Review of Finance Advance Access published July 31, 2016

Review of Finance, 2016, 1–38
doi: 10.1093/rof/rfw037

OXFORD

Hedge Fund Replication: A Model Combination Approach*

Michael S. O'Doherty[1], N. E. Savin[2], and Ashish Tiwari[2]

[1]University of Missouri and [2]University of Iowa

Abstract

Recent years have seen increased demand from institutional investors for passive replication products that track the performance of hedge fund strategies using liquid investable assets such as futures contracts. In practice, linear replication methods suffer from poor tracking performance and high turnover. We propose a model combination approach to index replication that pools information from a diverse set of pre-specified factor models. Compared with existing methods, 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.

JEL classification: C11, C51, C53, C58, G17

Keywords: Hedge Funds, Model Pooling, Model Combination, Hedge Fund Replication, Log Score

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 two plus twenty 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 twenty-four hedge funds and six hedge funds-of-funds.[1]

* We thank Scott Cederburg, Ryan Flugum, Amit Goyal (the editor), Paul Weller, Tong Yao, an anonymous referee, and seminar participants at Erasmus University–Rotterdam, the University of Iowa, the University of Nebraska, and VU University–Amsterdam for valuable comments and suggestions. We also credit Nicolas Bollen for helpful discussions on the modeling approach implemented in this article. Michael O'Doherty acknowledges financial support from the University of Missouri Research Board.

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

As a result, there is interest in 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 factor portfolios 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, 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, existing literature (e.g., Hasanhodzic and Lo, 2007; Amenc *et al.*, 2010; Bollen and Fisher, 2013) largely finds that factor-based replication products for individual hedge funds and hedge fund indexes tend to underperform 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. Whereas 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 segment of the investment universe. The factors in these models are returns on liquid futures contracts. To capture the wide range of potential investment strategies available to hedge fund managers, we consider factor models that span the following four market segments: domestic equity and fixed income, international equity, foreign currencies, and commodities including precious metals. The weights assigned to the four individual models are based on the log score criterion, which assesses the performance of various convex combinations of the individual models in

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).

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.

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 with alternative techniques. We also incorporate realistic leverage constraints and transaction costs into our clone designs.

Before proceeding, it is important to characterize the diverse set of strategies employed by hedge fund managers and consider what aspects of those strategies any factor-based replication procedure can hope to capture. The target investments for our clones are ten Dow Jones Credit Suisse hedge fund indexes. These indexes include the broad Credit Suisse Hedge Fund Index and individual indexes for the following nine categories: long/short equity, dedicated short bias, equity market neutral, event driven, convertible arbitrage, fixed income arbitrage, global macro, emerging markets, and managed futures. Pedersen (2015) provides a detailed summary of these strategies and classifies them into three groups: arbitrage, equity, and macro. Arbitrage funds (e.g., event driven, convertible arbitrage, and fixed income arbitrage) essentially make relative-value bets in related securities. As arbitrage profits are by definition idiosyncratic, we anticipate that linear clones will have limited value in replicating the returns of these indexes. Equity funds (e.g., long/short equity, dedicated short bias, and equity market neutral) engage in either discretionary trading of stocks based on fundamental analysis or the execution of quantitative, rules-based trading algorithms. Ex ante, the discretionary and quantitative features of these funds seem difficult to replicate given that the building blocks for the clone strategies are generic futures contracts. On the other hand, long/short and short bias funds typically maintain non-zero market exposures, so clones may find some success in tracking time-series patterns in these indexes. Similarly, macro-funds (e.g., global macro, emerging markets, and managed futures) make timely bets on equity markets, interest rates, currencies, and commodities using statistical models or insights on global markets and national economies. Ex ante, replication products appear most promising in this space of the hedge fund universe, as cloning methods may be able to successfully infer funds' asset allocation weights from their historical returns.

Based on the discussion above, we classify the nine strategy indexes into two groups: directional strategies (i.e., long/short equity, dedicated short bias, and the macro-indexes) and non-directional/arbitrage strategies (i.e., equity market neutral and the arbitrage indexes). Distinguishing between these types of investments allows us to better focus parts of our analysis on the directional strategies for which replication is more likely to offer economic benefits.

In our empirical tests, we compare the performance of our model combination approach to several existing alternatives in replicating the ten Dow Jones Credit Suisse hedge fund indexes over the period January 1994–December 2015. 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 available futures contracts. In contrast, our model combination method bases model weights and, ultimately, investment recommendations on replication performance over a pseudo out-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 with 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 with 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 and standard deviations for the 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 with the underlying hedge fund indexes over our sample period. These results are broadly consistent with managerial ability in the hedge fund industry that is difficult to match through passive, rules-based investment approaches. We note, however, that the target hedge fund indexes are not directly investable. Thus, despite their underperformance, the clones still offer the diversification benefits of hedge-fund-like returns, along with scalability, liquidity, and transparency. We also show that the pooled clones perform well in matching the out-of-sample volatilities of their targets, particularly among the directional indexes.

We examine the practical implementation details of the replication approaches in terms of required portfolio turnover and leverage. We find that pooled clones require reliably less trading activity when compared with the alternatives. The gross 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 issue, we analyze the ex-post mean-variance optimal asset allocations for such investors during the period January 1998–December 2015. We find that the implied optimal allocation to the hedge fund clone is often substantial. Furthermore, the utility costs of not being able to invest in the clone, expressed in terms of annualized certainty equivalent rates of return (CER), 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 non-linearities inherent in hedge fund returns.[3] The model combination clones directly alleviate these issues by diversifying

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

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 article is organized as follows. Section 2 describes the data on index returns and the returns on futures contracts used as factors in the replication models. 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 article 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-movements between the index and factors during a prior in-sample period. Section 2.1 discusses the index returns data used in our empirical tests, and Section 2.2 describes data on futures contract returns used as factors 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–December 2015.[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 nine strategy subcategories. 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 our empirical work, we convert the index returns to excess returns by subtracting the monthly London Interbank Offered Rate (LIBOR) rate. We obtain data on the LIBOR rate from the Federal Reserve Bank of St. Louis 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 0.43% per month with a standard deviation of only 2.02% over the full sample period. This index also has a correlation coefficient of 0.58 with the Chicago Mercantile Exchange (CME) S&P 500 Index futures contract, suggesting a reasonably high level of common variation between hedge funds and large-cap US stocks. There are considerable differences in performance

Jurek and Stafford (2015), 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.

