Hedge Fund Replication: A Model Combination Approach

Michael O'Doherty* N. E. Savin† Ashish Tiwari‡

May 27, 2015§

Abstract

Recent years have seen increased demand from institutional investors for passive replication products that track the performance of hedge fund strategies. We find that, in practice, linear replication methods routinely suffer from poor tracking performance and high turnover. To address these concerns, we propose a model combination approach to index replication that pools information from a diverse set of pre-specified factor models. Compared to existing approaches, the pooled clone strategies yield consistently lower tracking errors, generate less severe portfolio drawdowns, and require substantially smaller trading volume. The pooled hedge fund clones also provide economic benefits in a portfolio allocation context.

Keywords: Hedge funds, Model pooling, Model combination, Hedge fund replication, Log score
JEL classifications: C11, C51, C53, C58, G17

*Department of Finance, Trulaske College of Business, University of Missouri, 513 Cornell Hall, Columbia, MO 65211. Phone: (573) 882-2733. Email: odohertym@missouri.edu

†Department of Economics, Tippie College of Business, University of Iowa, 108 PBB, Iowa City, IA 52242-1000. Phone: (319) 335-0855. Email: gene-savin@uiowa.edu

‡Department of Finance, Tippie College of Business, University of Iowa, 108 PBB, Iowa City, IA 52242-1000. Phone: (319) 353-2185. Email: ashish-tiwari@uiowa.edu

§We thank Scott Cederburg, Ryan Flugum, Paul Weller, Tong Yao, and seminar participants at the University of Iowa for valuable comments and suggestions. We also credit Nicolas Bollen for helpful discussions on the modeling approach implemented in this paper. Michael O'Doherty acknowledges financial support from the University of Missouri Research Board.

1 Introduction

In recent years there has been a growing interest in alternative assets, including hedge funds, on the part of institutional investors. Some of the characteristic features of hedge funds, however, including the lack of transparency, illiquidity, high leverage, and the typical 2 plus 20 incentive fee structure, are not very appealing from the perspective of such investors. For example, in an effort to reduce complexity and costs in its investment program, the California Public Employees' Retirement System (CALPERS) recently decided to shed its entire $4 billion investment in 24 hedge funds and six hedge funds-of-funds.[1]

As a result there is a robust demand for hedge fund replication strategies that seek to "clone" hedge-fund-like returns by investing in liquid instruments such as futures contracts. A popular approach to hedge fund replication involves estimating a target fund's factor exposures via Sharpe's (1992) asset-class factor model framework (e.g., Hasanhodzic and Lo (2007)) and using the estimated coefficients to determine clone portfolio weights. Briefly, the factor-based replication strategy relies on estimating long/short positions in a set of pre-specified asset-class indices with a view to minimizing the tracking error of the clone strategy. This conventional approach to replication thus involves the identification of a "best fit" factor model.

Although this framework is intuitively appealing, the existing literature has identified several practical problems with the implementation of linear factor clones. First, given the inherent flexibility of hedge funds in terms of asset-class exposure and choice of markets in which to operate, it is difficult to identify an appropriate set of factors to include in the replication model. Parsimony is also a first-order concern, as models with a large number of factors often suffer from estimation error, overfitting, and poor out-of-sample performance. Second, to capture any dynamics in the underlying hedge fund investment strategies, factor clones are typically estimated using a rolling window of prior fund returns with investment positions updated on a monthly basis. The resulting coefficient estimates often suggest a high level of portfolio turnover, making cloning costly to implement in practice. Third, the existing literature (e.g., Hasanhodzic and Lo (2007), Amenc, Martellini, Meyfredi, and Ziemann (2010), and Bollen and Fisher (2013)) largely finds that factor-based replication products for individual hedge funds and hedge fund indexes tend to underperform

[1] In explaining the decision to end the hedge fund program, known as the Absolute Return Strategies (ARS) program, Ted Eliopoulos, CalPERS Interim Chief Investment Officer, stated, "Hedge funds are certainly a viable strategy for some, but at the end of the day, when judged against their complexity, cost, and the lack of ability to scale at CalPERS' size, the ARS program is no longer warranted." (source: `http://www.calpers.ca.gov/index.jsp?bc=/about/newsroom/news/eliminate-hedge-fund.xml`)

their target portfolios.

Motivated by these concerns, we propose a novel replication approach that relies on combining or pooling a set of diverse factor models to clone hedge fund index returns. The optimal combination of factor models is determined via a decision-theoretic framework. This method employs the log score criterion, which has a long history as a decision tool.[2] Our application is motivated by the notion that, given data limitations and the flexibility of hedge funds' strategies, there is no single "best fit" model that fully characterizes index returns. We argue that a more reasonable way to proceed is to start with a pre-specified set of potential factor models for a given index. While each of the individual models is likely to generate errors in index replication, an optimal combination of these models helps to diversify individual tracking errors and leads to better out-of-sample performance.

We specifically propose clone strategies for hedge fund index returns based on an optimal combination of factor models, each representing a distinct asset class. To capture the wide range of potential investment strategies in the hedge fund universe, we consider factor models that span five asset classes: domestic equity, international equity, domestic fixed income, international fixed income, and commodities including precious metals. The weights assigned to the five individual models are based on the log score criterion, which assesses the performance of various convex combinations of the individual models in replicating index returns. Conceptually, the model combination approach is designed to deliver a replication strategy that best tracks the return distribution of the target portfolio. The key innovations are the pooling of information across models and a focus on replicating the entire target return distribution rather than simply matching certain moments, as is the conventional practice. As we discuss below, these features result in significantly improved ability to replicate index returns compared to alternative techniques.

In our empirical tests, we compare the performance of our model combination approach to several existing alternatives in replicating ten Dow Jones Credit Suisse hedge fund indexes over the period January 1994 to October 2014. The competing replication methods are based on linear models either with a pre-specified set of factors or factors selected for each index via a stepwise regression algorithm. These models rely on in-sample fit to estimate passive clone investment positions in the tradable factor portfolios. In contrast, our model combination method bases model weights and, ultimately, investment recommendations on replication performance over a pseudo out-

[2] The log score criterion, first introduced by Good (1952), is widely used in a number of similar contexts. According to Gelman, Hwang, and Vehtari (2014), the log score is perhaps the most widely used scoring rule for the purpose of model evaluation and selection. Klein and Brown (1984), for example, propose a model selection criterion related to the log score.

of-sample period. This approach allows us to incorporate a wide range of potential factors in the clone estimation while simultaneously guarding against overfitting based on in-sample performance.

Our proposed replication method offers significant improvements in statistical and economic terms over the conventional, factor-based alternatives. We first show that the model combination clones are able to better track their target indexes compared to the "single-model" replication approaches. Specifically, the pooled replication strategies have lower out-of-sample replication errors and higher correlations with the realized returns of their targets compared to alternative methods. These results are generally consistent across the ten Credit Suisse index styles. An analysis of the cumulative sum of squared (replication) errors reveals that whereas the model combination clones offer consistently superior performance during the sample period, the advantages are particularly pronounced during the 2008 global financial crisis and other periods of economic turmoil.

We next consider the out-of-sample returns earned by our pooled clones and the competing replication strategies relative to their target indexes. We find that all of the cloning methods considered lead to inferior performance when compared to the underlying hedge fund indexes over our sample period. These results are broadly consistent with managerial ability in the hedge fund industry which is difficult to match through passive, rules-based investment approaches. We further demonstrate, however, that the performance differences between target and clone portfolios are much less pronounced in recent years, suggesting an increasing attractiveness for clones as alternatives to direct hedge fund investments.

We examine the practical implementation details of the replication approaches in terms of required leverage and portfolio turnover. We find that pooled clones require reliably less trading activity when compared to the alternatives. The net portfolio leverage needed to construct the pooled clone strategies also tends to be economically reasonable and relatively more stable over time.

A natural question is whether a hedge fund clone constructed using the model combination approach would be a valuable addition to the optimal asset mix of a typical institutional investor, e.g., a large university endowment. To address this question we analyze the ex-post mean-variance optimal asset allocations for such investors during the period January 1998 to October 2014. We find that in general, the implied optimal allocation to the hedge fund clone is often quite substantial. Furthermore, the utility costs of not being able to invest in the clone, expressed in terms of annualized certainty equivalent rates (CER) of return, tend to be large. This evidence suggests that hedge fund clones can play an economically significant role in the portfolios of large institutional investors.

Overall, our findings suggest that model combination methods can mitigate many of the practical problems with traditional replication products. In particular, pooled clones lead to superior out-of-sample tracking performance relative to existing factor-based methods. These improvements can be attributed to two important differences between the pooled and single-model clones. First, individual models suffer from both specification error in identifying an appropriate set of factors and estimation error in reliably determining an index's exposure to the factors. These problems are further compounded by the fact that factor models have difficulties capturing any nonlinearities inherent in hedge fund returns.[3] The model combination clones directly alleviate these issues by diversifying replication errors across models. Second, the model pooling method emphasizes replication performance over a pseudo out-of-sample period in determining optimal model weights and investment positions. This framework thus helps to guard against problems with spurious in-sample results and poor tracking performance that can plague single-model clones.

The paper is organized as follows. Section 2 describes the data on index and factor returns. Section 3 outlines the replication procedures for our model combination approach and the competing methods. Section 4 examines the properties of pooled clones, compares their out-of-sample replication performance to that of the other clone portfolios, and demonstrates the economic value of including hedge fund clones in an asset allocation context. Section 5 concludes.

2 Data

The objective of this paper is to demonstrate the advantages of following a model combination approach to constructing clones for hedge fund index portfolios. Linear cloning strategies in general are designed to mimic the out-of-sample properties of a target index by investing in a set of factor portfolios according to estimated co-movement between the index and factors during a prior in-sample period. Section 2.1 discusses the index return data used in our empirical tests, and Section

[3]Nonlinearities in hedge fund returns can arise from dynamics in factor exposures caused by, for example, high-frequency trading. Alternatively funds can realize nonlinearities in their return distributions by taking positions in securities with option-like payoffs. Mitchell and Pulvino (2001) and Jurek and Stafford (2014), for example, provide evidence that returns for hedge fund indexes resemble equity index put writing strategies. Fung and Hsieh (2001) and Agarwal and Naik (2004) also construct nonlinear factors that are widely used in the literature on benchmarking hedge fund performance. Most replication approaches do not directly include these types of factors given concerns over liquidity.

2.2 describes the factors used to build our clone strategies.

2.1 Hedge fund index returns

We collect monthly returns data for ten Dow Jones Credit Suisse indexes over the period January 1994 to October 2014.[4] In our empirical results, we evaluate the performance of the various cloning strategies in replicating the returns of the broad Credit Suisse Hedge Fund Index and individual indexes for the following nine categories: convertible arbitrage, dedicated short bias, emerging markets, equity market neutral, event driven, fixed income arbitrage, global macro, long/short equity, and managed futures. These indexes are constructed from the Credit Suisse Hedge Fund Database, which currently covers approximately 9,000 individual funds. Each index is asset weighted, rebalanced monthly, and reflects the net-of-fee performance of the underlying funds. Constituent funds are also required to have a minimum of \$50 million in assets under management, a 12-month history of prior performance, and audited financial statements. For many of our applications, we convert the index returns to excess returns by subtracting the monthly risk-free rate. We obtain data on the risk-free rate from Kenneth French's website.[5]

Table I reports summary statistics for the excess returns for each of the ten hedge fund indexes. The broad Credit Suisse Hedge Fund Index earns an average excess return of 0.48% per month with a standard deviation of only 2.06% over the full sample period. This index also has a correlation coefficient of 0.59 with the Russell 1000 Index, suggesting a reasonably high level of common variation between hedge funds and large-cap U.S. stocks. There are considerable differences in performance across the remaining nine hedge fund categories. For example, the Dedicated Short Bias Index earns the lowest average excess return of −0.60% per month with the highest standard deviation at 4.71%. As expected, however, the correlation between short bias funds and large-cap equities is strongly negative (correlation coefficient of −0.79). The Global Macro Index has the highest average excess return at 0.68% per month. Most of the indexes exhibit excess kurtosis, and several show meaningful correlation with the Russell 1000 Index.

Overall, the variation in index performance seen in Table I as well as in the strategies of the constituent hedge funds should provide us with a rich environment to examine the relative performance of the replication methods outlined in Section 3.

[4] The data and index construction details are available at `www.hedgeindex.com`.

[5] See `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/`. We thank Kenneth French for making this data available.

2.2 Factor returns

In building clones for each of the Credit Suisse indexes, we start with a diverse set of 14 factors and attempt to construct portfolios of these factors that closely match the historical returns of the given target. Given that hedge fund managers have considerable flexibility in terms of trading strategies, asset class exposure, and the choice of markets in which to operate, we specify a wide range of candidate factors to capture common risk exposures across the hedge fund sector. Following Hasanhodzic and Lo (2007), the factors are returns for index portfolios.[6] It is also important to note that cloning often requires short positions in the underlying factors, considerable portfolio turnover, and, in practical applications, some consideration of scalability. As such, we select a set of factors with returns that could be realized via liquid investments in securities such as futures contracts or exchange traded funds (ETFs).

We organize the factors according to the following five broadly defined categories: domestic equity, international equity, domestic fixed income, international fixed income, and commodities. We obtain the time series of returns for each factor from Bloomberg unless otherwise noted. For the majority of our cloning applications, we convert all factor returns to excess returns by subtracting the risk-free rate. Table II reports summary statistics for the factor excess returns.

