0.0012	0.0387	0.0000	1.0217
29	0.9926	0.0020	0.0469	0.0001	0.9838	0.9870	0.0012	0.0374	0.0000	1.0143
30	0.9926	0.0020	0.0471	0.0001	0.9842	0.9870	0.0012	0.0377	0.0000	1.0147

This table reports the in-sample and out-of-sample correlation of the tracking portfolio return with the index return ρ_{py}, daily tracking error $\sqrt{v^\star}$ and cumulative out-of-sample return differential between the tracking portfolio and the index RAR, and the intercept α_p and the slope β_p of a simple linear regression of the portfolio excess return on the benchmark index excess return. The estimation period covers 2 January 2018 until 31 December 2018. The out-of-sample test period starts in 2 January 2019 and ends on 31 December 2019. RAR is the cumulative realized active return given by $\Pi_{t=1}^{t=T}(1+r_{p,t}) - \Pi_{t=1}^{t=T}(1+r_{y,t})$.

The last important factor to include in the replication algorithm is the momentum effect. I present the findings regarding the performance of Algortihm 3B which uses the four-factor model of Carhart (1997) in Table 4. One potential reason for considering the momentum effect is to take advantage of component stocks that may have large exposure to the momentum factor, whose inclusion in the replicating portfolio might result in a better performance out-of-sample. This certainly appears to the case, though once again we needed to include at least half of the component stocks before any consistent outperformance began to take place. However, it must be emphasized that this algorithm's objective is to maximize correlation with the benchmark, and in that way it meets its objective very well both in-sample and out-of-sample.

Table 4. Optimal tracking portfolios for DJIA index: Algorithm 3B.

Number of Stocks	ρ_{py}	$\sqrt{v^\star}$	In-Sample RAR	α_p	β_p	ρ_{py}	$\sqrt{v^\star}$	Out-of-Sample RAR	α_p	β_p
1	0.8241	0.0075	−0.1261	−0.0006	1.1436	0.5928	0.0129	−0.3359	−0.0015	1.2636
2	0.9062	0.0048	0.0033	−0.0001	1.2329	0.8169	0.0066	0.0560	0.0000	1.2272
3	0.9445	0.0041	0.0496	0.0001	1.2875	0.8594	0.0056	−0.0452	−0.0003	1.2135
4	0.9468	0.0037	0.0674	0.0002	1.2612	0.8792	0.0054	−0.0734	−0.0005	1.2490
5	0.9425	0.0033	0.0845	0.0002	1.2394	0.9030	0.0044	0.0072	−0.0001	1.1925
6	0.9490	0.0031	0.0597	0.0001	1.2777	0.9260	0.0041	−0.0035	−0.0002	1.2269
7	0.9594	0.0029	0.0366	0.0001	1.2319	0.9432	0.0036	0.0299	−0.0001	1.2187
8	0.9621	0.0028	0.0408	0.0001	1.2036	0.9511	0.0033	0.0441	−0.0000	1.2067
9	0.9637	0.0027	0.0133	−0.0000	1.1995	0.9546	0.0033	0.0491	−0.0000	1.2228
10	0.9683	0.0026	0.0139	−0.0000	1.1720	0.9622	0.0029	0.0624	0.0000	1.1909
11	0.9755	0.0025	0.0254	0.0000	1.1515	0.9705	0.0025	0.0584	0.0000	1.1694
12	0.9785	0.0024	0.0221	0.0000	1.1380	0.9741	0.0024	0.0601	0.0000	1.1769
13	0.9792	0.0024	0.0196	−0.0000	1.1454	0.9760	0.0024	0.0945	0.0001	1.1938
14	0.9827	0.0023	0.0124	−0.0000	1.1205	0.9803	0.0021	0.0827	0.0001	1.1534
15	0.9852	0.0023	0.0137	−0.0000	1.1054	0.9814	0.0020	0.0603	0.0000	1.1449
16	0.9864	0.0022	0.0161	−0.0000	1.0887	0.9818	0.0019	0.0621	0.0001	1.1298
17	0.9858	0.0022	0.0184	−0.0000	1.0948	0.9808	0.0020	0.0648	0.0001	1.1353
18	0.9864	0.0022	0.0208	0.0000	1.0854	0.9811	0.0019	0.0560	0.0000	1.1249
19	0.9873	0.0021	0.0169	−0.0000	1.0816	0.9820	0.0018	0.0550	0.0000	1.1224
20	0.9879	0.0021	0.0225	0.0000	1.0641	0.9826	0.0017	0.0425	0.0000	1.1032
21	0.9884	0.0021	0.0382	0.0001	1.0467	0.9831	0.0016	0.0380	0.0000	1.0808
22	0.9896	0.0021	0.0376	0.0001	1.0362	0.9842	0.0015	0.0365	0.0000	1.0639
23	0.9900	0.0020	0.0414	0.0001	1.0352	0.9848	0.0015	0.0376	0.0000	1.0664
24	0.9903	0.0020	0.0402	0.0001	1.0313	0.9854	0.0015	0.0305	−0.0000	1.0666
25	0.9911	0.0020	0.0403	0.0001	1.0236	0.9862	0.0014	0.0315	−0.0000	1.0560
26	0.9921	0.0020	0.0423	0.0001	1.0137	0.9867	0.0013	0.0320	0.0000	1.0451
27	0.9923	0.0020	0.0423	0.0001	1.0010	0.9867	0.0013	0.0359	0.0000	1.0311
28	0.9926	0.0020	0.0453	0.0001	0.9918	0.9870	0.0012	0.0378	0.0000	1.0214
29	0.9926	0.0020	0.0466	0.0001	0.9837	0.9870	0.0012	0.0366	0.0000	1.0140
30	0.9926	0.0020	0.0467	0.0001	0.9838	0.9870	0.0012	0.0367	0.0000	1.0142

This table reports the in-sample and out-of-sample correlation of the tracking portfolio return with the index return ρ_{py}, daily tracking error $\sqrt{v^\star}$ and cumulative out-of-sample return differential between the tracking portfolio and the index RAR, and the intercept α_p and the slope β_p of a simple linear regression of the portfolio excess return on the benchmark index excess return. The estimation period covers 2 January 2018 until 31 December 2018. The out-of-sample test period starts in 2 January 2019 and ends on 31 December 2019. RAR is the cumulative realized active return given by $\Pi_{t=1}^{t=T}(1+r_{p,t}) - \Pi_{t=1}^{t=T}(1+r_{y,t})$.