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

5 See https://research.stlouisfed.org/fred2/.

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 return for the CME S&P 500 Index futures contract. The sample period is January 1994–December 2015.

Index	Mean	Standard deviation	Skewness	Excess kurtosis	Correlation with S&P 500 Index
Panel A: Credit Suisse Hedge Fund Index					
Hedge Fund	0.43	2.02	−0.31	3.03	0.58
Panel B: Directional strategies					
Long/Short Equity	0.52	2.67	−0.12	3.83	0.67
Dedicated Short Bias	−0.58	4.68	0.69	1.51	−0.76
Global Macro	0.62	2.60	−0.07	4.66	0.25
Emerging Markets	0.40	3.99	−0.86	6.24	0.53
Managed Futures	0.25	3.34	0.02	−0.09	−0.07
Panel C: Non-directional/arbitrage strategies					
Equity Market Neutral	0.18	2.74	−12.70	188.90	0.29
Event Driven	0.45	1.76	−2.12	10.70	0.63
Convertible Arbitrage	0.32	1.87	−2.65	17.95	0.36
Fixed Income Arbitrage	0.19	1.54	−4.60	35.09	0.33

across the remaining nine hedge fund categories. For example, the Dedicated Short Bias Index earns the lowest average excess return of −0.58% per month with the highest standard deviation at 4.68%. As expected, however, the correlation between short bias funds and large-cap equities is strongly negative (correlation coefficient of −0.76). The Global Macro Index has the highest average excess return at 0.62% per month. Most of the indexes exhibit excess kurtosis, and several show meaningful correlation with the S&P 500 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.

2.2 Factor Returns

In building clones for the Credit Suisse indexes, we start with a diverse set of twelve factors. 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 include a wide range of candidate factors to capture common risk exposures across the hedge fund sector. 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 follow Bollen and Fisher (2013) and specify the factors as returns for liquid futures contracts.

We organize the factors according to the following four broadly defined categories: domestic equity and fixed income, international, currencies, and commodities. We obtain a continuous time series of monthly prices for each factor from the Stevens Continuous Futures database.[6] The price series for each factor is generated from holding the front

6 See https://www.quandl.com/.

Table II. Summary statistics for factors

The table reports summary statistics for the twelve futures contracts that are used as factors in constructing the hedge fund clones. The summary statistics include average monthly excess return, standard deviation, skewness, excess kurtosis, and correlation with the return for the CME S&P 500 Index futures contract. The sample period is January 1994–December 2015.

Futures contract	Exchange	Mean	Standard deviation	Skewness	Excess kurtosis	Correlation with S&P 500 Index
Panel A: Domestic factors						
S&P 500 Index	CME	0.65	4.31	−0.66	1.16	1.00
S&P 400 Midcap Index	CME	0.90	4.94	−0.64	2.11	0.90
10-year US Treasury Note	CBOT	0.06	1.82	−0.42	2.75	−0.15
Panel B: International factors						
FTSE 100 Index	LIFFE	0.31	4.08	−0.57	0.63	0.82
Nikkei 225	CME	0.21	5.99	−0.19	0.47	0.59
US Dollar Index	ICE	0.03	2.34	0.32	0.68	−0.24
Panel C: Currency factors						
British Pound GBP	CME	0.03	2.37	−0.32	1.59	0.19
Swiss Franc CHF	CME	0.20	3.14	0.27	1.18	0.02
Japanese Yen JPY	CME	0.02	3.19	0.72	3.46	−0.04
Panel D: Commodity factors						
Gold	NYMEX	0.49	4.69	0.18	1.27	0.01
Corn	CBOT	0.39	8.14	0.05	0.35	0.21
WTI Crude Oil	NYMEX	0.79	9.21	−0.01	1.11	0.19

contract and rolling to the back contract on the first day of the delivery month of the expiring contract. These data series are based on a calendar-weighted method to smooth the price gaps between consecutive contracts using a weighted average of the front and back contracts during a four-day roll window as detailed by the data vendor. Following Bollen and Fisher (2013), we assume that the individual futures positions are fully collateralized such that the monthly return on factor k in excess of the risk-free rate is given by

$$f_{k,t} = \frac{F_{k,t} - F_{k,t-1}}{F_{k,t-1}}, \tag{1}$$

where $F_{k,t}$ is the futures price at the end of month t.[7] Table II reports summary statistics for the factor excess returns.

7 Briefly, assume that an investor allocates X_{t-1} of total margin capital to finance a position in a single futures contract with price $F_{k,t-1}$ and earns the risk-free rate, $R_{f,t}$, on the margin capital. The total return earned on the investment is

$$\frac{X_{t-1}(1+R_{f,t})+F_{k,t}-F_{k,t-1}-X_{t-1}}{X_{t-1}} = \frac{F_{k,t}-F_{k,t-1}}{X_{t-1}} + R_{f,t},$$

and the excess return is

$$f_{k,t} = \frac{F_{k,t}-F_{k,t-1}}{X_{t-1}}.$$

For a fully collateralized position with $X_{t-1} = F_{k,t-1}$, the equation above simplifies to Equation (1).

The first two factors in Panel A, the CME S&P 500 Index contract return and the CME S&P Midcap 400 Index contract return, are included to capture potential exposure to the domestic equity market. The Chicago Board of Trade (CBOT) 10-year Treasury note contract is also included in this group to reflect target exposure to interest rate movements. As shown in Panel A of Table II, the S&P 500 Index contract earns an average excess return of 0.65% per month over the January 1994–December 2015 sample period. This figure is slightly higher than the excess return of 0.43% earned by the Credit Suisse Hedge Fund Index but also comes with noticeably more volatility (i.e., 4.31% for the large-cap equity factor versus 2.02% for the Hedge Fund Index). The S&P 400 Midcap Index earns the highest average return across all factors at 0.90% per month.

We use three international factors. The London International Financial Futures and Options Exchange (LIFFE) Financial Times Stock Exchange (FTSE) 100 Index contract and CME Nikkei 225 contract account for fund exposure to UK and Japanese equities, respectively.[8] We also incorporate the Intercontinental Exchange (ICE) US Dollar Index contract return in this group to capture fund exposure to foreign exchange markets. Panel B of Table II shows that the FTSE 100 and Nikkei 225 contracts have reasonably high correlations with the S&P 500 Index but underperform large-cap US stocks in terms of average returns.