The first two domestic equity market factors are the Russell 1000 Index return and the Russell 2000 Index return, designed to capture potential exposure to large-cap and small-cap stocks, respectively. We also include the FTSE NAREIT All Equity REITs Index. As shown in Panel A of Table II, the Russell 1000 Index earns an average excess return of 0.62% per month over the January 1994 to October 2014 sample period. This figure is slightly higher than the excess return of 0.48% earned by the Credit Suisse Hedge Fund Index but also comes with noticeably more volatility (i.e., 4.38% for the large-cap equity factor versus 2.06% for the Hedge Fund Index).

We use two international equity factors. The Morgan Stanley Capital International (MSCI) World ex USA Index and the MSCI Emerging Markets Index account for fund exposure to developed and emerging equity markets, respectively. Panel B of Table II shows that each of these factors has reasonably high correlation with the Russell 1000 Index and slightly underperforms large-cap U.S. stocks in terms of average return and standard deviation.

Panels C and D provide summary statistics for the domestic and international fixed income factors, respectively. For the U.S. market, we use the Bank of America (BofA) Merrill Lynch 7-10

[6] Bollen and Fisher (2013) follow an alternative approach and use a set of futures contracts as the basis for their clone portfolios.

Year U.S. Treasury Index, the BofA Merrill Lynch 7-10 Year BBB U.S. Corporate Index, and the BofA Merrill Lynch U.S. Mortgage Backed Securities Index. The international factors include the returns for the BofA Merrill Lynch 7-10 Year Global Government Index and the BofA Merrill Lynch U.S. Dollar Emerging Markets Sovereign Index. We also incorporate the trade weighted U.S. Dollar Index return in this group to capture fund exposure to foreign exchange markets. The U.S. Dollar Index series is from the Federal Reserve Bank of St. Louis website.[7]

Finally, the three commodities factors are summarized in Panel E of Table II. We construct returns from the S&P GSCI Precious Metals Index, the S&P GSCI Grains Index, and the S&P GSCI Crude Oil Index. The crude oil factor earns the highest average return across all factors at 0.92% per month but also has the highest monthly standard deviation at 9.28%.

3 Replication methods

Our eventual objective is to propose a method for constructing hedge fund clone portfolios that optimally combines investment recommendations from several individual replication models. Much of the existing literature (e.g., Hasanhodzic and Lo (2007), Amenc, Gehin, Martellini, and Meyfredi (2008), Amenc, Martellini, Meyfredi, and Ziemann (2010), and Bollen and Fisher (2013)) has focused on linear clones constructed from individual factor models. In this setup, a target fund's factor loadings are iteratively estimated over rolling windows, and these estimated coefficients are used as the basis for allocation weights in subsequent periods.[8] Sections 3.1 and 3.2 review these modeling approaches for building linear factor clones. The resulting individual linear models are the building blocks for our model pooling approach, which is subsequently described in Section 3.3.

3.1 Linear factor models: Pre-specified factors

For a given index, a linear clone is constructed by estimating a regression model using a time series of prior index and factor returns and applying the resulting coefficient estimates to determine investment positions in the underlying factors. We follow Hasanhodzic and Lo (2007) and update

[7]See `http://research.stlouisfed.org/fred2/`.

[8]Amenc, Martellini, Meyfredi, and Ziemann (2010) examine non-linear replication models designed to capture dynamics in fund factor exposures. They estimate conditional models for hedge fund indexes using both Markov regime switching methods and a Kalman filter. They find that these non-linear approaches deliver little, if any, improvement in replication performance over linear clones. Amin and Kat (2003), Kat and Palaro (2005), and Kat and Palaro (2006) follow an alternative approach and propose distribution-based replication strategies that attempt to match properties of the historical return distribution of the target index rather than the time series of index returns.

the clone investment positions each month using a rolling 24-month window of data. This empirical design ensures that the resulting clones are not affected by look-ahead bias in estimating risk exposures and also allows the implied investment strategies to change over time. More formally, let r_t be the excess return for a given hedge fund index at time t and $f_{k,t}$ be the excess return for factor k. We estimate clone positions at time t via the following regression model:

$$r_{t-j} = \sum_{k=1}^{K} \beta_k f_{k,t-j} + \varepsilon_{t-j}, \quad j = 1, \ldots, 24. \tag{1}$$

As described in Section 2.2, the factors are specified as excess returns. The estimated coefficients, $\{\hat{\beta}_k\}_{k=1}^{K}$, can therefore be interpreted as implied weights in the corresponding factors and the associated portfolio positions will sum to zero. The remainder of the replication portfolio can be invested in cash, and the clone return is given by

$$R_t^C = R_{f,t} + \sum_{k=1}^{K} \hat{\beta}_k f_{k,t}, \tag{2}$$

where $R_{f,t}$ is the risk-free rate.

Note that the intercept term in equation (1) is set equal to zero. This is appropriate in the present context as we are interested in tracking the performance of various hedge fund indexes via a passive investment in the factor portfolios. Any intercept term in the linear regression specification would have a connotation of managerial skill and, by definition, can not be replicated.

For each of the ten Credit Suisse index portfolios, we report replication results for two versions of the linear factor clone specified in equations (1) and (2). For each model in each sample month, we estimate clone weights using the prior 24 months of index returns. We then compare the resulting clone returns to those for the underlying index. The first model is termed the "kitchen sink" regression model and includes all 14 of the factors listed in Section 2.2. Although this approach allows a given clone portfolio to gain exposure to a wide range of asset classes and risk factors, this model is also likely to suffer from overfitting (i.e., estimating a large number of parameters with a relatively limited sample of data), high turnover, and poor out-of-sample tracking performance. Our motivation for including the kitchen sink model is primarily to compare its performance with that of the model pooling method which also incorporates the full set of factor returns in modeling index returns.

For the second linear clone, we incorporate a smaller group of five factors designed to closely match the models used by Hasanhodzic and Lo (2007) and Amenc, Martellini, Meyfredi, and Ziemann (2010). We specifically include excess returns for the Russell 1000 Index, the Bank of America (BofA) Merrill Lynch 7-10 Year U.S. Treasury Index, the BofA Merrill Lynch 7-10 Year BBB U.S. Corporate Index, the U.S. Dollar Index, and the S&P GSCI Crude Oil Index.[9] This approach limits the potential asset class exposures relative to the kitchen sink model, but the parsimonious specification is likely to yield more reasonable investment strategies and better performance out of sample. This model is also relatively established in the replication literature and offers another natural benchmark for our pooling method.

We refer to the five-factor specification estimated according to equation (1) as the "unconstrained five-factor model." This functional form parallels the setup for the individual clones in our model pooling specification outlined below. We note, however, that Hasanhodzic and Lo (2007) follow a slightly modified approach to building their clone portfolios. Rather than using excess returns for the target and factor portfolios as in equation (1), they use raw returns directly and constrain the regression coefficients to sum to one:

$$R_{t-j} = \sum_{k=1}^{K} \beta_k F_{k,t-j} + \varepsilon_{t-j}, \quad j = 1, \ldots, 24, \tag{3}$$

subject to $\beta_1 + \cdots + \beta_K = 1$. Given that the factors in equation (3) are also raw returns, we can form an initial "in-sample" clone portfolio as

$$R^*_{t-j} = \sum_{k=1}^{K} \hat{\beta}_k F_{k,t-j}, \quad j = 1, \ldots, 24. \tag{4}$$

Hasanhodzic and Lo (2007) then propose incorporating a renormalization factor, γ_t, such that the in-sample volatilities of the clone portfolio and the underlying index are equal:

$$\gamma_t = \frac{\sqrt{\sum_{j=1}^{24} (R_{t-j} - \bar{R})^2 / 23}}{\sqrt{\sum_{j=1}^{24} (R^*_{t-j} - \bar{R}^*)^2 / 23}}, \tag{5}$$

[9] Hasanhodzic and Lo (2007) model individual hedge fund strategies using returns for the S&P 500 Index, the Lehman Corporate AA Intermediate Bond Index, the spread between the Lehman BAA Corporate Bond Index and the Lehman Treasury Index, the U.S. Dollar Index, and the Goldman Sachs Commodity Index.

where $\bar{R} = \frac{1}{24}\sum_{j=1}^{24} R_{t-j}$ and $\bar{R}^* = \frac{1}{24}\sum_{j=1}^{24} R^*_{t-j}$. The clone return is then given by

$$R_t^C = \delta_t R_{f,t} + \gamma_t \sum_{k=1}^{K} \hat{\beta}_k F_{k,t}, \tag{6}$$

where δ_t represents the clone investment in the risk-free asset and is set such that the overall weights sum to one: $\delta_t = 1 - \gamma_t(\hat{\beta}_1 + \cdots + \hat{\beta}_K)$. The renormalization factor, γ_t, intuitively reflects the amount of leverage required for the clone investment in the risky factor portfolios during period t. Given the popularity of the Hasanhodzic and Lo (2007) replication approach in the literature, we also present results for this model using the same subset of five factors listed above. We refer to this specification as the "constrained five-factor model."

3.2 Linear factor models: Factor selection methods

In practice, the linear five-factor clones should have several advantages over a kitchen sink approach in replicating hedge fund returns. The more parsimonious specification is likely to exhibit superior out-of-sample performance and, in many cases, will generate more stable investment recommendations, resulting in lower portfolio turnover and transactions costs. One obvious drawback is that these models incorporate a limited number of pre-specified factors. As such, these clones may have difficulty reflecting the diverse sets of trading strategies and asset class exposures followed by hedge fund managers. The same set of five component factors is also unlikely to be optimal for each of the ten index strategies considered in this paper.

One approach introduced in the hedge fund literature to tailor factor models to individual funds is to employ a stepwise regression algorithm (e.g., Liang (1999), Fung and Hsieh (2000), Agarwal and Naik (2004), and Bollen and Whaley (2009)). Starting from the initial list of 14 potential factors described in Section 2.2, we estimate stepwise clones as follows. For each index portfolio in each month t, we estimate the model in equation (1) by choosing the optimal subset of factors that best explains the index excess returns in sample over the prior 24 months. We employ a standard stepwise regression process with an entry significance level of 25% and a retaining significance level of 50%. Once the relevant factors and coefficient estimates have been determined, we construct the clone portfolio according to equation (2). Note that for each index, the stepwise model is updated each month so that the identity of the factors and the corresponding clone weights are allowed to change over time.

Stepwise regression techniques are intuitively appealing for allowing a broad set of factors to be included into the clone design for a given index portfolio. As noted by O'Doherty, Savin, and Tiwari (2014), however, stepwise approaches to hedge fund replication can also be problematic. Given the relatively short return histories (i.e., 24 months) used to estimate clone weights, only a small subset of the available factors is typically included in the clone strategy for each period. Moreover, because factors are selected based on in-sample fit, stepwise regression approaches can ultimately lead to poor tracking performance.

3.3 Model combination

The problems with existing approaches to hedge fund replication (e.g., pre-specified factor models and stepwise regression models) are thus largely attributed to model specification error, estimation error in factor loadings given relatively short samples of data, and in-sample overfitting. Many of these issues are further magnified by a desire to incorporate a wide range of potential factors and simultaneously capture any dynamics in index factor exposures over time. We argue that an effective way to mitigate these challenges is to follow a model combination approach.

Our model combination algorithm for hedge fund replication proceeds as follows. We first specify a pool of individual linear clone models, each with a limited number of factors. For our empirical applications, we divide the 14 factors presented in Table II into five separate models, each corresponding to a broad asset class. For example, the three domestic equity factors listed in Panel A are combined to form the "domestic equity three-factor model," the global factors in Panel B form the "international equity two-factor model," and so on. For a given index portfolio, we then estimate clone positions at the beginning of each sample month following equation (1) separately for each of the five models. We subsequently evaluate weighted combinations (pools) of these clone positions across models over a pseudo out-of-sample period to determine a set of model weights that best tracks the underlying hedge fund index. These individual model weights, in combination with the estimated factor loadings for each model, determine the out-of-sample investment recommendation for the pooled clone. The model weights and factor loadings are also allowed to adjust over time to capture any dynamics in the underlying index investment strategy.

Our proposed model pooling method to hedge fund replication has several notable advantages over existing alternatives. First, we are able to incorporate a large number of factors by optimally combining a set of parsimonious linear clones. This framework allows us to capture the exposure of each index to a broad range of markets and asset classes. Second, any pre-specified individual clone will likely lead to some tracking error in capturing index performance out of sample. To the extent that the replication errors are not perfectly correlated across our five constituent models, however, the pooled clone should yield considerable diversification benefits. Finally, because the optimal model weights for our combination approach are based on the pool's ability to replicate index returns over a pseudo out-of-sample period, we implicitly guard against problems arising from spurious in-sample relations.

The optimal pooled clone corresponding to a given index is constructed by combining conditional return densities from the individual constituent models. Section 3.3.1 outlines the estimation of these individual densities, Section 3.3.2 discusses the determination of optimal weights and clone construction, and Section 3.3.3 covers some of the empirical design issues.

3.3.1 Conditional return distributions

The building blocks for the pooled clone portfolios are the individual linear factors models, which are specified in accord with equation (1) using index and factor excess returns

$$r_{t-j} = \sum_{k=1}^{K} \beta_k f_{k,t-j} + \varepsilon_{t-j}, \quad j = 1, \ldots, 24, \quad \varepsilon_{t-j} \sim \text{i.i.d. } N\left(0, \sigma^2\right). \tag{7}$$

The key distinction between the clone model above and equation (1) is the assumed distribution for the error term. The normal assumption is convenient in the present context for deriving some of the analytical results below. The overall modeling approach is flexible enough, however, to accommodate non-normal and autocorrelated disturbances. We also note that the assumed normal distribution for the error term does not imply that the index returns themselves are normally distributed. For example, we could incorporate non-linearities or fat tails into the model via an appropriate choice of factors (e.g., Fung and Hsieh (2001), Agarwal and Naik (2004), and Amenc, Martellini, Meyfredi, and Ziemann (2010)).