Next, I turn toward the outperformance algorithm. I consider three versions of the algorithm depending on which factor model is used to describe security returns, variances and covariances. Using the market model of stock return gives rise to Algorithm 4. Choosing the three-factor Fama and French model produces Algorithm 5A, while the four-factor Carhart model I denote Algorithm 5B. I report the performance of Algorithm 4 in Table 5. This algorithm by design tries to maximize ex ante the active portfolio relative to the benchmark. It is worth noting that the algorithm reached a very high correlation with the benchmark very quickly and performed quite well in terms of the realized active return and the abnormal return both in-sample and out-of-sample. It is also interesting that the outperformance was similar to that of Algorithm 2 but was achieved at the cost of less systematic risk, as the active portfolio beta was slightly lower than 1, as opposed to being slightly greater than 1 in Table 2.

Table 5. Optimal tracking portfolios for DJIA index: Algorithm 4.

Number of Stocks	In-Sample ρ_{py}	$\sqrt{v^\star}$	RAR	α_p	β_p	Out-of-Sample ρ_{py}	$\sqrt{v^\star}$	RAR	α_p	β_p
1	0.6103	0.0080	0.4661	0.0015	0.6779	0.3064	0.0117	−0.0810	−0.0000	0.4765
2	0.8447	0.0059	0.2857	0.0009	0.9366	0.6317	0.0074	0.1605	0.0006	0.7987
3	0.8708	0.0048	0.2446	0.0008	1.0658	0.7291	0.0066	0.3018	0.0009	0.9443
4	0.8998	0.0043	0.2221	0.0007	1.0369	0.7626	0.0059	0.2299	0.0007	0.9330
5	0.9373	0.0039	0.2125	0.0007	1.1012	0.8379	0.0047	0.1303	0.0004	0.9696
6	0.9383	0.0037	0.2071	0.0007	1.0411	0.8355	0.0045	0.0751	0.0002	0.9188
7	0.9399	0.0034	0.1932	0.0006	1.0636	0.8672	0.0042	0.0382	0.0001	0.9892
8	0.9437	0.0033	0.1958	0.0006	1.0241	0.8703	0.0040	0.0464	0.0001	0.9519
9	0.9459	0.0032	0.1973	0.0006	1.0220	0.8857	0.0038	0.0497	0.0001	0.9651
10	0.9501	0.0032	0.1947	0.0006	0.9917	0.8838	0.0037	0.0511	0.0001	0.9347
11	0.9572	0.0030	0.1817	0.0006	0.9915	0.9017	0.0034	0.0630	0.0002	0.9561
12	0.9567	0.0030	0.1809	0.0006	0.9655	0.9031	0.0033	0.0578	0.0002	0.9332
13	0.9618	0.0029	0.1740	0.0006	0.9592	0.9136	0.0031	0.0427	0.0001	0.9281
14	0.9643	0.0028	0.1647	0.0005	0.9461	0.9204	0.0030	0.0465	0.0001	0.9236
15	0.9651	0.0027	0.1623	0.0005	0.9591	0.9300	0.0028	0.0513	0.0001	0.9388
16	0.9652	0.0027	0.1561	0.0005	0.9316	0.9268	0.0028	0.0599	0.0002	0.9095
17	0.9675	0.0026	0.1478	0.0005	0.9265	0.9368	0.0026	0.0340	0.0001	0.9135
18	0.9684	0.0026	0.1437	0.0005	0.9162	0.9391	0.0026	0.0357	0.0001	0.8994
19	0.9682	0.0025	0.1379	0.0004	0.9068	0.9386	0.0026	0.0337	0.0001	0.8817
20	0.9718	0.0024	0.1295	0.0004	0.9071	0.9465	0.0024	0.0435	0.0001	0.8836
21	0.9713	0.0024	0.1255	0.0004	0.9050	0.9505	0.0023	0.0331	0.0001	0.8874
22	0.9739	0.0024	0.1200	0.0004	0.9183	0.9573	0.0021	0.0600	0.0002	0.9130
23	0.9740	0.0024	0.1182	0.0004	0.9143	0.9572	0.0021	0.0573	0.0002	0.9081
24	0.9781	0.0023	0.1110	0.0003	0.9168	0.9631	0.0020	0.0586	0.0002	0.9277
25	0.9787	0.0023	0.1037	0.0003	0.9177	0.9654	0.0019	0.0571	0.0002	0.9312
26	0.9836	0.0022	0.0950	0.0003	0.9400	0.9732	0.0017	0.0515	0.0001	0.9508
27	0.9858	0.0021	0.0849	0.0002	0.9322	0.9709	0.0018	0.0448	0.0001	0.9341
28	0.9880	0.0021	0.0751	0.0002	0.9332	0.9768	0.0016	0.0552	0.0001	0.9476
29	0.9899	0.0020	0.0625	0.0002	0.9449	0.9825	0.0014	0.0308	0.0000	0.9637
30	0.9914	0.0020	0.0506	0.0001	0.9505	0.9848	0.0013	0.0333	0.0000	0.9786

This table reports the in-sample and out-of-sample correlation of the tracking portfolio return with the index return ρ_{py}, daily tracking error $\sqrt{v^\star}$ and cumulative out-of-sample return differential between the tracking portfolio and the index RAR, and the intercept α_p and the slope β_p of a simple linear regression of the portfolio excess return on the benchmark index excess return. The estimation period covers 2 January 2018 until 31 December 2018. The out-of-sample test period starts in 2 January 2019 and ends on 31 December 2019. RAR is the cumulative realized active return given by $\Pi_{t=1}^{t=T}(1+r_{p,t}) - \Pi_{t=1}^{t=T}(1+r_{y,t})$.