Panel C provides summary statistics for the three currency factors. We use the contracts for the British Pound, Swiss Franc, and Japanese Yen traded on the CME. Each of these factors exhibits low average return and modest correlation with the S&P 500 Index.

Finally, the three commodity factors are summarized in Panel D of Table II. We construct returns from the New York Mercantile Exchange (NYMEX) Gold contract, the CBOT Corn contract, and the NYMEX West Texas Intermediate (WTI) Crude Oil contract. The crude oil factor earns a relatively high average return of 0.79% per month but also has the highest monthly standard deviation across all factors at 9.21%.

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. Existing literature (e.g., Hasanhodzic and Lo, 2007; Amenc *et al.*, 2008; Amenc *et al.*, 2010; 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.[9] Sections 3.1 and 3.2 review these modeling approaches for

8 The FTSE 100 Index contract is traded on the LIFFE and denominated in British Pounds. We follow the approach in Koijen *et al.* (2015) to compute excess returns for this contract in US dollars as

$$f_{k,t} = \frac{F_{k,t}-F_{k,t-1}}{F_{k,t-1}} + \frac{e_t-e_{t-1}}{e_{t-1}}\frac{F_{k,t}-F_{k,t-1}}{F_{k,t-1}},$$

where e_t is the exchange rate at time t measured in USD per GBP. The exchange rate data are from the Federal Reserve Bank of St. Louis website.

9 Amenc *et al.* (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. Giamouridis and Paterlini (2010) use Least Absolute Shrinkage and Selection Operator (LASSO) regression methods and find

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 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.

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:

$$r_{t-j} = \sum_{k=1}^{K} \beta_k f_{k,t-j} + \epsilon_{t-j}, \quad j = 1, \ldots, 24. \tag{2}$$

As described in Section 2.2, the factors are specified as excess returns on liquid futures contracts. The estimated coefficients, $\{\hat{\beta}_k\}_{k=1}^{K}$, can be interpreted as implied weights in the corresponding notional futures positions. We assume that each of the individual futures positions is fully collateralized, such that a dollar of cash is set aside in an interest bearing account earning LIBOR for every dollar of notional exposure. A positive (negative) value for $\hat{\beta}_k$ corresponds to a long (short) position in factor k. The total proportion of clone capital that is allocated to futures positions is, thus, given by the quantity

$$\gamma_t = \sum_{k=1}^{K} |\hat{\beta}_k|. \tag{3}$$

Following the convention in Bollen and Fisher (2013), if this quantity is less than 100%, we assume that the remaining clone capital can be invested at LIBOR. If γ_t exceeds 100%, we assume that the excess capital is borrowed at LIBOR plus 100 basis points per year. Formally, the monthly clone return is given by

$$R_t^C = \begin{cases} R_{f,t} + \sum_{k=1}^{K} \hat{\beta}_k f_{k,t} & \text{if } \gamma_t \leq 1 \\ R_{f,t} + \sum_{k=1}^{K} \hat{\beta}_k f_{k,t} + \frac{0.01}{12}(1-\gamma_t) & \text{if } \gamma_t > 1, \end{cases} \tag{4}$$

where $R_{f,t}$ in the one-month LIBOR rate.

We make two additional adjustments to the clone returns to ensure that they provide a reasonable representation of investment performance for these strategies. First, we incorporate realistic constraints on leverage. Following Ang, Gorovyy, and van Inwegen (2011),

that this approach offers advantages in replicating hedge fund index returns. 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. Payne and Tresl (2015) use a genetic algorithm technique that combines factor-based and distribution-based methods to replicate hedge fund returns.

we define gross leverage for the clone portfolio as the proportion of capital deployed as notional futures exposure (i.e., γ_t). In constructing each of our clone series, we restrict maximum gross leverage to 4×. That is, if the quantity in Equation (3) exceeds four in a given period, we adjust the clone position in each factor k to be $\frac{4}{\gamma_t} \times \hat{\beta}_k$, such that the adjusted gross leverage is equal to four. Under the Federal Reserve Board's Regulation T, which governs the provision of credit by securities brokers and dealers, the maximum amount an investor may borrow is restricted to 50% of the value of the security, implying 2× leverage. Sophisticated investors like hedge funds are able to achieve higher leverage multiples through portfolio margining arrangements or the use of offshore prime broker facilities. However, we would expect that the typical institutional investor who chooses to forego direct hedge fund investments in favor of a passive replication strategy would also avoid extreme leverage. By limiting gross leverage, we are able to account for the implicit leverage constraint faced by such an investor.[10]

Second, we estimate transaction costs and adjust the clone returns to reflect performance net of these costs. In general, direct transaction costs in futures markets tend to be quite low. Although the costs vary by contract and transaction size, in our analysis we employ cost estimates provided by the CME Group in a recent report.[11] The report assesses the total cost of transacting in E-mini S&P 500 futures contracts for a hypothetical mid-sized institutional investor transacting through a broker. According to the analysis, the commission and market impact costs together amount to 1.50 basis points (bps). We view this estimate as a representative figure and, accordingly, use a round-trip transaction cost of 3 bps in assessing the performance of our clone strategies.[12] Implementing a hedge fund replication program would likely incur additional fixed costs in the form of salaries of the traders and other administrative personnel involved with the program. We do not explicitly account for these fixed costs in our analysis as they are likely to vary considerably across institutions. Accordingly, the performance of the clone strategies may be interpreted as a best-case scenario.

Note that the intercept term in Equation (2) 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, cannot be replicated. A relevant concern with this specification is that if the true conditional expectation function for the hedge fund index excess return has a non-zero intercept, then forcing the intercept to zero in the estimation ("regression through the origin") would result in biased estimates of betas. This, in turn, may adversely affect the ability of the clone to track the target index return. We also note that this concern is mitigated to some degree by the model combination approach to hedge fund replication described below. Specifically, our replication methodology penalizes models that do a poor job of tracking the target hedge fund index returns out of sample. As we subsequently show, the model

10 The article's conclusions are qualitatively similar if we either restrict gross leverage to 2× or eliminate the upper bound on the leverage allowed.