To derive the conditional excess return distribution implied by equation (7), it is convenient to rewrite the model in vector notation:

$$\mathcal{R}_{t-1} = \mathcal{F}_{t-1}\mathcal{B} + \mathcal{E}_{t-1}, \tag{8}$$

where $\mathcal{R}_{t-1} = [r_{t-1} \ \ldots \ r_{t-24}]'$ is a 24×1 vector of index excess returns, $\mathcal{F}_{t-1}$ is a $24 \times K$ vector of factor excess returns, $\mathcal{B} = [\beta_1 \ \ldots \ \beta_K]'$ is a $K \times 1$ vector of factor loadings, and $\mathcal{E}_{t-1}$ is a 24×1 vector of disturbance terms. By the earlier assumption on ε_{t-j}, the elements of $\mathcal{E}_{t-1}$ are i.i.d. $N\left(0, \sigma^2\right)$.

For any given index, we use the replication model specified in equation (8) to derive a density for the index excess return in period t conditional on index and factor excess returns through period $t-1$ and the contemporaneous factor returns in period t. We specifically follow a Bayesian approach to obtain these conditional densities for index returns. Assuming that an investor has a standard uninformative prior with respect to the index's factor exposures, $\mathcal{B}$,

$$p(\mathcal{B}, \sigma^2 | \mathcal{R}_{t-1}, \mathcal{F}_{t-1}) \propto \frac{1}{\sigma^2}, \tag{9}$$

the posterior distribution of the parameters of interest is given by

$$\sigma^2 | \mathcal{R}_{t-1}, \mathcal{F}_{t-1} \sim IG(24 - K, S), \tag{10}$$

$$\mathcal{B} | \mathcal{R}_{t-1}, \mathcal{F}_{t-1}, \sigma^2 \sim N\left(\hat{\mathcal{B}}, \sigma^2 (\mathcal{F}'_{t-1} \mathcal{F}_{t-1})^{-1}\right), \tag{11}$$

where $S = (\mathcal{R}_{t-1} - \mathcal{F}_{t-1}\hat{\mathcal{B}})'(\mathcal{R}_{t-1} - \mathcal{F}_{t-1}\hat{\mathcal{B}})$, and $\hat{\mathcal{B}}$ is the matrix of maximum likelihood (ML) parameter estimates.

The marginal posterior probability density function for σ^2 has an inverse gamma (IG) distribution, and the conditional posterior for $\mathcal{B}$ is multivariate normal. Following Zellner (1971), the conditional density for the index excess return in period t, r_t, is in the form of the univariate Student t distribution:

$$p(r_t | \mathcal{R}_{t-1}, \mathcal{F}_{t-1}, \tilde{f}_t) = \frac{\Gamma\left[(\nu + 1)/2\right]}{\Gamma(1/2)\,\Gamma(\nu/2)} \left(\frac{h}{\nu}\right)^{1/2} \left[1 + \frac{h}{\nu}\left(r_t - \hat{\mathcal{B}}'\tilde{f}_t\right)^2\right]^{-(\nu+1)/2}, \tag{12}$$

where $\tilde{f}_t = [f_{1,t} \ \ldots \ f_{K,t}]'$ is a $K \times 1$ vector of factor realizations, Γ denotes the gamma function, $h = g\nu / \left(r_t - \hat{\mathcal{B}}'\tilde{f}_t\right)^2$, $g = 1 - \tilde{f}'_t(\bar{F}'\bar{F} + \tilde{f}'_t\tilde{f}_t)^{-1}\tilde{f}_t$, $\bar{F} = \{\tilde{f}_{t-24} \ \ldots \ \tilde{f}_{t-1}\}$, and the degrees of freedom $\nu = 24 - K$.

For each combination of index portfolio, sample month, and one of the five replication models, we can derive a conditional return density in the form of equation (12) above. Our eventual objective is to then derive an optimal set of index-specific model weights by combining these densities across models and evaluating them at observed index and factor excess returns. We provide the implementation details for construction of the optimal model pools in the following section.

3.3.2 Optimal weights and clone construction

We derive a series of optimal model weights for each index following the model combination methods outlined in Geweke and Amisano (2011) and O'Doherty, Savin, and Tiwari (2014).[10] Specifically, we rely on the log scoring rule to evaluate linear combinations of the conditional densities implied by the various individual factor models and determine optimal model weights. Generally speaking, the log score for a given individual model reflects the degree of agreement between the model-implied conditional return distribution and the realized return distribution. A model that suggests a high probability for the return outcome that is subsequently realized receives a relatively high score. Hedge fund return replication is therefore a natural application for the log score approach, as we are ultimately concerned with the ability of a model to characterize returns on an out-of-sample basis.

As we will outline below, the log score method can also be easily applied to evaluate combinations of models rather than individual models in isolation. This approach allows us to incorporate information from several replication models and determine the optimal model weights for replicating each index. Interestingly, the use of the log score rule often results in several of the models in the pool under consideration receiving positive weights, a desirable feature in replication contexts in which the model space likely does not include the "correct" individual model.

In the present context, the log score for a particular replication model A_i is computed based on the model-implied conditional densities and the time-series realizations of the index and factor excess returns. Given our 24-month rolling window design for estimating factor loadings (see, e.g., equation (8)) and conditional densities (see, e.g., equation (12)), we can define the log score for a given index at the end of period $t = T$ using conditional densities from $t = 25$ to $t = T$ as

$$LS(r_t^o, \tilde{f}_t^o, \mathcal{R}_{t-1}^o, \mathcal{F}_{t-1}^o, A_i) = \sum_{t=25}^{T} \log \left[p(r_t^o | \mathcal{R}_{t-1}^o, \mathcal{F}_{t-1}^o, \tilde{f}_t^o, A_i) \right], \tag{13}$$

where the superscript o denotes observed values of the relevant time series. The intuition for using

[10] Our framework is related to the forecast combination methodology of Bates and Granger (1969). Clemen (1989), Diebold and Lopez (1996), Newbold and Harvey (2001), and Timmermann (2006) provide surveys of this literature. Other applications of this approach include the use of combination forecasts for predicting macroeconomic variables (Stock and Watson (2003, 2004) and Guidolin and Timmermann (2009)) and the equity premium (Rapach, Strauss, and Zhou (2010)). Whereas much of this earlier work focuses on point forecasts, our model pooling framework emphasizes combining predictive densities.

the log score function to evaluate replication models is straightforward: any model that attaches a high probability to the index returns that eventually materialize, $\{r_t^o\}_{t=25}^T$, achieves a high log score value. This approach is particularly relevant for assessing replication models because it rewards model performance during a pseudo out-of-sample period over in-sample fit. The log score can therefore be directly interpreted as a measure of how well a model tracks the performance of the underlying index conditional on the model's factor realizations.

The log score function can also be used to evaluate combinations of models, and we pursue this application below. In particular, we start from an initial pool of five replication models $A_1, \ldots, A_5$ for r_t conditional on $\tilde{f}_t$. We then apply the log scoring rule to evaluate linear pools of model-implied conditional densities of the form

$$\sum_{i=1}^{5} w_i p(r_t^o | \mathcal{R}_{t-1}^o, \mathcal{F}_{t-1}^o, \tilde{f}_t^o, A_i); \quad \sum_{i=1}^{5} w_i = 1; \quad w_i \geq 0 \quad (i = 1, \ldots, 5), \tag{14}$$

where w_i denotes the weight assigned to model A_i in a given pool. The restrictions placed on the model weights guarantee that the linear combination of densities in equation (14) is also a valid density function for all values of the weights and all argument of any individual density. We can then define the log score function for a five-model pool analogously to the single-model case:

$$LS_T(w) = \sum_{t=25}^{T} \log \left[\sum_{i=1}^{5} w_i p(r_t^o | \mathcal{R}_{t-1}^o, \mathcal{F}_{t-1}^o, \tilde{f}_t^o, A_i) \right], \tag{15}$$

where $w = [w_1 \; \ldots \; w_5]'$, $w_i \geq 0$ for $i = 1, \ldots, 5$ and $\sum_{i=1}^5 w_i = 1$. We define an optimal pool as one where the weights are chosen to maximize $LS_T(w)$ subject to the constraints noted above. The optimal pool thus corresponds to $w_T^* = \arg\max_w LS_T(w)$. It is also important to note that, for a given index portfolio, the optimal set of model weights is allowed to vary over time based on the entire prior history of conditional densities.

Once we have determined the optimal model weights $w_T^* = \left[w_{1,T}^* \; \ldots \; w_{5,T}^* \right]'$ for a given hedge fund index at the end of period $t = T$, the clone return for the optimal pool in the subsequent period is given by

$$R_{T+1}^C = R_{f,T+1} + \sum_{i=1}^{5} w_{i,T}^* \left[\sum_{k=1}^{K_i} \hat{\beta}_{k,A_i} f_{k,A_i,T+1} \right], \tag{16}$$

where K_i is the number of factors in replication model A_i and, as before, the factor loadings for each model $\{\hat{\beta}_k\}_{k=1}^{K_i}$ are estimated using the prior 24 months of index and factor excess returns.

Thus, the clone investment positions in each factor are determined by both the individual model weights and the estimated index exposures for each replication model.

3.3.3 Empirical design

As outlined above, our pooled clone portfolios are constructed using model weights and individual factor loadings that are allowed to change over time. This empirical design should allow us to capture some of the dynamics in risk exposures for the individual indexes throughout the sample period. Following the prior literature (e.g, Hasanhodzic and Lo (2007) and Amenc, Martellini, Meyfredi, and Ziemann (2010)), we estimate individual model factor loadings and implied conditional densities using a 24-month rolling window of data.[11] As shown in equation (15), the optimal model weights to apply at the beginning of a given period are estimated using the entire prior history of model-implied conditional densities. In our results below, we require a minimum of 24 months of conditional densities to determine the initial set of model weights for each index.

For our sample of index and factor excess returns extending from January 1994 to October 2014, we are able to calculate the first set of realized conditional densities for each index for January 1996 (i.e., using the prior 24 months of data to estimate factor loadings and model-implied densities according to equation (12)). We can then use the 24 months of conditional density realizations over the period January 1996 to December 1997 to compute the first set of model weights for each index via the maximization of the log score objective function in equation (15). These optimal weights are subsequently used to construct clone returns for January 1998. The model weights to be applied in subsequent periods are updated using an expanding window of model conditional densities, and we ultimately have a time series of pooled clone returns for each index covering the period January 1998 to October 2014. We also note that, in all cases, the model weights and factor loadings used to build our clone portfolios following equation (16) are known ex ante, thus avoiding concerns over look-ahead bias.

As discussed previously, the pooled clone portfolios are constructed from a set of five constituent factor models, each defined based on one of the following five asset classes: domestic equity, international equity, domestic fixed income, international fixed income, and commodities. Our objective in this empirical design is to allow for a wide range of potential risk exposures to capture the diverse set of underlying strategies pursued by hedge fund managers. We also attempt to select factors

[11] The empirical results presented below are robust to using 12, 36, or 48 months of prior data to estimate clone positions in a given month.

for each model with returns that are realizable via liquid investments such as futures contracts or ETFs. Generally speaking, the exact identity of the models to be included in the pool should depend on the relevant application. Amenc, Martellini, Meyfredi, and Ziemann (2010), for example, demonstrate substantial benefits in replication from tailoring factor models to each specific index based on an economic analysis. It is also desirable to include models in the pool with sufficient diversity to realize the benefits from model combination methods.

We have chosen to use a common pool of five models for each index largely to mitigate any concerns over data mining. The estimated model weights, of course, will vary considerably across indexes. For a given index, the weight applied to a particular model reflects both the model's replication performance during the pseudo out-of-sample period as well as that model's diversification benefits in the pool. Given our empirical design, the weights also have a natural interpretation as reflecting a given index's exposure to each of the five broadly defined asset classes.

4 Results

We characterize the performance of the optimal pooled clones in capturing the time-series properties of the underlying indexes over an out-of-sample period. To benchmark our results to the prior literature, we also compare the replication ability of the five-model pool to alternative methods based on pre-specified factors (i.e., the kitchen sink model, the unconstrained five-factor model, and the constrained five-factor model) or factor selection (i.e., the stepwise regression model). As a starting point, Section 4.1 outlines the general properties of the optimal five-model pools. Section 4.2 examines the out-of-sample performance of the various clone models in terms of tracking ability, average returns, and required portfolio turnover. Finally, Section 4.3 demonstrates the economic benefits of including hedge fund clones in the context of an asset allocation exercise.

4.1 Characteristics of optimal pooled clones

Table III reports summary statistics for the optimal five-model pools for each of the Credit Suisse hedge fund indexes. As described in Section 3.3.3, we estimate model weights in order to replicate each index series starting at the beginning of January 1998. These weights are then updated each month using the entire prior history of conditional densities for index returns. Table III shows, for each index, the time-series average weight assigned to each of the five constituent models as well as mean and median number of models included in the optimal pool.