In Table 6 I document the performance of the outperformance Algorithm 5A which uses the Fama and French (1992) three-factor model of stock returns. This allows the algorithm to control for differences in the loadings of the component stocks on the market return, the SMB factor and the HML factor. Thus, we have an improvement on the single-factor model of the empirical implementation previously reported in Table 5. The performance here was in line with what was reported in Table 5 previously, namely, the portfolios were once again more concentrated on lower beta stocks which also have lower idiosyncratic risks and that tends to deliver good returns with less systematic risk. When comparing the performance of Algorithm 5A to that of Algorithm 4, we notice that the realized active return and the abnormal return become slightly more volatile after the addition of every new stock. This is perhaps due to the additional parameters that need to be estimated—namely, the three factor exposures for each additional stock—and that appear to add some estimation noise to the performance results.

Table 6. Optimal tracking portfolios for DJIA index: Algorithm 5A.

Number of Stocks	ρ_{py}	$\sqrt{v^\star}$	In-Sample RAR	α_p	β_p	ρ_{py}	$\sqrt{v^\star}$	Out-of-Sample RAR	α_p	β_p
1	0.6104	0.0088	0.4658	0.0015	0.6780	0.3065	0.0117	−0.0810	−0.0000	0.4766
2	0.7650	0.0068	0.2276	0.0008	0.7662	0.6432	0.0068	−0.2124	−0.0007	0.7376
3	0.7917	0.0063	0.1872	0.0006	0.7411	0.7064	0.0056	−0.1474	−0.0005	0.6937
4	0.8061	0.0059	0.1537	0.0005	0.7535	0.7465	0.0053	−0.1810	−0.0006	0.7435
5	0.8013	0.0058	0.1545	0.0005	0.7142	0.7472	0.0051	−0.1693	−0.0006	0.7047
6	0.9077	0.0048	0.1542	0.0005	0.8483	0.8204	0.0044	−0.0137	−0.0000	0.8094
7	0.9231	0.0043	0.1543	0.0005	0.8689	0.8242	0.0044	−0.0080	−0.0000	0.8268
8	0.9371	0.0040	0.1120	0.0004	0.8568	0.8144	0.0045	−0.0185	−0.0000	0.7972
9	0.9537	0.0037	0.0948	0.0003	0.8706	0.8696	0.0038	−0.0047	−0.0000	0.8635
10	0.9521	0.0037	0.1016	0.0003	0.8497	0.8681	0.0038	0.0035	0.0000	0.8422
11	0.9458	0.0035	0.1078	0.0003	0.8344	0.8615	0.0038	−0.0200	−0.0001	0.8199
12	0.9500	0.0034	0.1139	0.0004	0.8429	0.8780	0.0036	−0.0131	−0.0000	0.8390
13	0.9576	0.0031	0.1195	0.0004	0.9018	0.8923	0.0034	0.0562	0.0002	0.8888
14	0.9680	0.0029	0.0860	0.0003	0.9316	0.9228	0.0029	0.0024	−0.0000	0.9318
15	0.9690	0.0029	0.0883	0.0003	0.9171	0.9208	0.0029	0.0046	−0.0000	0.9158
16	0.9675	0.0028	0.0868	0.0003	0.8984	0.9167	0.0030	0.0124	0.0000	0.8954
17	0.9737	0.0027	0.0708	0.0002	0.9020	0.9364	0.0026	0.0322	0.0001	0.9204
18	0.9721	0.0027	0.0686	0.0002	0.8934	0.9381	0.0026	0.0304	0.0001	0.9023
19	0.9732	0.0026	0.0696	0.0002	0.9068	0.9450	0.0024	0.0351	0.0001	0.9162
20	0.9742	0.0025	0.0626	0.0002	0.9082	0.9481	0.0024	0.0344	0.0001	0.9212
21	0.9772	0.0025	0.0594	0.0002	0.9087	0.9532	0.0022	0.0439	0.0001	0.9206
22	0.9807	0.0024	0.0421	0.0001	0.9188	0.9607	0.0021	0.0471	0.0001	0.9455
23	0.9804	0.0023	0.0431	0.0001	0.9123	0.9620	0.0020	0.0492	0.0001	0.9415
24	0.9804	0.0023	0.0433	0.0001	0.9128	0.9620	0.0020	0.0495	0.0001	0.9422
25	0.9816	0.0023	0.0448	0.0001	0.9112	0.9642	0.0020	0.0416	0.0001	0.9389
26	0.9808	0.0022	0.0491	0.0001	0.9283	0.9656	0.0019	0.0296	0.0000	0.9641
27	0.9844	0.0021	0.0420	0.0001	0.9528	0.9710	0.0018	0.0247	0.0000	0.9846
28	0.9910	0.0021	0.0480	0.0001	0.9744	0.9849	0.0013	0.0089	−0.0001	0.9918
29	0.9911	0.0020	0.0495	0.0001	0.9751	0.9851	0.0013	0.0158	−0.0000	0.9980
30	0.9926	0.0020	0.0468	0.0001	0.9843	0.9870	0.0012	0.0377	0.0000	1.0148

This table reports the in-sample and out-of-sample correlation of the tracking portfolio return with the index return ρ_{py}, daily tracking error $\sqrt{v^\star}$ and cumulative out-of-sample return differential between the tracking portfolio and the index RAR, and the intercept α_p and the slope β_p of a simple linear regression of the portfolio excess return on the benchmark index excess return. The estimation period covers 2 January 2018 until 31 December 2018. The out-of-sample test period starts in 2 January 2019 and ends on 31 December 2019. RAR is the cumulative realized active return given by $\Pi_{t=1}^{t=T}(1 + r_{p,t}) - \Pi_{t=1}^{t=T}(1 + r_{y,t})$.