11 The report titled "The Big Picture: A Cost Comparison of Futures and ETFs, 2nd edition" is available at http://www.cmegroup.com/trading/equity-index/files/a-cost-comparison-of-futures-and-etfs.pdf.

12 We specifically compute monthly turnover as 0.5 times the sum of the absolute values of the change in portfolio weights in each underlying futures contract. We then subtract (0.03% × turnover) from the clone return in Equation (4) to compute net monthly performance.

combination approach compares favorably with the competing factor-based replication techniques when judged by the ability to track the target hedge fund returns out of sample.

For each of the ten Credit Suisse index portfolios, we report replication results for two versions of the linear factor clone specified in Equations (2–4). 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 twelve 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 model used by Bollen and Fisher (2013). We specifically include returns for the S&P 500 Index, 10-year T-Note, US Dollar Index, gold, and crude oil contracts.[13] 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 well 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 (2) as the "unconstrained five-factor model". This functional form parallels the setup for the individual clones in our model pooling specification outlined below. Another approach in the literature (e.g., Hasanhodzic and Lo, 2007) is to construct clone portfolios that match their targets in terms of in-sample return volatility. In this case, we proceed as before and first estimate the model in Equation (2). We can then form an "in-sample" series of excess returns for the clone portfolio as

$$r^*_{t-j} = \sum_{k=1}^{K} \hat{\beta}_k f_{k,t-j}, \quad j = 1, \ldots, 24. \tag{5}$$

In the spirit of Hasanhodzic and Lo (2007), we incorporate a renormalization factor, δ_t, such that the in-sample volatilities of the clone portfolio and the underlying index are equal:

$$\delta_t = \frac{\sqrt{\sum_{j=1}^{24} (r_{t-j} - \bar{r})^2 / 23}}{\sqrt{\sum_{j=1}^{24} (r^*_{t-j} - \bar{r}^*)^2 / 23}}, \tag{6}$$

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 weight in factor k is then given by $\delta_t \hat{\beta}_k$. The renormalization factor, δ_t, intuitively reflects an adjustment to clone investment

13 Hasanhodzic and Lo (2007) and Amenc *et al.* (2010) also consider replication models with a limited number of factors. These papers, however, use index returns rather than futures contract returns as the building blocks for their clone strategies.

positions designed to better match the volatility properties of the target portfolio. Given the popularity of the Hasanhodzic and Lo (2007) renormalization 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 "renormalized 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 transaction 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 article.

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; Bollen and Whaley, 2009). Starting from the initial list of twelve 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 (2) 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 10% and a retaining significance level of 25%. Once the relevant factors and coefficient estimates have been determined, we construct the clone portfolio according to Equations (2–4) and incorporate adjustments for leverage constraints and transaction costs. 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 (2016), 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 positions, 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 set of individual models, each with a limited number of factors. For our empirical applications, we divide the twelve factors presented in Table II into four separate models, each corresponding to a broad segment of the investment universe. For example, the three equity and fixed income factors listed in Panel A are combined to form the "domestic three-factor model", the global factors in Panel B form the "international three-factor model", and so on. For a given index portfolio, we then estimate clone positions at the beginning of each sample month following Equation (2) separately for each of the four models. We subsequently evaluate weighted combinations (pools) of these clone positions across models over a pseudo out-of-sample period to determine an optimal 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 four constituent models, however, the pooled clone should yield error 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.a outlines the estimation of these individual densities, Section 3.3.b discusses the determination of optimal weights and clone construction, and Section 3.3.c covers some of the empirical design issues.

3.3.a Conditional return distributions

The building blocks for the pooled clone portfolios are the individual linear factors models, which are specified in accord with Equation (2) using index and factor excess returns,

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

The key distinction between the clone model above and Equation (2) 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.

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 $24 \times K$ matrix 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. $\mathrm{N}(0, \sigma^2)$.

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

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

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

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:

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

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

For each combination of index portfolio, sample month, and one of the four 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.b 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 (2016).[14] 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

14 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; Stock and Watson, 2004; 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.

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.

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 model 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 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 four replication models $A_1, \ldots, A_4$ 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

$$\sum_{i=1}^{4} w_i p(r_t^o | \mathcal{R}_{t-1}^o, \mathcal{F}_{t-1}^o, \tilde{f}_t^o, A_i); \quad \sum_{i=1}^{4} w_i = 1; \quad w_i \geq 0 \quad (i = 1, \ldots, 4), \tag{14}$$

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 arguments of any individual density. We can then define the log score function for a four-model pool analogously to the single-model case:

$$LS_T(w) = \sum_{t=25}^{T} \log\left[\sum_{i=1}^{4} w_i p(r_t^o | \mathcal{R}_{t-1}^o, \mathcal{F}_{t-1}^o, \tilde{f}_t^o, A_i)\right], \tag{15}$$

where $w = [w_1 \ \ldots \ w_4]$, $w_i \geq 0$ for $i = 1, \ldots, 4$ and $\sum_{i=1}^{4} w_i = 1$. We define an optimal pool as one for which the weights are chosen to maximize $LS_T(w)$ subject to the constraints noted above. The optimal pool thus corresponds to $w_T^* = \text{argmax}_w LS_T(w)$. 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 $w_T^* = \left[w_{1,T}^* \ \dots \ w_{4,T}^*\right]$ for a given hedge fund index at the end of period $t = T$, the clone investment positions, $\hat{\beta}_{k,A_i}^*$, are determined by both the model weights and the estimated index exposures from the individual replication models. That is

$$\hat{\beta}_{k,A_i}^* = w_{i,T}^* \hat{\beta}_{k,A_i}, \tag{16}$$

where, as before, the factor loadings for each model A_i are estimated using the prior 24 months of index and factor excess returns. To construct the clone return for the optimal pool in month $T+1$, we first compute gross leverage as

$$\gamma_{T+1} = \sum_{i=1}^{4} \sum_{k=1}^{K_i} |\hat{\beta}_{k,A_i}^*|, \tag{17}$$

where K_i is the number of factors in replication model A_i. Gross clone returns are then given by

$$R_{T+1}^C = \begin{cases} R_{f,T+1} + \sum_{i=1}^{4} \sum_{k=1}^{K_i} \hat{\beta}_{k,A_i}^* f_{k,A_i,T+1} & \text{if } \gamma_{T+1} \leq 1 \\ R_{f,T+1} + \sum_{i=1}^{4} \sum_{k=1}^{K_i} \hat{\beta}_{k,A_i}^* f_{k,A_i,T+1} + \frac{0.01}{12}\left(1 - \gamma_{T+1}\right) & \text{if } \gamma_{T+1} > 1. \end{cases} \tag{18}$$