The first set of results in the table correspond to the broad Credit Suisse Hedge Fund Index. None of the individual models appears to dominate the pool, as the largest average weight is assigned to the domestic equity three-factor model at 39%. The international fixed income and international equity models receive average weights of 29% and 27%, respectively. In contrast, the commodity three-factor model is excluded from the optimal pool in all periods. Across the full out-of-sample period, the average number of models included in the pool for the Hedge Fund Index is 2.98, and the median is three.

The averages reported in Table III do, however, mask some of the interesting variation in model weights throughout the sample period. As an example, Figure 1 shows the evolution of model weights for the Hedge Fund Index. The domestic equity model is particularly influential in the early part of the sample period, achieving a weight as high as 74% in 2002. The domestic fixed income model is featured in the optimal pool only prior to 2006. The weight on the international equity model tends to increase in the later half of the sample period, largely at the expense of the domestic equity model.

Further analysis of Table III reveals additional features of the optimal pools. First, a single replication model seems insufficient for describing the majority of the index strategies. Across the ten hedge fund indexes, the median number of models included in the optimal pool is three or higher for eight of the strategies. This result suggests that even strategy-distinct indexes designed to reflect a specific segment of the hedge fund market are exposed to a wide range of asset classes and risk factors. Second, there is some noticeable variation in the number of models needed to characterize the various indexes. The average size of the optimal pool ranges from 1.79 models for the Dedicated Short Bias Index to 4.00 models for the Equity Market Neutral Index.

Third, and perhaps most importantly, the average model weights reported in Table III suggest that the log score approach leads to an economically meaningful characterization of index returns. That is, the model weights for each index seem to reflect an intuitively reasonable set of underlying asset classes. For example, the pool for the Emerging Markets Index assigns an average weight of 86% to the international equity model, and the pool for the Global Macro Index allocates 62% weight on average to the international fixed income model. Similarly, the pools for the Dedicated Short Bias and Long/Short Equity Indexes tend to be dominated by the domestic equity three-factor model. The domestic and international fixed income models also have considerable influence in explaining the returns for convertible arbitrage and fixed income arbitrage hedge funds.

The value of the model combination approach to hedge fund replication ultimately depends on the ability of the pooled clones constructed according to equation (16) to explain returns for the target indexes on an out-of-sample basis. Figure 2 compares the time series of clone and target returns for the Credit Suisse Hedge Fund Index. Panel A shows that the model pooling method for index replication captures much of the time-series variation in returns for the target index. Panel B plots the time series of drawdowns for the target and clone portfolios and suggests that the pooled replicator effectively matches losses in the broad Hedge Fund Index, particularly during the 2008 crisis period. Panel C, however, shows that the clone strategy generally underperforms the target portfolio, leading to a noticeable shortfall in the terminal wealth level. This feature is shared by many of the competing replication approaches, and we discuss reasons for the differences in clone and target returns below. In the following section, we formally compare the performance of pooled replicators for each index to several existing alternatives.

4.2 Comparison of model pooling and single-model alternatives

Section 4.2.1 compares the performance of pooled replicators with that of replicators based on either pre-specified factors or stepwise factor selection in tracking index returns out of sample. Section 4.2.2 examines the ability of each cloning strategy to match the level of average index returns, and Section 4.2.3 considers the difference in portfolio turnover across the replication methods.

4.2.1 Tracking

In addition to its usefulness in selecting weights for the optimal model pools, the log score function can be applied to assess the out-of-sample performance of the various replication approaches outlined in Section 3. That is, we can compare the log score values for the optimal pool with the corresponding values for, say, the kitchen sink replication model for each index portfolio. In this case, the log score for the pooled replicator ($LS_{\mathcal{T}}^{OP}$) is defined using the time-varying, ex ante model weights that are applied in constructing the out-of-sample clone:

$$LS_{\mathcal{T}}^{OP} = \sum_{t=49}^{\mathcal{T}} \log \left[\sum_{i=1}^{5} w_{i,t-1}^{*} p(r_t^o | \mathcal{R}_{t-1}^o, \mathcal{F}_{t-1}^o, \tilde{f}_t^o, A_i) \right]. \tag{17}$$

Note that, as discussed above, the out-of-sample period for clone evaluation extends from January 1998 (i.e., $t = 49$) to October 2014 (i.e., $t = \mathcal{T}$). For each index we also compute log score values for the kitchen sink clone, the unconstrained five-factor clone, and the stepwise regression clone and compare these figures to the log scores for the pooled replicators given by equation (17). For the kitchen sink and unconstrained five-factor replicators, the log score function is based on conditional densities estimated via equation (12) and evaluated at the out-of-sample index return realizations. The log score for the stepwise clone is also based on equation (12), but the identity of the factors is allowed to change each month based on the stepwise regression algorithm.[12]

Table IV reports the log score values for the optimal pool and the three alternative replication strategies for the period January 1998 to October 2014. The relative log score values provide a straightforward indication of the out-of-sample replication performance of each method across the ten indexes. The results suggest that the model combination approach to index replication yields substantial improvements in tracking returns on an out-of-sample basis. For each of the ten Credit Suisse indexes, the log score value for the optimal pool is higher than the corresponding log score achieved by any of the three single-model alternatives. If we remove the model pooling approach from consideration, the highest log score value is achieved by the unconstrained five-factor model in eight cases, the stepwise approach in two cases (i.e., the Emerging Markets and Long/Short Equity Indexes), and the kitchen sink method in zero cases. Given that the kitchen sink model includes the full set of 14 factors, it does not appear that simply incorporating a large number of risk exposures leads to superior replication performance. The tracking ability of the optimal model pool is thus largely driven by its ability to diversify replication errors across the constituent models.

We can also assess the statistical significance of the results in Table IV following the formal statistical approach to comparing log scores from competing models outlined in Giacomini and White (2006), Amisano and Giacomini (2007), and O'Doherty, Savin, and Tiwari (2014). This test is based on a standard difference-in-means approach, using the time-series variability of the difference in monthly log scores for two models to assess statistical significance.[13] Each of the p-values reported in Table IV corresponds to a test of the null hypothesis that the log scores for the optimal pool and the given model are equal against the alternative that the log score for the optimal pool is greater than that for the individual model. The table thus includes results for 30 separate tests. As noted previously, the optimal pooled replicator yields superior tracking performance over the alternative models in all 30 cases. These differences are statistically significant at the 10% (5%)

[12]The constrained five-factor clone does not imply a conditional density in the form of equation (12). We therefore omit this model in the log score analysis.

[13]See Amisano and Giacomini (2007) for further details on implementation as well as the size and power properties of the test.

level in 23 (21) instances.

Following Amenc, Martellini, Meyfredi, and Ziemann (2010), we further examine the tracking performance of the various clone portfolios using root mean square replication errors and out-of-sample correlations between the index and clone strategies. For each index and each of the five replication approaches (i.e., optimal pool, kitchen sink model, constrained five-factor model, unconstrained five-factor model, and stepwise regression), we compute root mean square replication errors as follows:

$$\mathrm{RMSE}_{\mathcal{T}} = \sqrt{\frac{1}{\mathcal{T}-48}\sum_{t=49}^{\mathcal{T}}\left(R_t - R_t^C\right)^2}, \tag{18}$$

where R_t and R_t^C are the index and clone returns, respectively. Panel A of Table V reports the results. For the broad Credit Suisse Hedge Fund Index, the optimal pool exhibits the best tracking performance with a root mean square replication error of 1.37% per month over the full January 1998 to October 2014 out-of-sample period. For the remaining four models, the replication errors range from 1.62% per month for the stepwise approach to 2.01% for the constrained five-factor method. The replication errors vary substantially across the remaining nine hedge fund indexes. For example, the optimal pool yields a root mean square error of only 1.23% for the Fixed Income Arbitrage Index, but the corresponding figure for the Managed Futures Index is 3.44%. More importantly, however, for each index strategy the optimal pool leads to the lowest out-of-sample replication errors among all of the alternative cloning approaches.

Figure 3 provides a more in-depth analysis of the benefits of model combination. Panels A through J of the figure each correspond to one of the Credit Suisse indexes. Each panel shows the difference in cumulative sum of squared replication errors for the four individual replication models (i.e., kitchen sink model, unconstrained five-factor model, constrained five-factor model, stepwise model) relative to the optimal pool. Thus, any positive value shown on the plots indicates that the optimal pool has outperformed the given individual model from the beginning of the out-of-sample period through the associated date. The plots can be used to assess the nature of the benefits to following a pooled replication strategy over any of the alterative models. For example, any gradual increase over time in the plotted values implies that pooling delivers consistently lower replication errors over the relevant period. Upward spikes, in contrast, highlight individual months in which pooled clones generate superior replication performance.

Focusing on the results for the Credit Suisse Hedge Fund Index in Panel A of Figure 3, we see that each of the four cumulative differences in clone performance trends upward over the January 1998 to October 2014 sample period. There are pronounced jumps in all four of the plots, however, in September 2001 (i.e., around the September 11th attacks) and the last quarter of 2008 (i.e., during the recent global financial crisis). Many of these general patterns are reflected in the remaining panels in Figure 3. There is a noticeable upward trend in most of the plots, suggesting that the advantages of pooling are not limited to only a few sample months. Most panels also show spikes during the later part of 2008 and other periods of economic turmoil. Several of the alternative clones for the Emerging Markets Index (Panel D), for example, generate substantial replication errors relative to the model pooling strategy surrounding the Russian debt default in August 1998.

The results in Figure 3 support the building consensus in the literature that the benefits to model combination approaches tend to be most pronounced during periods of economic stress. For example, in modeling the cross section of stock returns, O'Doherty, Savin, and Tiwari (2012) show that the improvements from model pooling relative to the best individual models are largely concentrated in periods of recession as defined by the National Bureau of Economic Research. Geweke and Amisano (2014) further argue that the adverse impact of parameter uncertainty on forecasting is systematically greater in volatile economic times, implying larger gains from model pooling during these periods.

Panel B of Table V provides further evidence on the out-of-sample performance of the various replication strategies by reporting correlations between the index and clone portfolios. The results are again favorable for the optimal pool, with this method achieving the highest out-of-sample correlation for seven of the ten indexes.

4.2.2 Average returns

An obvious concern for investors is the extent to which clone strategies produce returns and Sharpe ratios in accord with the underlying indexes. The direct objective of clone construction, however, is not to "outperform" the target portfolio, but to generate a passive investment strategy that tracks the time-series behavior of the target over time. Given the inherent advantages for clone strategies in terms of liquidity, transparency, scalability, and fees, it seems plausible that investors would demand higher returns from the underlying hedge funds constituting a given index. Perhaps more importantly, Bollen and Fisher (2013) note that it is impossible to capture managerial ability in security selection via a passive replication approach. Thus, any observed underperformance for clones relative to the corresponding indexes can be partially attributed to managerial skill.

The existing literature generally suggests that the performance of clones tends to be inferior to that of their hedge fund counterparts. For example, Hasanhodzic and Lo (2007) develop replication strategies for individual hedge funds and find that the clones yield lower returns and Sharpe ratios. Amenc, Gehin, Martellini, and Meyfredi (2008), Amenc, Martellini, Meyfredi, and Ziemann (2010), Bollen and Fisher (2013), and Jurek and Stafford (2014) consider factor-based replication approaches for hedge fund indexes and reach similar conclusions on the underperformance of clones relative to their targets.

Table VI presents average returns, standard deviations, and minimum drawdowns for the ten Credit Suisse indexes and each of the corresponding clones portfolios. For each clone, the table also reports a wealth relative (W_C/W_I) comparing the growth of one dollar initial investments in the clone and underlying index over the out-of-sample period. The results are broadly consistent with the prior literature. There are subtle differences in performance across the five clone strategies, but the hedge fund indexes reliably outperform the replication products. For example, the clones for the optimal pool, kitchen sink model, and unconstrained five-factor model deliver lower average returns and wealth relatives below one for nine of the ten hedge fund indexes (the exception is the Dedicated Short Bias Index in each case). The constrained five-factor and stepwise clones each outperform in only two categories. Focusing only on the clone results, none of the strategies appears to deliver reliably superior performance. The highest average return is earned by the constrained five-factor clone in seven cases. Note, however, that this approach consistently leads to investment strategies with higher volatility than the corresponding pooled replicators. Across replication strategies, the pooled clones also generate the least severe portfolio drawdowns in nine out of ten cases.

The underperformance of clone portfolios shown in Table VI should not be interpreted as a failure of hedge fund replication for several reasons. First, as noted by Jurek and Stafford (2014), the Credit Suisse indexes are not investable and likely produce an upward biased reflection of hedge fund performance due to backfill and survivorship. Malkiel and Saha (2005), for example, examine the TASS database over the period 1994 to 2003 and find both of these biases to be substantial. Second, a key motivation for passive replication strategies is to reap the diversification benefits of hedge-fund-like returns without the associated fees and lock-up requirements. Finally, if we view the performance differences between indexes and clones as evidence of managerial skill, we should also note that these abnormal returns may be more difficult to realize as funds become capacity constrained (e.g., Naik, Ramadorai, and Stromqvist (2007) and Getmansky, Liang, Schwarz, and Wermers (2010)) and the overall industry becomes more competitive. There is, in fact, some evidence that these trends have already started. Figure 4 shows the difference in average returns between each Credit Suisse index and its corresponding clone portfolios for two subperiods: 1998 to 2006 and 2007 to 2014. The start of the second period is roughly set to correspond with the start of the recent global financial crisis. Panel A shows that the difference in average returns between the target and its corresponding pooled clone is more pronounced in the early period for eight of the ten index portfolios. This improvement in relative performance for the clone strategies also tends to hold across the other replication methods in Panels B to E. Most notably, the constrained five-factor clone (Panel D) outperforms its target index during the 2007 to 2014 period in eight out of ten cases.