Table 7 presents the in-sample and out-of-sample performance of Algorithm 5B which uses the Carhart (1997) four-factor model of security returns. This allows the algorithm to control for exposure to the Fama and French factors and loadings on the momentum factor for each component stock. The performance of Algorithm 5B was largely in line that of Algorithm 5A. One thing that they both have in common that is different from the market model is that it takes far less concentrated portfolios in order for the out-of-sample performance to improve. Namely, very concentrated portfolios with up to 10 component stocks tended to underperform out-of-sample. However, once all 30 component stocks were included, both Algorithms 5A and 5B delivered a steady out-of-sample performance.

Table 7. Optimal tracking portfolios for DJIA index: Algorithm 5B.

Number of Stocks	ρ_{py}	$\sqrt{v^\star}$	In-Sample RAR	α_p	β_p	ρ_{py}	$\sqrt{v^\star}$	Out-of-Sample RAR	α_p	β_p
1	0.6104	0.0090	0.4655	0.0015	0.6780	0.3065	0.0117	−0.0810	−0.0000	0.4768
2	0.7650	0.0068	0.2276	0.0008	0.7662	0.6432	0.0068	−0.2124	−0.0007	0.7376
3	0.7917	0.0063	0.1872	0.0006	0.7411	0.7064	0.0056	−0.1474	−0.0005	0.6937
4	0.8061	0.0059	0.1536	0.0005	0.7535	0.7466	0.0053	−0.1810	−0.0006	0.7436
5	0.8014	0.0058	0.1545	0.0005	0.7144	0.7473	0.0051	−0.1693	−0.0006	0.7049
6	0.9077	0.0048	0.1542	0.0005	0.8484	0.8204	0.0044	−0.0137	−0.0000	0.8095
7	0.9231	0.0043	0.1543	0.0005	0.8691	0.8243	0.0044	−0.0079	−0.0000	0.8269
8	0.9370	0.0040	0.1113	0.0004	0.8564	0.8142	0.0045	−0.0192	−0.0000	0.7970
9	0.9535	0.0037	0.0943	0.0003	0.8702	0.8693	0.0038	−0.0053	−0.0000	0.8631
10	0.9519	0.0036	0.1011	0.0003	0.8494	0.8678	0.0038	0.0029	0.0000	0.8419
11	0.9457	0.0035	0.1072	0.0003	0.8340	0.8609	0.0038	−0.0205	−0.0001	0.8193
12	0.9498	0.0034	0.1135	0.0004	0.8424	0.8776	0.0036	−0.0139	−0.0000	0.8385
13	0.9574	0.0031	0.1193	0.0004	0.9008	0.8922	0.0034	0.0547	0.0002	0.8880
14	0.9681	0.0029	0.0852	0.0003	0.9315	0.9226	0.0029	0.0013	−0.0000	0.9315
15	0.9690	0.0029	0.0875	0.0003	0.9171	0.9206	0.0029	0.0033	−0.0000	0.9158
16	0.9674	0.0028	0.0860	0.0003	0.8984	0.9165	0.0030	0.0113	0.0000	0.8953
17	0.9736	0.0027	0.0702	0.0002	0.9019	0.9361	0.0026	0.0307	0.0001	0.9203
18	0.9720	0.0027	0.0680	0.0002	0.8934	0.9379	0.0026	0.0289	0.0001	0.9024
19	0.9731	0.0026	0.0690	0.0002	0.9066	0.9447	0.0024	0.0333	0.0001	0.9162
20	0.9741	0.0025	0.0621	0.0002	0.9080	0.9478	0.0024	0.0326	0.0001	0.9212
21	0.9772	0.0025	0.0589	0.0002	0.9086	0.9530	0.0023	0.0430	0.0001	0.9205
22	0.9806	0.0024	0.0417	0.0001	0.9186	0.9605	0.0021	0.0459	0.0001	0.9453
23	0.9803	0.0023	0.0428	0.0001	0.9122	0.9618	0.0020	0.0483	0.0001	0.9413
24	0.9803	0.0023	0.0429	0.0001	0.9124	0.9619	0.0020	0.0484	0.0001	0.9416
25	0.9816	0.0023	0.0444	0.0001	0.9110	0.9640	0.0020	0.0409	0.0001	0.9385
26	0.9808	0.0022	0.0487	0.0001	0.9275	0.9655	0.0019	0.0283	0.0000	0.9632
27	0.9844	0.0021	0.0416	0.0001	0.9526	0.9709	0.0018	0.0238	0.0000	0.9841
28	0.9910	0.0021	0.0476	0.0001	0.9741	0.9848	0.0013	0.0082	−0.0001	0.9914
29	0.9911	0.0020	0.0491	0.0001	0.9746	0.9850	0.0013	0.0148	−0.0000	0.9974
30	0.9926	0.0020	0.0465	0.0001	0.9839	0.9870	0.0012	0.0367	0.0000	1.0143

This table reports the in-sample and out-of-sample correlation of the tracking portfolio return with the index return ρ_{py}, daily tracking error $\sqrt{v^\star}$ and cumulative out-of-sample return differential between the tracking portfolio and the index RAR, and the intercept α_p and the slope β_p of a simple linear regression of the portfolio excess return on the benchmark index excess return. The estimation period covers January 2, 2018 until December 31, 2018. The out-of-sample test period starts in 2 January 2019 and ends on 31 December 2019. RAR is the cumulative realized active return given by $\Pi_{t=1}^{t=T}(1+r_{p,t}) - \Pi_{t=1}^{t=T}(1+r_{y,t})$.