Finally, we compute net returns for the optimal pooled clone by incorporating round trip transaction costs of 3 bps.[15]

3.3.c 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; Amenc *et al.*, 2010), we estimate individual model factor loadings and implied conditional densities using a 24-month rolling window of data. 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 returns extending from January 1994 to December 2015, 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

15 As with each of the other replication methods, we specify a maximum gross leverage of 4×. We demonstrate below, however, that none of the pooled clones reaches this level of leverage over our sample period.

conditional density realizations over the period January 1996–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–December 2015. We also note that, in all cases, the model weights and factor loadings used to build our clone portfolios following Equations (16–18) are known *ex ante*, thus avoiding concerns over look-ahead bias.

As discussed previously, the pooled clone portfolios are constructed from a set of four constituent factor models, each defined based on one of the following asset groups: domestic equity and fixed income, international, currencies, and commodities. Our objective in this empirical design is to allow for a wide range of potential asset class exposures to capture the diverse set of underlying strategies pursued by hedge fund managers. Generally speaking, the exact identity of the models to be included in the pool should depend on the relevant application. Amenc *et al.* (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 four 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 four broadly defined asset groups.

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 four-model pool to alternative methods based on pre-specified factors (i.e., the kitchen sink model, the unconstrained five-factor model, and the renormalized 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 four-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 four-model pool for each of the Credit Suisse hedge fund indexes. As described in Section 3.3.c, 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 presents, for each index, the time-series average weight assigned to each

Table III. Summary statistics for optimal four-model pools

The table reports summary statistics for the optimal four-model pool corresponding to each of the Credit Suisse hedge fund indexes. The following models are considered for each pool and are identified based on their associated factors: the domestic three-factor model, the international three-factor model, the currency 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 presents 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 over the period January 1998–December 2015.

Index	Average model weights				Number of models	
	Domestic model	International model	Currency model	Commodity model	Mean	Median
Panel A: Credit Suisse Hedge Fund Index						
Hedge Fund	0.57	0.39	0.04	0.00	2.80	3.00
Panel B: Directional strategies						
Long/Short Equity	0.72	0.26	0.00	0.02	2.18	2.00
Dedicated Short Bias	0.96	0.03	0.01	0.00	2.32	2.00
Global Macro	0.28	0.38	0.17	0.18	3.50	4.00
Emerging Markets	0.39	0.53	0.05	0.02	2.86	3.00
Managed Futures	0.59	0.19	0.16	0.05	3.10	3.00
Panel C: Non-directional/arbitrage strategies						
Equity Market Neutral	0.51	0.25	0.14	0.10	3.50	4.00
Event Driven	0.70	0.30	0.00	0.00	2.00	2.00
Convertible Arbitrage	0.23	0.53	0.20	0.05	3.40	3.00
Fixed Income Arbitrage	0.13	0.43	0.41	0.04	3.29	3.00

of the four constituent models as well as mean and median number of models included in the optimal pool.

The results in Panel A of 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 three-factor model at 57%. The international and currency models receive average weights of 39% and 4%, 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 2.80, 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 model is particularly influential in the early part of the sample period, achieving a weight as high as 80% in 1998. The currency model is featured materially in the optimal pool only prior to 2006. The weight on the international model tends to increase in the later half of the sample period, largely at the expense of the domestic model.

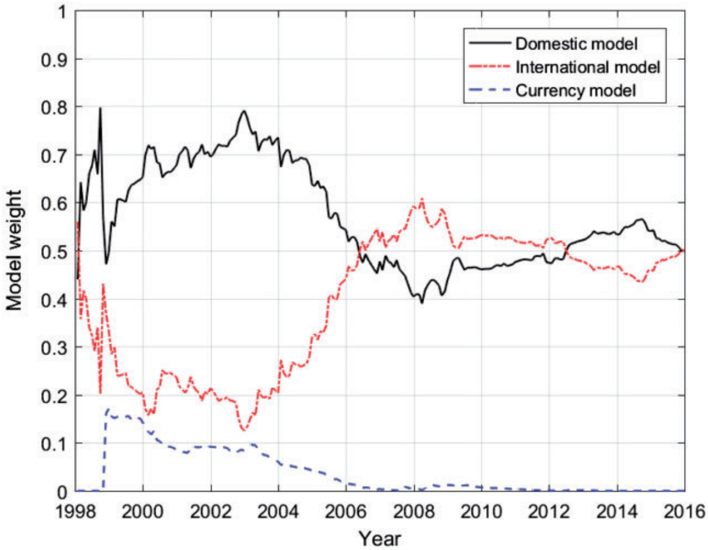

Figure 1. Model weights in the optimal four-model pool. The figure shows the time series of model weights in the optimal four-model pool for the broad Credit Suisse Hedge Fund Index. The following four models are considered for the pool and are identified based on their associated factors: the domestic three-factor model, the international three-factor model, the currency 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–December 2015. The commodity three-factor model is assigned a weight of zero throughout the period shown in the plot.

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 seven 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 noticeable variation in the number of models needed to characterize the various indexes. The average size of the optimal pool ranges from 2.00 models for the Event Driven Index to 3.50 models for the Global Macro and Equity Market Neutral Indexes.

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 exposures. For example, the pool for the Emerging Markets Index assigns an average weight of 53% to the international model, reflecting significant exposure to international equity and foreign exchange markets. Similarly, the pools for the Long/Short Equity and Dedicated Short Bias Indexes tend to be dominated by the domestic three-factor model. The pool for the Global Macro Index allocates at least 17% weight on average to each of the four models, which is consistent with the broad array of strategies followed by these funds.