4.2.3 Turnover and leverage

In practice, clone investors would also be concerned with the implementation details of the various replication approaches. For example, Bollen and Fisher (2013) find that linear cloning procedures can imply considerable levels of portfolio turnover caused by time-series volatility in the factor loadings estimated via rolling window regressions. Intuitively, turnover tends to be higher with shorter regression windows. There is a tradeoff, however, as longer windows generate more precise estimates of factor loadings but may miss some of the underlying dynamics of the index strategy. We do not present an in-depth analysis of these design issues, but instead focus on the required levels of turnover across the five replication approaches using the 24-month rolling-window regression design.

The results are reported in Figure 5. For each of the passive replication strategies, we compute monthly turnover as 0.5 times the sum of the absolute values of the change in portfolio weights in each underlying asset. The figure presents annual turnover, which is computed by multiplying the average monthly value by 12. Also recall that the unconstrained five-factor clone and constrained five-factor clone are limited to positions in the five underlying factors plus the risk-free asset.[14] In contrast, the optimal pool, kitchen sink, and stepwise clones can generate non-zero portfolio positions in as many as 15 assets (i.e., the 14 factors in Table II plus the risk-free asset) in any given month.

From Figure 5, the highest levels of turnover are associated with the kitchen sink clone in all

[14] See equations (2) and (6), respectively.

ten cases. The required levels of trading are economically large, with turnover ranging from 940% per year for the Fixed Income Arbitrage Index to greater than 3,000% for the Managed Futures Index. The stepwise replication procedure also results in substantial levels of turnover, ranging from 730% to 2,232%. These results are potentially not surprising, as each of these methods attempts to incorporate a large number of factors into the replication process using a relatively short, 24-month estimation window. These empirical designs can lead to considerable instability in prescribed investment positions attributable to estimation error and in-sample overfitting.

Limiting the number of factors can lead to a reduction in clone turnover. Figure 5 show that the unconstrained and constrained five-factor clones require less trading than the kitchen sink and stepwise procedures in all cases. Focusing only on the five-factor clones, the constrained version tends to require slightly less turnover, with values ranging from 184% for the Event Driven Index to 662% for the Managed Futures Index.

More importantly, the optimal pooled clone implies lower turnover than any of the four competing methods for nine of the ten index strategies. The pooled replication approach for the broad Hedge Fund Index has an average annual turnover of 125%, compared to values between 221% and 995% for the four alternatives. These improvements are even more pronounced within several of the strategy-specific indexes. For example, the pooled clone for the Long/Short Equity Index requires portfolio turnover of only 68% per year, whereas the other four models lead to turnover ranging from 251% to 1,205%. These results are surprising given that, like the kitchen sink and stepwise methods, the optimal pool allows for investment in a wide range of factor portfolios. The key difference is that the investment weights for the pooled clone depend not only on estimated factor loadings, but also on the individual model weights determined from an out-of-sample evaluation period. This empirical design leads to more stable investment recommendations and, thus, lower transactions costs.

Another important consideration for replication strategies is the magnitude of leverage implied by the portfolio positions. As noted by Hasanhodzic and Lo (2007), certain investors may have insufficient credit to support high levels of portfolio leverage. Table VII summarizes the levels of net leverage required by each of the clone portfolios. Consider an investment strategy that combines investments in risky assets with positions in cash (i.e., the risk-free asset). Both the risky assets and cash can be held as long or short positions. Net leverage is simply defined as the difference between the portfolio's long position (i.e., weight) in all risky assets minus the portfolio's short position in risky assets divided by 100%.[15] For the constrained five-factor clone, net leverage in a particular month is simply given by the value of γ_t in equation (6). Following equations (2) and (16), the clones for the kitchen sink, unconstrained five-factor, stepwise, and optimal pool strategies are constructed via long positions in cash and long/short investments in the factor portfolios. We aggregate the implied long and short positions in the risky assets each month and compute net leverage for each clone as defined above.

For each clone, Table VII reports the mean, minimum, and maximum net leverage over the January 1998 to October 2014 sample period. The average leverage tends to be similar for the optimal pool, kitchen sink, unconstrained five-factor, and stepwise clones within each target index. Each of these four approaches routinely requires levels of net leverage below one. Leverage levels for the constrained five-factor clones tend to be higher on average. The optimal pool and the five-factor clones also show relatively less volatility in net leverage over time, suggesting that these approaches require comparatively less extreme investment positions and would be much more practical to implement.

4.3 Economic benefits of hedge fund clones

The use of a hedge fund clone developed according to the replication method described in this paper is most likely to be of interest to large institutional investors. In this section, we assess the economic benefits of adding the optimal pooled clone for the broad Hedge Fund Index to a standard institutional portfolio similar to one managed by a large university endowment. As a starting point, we obtain aggregate data on university endowment allocations from the National Association of College and University Business Officers (NACUBO). We then consider the mean-variance optimal portfolio allocation for investors assuming that several traditional asset classes as well as the hedge fund clone are available for investment. Specifically, the assets available for investment include the Russell 1000 Index (RIY), Russell 2000 Index (RTY), U.S. Treasury Index (TB), U.S. Corporate Index (CB), MSCI World ex USA Index ($MXWOU$), MSCI Emerging Markets Index ($MXEF$), and FTSE NAREIT All Equity REITs Index ($FNER$), in addition to the pooled clone for the broad Hedge Fund Index.

Investors' expected utility is defined over the expected return, $E(r_p)$, and variance, σ_p^2, of their

[15]See Ang, Gorovyy, and van Inwegen (2011) for a discussion of various definitions of hedge fund leverage and a list of margin requirements by asset class.

portfolio:

$$E(U) = E(R_p) - \frac{1}{2}\lambda\sigma_p^2, \tag{19}$$

where λ represents the investors' degree of relative risk aversion. We consider investors with low, moderate, and high degrees of risk aversion, corresponding to relative risk aversion (λ) values of 2, 5, and 10, respectively. The ex-post utility maximizing portfolio allocations are determined using data on the various asset classes for the period January 1998 to October 2014.

Panels A and B of Table VIII show the constrained ex-post optimal risky portfolio compositions. Panel A of the table reports the optimal asset allocation when the weights are constrained to be approximately consistent with the empirically observed weights based on the NACUBO data for endowments in excess of one billion dollars.[16] The restrictions are in the form of upper bounds on the weights for the following asset classes: U.S. stocks ($RIY + RTY \leq 15\%$), U.S. bonds ($TB + CB \leq 10\%$), and emerging market stocks ($MXEF \leq 10\%$). All other asset classes are capped at a maximum of 50%. Furthermore, short positions are ruled out. We also report the optimal allocations for cases in which the hedge fund clone is unavailable for investment. As may be seen from Panel A of Table VIII, for investors with moderate to high risk aversion, the constrained optimal allocation to the hedge fund clone equals the maximum permissible value of 50%. Hence, the hedge fund clone plays an economically meaningful role in the asset allocation choices of such investors. The right-most column of the table reports the utility losses, in terms of annualized certainty equivalent rates of return (CER), when the investors are unable to invest in the hedge fund clone. The utility losses for investors with moderate (CER loss of 1.29% per year) and high risk aversion (CER loss of 4.68% per year) are quite substantial. Of course, it should be kept in mind that these results may be specific to the time period examined.

Panel B of the table reports the asset allocation weights that are optimal when the weights are broadly consistent with the traditional 60/40 endowment model. For this case we impose the following restrictions, in the form of upper bounds on the weights: U.S. stocks (60%), U.S. bonds (40%), and emerging market stocks (10%). All other asset classes are again capped at a maximum of 50%, and short positions are ruled out. Panel B of Table VIII shows that for investors with moderate risk aversion, the implied allocation to the hedge fund clone is modest (5.41%). On the other hand, for investors with high risk aversion the allocation to the hedge fund clone (36.96%) continues to be substantial. The corresponding utility loss from excluding the hedge fund clone in

[16]Source: 2013 NACUBO-COMMONFUND Study of Endowments.

the optimal asset mix is nearly 1 percent per year.

In summary, the evidence presented in this section shows that the hedge fund clone can play an economically significant role in the portfolio of a large institutional investor.

5 Conclusion

Demand for hedge fund replication products is likely to see continued growth as investors seek alternative investments that provide the benefits of hedge fund strategies, while avoiding their opacity, illiquidity, and fee structures. In this paper, we propose a novel method for clone construction and evaluation based on pooling or combining several factor models. The motivation for our approach is straightforward. In characterizing hedge fund returns, model specification and estimation errors are likely to be significant concerns given data limitations, the flexibility offered to fund managers, and potential dynamics in their underlying investment strategies. Given this economic environment, combining information from a pool of linear factor models helps to diversify replication errors across models and leads to better out-of-sample performance.

We demonstrate that our pooled clone strategies have superior tracking ability when compared to their single-model counterparts and consistently require less turnover to implement. Although all of the replication products considered tend to underperfrom their target indexes over the 1998 to 2014 sample period, we also present evidence that these performance gaps have narrowed in recent years. These results speak to the attractiveness of pooled replication strategies as additions to investment portfolios. We also note that the pooling approach outlined in this paper has potentially valuable applications in hedging systematic risk inherent in fund strategies and benchmarking managerial performance (e.g., O'Doherty, Savin, and Tiwari (2014)).

References

Agarwal, Vikas, and Narayan Y. Naik, 2004, Risks and portfolio decisions involving hedge funds, *Review of Financial Studies* 17, 63–98.

Amenc, Noel, Walter Gehin, Lionel Martellini, and Jean-Christophe Meyfredi, 2008, Passive hedge fund replication: A critical assessment of existing techniques, *Journal of Alternative Investments* 11, 69–83.

Amenc, Noel, Lionel Martellini, Jean-Christophe Meyfredi, and Volker Ziemann, 2010, Passive hedge fund replication – Beyond the linear case, *European Financial Management* 16, 191–210.

Amin, Gaurav S., and Harry M. Kat, 2003, Hedge fund performance 1990–2000: Do the "money machines" really add value?, *Journal of Financial and Quantitative Analysis* 38, 251–274.

Amisano, Gianni, and Raffaella Giacomini, 2007, Comparing density forecasts via weighted likelihood ratio tests, *Journal of Business and Economic Statistics* 25, 177–190.

Ang, Andrew, Sergiy Gorovyy, and Gregory B. van Inwegen, 2011, Hedge fund leverage, *Journal of Financial Economics* 102, 102–126.

Bates, J.M., and C.W.J. Granger, 1969, The combination of forecasts, *Operational Research Quarterly* 20, 451–468.

Bollen, Nicolas P.B., and Gregg S. Fisher, 2013, Send in the clones? Hedge fund replication using futures contracts, *Journal of Alternative Investments* 16, 80–95.

Bollen, Nicolas P.B., and Robert E. Whaley, 2009, Hedge fund risk dynamics: Implications for performance appraisal, *Journal of Finance* 64, 985–1035.

Clemen, Robert T., 1989, Combining forecasts: A review and annotated bibliography, *International Journal of Forecasting* 5, 559–581.

Diebold, Francis X., and Jose A. Lopez, 1996, Forecast evaluation and combination, in G.S. Maddala, and C.R. Rao, eds.: *Statistical Methods in Finance* (Elsevier Science B.V., Amsterdam), vol. 14, pp. 241–268.

Fung, William, and David A. Hsieh, 2000, Performance characteristics of hedge funds and commodity funds: Natural versus spurious biases, *Journal of Financial and Quantitative Analysis* 35, 291–307.

———, 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

Gelman, Andrew, Jessica Hwang, and Aki Vehtari, 2014, Understanding predictive information criteria for Bayesian models, *Statistics and Computing* 24, 997–1016.

Getmansky, Mila, Bing Liang, Chris Schwarz, and Russ Wermers, 2010, Share restrictions and investor flows in the hedge fund industry, Working paper, University of Massachusetts.

Geweke, John, and Gianni Amisano, 2011, Optimal prediction pools, *Journal of Econometrics* 164, 130–141.

——— , 2014, Analysis of variance for Bayesian inference, *Econometric Reviews* 33, 270–288.

Giacomini, Raffaella, and Halbert White, 2006, Tests of conditional predictive ability, *Econometrica* 74, 1545–1578.

Good, I. J., 1952, Rational decisions, *Journal of the Royal Statistical Society Series B* 14, 107–114.

Guidolin, Massimo, and Allan Timmermann, 2009, Forecasts of U.S. short-term interest rates: A flexible forecast combination approach, *Journal of Econometrics* 150, 297–311.

Hasanhodzic, Jasmina, and Andrew W. Lo, 2007, Can hedge-fund returns be replicated?: The linear case, *Journal of Investment Management* 5, 5–45.

Jurek, Jakub W., and Erik Stafford, 2014, The cost of capital for alternative investments, Forthcoming in *Journal of Finance*.

Kat, Harry M., and Helder P. Palaro, 2005, Who needs hedge funds? A copula-based approach to hedge fund return replication, Working paper, City University.

——— , 2006, Superstars or average Joes? A replication-based performance evaluation of 1917 individual hedge funds, Working paper, City University.

Klein, Roger W., and Stephen J. Brown, 1984, Model selection when there is "minimal" prior information, *Econometrica* 52, 1291–1312.

Liang, Bing, 1999, On the performance of hedge funds, *Financial Analysts Journal* 55, 72–85.