4. Conclusions

In this paper I have addressed the question of how best to build a replicating portfolio of a component set of securities with which to efficiently track a stock market benchmark index. The analysis is by no means limited to an equity index or a traded portfolio of securities but could be applied to ways of tracking various state variables of interest, such as commodities prices or macroeconomic time series, among others. I applied this analysis for a general specification of security returns and for single and multi-factor models of security returns. The results are quite intuitive and suggest that the securities that make it into the tracking portfolio have the highest ratios of factor beta to residual error variance. This is a new and interesting result. This ratio has important economic significance, as tangency portfolio weights under a single-factor model are proportional to it. Within the context of factor models for securities returns, the replicating portfolio is proportional to the highest Sharpe ratio (tangent) portfolio scaled by the benchmark beta. This result holds regardless of the number of factors employed. It is also straightforward to extend the results to the case wherein the parameters are time-varying along the lines of a multivariate GARCH model, for example, which then leads to conditional time-varying tracking portfolio weights.

Concerning the practical implications of the theoretical results of this paper, the main advantage to the use of a factor model of stock returns is that the computational effort is reduced to a linear function of the number of stocks in the replicating portfolio. This can be of particular importance for large portfolios consisting of hundreds or even thousands of stocks. Furthermore, the algorithm using factor models is not as data-hungry as the algorithms that use only variance-covariance matrices. Note that the latter require a longer time series than the number of assets in order for a positive-definite covariance matrix to be obtained. At the same time, estimating the parameters of the factor model requires a fixed amount of historical returns, regardless of the size of the cross-section under consideration. This can be of great convenience to both individual investors and delegated portfolio managers alike.

In regard to future work, several open questions remain to be addressed. In particular, constructing fixed-weight replicating portfolios with rebalancing and proportional transaction costs will change the performance results of the replicating portfolio relative to the benchmark. Similarly, the issue of time-variation in betas and idiosyncratic volatilities would suggest that the investor should periodically re-estimate these parameters and re-balance the tracking portfolio. The optimal frequency at which this should be done remains an open question.

Several extensions to the analysis in this paper are possible. First, the inter-temporal optimal strategy over a preset horizon would be of considerable practitioner interest. Second, the explicit incorporation of transaction costs and capital gains taxes and their effect on the optimal strategy is another possibility. Numerical solutions for small sets of stocks should be feasible, though time-consuming. Similarly, the question of what the minimum number of stocks depends on is of further interest and is left for future work.

Funding: This research received no external funding.

Acknowledgments: I would like to thank Michael Chng, Daisy Doan, Steinar Ekern, Domenico Giannone, Jørgen Haug, Thore Johnsen, Edward Lin, Alfred Yawson, Syed Zamin Ali, Leon Zolotoy and seminar participants at Deakin University, the Norwegian School of Economics and Business Administration and the University of Adelaide, along with conference participants at the 2011 Econometric Society Australasian Meetings in Adelaide. Any remaining errors are my own.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Proof of Proposition 1. Using the notation in the paper we can express the variance of the tracking error as:

$$\sigma_\epsilon^2 = \sigma_y^2 + w'\Sigma_{rr}w - 2w'\sigma_{ry} \tag{A1}$$

The first-order necessary condition for optimality is:

$$2\Sigma_{rr}w - 2\sigma_{ry} = 0 \tag{A2}$$

which yields the optimal portfolio weights as:

$$w^\star = \Sigma_{rr}^{-1}\sigma_{ry}. \tag{A3}$$

The second-order condition for obtaining a minimum of the objective function is $\Sigma_{rr} > 0$, which is trivially satisfied by any positive-definite variance-covariance matrix. □

Proof of Corollary 1. Let us re-write the regression equation in (6) more compactly as follows:

$$r_y = [1_T, r]\begin{bmatrix}\theta_0 \\ \theta_1\end{bmatrix} + u. \tag{A4}$$

Let $\hat{\mu}_y = \frac{1}{T}\sum_{t=1}^{t=T} r_{y,t}$, $\hat{\mu} = \frac{1}{T}\sum_{t=1}^{t=T} r_t'$, $\hat{\Sigma}_{rr} = \frac{1}{T}\sum_{t=1}^{t=T}(r_t - \hat{\mu}')(r_t - \hat{\mu})'$, and $\hat{\sigma}_{ry} = \frac{1}{T}\sum_{t=1}^{t=T}(r_{y,t} - \hat{\mu}_y')(r_t - \hat{\mu})'$. The ordinary least squares (OLS) estimate $\hat{\theta}$ is given by:

$$\begin{aligned}
\hat{\theta} &= \left([1_T, r]' [1_T, r]\right)^{-1} [1_T, r]' r_y, \qquad \text{(A5)} \\
&= \left(T \begin{bmatrix} 1 & \hat{\mu}' \\ \hat{\mu} & \hat{\mu}\hat{\mu}' + \hat{\Sigma}_{rr} \end{bmatrix}\right)^{-1} T \begin{bmatrix} \hat{\mu}_y \\ \hat{\mu}_y \hat{\mu} + \hat{\sigma}_{ry} \end{bmatrix}, \\
&= \frac{1}{T} \begin{bmatrix} 1 + \hat{\mu}'\hat{\Sigma}_{rr}^{-1}\hat{\mu} & -\hat{\mu}'\hat{\Sigma}_{rr}^{-1} \\ -\hat{\Sigma}_{rr}^{-1} & \hat{\Sigma}_{rr}^{-1} \end{bmatrix} T \begin{bmatrix} \hat{\mu}_y \\ \hat{\mu}_y \hat{\mu} + \hat{\sigma}_{ry} \end{bmatrix}, \\
&= \begin{bmatrix} \hat{\mu}_y - \hat{\mu}'\hat{\Sigma}_{rr}^{-1}\hat{\sigma}_{ry} \\ \hat{\Sigma}_{rr}^{-1}\hat{\sigma}_{ry} \end{bmatrix},
\end{aligned}$$

or $\hat{\theta}_0 = \hat{\mu}_y - \hat{\mu}'\hat{\theta}_1$ and $\hat{\theta}_1 = \hat{\Sigma}_{rr}^{-1}\hat{\sigma}_{ry}$. □