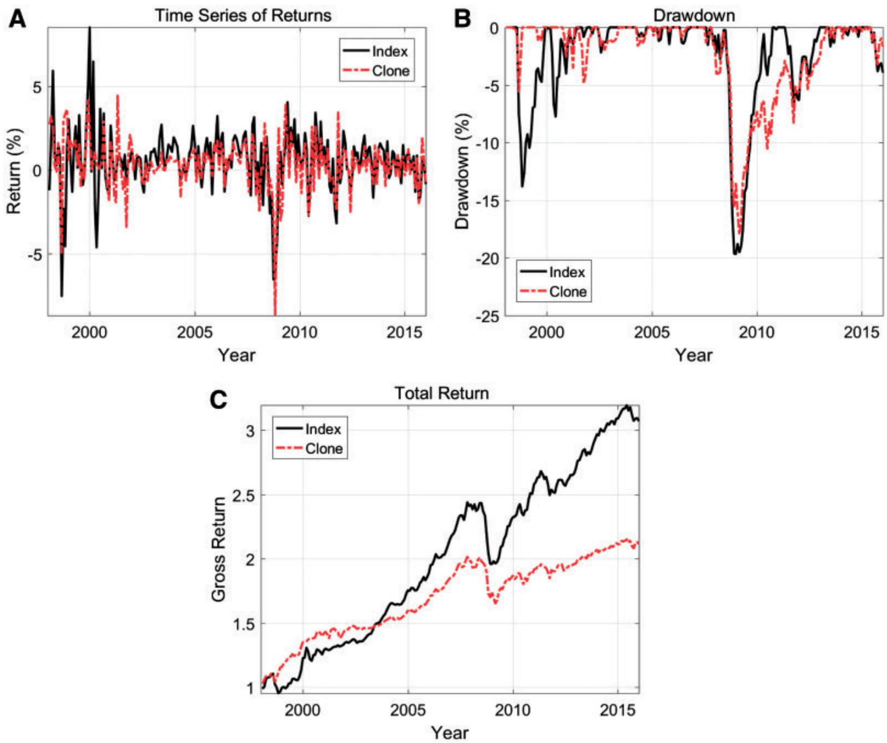

Figure 2. Comparison of index and clone returns. The figure compares the performance of the broad Credit Suisse Hedge Fund Index and the corresponding replication strategy based on the optimal four-model pool. Panel A shows the time series of index and clone returns. Panel B shows the monthly portfolio DD for the index and clone portfolios. Panel C shows the compounded gross returns representing the growth of $1 invested in each strategy. The following four models are considered for the pool and are identified based on their associated factors: the domestic three-factor model, the international three-factor model, the currency 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–December 2015.

The value of the model combination approach to hedge fund replication ultimately depends on the ability of the pooled clones constructed according to Equation (18) 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 (DDs) for the target and clone portfolios and suggests that the pooled replicator effectively matches losses in the broad Hedge Fund Index, particularly during the 2008–09 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.a 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.b examines the ability of each cloning strategy to match the level of average index returns, and Section 4.2.c considers the differences in portfolio turnover and leverage across the replication methods.

4.2.a Tracking performance

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* optimal model weights that are applied in constructing the out-of-sample clone:

$$LS_{\mathcal{T}}^{OP} = \sum_{t=49}^{\mathcal{T}} \log \left[\sum_{i=1}^{4} w_{i,t-1}^{*} p(r_t^o | \mathcal{R}_{t-1}^o, \mathcal{F}_{t-1}^o, \tilde{f}_t^o, A_i) \right]. \tag{19}$$

Note that, as discussed above, the out-of-sample period for clone evaluation extends from January 1998 (i.e., $t = 49$) to December 2015 (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 (19). 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.[16]

Table IV reports the log score values for the optimal pool and the three alternative replication strategies for the period January 1998–December 2015. 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 Long/Short Equity and Dedicated Short Bias Indexes), and the kitchen sink method in zero cases. Given that the kitchen sink model includes the full set of twelve factors, it does not appear that simply incorporating a large number of asset class 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.

16 The renormalized 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.

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 four-model pool, the kitchen sink regression model, the unconstrained 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–December 2015 sample period. For each of the alternative replication approaches to the optimal model pool, the table also reports a *p*-value corresponding to a one-sided test of the null hypothesis that the log score for the optimal pool is less than or equal to the log score for the individual model.

Index	Optimal pool	Alternative replication methods					
		Kitchen sink		Unconstrained five-factor		Stepwise	
	Log score	Log score	*p*-value	Log score	*p*-value	Log score	*p*-value
Panel A: Credit Suisse Hedge Fund Index							
Hedge Fund	649.4	617.5	0.007	639.5	0.159	608.1	0.012
Panel B: Directional strategies							
Long/Short Equity	601.2	580.2	0.029	593.9	0.238	596.3	0.359
Dedicated Short Bias	466.9	416.1	0.000	440.5	0.002	449.1	0.019
Global Macro	580.3	531.4	0.000	567.1	0.113	536.2	0.021
Emerging Markets	519.1	490.6	0.001	512.7	0.203	508.4	0.197
Managed Futures	431.9	373.1	0.000	414.7	0.009	396.2	0.002
Panel C: Non-directional/arbitrage strategies							
Equity Market Neutral	637.4	608.9	0.152	637.0	0.481	598.2	0.015
Event Driven	626.1	579.5	0.000	611.5	0.036	589.9	0.004
Convertible Arbitrage	594.5	559.7	0.004	588.5	0.214	564.1	0.017
Fixed Income Arbitrage	658.5	614.4	0.001	647.6	0.124	613.8	0.007

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 (2016). 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.[17] Each of the *p*-values reported in Table IV corresponds to a one-sided test of the null hypothesis that the log score for the optimal pool is less than or equal to the log score for the individual model. The table thus includes results for thirty separate tests. As noted previously, the optimal pooled replicator yields superior tracking performance over the alternative models in all thirty cases. These differences are statistically significant at the 5% level in twenty instances.

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

Following Amenc *et al.* (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, unconstrained five-factor model, renormalized five-factor model, and stepwise regression), we compute root mean square replication errors as follows:

$$\mathrm{RMSE}_{\mathcal{T}} = \sqrt{\frac{1}{\mathcal{T}-48}\sum_{t=49}^{\mathcal{T}}\left(R_t^I - R_t^C\right)^2}, \tag{20}$$

where R_t^I 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 1.50% per month over the full January 1998–December 2015 out-of-sample period. For the remaining four models, the replication errors range from 1.64% per month for unconstrained five-factor approach to 1.93% for the renormalized 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 1.44% for the Event Driven Index, but the corresponding figure for the Managed Futures Index is 3.29%. 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, renormalized five-factor model, and 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 sample period. Much of the superior performance is concentrated in the 2000–01 period around the dot-com collapse. There are also pronounced jumps in three of the plots during 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 Global Macro Index (Panel D), for example, generate substantial replication errors relative to the model pooling strategy following 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

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 four-model pool, the kitchen sink regression model, the unconstrained five-factor regression model, the renormalized 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–December 2015.