Malkiel, Burton G., and Atanu Saha, 2005, Hedge funds: Risk and return, *Financial Analysts Journal* 61, 80–88.

Mitchell, Mark, and Todd Pulvino, 2001, Characteristics of risk and return in risk arbitrage, *Journal of Finance* 56, 2135–2175.

Naik, Narayan Y., Tarun Ramadorai, and Maria Stromqvist, 2007, Capacity constraints and hedge fund strategy returns, *European Financial Management* 13, 239–256.

Newbold, Paul, and David I. Harvey, 2001, Forecast combination and encompassing, in Michael P. Clements, and David F. Hendry, eds.: *A Companion to Economic Forecasting* (Blackwell, Oxford).

O'Doherty, Michael, N.E. Savin, and Ashish Tiwari, 2012, Modeling the cross section of stock returns: A model pooling approach, *Journal of Financial and Quantitative Analysis* 47, 1331–1360.

——— , 2014, Evaluating hedge fund performance with pooled benchmarks, Forthcoming in *Management Science*.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *Review of Financial Studies* 23, 821–862.

Sharpe, William F., 1992, Asset allocation: Management style and performance measurement, *Journal of Portfolio Management* 18, 7–19.

Stock, James H., and Mark W. Watson, 2003, Forecating output and inflation: The role of asset prices, *Journal of Economic Literature* 41, 788–829.

——— , 2004, Combination forecasts of output growth in a seven-country data set, *Journal of Forecasting* 23, 405–430.

Timmermann, Allan, 2006, Forecast combination, in G. Elliott, C.W.J. Granger, and A. Timmermann, eds.: *Handbook of Economic Forecasting* (North–Holland, Amsterdam), pp. 135–196.

Zellner, Arnold, 1971, *An introduction to Bayesian inference in econometrics*. Wiley, Hoboken, NJ.

Table I: Summary statistics for monthly excess returns of hedge fund indexes
The table reports summary statistics for the ten Credit Suisse hedge fund indexes. The summary statistics include average monthly excess return, standard deviation, skewness, excess kurtosis, and correlation with the excess return for the Russell 1000 Index. The sample period is January 1994 to October 2014.

Index	Mean	St. Dev.	Skewness	Excess Kurtosis	Correl. w/ Russell 1000
Hedge Fund	0.48	2.06	-0.30	2.88	0.59
Convertible Arbitrage	0.37	1.89	-2.64	17.39	0.38
Dedicated Short Bias	-0.60	4.71	0.72	1.59	-0.79
Emerging Markets	0.44	4.07	-0.85	5.94	0.55
Equity Market Neutral	0.22	2.80	-12.50	181.20	0.30
Event Driven	0.53	1.76	-2.24	11.65	0.65
Fixed Income Arbitrage	0.23	1.57	-4.47	32.93	0.34
Global Macro	0.68	2.64	-0.06	4.53	0.24
Long/Short Equity	0.56	2.73	-0.11	3.63	0.70
Managed Futures	0.26	3.31	0.00	-0.01	-0.08

Table II: Summary statistics for factors
The table lists the following summary statistics for factor excess returns: average monthly return, standard deviation, skewness, excess kurtosis, and correlation with the excess return for the Russell 1000 Index. The factors are defined in the text. The sample period is January 1994 to October 2014.

Factor	Index Description	Mean	St. Dev.	Skewness	Excess Kurtosis	Correl. w/ Russell 1000
	Panel A: Domestic equity factors					
RIY	Russell 1000 Index	0.62	4.38	-0.73	1.30	1.00
RTY	Russell 2000 Index	0.65	5.65	-0.51	1.07	0.84
$FNER$	FTSE All Equity REITS	0.48	5.74	-0.65	8.06	0.57
	Panel B: International equity factors					
$MXWOU$	MSCI World Excluding US	0.39	4.77	-0.64	1.64	0.84
$MXEF$	MSCI Emerging Markets	0.46	6.80	-0.72	2.00	0.75
	Panel C: Domestic fixed income factors					
TB	BofA ML 7-10 Year US Treasury	0.29	1.81	0.13	1.29	-0.17
CB	BofA ML 7-10 Year BBB US Corporate	0.36	1.90	-1.64	11.93	0.33
MBS	BofA ML US Mortgage Backed Securities	0.25	0.83	0.02	1.75	0.00
	Panel D: International fixed income factors					
IG	BofA ML 7-10 Year Global Government	0.24	1.11	0.03	0.38	-0.13
EM	BofA ML Emerging Markets Sovereign	0.73	4.23	-2.60	16.42	0.57
USD	Trade Weighted U.S. Dollar	-0.25	2.01	0.36	1.54	-0.35
	Panel E: Commodity factors					
PM	S&P GSCI Precious Metals	0.36	4.95	0.07	1.03	0.08
GR	S&P GSCI Grains	-0.33	6.93	0.21	0.58	0.26
CL	S&P GSCI Crude Oil	0.92	9.28	-0.05	0.89	0.21

Table III: Summary statistics for optimal five-model pools

The table reports summary statistics for the optimal five-model pool corresponding to each of the Credit Suisse hedge fund indexes. The following five models are considered for each pool and are identified based on their associated factors: the domestic equity three-factor model, the international equity two-factor model, the domestic fixed income three-factor model, the international fixed income three-factor model, and the commodity three-factor model. For a given index-month, the log score and optimal weights are computed from conditional densities based on the maximization of the log score objective function using the entire prior history of index returns. The conditional densities for a given month are based on the prior two years of index and factor returns. For each hedge fund index, the table shows the time-series average weight assigned to each model, as well as the mean and median number of models assigned a positive weight in the optimal pool.

	Average model weights					Number of models	
Index	Domestic Equity	International Equity	Domestic Fixed Income	International Fixed Income	Commodity	Mean	Median
Hedge Fund	0.39	0.27	0.04	0.29	0.00	2.98	3.00
Convertible Arbitrage	0.00	0.06	0.56	0.36	0.02	3.00	3.00
Dedicated Short Bias	0.98	0.01	0.00	0.00	0.01	1.79	2.00
Emerging Markets	0.00	0.86	0.03	0.11	0.00	2.92	3.00
Equity Market Neutral	0.49	0.13	0.05	0.20	0.13	4.00	4.00
Event Driven	0.32	0.24	0.27	0.18	0.00	3.37	3.00
Fixed Income Arbitrage	0.13	0.09	0.35	0.43	0.00	3.70	4.00
Global Macro	0.07	0.15	0.08	0.62	0.07	3.38	4.00
Long/Short Equity	0.75	0.25	0.00	0.00	0.00	1.88	2.00
Managed Futures	0.11	0.02	0.33	0.48	0.05	3.19	3.00

Table IV: Comparison of log score values

For each Credit Suisse hedge fund index, the table compares log scores for the replication strategies based on the optimal five-model pool, the kitchen sink regression model, the unconstrained (UC) five-factor regression model, and the stepwise regression model. Each of the approaches to constructing conditional densities for the hedge fund indexes is described in the text. For each of the optimal pools, the log score value in a given month is based on ex ante weights computed using the entire prior history of index conditional densities. For each strategy and each index portfolio, the table reports the cumulative log score value (Log score) over the full January 1998 to October 2014 sample period. For each of the alternative replication approaches to the optimal model pool, the table also reports a p-value corresponding to the test of the null hypothesis that the log score for the optimal pool is equal to the log score for the given model against the alternative that the log score for the optimal pool is greater than the log score for the given model.

		Alternative replication strategies					
	Optimal pool	Kitchen sink model		5-factor model (UC)		Stepwise model	
Index	Log score	Log score	p-value	Log score	p-value	Log score	p-value
Hedge Fund	610.9	576.7	0.007	596.3	0.125	548.1	0.002
Convertible Arbitrage	589.2	538.7	0.000	581.9	0.201	507.0	0.000
Dedicated Short Bias	451.3	376.2	0.000	420.0	0.001	394.1	0.000
Emerging Markets	544.5	475.2	0.000	483.7	0.000	488.4	0.003
Equity Market Neutral	586.0	542.9	0.080	583.8	0.418	535.5	0.003
Event Driven	600.8	546.8	0.000	595.8	0.292	552.9	0.001
Fixed Income Arbitrage	628.9	584.2	0.003	622.7	0.312	542.1	0.000
Global Macro	534.4	479.7	0.000	525.6	0.244	491.1	0.011
Long/Short Equity	585.1	555.0	0.009	547.6	0.001	559.5	0.062
Managed Futures	404.4	332.6	0.000	396.2	0.122	345.1	0.000

Table V: Comparison of root mean square replication errors and out-of-sample correlation coefficients

For each Credit Suisse hedge fund index, the table compares the out-of-sample performance of index replication strategies based on the optimal five-model pool, the kitchen sink regression model, the unconstrained (UC) five-factor regression model, the constrained (C) five-factor model, and the stepwise regression model. Each of the approaches to building clones for the hedge fund indexes is described in the text. Panel A reports root mean square replication errors in percentage per month, and Panel B reports out-of-sample correlation coefficients between index returns and clone returns. The out-of-sample period is January 1998 to October 2014.

	Replication strategy				
Index	Optimal pool	Kitchen sink	5-factor (UC)	5-factor (C)	Stepwise
Panel A: Root mean square replication errors					
Hedge Fund	1.37	1.89	1.69	2.01	1.62
Convertible Arbitrage	1.51	1.96	1.66	1.93	1.84
Dedicated Short Bias	2.79	4.89	3.38	4.15	3.60
Emerging Markets	1.94	3.03	3.23	4.06	2.59
Equity Market Neutral	3.03	4.26	3.57	4.29	3.87
Event Driven	1.30	1.97	1.38	1.66	1.48
Fixed Income Arbitrage	1.23	1.96	1.48	1.61	1.68
Global Macro	2.23	3.28	2.52	3.05	2.55
Long/Short Equity	1.51	2.39	2.24	2.66	1.81
Managed Futures	3.44	4.93	3.64	4.33	4.01
Panel B: Out-of-sample correlation between index and clone					
Hedge Fund	0.72	0.70	0.68	0.61	0.69
Convertible Arbitrage	0.73	0.68	0.74	0.71	0.65
Dedicated Short Bias	0.83	0.56	0.73	0.69	0.73
Emerging Markets	0.85	0.75	0.73	0.72	0.77
Equity Market Neutral	0.25	-0.02	0.19	0.03	0.07
Event Driven	0.74	0.53	0.73	0.65	0.67
Fixed Income Arbitrage	0.69	0.49	0.68	0.68	0.59
Global Macro	0.34	0.42	0.37	0.27	0.38
Long/Short Equity	0.86	0.73	0.69	0.65	0.80
Managed Futures	0.25	0.17	0.25	0.29	0.26

Table VI: Comparison of index and clone performance measures

The table compares the performance of each of the Credit Suisse hedge fund indexes to several corresponding out-of-sample replication strategies. The replication strategies are based on the optimal five-model pool, the kitchen sink regression model, the unconstrained (UC) five-factor regression model, the constrained (C) five-factor model, and the stepwise regression model. Each of these approaches to building clones for the hedge fund indexes is described in the text. For each index portfolio and each cloning strategy, the table reports the average monthly return (μ), standard deviation (σ), and minimum portfolio drawdown (DD) over the out-of-sample period, January 1998 to October 2014. The figures in parentheses are p-values for tests that the mean return for a given clone and the corresponding index are equal. For each clone portfolio, the table also reports a wealth relative (W_C/W_I) comparing the growth of a \$1 initial investment in the clone portfolio to the growth of a \$1 initial investment in the corresponding index over the out-of-sample period.