Proof of Corollary 2. For a detailed proof see, for example, Huberman and Kandel (1987). When $\hat{\theta}_0 = 0$ and $1_N'\hat{\theta}_1 = 1$, the mean-variance frontier generated by the replicating securities coincides with the augmented mean-variance frontier that includes the benchmark index *and* the replicating securities. □

Proof of Corollary 3. From the OLS estimates in the proof of Corollary 1 above, it follows that if $\hat{\theta}_0 < 0$ then

$$\hat{\mu}_y < \hat{\mu}'\hat{\Sigma}_{rr}^{-1}\hat{\sigma}_{ry} = \hat{\mu}'w^\star = \hat{\mu}_p, \qquad \text{(A6)}$$

which implies that the replicating portfolio has a higher excess return in-sample than the benchmark index. □

Proof of Proposition 2. The inverse of the variance-covariance matrix of securities' excess return in (17) is given by:

$$\Sigma_{rr}^{-1} = D^{-1} - D^{-1}\beta_r \left(\frac{1}{\sigma_m^2} + \beta_r' D^{-1}\beta_r\right)^{-1} \beta_r' D^{-1}. \qquad \text{(A7)}$$

Substituting (A7) and (18) into (5) yields:

$$w^\star = D^{-1}\beta_r\beta_y\sigma_m^2 - \frac{D^{-1}\beta_r(\beta_r' D^{-1}\beta_r)\beta_y\sigma_m^2}{\frac{1}{\sigma_m^2} + \beta_r' D^{-1}\beta_r}. \qquad \text{(A8)}$$

Collecting terms in the previous line produces the final result in (21). The expression for the variance of the tracking error under the optimal weights follows immediately by substituting the latter into $v(w) = \sigma_y^2 + w'\Sigma_{rr} - 2w'\sigma_{ry}$ and collecting terms. □

Proof of Corollary 4. The beta of the replicating portfolio can be obtained by using the optimal portfolio weights in (21):

$$\begin{aligned}
\beta_p &= \beta_r' w^\star, \qquad \text{(A9)} \\
&= \frac{\beta_r' D^{-1}\beta_r}{\frac{1}{\sigma_m^2} + \beta_r' D^{-1}\beta_r}\beta_y, \\
&< \beta_y.
\end{aligned}$$

□

Proof of Corollary 5. The tangent portfolio weights in the framework of an exact single-factor model are:

$$\begin{aligned} w_{tg} &= \Sigma_{rr}^{-1}\mu_r, \\ &= \left[D^{-1} - D^{-1}\beta_r\left(\frac{1}{\sigma_m^2} + \beta_r' D^{-1}\beta_r\right)^{-1}\beta_r' D^{-1}\right]\beta_r\mu_m, \\ &= \frac{D^{-1}\beta_r}{\frac{1}{\sigma_m^2} + \beta_r' D^{-1}\beta_r} \cdot \frac{\mu_m}{\sigma_m^2}. \end{aligned} \tag{A10}$$

Clearly the solution (21) differs from w_{tg} only up to the beta of the benchmark index, β_y, and the factor premium per unit of factor risk, μ_m/σ_m^2:

$$w^\star = w_{tg}\beta_y\frac{\sigma_m^2}{\mu_m}, \tag{A11}$$

which demonstrates the claim made in the corollary. □

Proof of Proposition 3. Following the same steps as in the proof of proposition 2 above, the inverse of the variance-covariance matrix in (24) is:

$$\Sigma_{rr}^{-1} = D^{-1} - D^{-1}B_r(V_f^{-1} + B_r' D^{-1}B_r)^{-1}B_r' D^{-1}. \tag{A12}$$

Substituting (A12) and (25) into (5) we get:

$$\begin{aligned} w^\star &= D^{-1}B_rV_fb_y - D^{-1}B_r(V_f^{-1} + B_r'D^{-1}B_r)^{-1}B_rD^{-1}B_rV_fb_y, \\ &= D^{-1}B_r\left[I_K - (V_f^{-1} + B_r'D^{-1}B_r)^{-1}B_rD^{-1}B_r\right]V_fb_y, \\ &= D^{-1}B_r\left[(V_f^{-1} + B_r'D^{-1}B_r)^{-1}(V_f^{-1} + B_r'D^{-1}B_r) - (V_f^{-1} + B_r'D^{-1}B_r)^{-1}B_rD^{-1}B_r\right]V_fb_y, \\ &= D^{-1}B_r(V_f^{-1} + B_r'D^{-1}B_r)^{-1}V_f^{-1}V_fb_y, \\ &= D^{-1}B_r(V_f^{-1} + B_r'D^{-1}B_r)^{-1}b_y, \end{aligned} \tag{A13}$$

which is the proposed solution. The variance of the tracking error under the optimal portfolio weights is obtained in the same way as in the proof of proposition 2 above. □

Proof of Proposition 4. Using the notation in the paper we can express the Lagrangian as:

$$\mathcal{L} = \sigma_y^2 + w'\Sigma_{RR}w - 2w'\sigma_{Ry} + \lambda\left(1 - w'1_N\right) \tag{A14}$$

The first-order necessary condition for optimality is:

$$2\Sigma_{RR}w - 2\sigma_{Ry} - \lambda 1_N = 0 \tag{A15}$$

which yields the optimal portfolio weights as:

$$w^\star(\lambda) = \Sigma_{RR}^{-1}\sigma_{Ry} + (\lambda/2)\Sigma_{RR}^{-1}1_N. \tag{A16}$$