Index	Replication method				
	Optimal pool	Kitchen sink	Unconstrained five-factor	Renormalized five-factor	Stepwise
Panel A: Root mean square replication errors					
A.1: Credit Suisse Hedge Fund Index					
Hedge Fund	1.50	1.86	1.64	1.93	1.69
A.2: Directional strategies					
Long/Short Equity	1.94	2.25	2.02	2.35	2.00
Dedicated Short Bias	3.05	3.79	3.34	3.62	3.26
Global Macro	2.18	2.74	2.52	3.15	2.36
Emerging Markets	2.74	3.03	3.00	3.27	2.97
Managed Futures	3.29	3.87	3.63	4.42	3.58
A.3: Non-directional/arbitrage strategies					
Equity Market Neutral	2.94	3.07	3.30	3.36	3.38
Event Driven	1.44	1.95	1.62	1.84	1.53
Convertible Arbitrage	1.80	2.10	1.99	2.35	2.01
Fixed Income Arbitrage	1.46	1.79	1.60	1.86	1.69
Panel B: Out-of-sample correlation between index and clone					
B.1: Credit Suisse Hedge Fund Index					
Hedge Fund	0.62	0.60	0.61	0.59	0.59
B.2: Directional strategies					
Long/Short Equity	0.73	0.70	0.72	0.70	0.73
Dedicated Short Bias	0.78	0.67	0.74	0.74	0.75
Global Macro	0.28	0.30	0.19	0.15	0.40
Emerging Markets	0.66	0.62	0.66	0.68	0.66
Managed Futures	0.33	0.28	0.30	0.30	0.31
B.3: Non-directional/arbitrage strategies					
Equity Market Neutral	0.23	0.18	0.19	0.20	0.13
Event Driven	0.65	0.53	0.61	0.60	0.65
Convertible Arbitrage	0.47	0.39	0.43	0.43	0.38
Fixed Income Arbitrage	0.47	0.44	0.45	0.47	0.46

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.

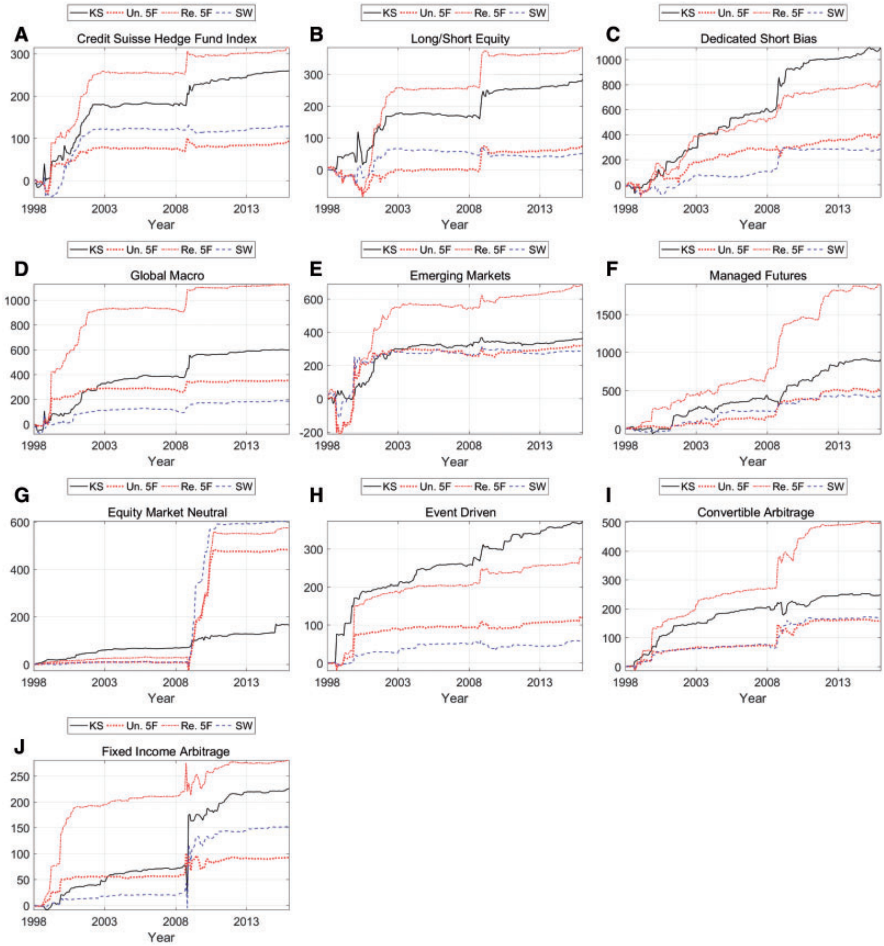

Figure 3. Difference in cumulative sum of squared replication errors. The figure shows the difference in cumulative sum of squared replication errors for each of the four alternative replication strategies (i.e., the kitchen sink regression model (KS), the unconstrained five-factor regression model (Un. 5F), the renormalized five-factor model (Re. 5F), and the stepwise regression model (SW)) relative to the optimal four-model pool. 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–December 2015.

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 eight of the ten indexes. Broadly speaking, the clones are more reliable in capturing the time-series patterns in returns for the directional indexes. The largest three out-of-sample correlations among the pooled replicators, for example, are achieved by the clones for the Dedicated Short Bias, Long/Short Equity, and Emerging Markets Indexes (Panel B.2). In contrast, the correlation for the Equity Market Neutral Index in Panel B.3 is just 0.23, suggesting that the value of replication is more limited for non-directional strategies.