					Replication strategy							
	Index				Optimal pool				Kitchen sink			
Index	μ (p-value)	σ	DD	$\frac{W_C}{W_I}$	μ (p-value)	σ	DD	$\frac{W_C}{W_I}$	μ (p-value)	σ	DD	$\frac{W_C}{W_I}$
Hedge Fund	0.57 (n/a)	1.90	-19.7	n/a	0.26 (0.00)	1.57	-19.3	0.54	0.22 (0.01)	2.60	-32.4	0.48
Convertible Arbitrage	0.55 (n/a)	2.04	-32.9	n/a	0.27 (0.01)	1.96	-30.1	0.57	0.45 (0.49)	2.64	-38.4	0.80
Dedicated Short Bias	-0.49 (n/a)	4.86	-76.3	n/a	-0.18 (0.12)	4.64	-57.1	1.89	-0.32 (0.63)	5.47	-73.7	1.30
Emerging Markets	0.56 (n/a)	3.70	-39.2	n/a	0.34 (0.11)	3.29	-32.9	0.67	0.18 (0.07)	4.57	-46.8	0.43
Equity Market Neutral	0.36 (n/a)	3.09	-45.1	n/a	0.11 (0.24)	1.22	-23.9	0.67	-0.18 (0.07)	2.83	-59.6	0.35
Event Driven	0.65 (n/a)	1.86	-19.1	n/a	0.31 (0.00)	1.33	-17.2	0.51	0.38 (0.05)	2.13	-32.6	0.57
Fixed Income Arbitrage	0.38 (n/a)	1.68	-29.0	n/a	0.26 (0.17)	1.36	-19.6	0.79	0.37 (0.93)	2.13	-22.5	0.96
Global Macro	0.73 (n/a)	2.16	-26.8	n/a	0.28 (0.00)	1.54	-13.7	0.42	0.38 (0.14)	3.53	-35.6	0.46
Long/Short Equity	0.72 (n/a)	2.81	-22.0	n/a	0.36 (0.00)	2.76	-27.8	0.48	0.22 (0.00)	3.36	-40.8	0.35
Managed Futures	0.50 (n/a)	3.31	-17.4	n/a	0.06 (0.07)	2.03	-25.3	0.45	0.09 (0.25)	4.27	-50.1	0.41

	Replication strategy											
	5-factor (UC)				5-factor (C)				Stepwise			
Index	μ (p-value)	σ	DD	$\frac{W_C}{W_I}$	μ (p-value)	σ	DD	$\frac{W_C}{W_I}$	μ (p-value)	σ	DD	$\frac{W_C}{W_I}$
Hedge Fund	0.27 (0.01)	2.18	-25.4	0.54	0.43 (0.30)	2.49	-26.3	0.72	0.41 (0.17)	2.16	-22.1	0.72
Convertible Arbitrage	0.29 (0.03)	2.39	-36.9	0.59	0.37 (0.19)	2.73	-38.1	0.67	0.37 (0.17)	2.29	-31.5	0.69
Dedicated Short Bias	-0.18 (0.20)	4.25	-57.8	1.97	0.21 (0.02)	5.39	-49.8	3.89	-0.07 (0.10)	4.97	-58.7	2.30
Emerging Markets	0.20 (0.11)	4.70	-52.7	0.42	0.27 (0.31)	5.82	-65.1	0.40	0.33 (0.21)	3.94	-43.2	0.62
Equity Market Neutral	0.01 (0.16)	2.45	-40.6	0.52	0.29 (0.82)	3.09	-33.6	0.90	-0.15 (0.06)	2.53	-59.5	0.38
Event Driven	0.33 (0.00)	1.79	-22.8	0.53	0.42 (0.05)	2.07	-25.7	0.62	0.30 (0.00)	1.69	-20.1	0.49
Fixed Income Arbitrage	0.28 (0.34)	1.95	-27.0	0.81	0.31 (0.55)	2.18	-28.3	0.85	0.42 (0.77)	1.99	-26.1	1.06
Global Macro	0.30 (0.02)	2.29	-21.5	0.42	0.58 (0.51)	2.84	-21.2	0.73	0.52 (0.25)	2.39	-19.8	0.66
Long/Short Equity	0.42 (0.05)	2.86	-36.4	0.54	0.52 (0.27)	3.41	-37.6	0.63	0.44 (0.02)	2.87	-32.9	0.56
Managed Futures	0.20 (0.24)	2.56	-28.2	0.57	0.55 (0.85)	3.91	-35.5	1.07	0.12 (0.19)	3.26	-44.0	0.47

Table VII: Cumulative portfolio weight in risky assets

For each Credit Suisse hedge fund index, the table summarizes the implied net leverage over the out-of sample period (January 1998 to October 2014) for each of the following index replication strategies: the optimal five-model pool, the kitchen sink regression model, the unconstrained (UC) five-factor regression model, the constrained (C) five-factor model, and the stepwise regression model. For a portfolio that invests in risky assets and cash, net leverage is defined as the difference between the weights in the long and short positions in the risky assets divided by 100%. Each of the approaches to building clones for the hedge fund indexes is described in the text. For each index portfolio and each cloning strategy, the table shows the average net leverage, as well as the minimum and maximum leverage across all sample months.

	Replication strategy									
	Optimal pool		Kitchen sink		5-factor (UC)		5-factor (C)		Stepwise	
Index	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Hedge Fund	0.21	[-0.20, 2.02]	0.20	[-2.82, 2.94]	0.35	[-0.99, 1.86]	1.08	[0.74, 1.69]	0.33	[-0.99, 3.54]
Convertible Arbitrage	0.17	[-0.71, 1.59]	0.07	[-2.13, 2.24]	0.25	[-0.58, 1.15]	1.10	[0.51, 1.70]	0.22	[-1.95, 1.56]
Dedicated Short Bias	-0.83	[-1.35, -0.37]	-1.54	[-6.16, 4.01]	-1.04	[-4.62, 0.68]	1.19	[0.93, 1.64]	-1.12	[-5.32, 2.94]
Emerging Markets	0.40	[0.05, 0.99]	0.98	[-3.17, 7.81]	0.61	[-1.49, 3.34]	1.19	[0.99, 1.60]	0.89	[-0.55, 7.40]
Equity Market Neutral	-0.21	[-2.13, 0.48]	-1.09	[-12.38, 1.74]	-0.63	[-5.89, 0.41]	0.96	[0.51, 1.95]	-0.61	[-6.08, 1.54]
Event Driven	0.17	[-0.23, 0.77]	0.13	[-2.27, 2.72]	0.12	[-1.01, 0.75]	1.01	[0.66, 1.29]	0.25	[-1.21, 2.94]
Fixed Income Arbitrage	0.12	[-0.47, 0.64]	0.52	[-1.00, 2.19]	0.19	[-0.72, 0.80]	0.95	[0.54, 1.38]	0.35	[-0.83, 1.66]
Global Macro	0.47	[-0.93, 3.96]	0.64	[-2.82, 2.97]	0.63	[-1.09, 3.42]	1.30	[0.72, 2.50]	0.81	[-2.08, 3.76]
Long/Short Equity	0.42	[0.10, 0.69]	-0.04	[-4.57, 2.27]	0.29	[-1.23, 2.61]	1.10	[0.94, 1.77]	0.23	[-2.08, 1.69]
Managed Futures	0.18	[-1.45, 1.75]	0.89	[-3.03, 5.83]	0.36	[-1.66, 2.26]	1.63	[1.19, 3.01]	0.61	[-1.92, 3.40]

Table VIII: Optimal asset allocation with hedge fund clone investments

The table reports the ex-post optimal portfolio allocations and the corresponding certainty equivalent rates of return (CER) in percentage per year for mean-variance optimizing investors during the period January 1998 to October 2014. Investors' expected utility is defined over the expected return, $E(R_p)$, and variance, σ_p^2, of their portfolio: $E(U) = E(R_p) - \frac{1}{2}\lambda\sigma_p^2$, where λ represents the investors' degree of relative risk aversion. Portfolio allocations are reported for investors with low, moderate, and high degrees of risk aversion, corresponding to relative risk aversion (λ) values of 2, 5, and 10, respectively. The assets available for investment include the Russell 1000 Index (RIY), Russell 2000 Index (RTY), U.S. Treasury Index (TB), U.S. Corporate Index (CB), MSCI World ex USA Index ($MXWOU$), MSCI Emerging Markets Index ($MXEF$), and FTSE NAREIT All Equity REITs Index ($FNER$), in addition to the pooled clone for the broad Hedge Fund Index (HF Clone). In each panel the allocations to domestic stocks, bonds, and emerging markets equity are capped at the indicated levels. All other individual asset class weights are capped at 0.50. The stock and bond restrictions in Panel A are broadly consistent with the observed average portfolio allocations for large ($\geq$ \$1 billion) college endowments as per the NACUBO statistics. The restrictions in Panel B are consistent with the traditional 60/40 endowment model. The right-most column reports the utility loss in percent per year if investors are forced to exclude the hedge fund clone from their portfolios.

	Portfolio weights								Certainty equivalent returns	
Risk aversion	RIY	RTY	TB	CB	$MXWOU$	$MXEF$	$FNER$	HF Clone	CER (%)	CER Difference (%)
Panel A: Institutional asset-class restrictions – (i) $RIY + RTY \leq 0.15$; (ii) $TB + CB \leq 0.10$; (iii) $MXEF \leq 0.10$										
Including clone investments										
2	0.000	0.150	0.100	0.000	0.340	0.100	0.310	0.000	5.68	0.00
5	0.150	0.000	0.100	0.000	0.009	0.100	0.142	0.500	3.55	1.29
10	0.150	0.000	0.100	0.000	0.134	0.000	0.116	0.500	1.60	4.68
Excluding clone investments										
2	0.000	0.150	0.100	0.000	0.338	0.100	0.312	n/a	5.68	n/a
5	0.150	0.000	0.100	0.000	0.393	0.100	0.257	n/a	2.26	n/a
10	0.150	0.000	0.100	0.000	0.500	0.000	0.250	n/a	-3.08	n/a
Panel B: Traditional asset-class restrictions – (i) $RIY + RTY \leq 0.60$; (ii) $TB + CB \leq 0.40$; (iii) $MXEF \leq 0.10$										
Including clone investments										
2	0.000	0.482	0.400	0.000	0.000	0.100	0.018	0.000	6.85	0.00
5	0.290	0.092	0.400	0.000	0.000	0.100	0.063	0.054	5.24	0.01
10	0.163	0.014	0.400	0.000	0.000	0.034	0.019	0.370	4.02	0.97
Excluding clone investments										
2	0.000	0.482	0.400	0.000	0.000	0.100	0.018	n/a	6.85	n/a
5	0.341	0.086	0.400	0.000	0.000	0.100	0.073	n/a	5.23	n/a
10	0.498	0.000	0.400	0.000	0.000	0.021	0.081	n/a	3.05	n/a

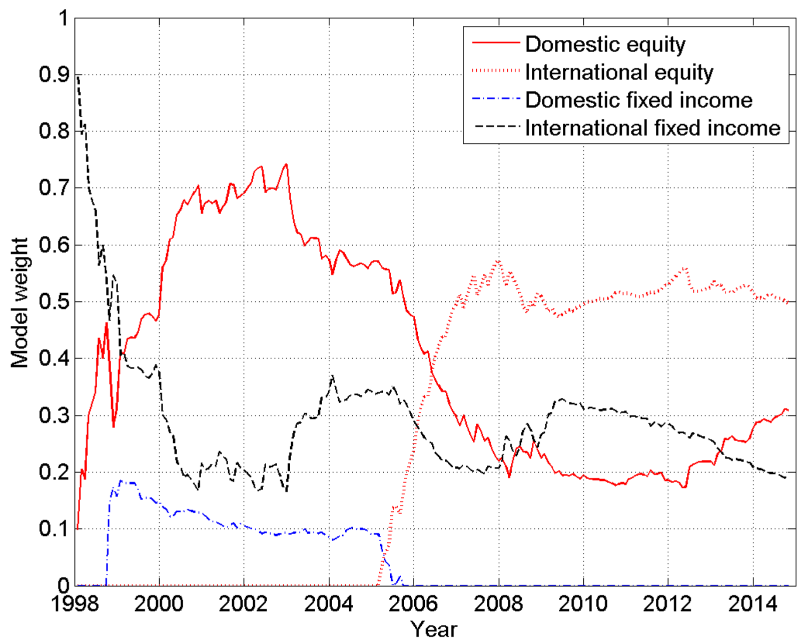

Figure 1: Model weights in the optimal five-model pool
The figure shows the time-series of model weights in the optimal five-model pool for the broad Credit Suisse Hedge Fund Index. The following five models are considered for the pool and are identified based on their associated factors: the domestic equity three-factor model, the international equity two-factor model, the domestic fixed income three-factor model, the international fixed income three-factor model, and the commodity three-factor model. For a given month, the log score and optimal weights are computed from conditional densities based on the maximization of the log score objective function using the entire prior history of index returns. The conditional densities for a given month are based on the prior two years of index and factor returns. The out-of-sample period is January 1998 to October 2014. The three-factor commodity model is assigned a weight of zero throughout the period shown in the plot.

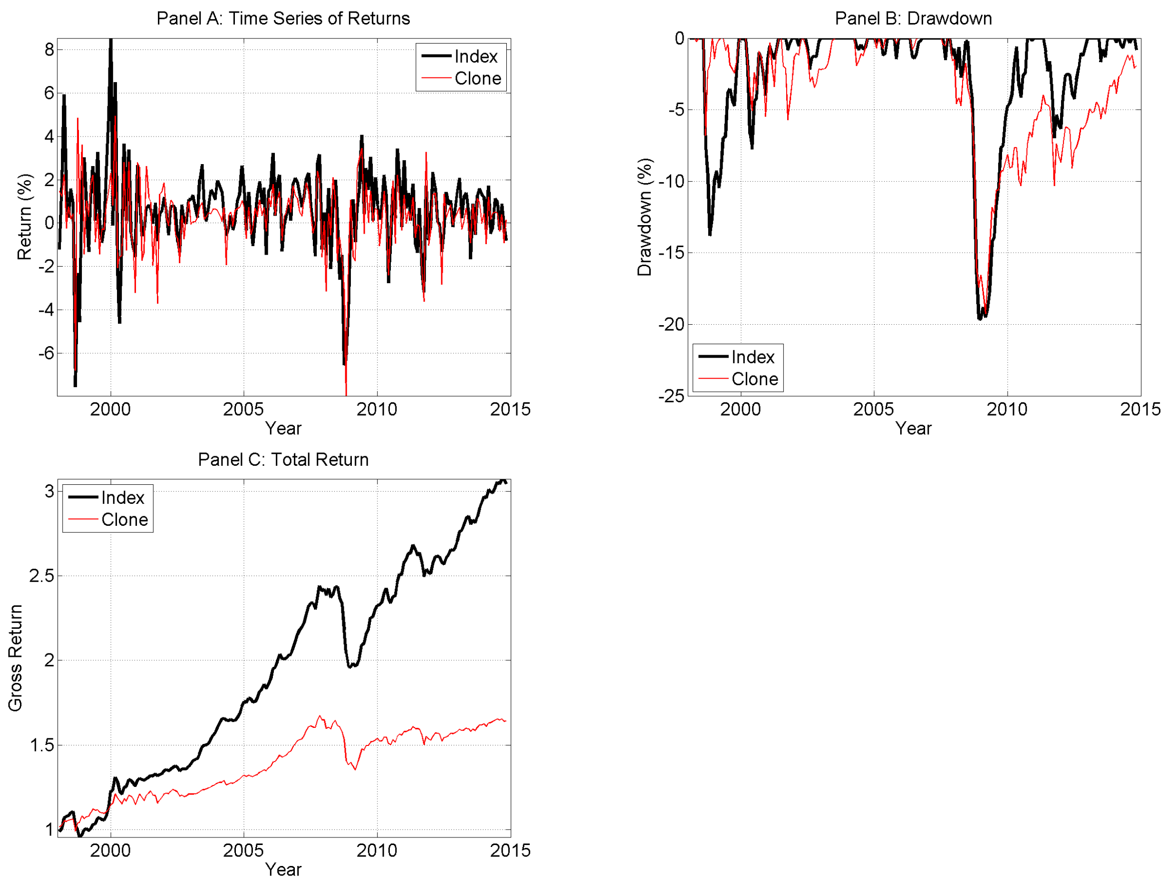

Figure 2: Comparison of index and clone returns

The figure compares the performance of the broad Credit Suisse Hedge Fund Index and the corresponding replication strategy based on the optimal five-model pool. Panel A shows the time series of index and clone returns. Panel B shows the monthly portfolio drawdown for the index and clone portfolios. Panel C shows the compounded gross returns representing the growth of $1 invested in each strategy. The following five models are considered for the pool and are identified based on their associated factors: the domestic equity three-factor model, the international equity two-factor model, the domestic fixed income three-factor model, the international fixed income three-factor model, and the commodity three-factor model. For a given month, the optimal weights are computed from conditional densities based on the maximization of the log score objective function using the entire prior history of index returns. The conditional densities for a given month are based on the prior two years of index and factor returns. The out-of-sample period is January 1998 to October 2014.