Substituting $w^\star(\lambda)$ into the budget constraint we can solve for the shadow price:

$$\begin{aligned} 1 &= 1_N'\Sigma_{RR}^{-1}\sigma_{Ry} + (\lambda/2)1_N'\Sigma_{RR}^{-1}1_N, \\ (\lambda/2) &= \frac{1 - 1_N'\Sigma_{RR}^{-1}\sigma_{Ry}}{1_N'\Sigma_{RR}^{-1}1_N}. \end{aligned}$$

Substituting the above into (A16) produces the stated result. The second-order condition for obtaining a minimum of the objective function is $\Sigma_{RR} > 0$, which is trivially satisfied by any positive-definite variance-covariance matrix. □

Proof of Proposition 5. Using the notation in the paper we can express the Lagrangian as:

$$\mathcal{L} = \sigma_y^2 + w'\Sigma_{rr}w - 2w'\sigma_{ry} + \lambda\left(\sigma_0^2 - w'\Sigma_{rr}w\right) \tag{A17}$$

The first-order necessary condition for optimality is:

$$2\Sigma_{rr}w - 2\sigma_{ry} - 2\lambda\Sigma_{rr}w = 0 \tag{A18}$$

which yields the optimal portfolio weights as:

$$w^\star(\lambda) = \left(\frac{1}{1-\lambda}\right)\Sigma_{rr}^{-1}\sigma_{ry}. \tag{A19}$$

By substituting $w^\star(\lambda)$ into the variance constraint we can solve for the shadow price:

$$\begin{aligned} \sigma_0^2 &= \left(\frac{1}{1-\lambda}\right)^2 (\sigma_{ry}'\Sigma_{rr}^{-1}\sigma_{ry}), \\ \left(\frac{1}{1-\lambda}\right) &= \frac{\sigma_0}{\sqrt{\sigma_{ry}'\Sigma_{rr}^{-1}\sigma_{ry}}}. \end{aligned}$$

Substituting the above into (A19) produces the stated result. The second-order condition for obtaining a minimum of the objective function is $(1-\lambda)\Sigma_{rr} > 0$, which is trivially satisfied by any positive-definite variance-covariance matrix *and* $\sigma_0 > 0$. □

Proof of Proposition 6. Using the notation in the paper we can express the Lagrangian as:

$$\mathcal{L} = \sigma_y^2 + w'\Sigma_{RR}w - 2w'\sigma_{Ry} + \lambda_1(1 - w'1_N) + \lambda_2\left(\sigma_0^2 - w'\Sigma_{RR}w\right). \tag{A20}$$

The first-order necessary condition for optimality is:

$$2\Sigma_{RR}w - 2\sigma_{Ry} - \lambda_1 1_N - 2\lambda_2\Sigma_{RR}w = 0. \tag{A21}$$

Substituting the above in the two constraints leads to the following system of equations in the shadow prices λ_1 and λ_2:

$$\begin{aligned} \left(\frac{1}{1-\lambda_2}\right)\left[(1_N'\Sigma_{RR}^{-1}\sigma_{Ry}) + (\lambda_1/2)(1_N\Sigma_{RR}^{-1}1_N)\right] &= 1, \\ \left(\frac{1}{1-\lambda_2}\right)^2\left[(\sigma_{Ry}'\Sigma_{RR}^{-1}\sigma_{Ry}) + 2\sigma_{Ry}'\Sigma_{RR}^{-1}1_N(\lambda_1/2) + (\lambda_1/2)^2(1_N'\Sigma_{RR}^{-1}1_N)\right] &= \sigma_0^2. \end{aligned}$$

The solution to the above system of equations is:

$$\begin{aligned} \left(\frac{1}{1-\lambda_2}\right) &= \frac{\sigma_0^2 - \sigma_{mv}^2}{Q}, \\ \lambda_1/2 &= \sigma_{mv}^2\left[\sqrt{\frac{Q}{\sigma_0^2 - \sigma_{mv}^2}} - (1_N'\Sigma_{RR}^{-1}\sigma_{Ry})\right], \end{aligned}$$

where

$$Q = (\sigma'_{Ry}\Sigma_{RR}^{-1}\sigma_{Ry}) - \frac{(1'_N\Sigma_{RR}^{-1}\sigma_{Ry})}{(1'_N\Sigma_{RR}^{-1}1_N)},$$
$$\sigma_{mv}^2 = \frac{1}{1'_N\Sigma_{RR}^{-1}1_N}.$$

For substituting λ_1 and λ_2 into the first order condition in (A21) above and simplifying the yield, the result is stated in the text. The second-order condition for obtaining a minimum of the objective function is $(1-\lambda_2)\Sigma_{RR} > 0$ which is trivially satisfied by any positive-definite variance-covariance matrix *and* $\sigma_0^2 > \sigma_{mv}^2$. □

Proof of Proposition 7. Using the notation in the paper we can express the Lagrangian as:

$$\mathcal{L} = \sigma_y^2 + w'\Sigma_{rr}w - 2w'\sigma_{ry} + \lambda(\beta_0 - w'\beta_r). \tag{A22}$$

The first-order necessary condition for optimality is:

$$2\Sigma_{rr}w - 2\sigma_{ry} - \lambda\beta_r = 0. \tag{A23}$$

Substituting the above in the factor loading constraint leads to the following equation for the shadow price $\lambda/2$:

$$\beta_0 = \beta'_r\Sigma_{rr}^{-1}\sigma_{ry} + (\lambda/2)(\beta'_r\Sigma_{rr}^{-1}\beta_r).$$

The solution to the above yields the following value for the Lagrangian multiplier $\lambda/2$:

$$\lambda/2 = \frac{\beta_0 - \beta'_r\Sigma_{rr}^{-1}\sigma_{ry}}{\beta'_r\Sigma_{rr}^{-1}\beta_r}.$$

Substituting $\lambda/2$ into the first order condition in (A23) above and using the factor structure of σ_{ry} and Σ_{rr} yields the stated result after collecting terms. The second-order condition for obtaining a minimum of the objective function is $\Sigma_{rr} > 0$ which is trivially satisfied by any positive-definite variance-covariance matrix. □

Proof of Proposition 8. Using the notation in the paper we can express the Lagrangian as:

$$\mathcal{L} = \sigma_y^2 + w'\Sigma_{rr}w - 2w'\sigma_{ry} + \lambda'(b_0 - B'_r w). \tag{A24}$$

The first-order necessary condition for optimality is:

$$2\Sigma_{rr}w - 2\sigma_{ry} - B_r\lambda = 0. \tag{A25}$$

Substituting the above in the beta constraint leads to the following equation in the shadow price vector $\lambda/2$:

$$b_0 = B'_r\Sigma_{rr}^{-1}\sigma_{ry} + (B'_r\Sigma_{rr}^{-1}B_r)(\lambda/2).$$

Using the factor structure of σ_{ry} and Σ_{rr}, the solution to the above yields the following value for the Lagrangian multiplier vector $\lambda/2$:

$$\lambda/2 = (B'_r D^{-1}B_r)^{-1}b_0 + V_f(b_0 - b_y).$$

Substituting $\lambda/2$ into the first order condition in (A25) above yields the stated result after collecting terms. The second-order condition for obtaining a minimum of the objective function is $\Sigma_{rr} > 0$, which is trivially satisfied by any positive-definite variance-covariance matrix. □

References

Amenc, Noel, Lionel Martellini, Jean-Christophe Meyfredi, and Volker Zieman. 2010. Passive Hedge Fund Replication—Beyond the Linear Case. *European Financial Management* 16: 191–200. [CrossRef]

Burmeister, Curt, Helmut Mausser, and Rafael Mendoza. 2004. Actively managing tracking error. *Journal of Asset Management* 5: 410–22. [CrossRef]

Carhart, Mark M. 1997. On Persistence in Mutual Fund Performance. *Journal of Finance* 52: 57–82. [CrossRef]

Chan, Louis K. C., Jason Karceski, and Josef Lakonishok. 1999. On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *Review of Financial Studies* 12: 937–74. [CrossRef]

Corielli, Francesco, and Massimiliano Marcellino. 2006. Factor based index tracking. *Journal of Banking and Finance* 30: 2215–33. [CrossRef]

Dunis, Christian L., and Richard Ho. 2005. Cointegration portfolio of European equities for index tracking and market neutral strategies. *Journal of Asset Management* 6: 33–52. [CrossRef]

El-Hassan, Nadima, and Paul Kofman. 2003. Tracking Error and Active Portfolio Management. *Australian Journal of Management* 28: 183–207. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 1992. The Cross-Section of Expected Stock Returns. *Journal of Finance* 47: 427–65. [CrossRef]

Focardi, Sergio F., and Frank J. Fabozzi. 2004. A methodology for index tracking based on time-series clustering. *Quantitative Finance* 4: 417–25. [CrossRef]

Glabadanidis, Paskalis, and Leon Zolotoy. 2013. Benchmark Replication Portfolio Strategies. *Journal of Asset Management* 14: 95–110. [CrossRef]

Glabadanidis, Paskalis. 2014. Tangent Portfolio Weights Without Explicitly Specified Expected Returns. *Journal of Asset Management* 15: 177–90. [CrossRef]

Glabadanidis, Paskalis. 2020. Actively Managed Mean-Variance Portfolios. *Journal of Wealth Management*. [CrossRef]

Hasanhodzic, Jasmina, and Andrew Lo. 2007. Can Hedge Fund Returns be Replicated? The Linear Case. *Journal of Investment Management* 5: 5–45. [CrossRef]

Heaton, James B., Nick G. Polson, and Jan Hendrik Witte. 2017. Deep learning for finance: Deep portfolios. *Applied Stochastic Models for Business and Industry* 33: 3–12. [CrossRef]

Huberman, Gur, and Shmuel Kandel. 1987. Mean-Variance Spanning. *Journal of Finance* 42: 873–88. [CrossRef]

Ingersoll, Jonathan E. 1987. *Theory of Financial Decision Making*. Lanham: Rowman & Littlefield.

Israelsen, Craig L., and Gary F. Cogswell. 2006. The error of tracking error. *Journal of Asset Management* 7: 419–24. [CrossRef]

Jorion, Philippe. 2003. Portfolio Optimization with Tracking-Error Constraints. *Financial Analysts Journal* 59: 70–82. [CrossRef]

Merton, Robert C. 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41: 867–87. [CrossRef]

Ouyang, Hongbing, Xiaowei Zhang, and Hongju Yan. 2019. Index tracking based on deep neural network. *Congnitive Systems Research* 57: 107–14. [CrossRef]

Roll, Richard. 1992. A Mean/Variance Analysis of Tracking Error. *Journal of Portfolio Management* 18: 13–22. [CrossRef]

Rudolf, Marcus, Hans-Jurgen Wolter, and Heinz Zimmermann. 1999. A Linear Model for Tracking Error Minimization. *Journal of Banking and Finance* 23: 85–103. [CrossRef]

Sharpe, William. 1963. A Simplified Model for Portfolio Analysis. *Management Science* 9: 277–93. [CrossRef]

Stoyanov, Stoyan V., Svetlozar T. Rachev, Sergio Ortobelli, and Frank J. Fabozzi. 2008. Relative Deviation Metrics and the Problem of Strategy Replication. *Journal of Banking and Finance* 32: 199–206. [CrossRef]

Stutzer, Michael. 2003. Fund Managers May Cause Their Benchmarks to be Priced Risks. *Journal of Investment Management* 1: 1–13. [CrossRef]