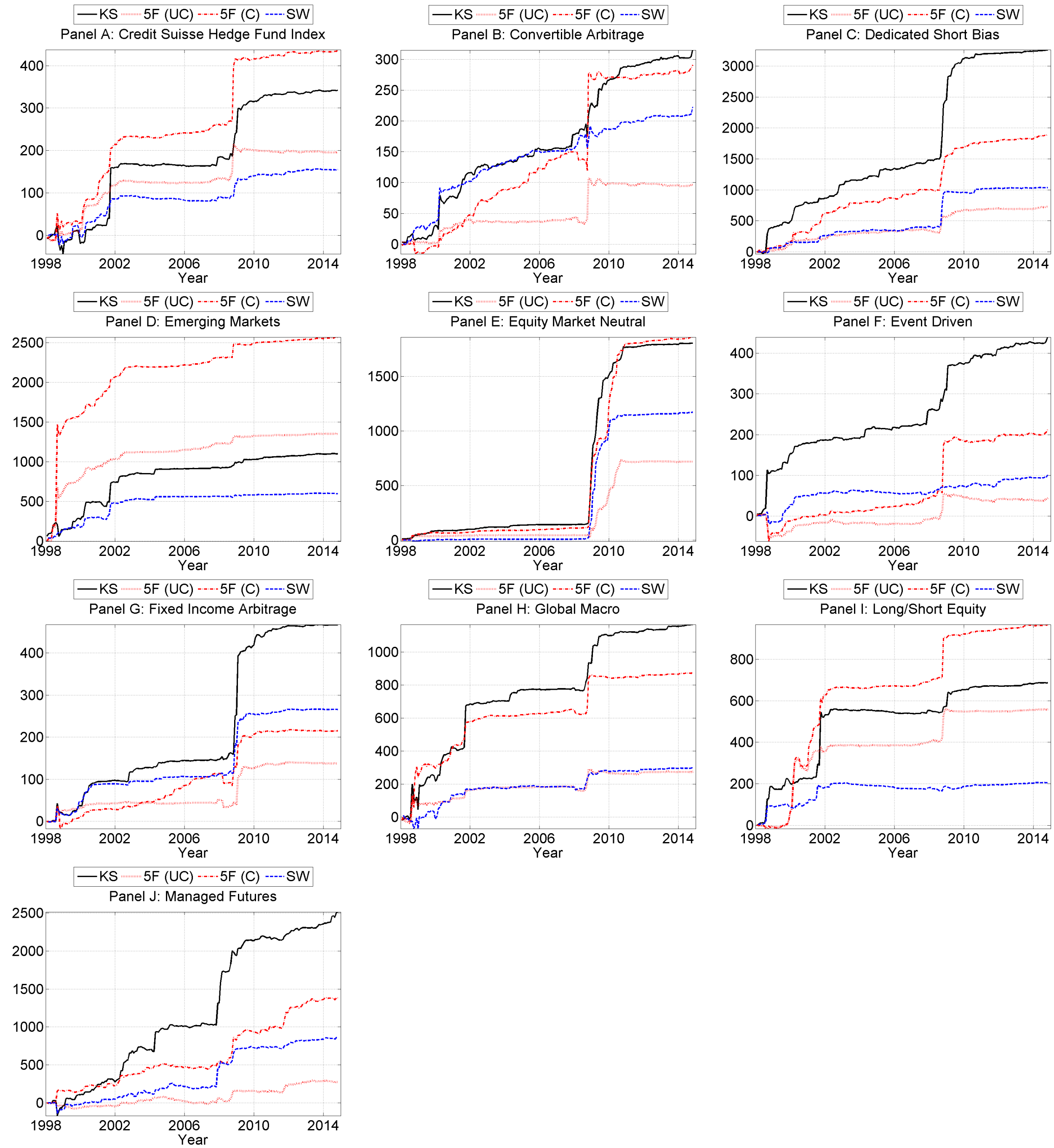

Figure 3: Difference in cumulative sum of squared replication errors

The figures shows the difference in cumulative sum of squared replication errors for the optimal five-model pool relative to each of the following four alternative replication strategies: the kitchen sink regression model (KS), the unconstrained five-factor regression model (5F (UC)), the constrained five-factor model (5F (C)), and the stepwise regression model (SW). Each of the approaches to building clones for the hedge fund indexes is described in the text. Each panel corresponds to one of the Credit Suisse hedge fund indexes. The out-of-sample period is January 1998 to October 2014.

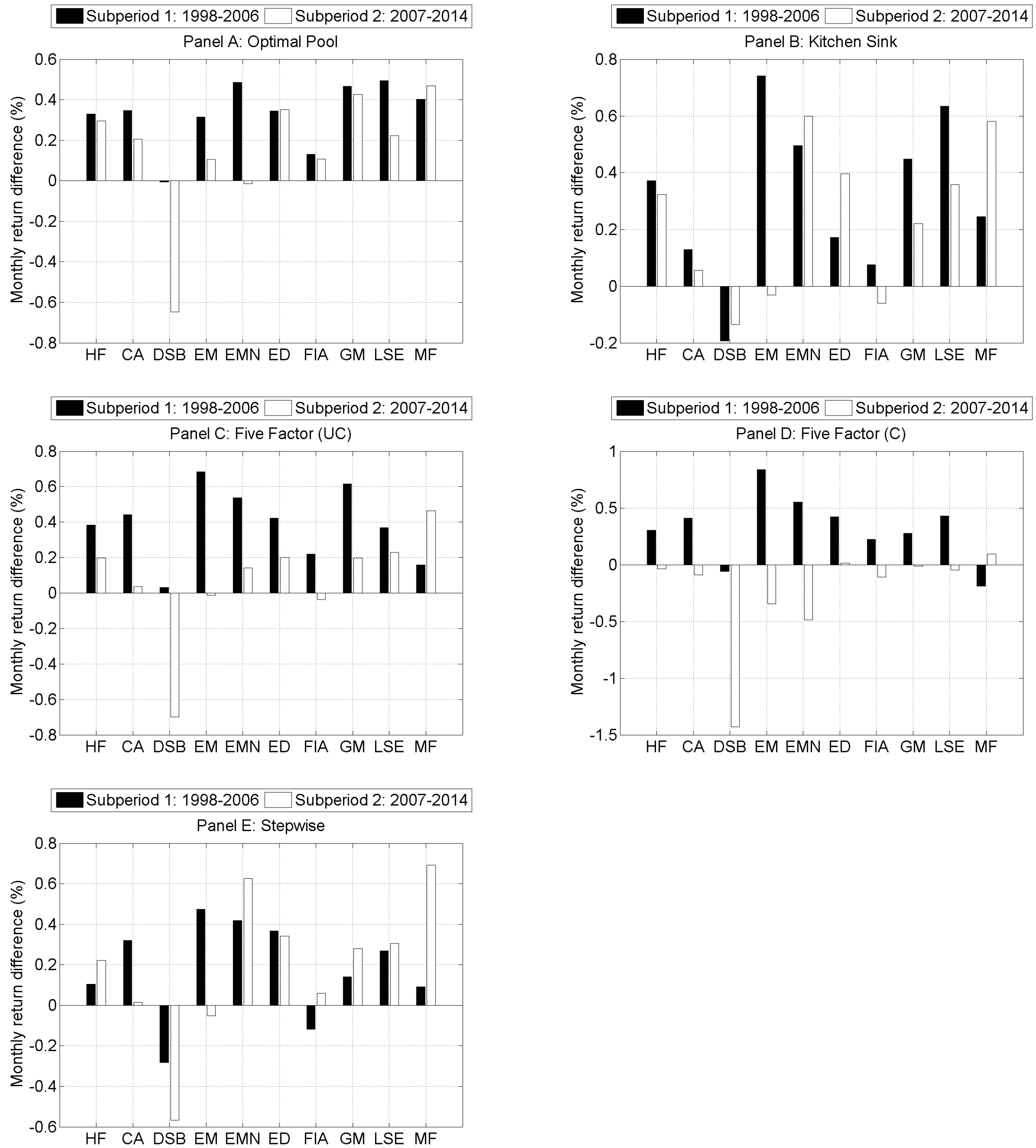

Figure 4: Comparison of differences in average returns by subperiod
For each replication strategy and each of the Credit Suisse hedge fund indexes, the figure shows the average monthly difference between index and clone returns for the following two subperiods: (1) 1998–2006 and (2) 2007–2014. Each panel corresponds to one of the following replication strategies: the optimal five-model pool (Panel A), the kitchen sink regression model (Panel B), the unconstrained (UC) five-factor regression model (Panel C), the constrained (C) five-factor model (Panel D), and the stepwise regression model (Panel E). Each of these approaches to building clones for the hedge fund indexes is described in the text. In each panel, the indexes considered are the broad Hedge Fund Index (HF), the Convertible Arbitrage Index (CA), the Dedicated Short Bias Index (DSB), the Emerging Markets Index (EM), the Equity Market Neutral Index (EMN), the Event Driven Index (ED), the Fixed Income Arbitrage Index (FIA), the Global Macro Index (GM), the Long/Short Equity Index (LSE), and the Managed Futures Index (MF).

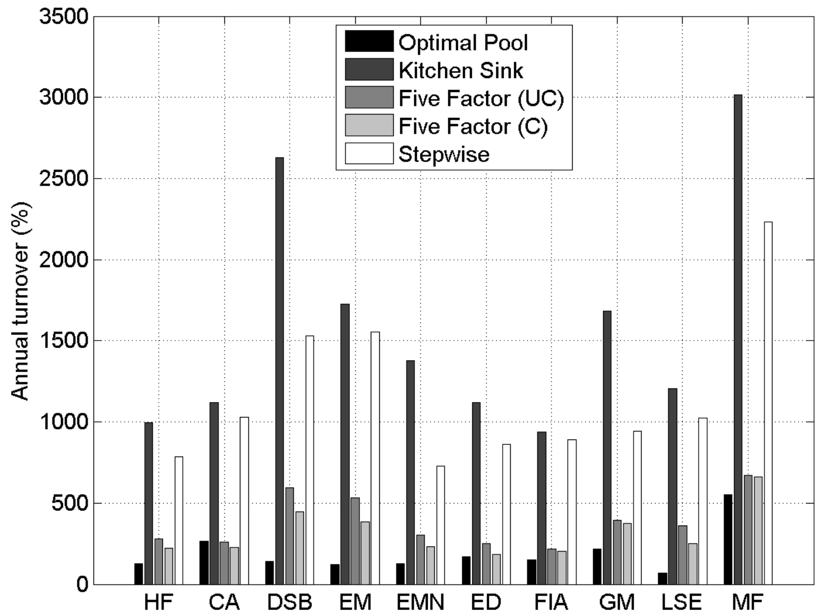

Figure 5: Comparison of annual turnover
For each of the Credit Suisse hedge fund indexes, the figure shows the average annual portfolio turnover for each of the following replication strategies: the optimal five-model pool, the kitchen sink regression model, the unconstrained (UC) five-factor regression model, the constrained (C) five-factor model, and the stepwise regression model. The indexes considered are the broad Hedge Fund Index (HF), the Convertible Arbitrage Index (CA), the Dedicated Short Bias Index (DSB), the Emerging Markets Index (EM), the Equity Market Neutral Index (EMN), the Event Driven Index (ED), the Fixed Income Arbitrage Index (FIA), the Global Macro Index (GM), the Long/Short Equity Index (LSE), and the Managed Futures Index (MF). Monthly portfolio turnover is computed as 0.5 times the sum of the absolute values of the change in portfolio weights in each underlying asset. The annual turnover figures are computed by multiplying the monthly turnover values by 12. The sample period is January 1998 to October 2014.