Finally, for each index and cloning method, we also assess replication performance by regressing index returns on clone returns over the out-of-sample period:

$$R_t^I = \alpha + \beta R_t^C + \eta_t. \tag{21}$$

The estimated coefficients are a convenient way to measure the match between the clone and target portfolios, with the ideal replicator corresponding to $\alpha = 0$ and $\beta = 1$. The regression R^2s also provide an indication of the amount of time-series variability in target returns captured by the clones. The results are presented in Table VI. Focusing initially on the pooled clones, we see considerable variation in replication performance. The R^2 values range from just 6.3% for the Equity Market Neutral Index to 60.8% for the Dedicated Short Bias Index. As expected, the R^2s are generally higher for the directional indexes in Panel B compared with the non-directional/arbitrage strategies in Panel C. There are exceptions to this result, however, as the out-of-sample R^2s for the Global Macro and Managed Futures Indexes are only 16.3% and 12.7%, respectively. Across the ten indexes, the estimates of β range from 0.49 to 0.90, and five of these coefficients are more than two standard errors less than one. The indexes also tend to outperform their clone portfolios, as nine of the α estimates are positive, and four are more than two standard errors above zero.

The model pooling approach does, however, perform well in comparison to the alternative methods. The regression R^2 for the broad Hedge Fund Index's clone is 43.4%, whereas the corresponding values for the other clones are between 40.1% and 41.9%. The pooled clone achieves the highest out-of-sample R^2 for six of the nine individual strategies in Panels B and C of Table VI. The estimates of β for the pooled clones also tend to be larger and reliably closer to one. For example, the slope coefficient is more than two standard errors less than one in thirty-nine out of the forty cases spanning the alternative cloning approaches.

4.2.b 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 or the execution of arbitrage strategies via a passive replication approach. Thus, any observed underperformance for clones relative to the corresponding indexes can be partially attributed to managerial skill.

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 *et al.* (2008), Amenc *et al.* (2010), Bollen and Fisher (2013), and Jurek and Stafford (2015) consider factor-based replication approaches for hedge fund indexes and reach similar conclusions on the underperformance of clones relative to their targets.

Table VI. Regressions of target returns on clone returns

For each Credit Suisse hedge fund index and replication strategy (i.e., the optimal four-model pool, the kitchen sink regression model, the unconstrained five-factor regression model, the renormalized five-factor model, and the stepwise regression model), the table reports results from the following regression of index returns (R_t^I) on clone returns (R_t^C): $R_t^I = \alpha + \beta R_t^C + \eta_t$. For each regression, the table presents estimates of α (in percentage per month), β, and the regression R^2 value. The numbers in parentheses are Newey–West (1987) standard errors with a lag length equal to five. The sample period for each model is January 1998–December 2015.

Index	Replication method														
	Optimal pool			Kitchen sink			Unconstrained five-factor			Renormalized five-factor			Stepwise		
	α	β	R^2	α	β	R^2	α	β	R^2	α	β	R^2	α	β	R^2
Panel A: Credit Suisse Hedge Fund Index															
Hedge Fund	0.26	0.77	43.4	0.33	0.50	40.7	0.27	0.61	41.9	0.31	0.48	40.1	0.33	0.59	40.3
	(0.10)	(0.10)		(0.13)	(0.06)		(0.11)	(0.08)		(0.12)	(0.07)		(0.13)	(0.07)	
Panel B: Directional strategies															
Long/Short Equity	0.35	0.84	55.6	0.43	0.64	52.3	0.36	0.74	55.3	0.42	0.60	52.2	0.38	0.75	56.1
	(0.12)	(0.08)		(0.14)	(0.06)		(0.13)	(0.08)		(0.13)	(0.08)		(0.12)	(0.08)	
Dedicated Short Bias	−0.20	0.90	60.8	−0.20	0.72	44.9	−0.20	0.81	54.9	−0.20	0.68	55.9	−0.29	0.83	57.0
	(0.23)	(0.09)		(0.25)	(0.09)		(0.24)	(0.08)		(0.24)	(0.07)		(0.23)	(0.10)	
Global Macro	0.55	0.49	16.3	0.59	0.26	17.6	0.60	0.23	12.8	0.64	0.12	11.3	0.58	0.40	24.0
	(0.16)	(0.19)		(0.17)	(0.07)		(0.17)	(0.15)		(0.17)	(0.10)		(0.16)	(0.08)	
Emerging Markets	0.19	0.86	44.5	0.30	0.67	40.3	0.11	0.65	44.5	0.18	0.56	47.3	0.16	0.66	45.2
	(0.26)	(0.10)		(0.26)	(0.09)		(0.27)	(0.10)		(0.26)	(0.08)		(0.26)	(0.09)	
Managed Futures	0.44	0.57	12.7	0.45	0.30	9.9	0.46	0.37	10.8	0.49	0.25	10.8	0.46	0.39	11.8
	(0.19)	(0.16)		(0.19)	(0.08)		(0.19)	(0.10)		(0.19)	(0.07)		(0.19)	(0.10)	

(continued)

Table VI. Continued

Index	Replication method														
	Optimal pool			Kitchen sink			Unconstrained five-factor			Renormalized five-factor			Stepwise		
	α	β	R^2	α	β	R^2	α	β	R^2	α	β	R^2	α	β	R^2
Panel C: Non-directional/arbitrage strategies															
Equity Market Neutral	0.16	0.64	6.3	0.24	0.39	4.5	0.31	0.27	4.7	0.31	0.27	5.1	0.29	0.20	3.1
	(0.24)	(0.17)		(0.23)	(0.09)		(0.22)	(0.07)		(0.22)	(0.06)		(0.23)	(0.07)	
Event Driven	0.24	0.80	47.6	0.40	0.46	35.0	0.25	0.62	42.6	0.29	0.51	41.6	0.30	0.68	47.0
	(0.16)	(0.12)		(0.16)	(0.06)		(0.17)	(0.12)		(0.16)	(0.10)		(0.16)	(0.10)	
Convertible Arbitrage	0.31	0.75	27.1	0.39	0.44	20.4	0.28	0.50	23.3	0.30	0.36	23.9	0.38	0.50	19.7
	(0.20)	(0.21)		(0.21)	(0.14)		(0.20)	(0.12)		(0.20)	(0.09)		(0.20)	(0.16)	
Fixed Income Arbitrage	0.17	0.74	25.7	0.21	0.41	22.8	0.12	0.52	23.8	0.14	0.40	25.4	0.25	0.46	24.8
	(0.20)	(0.31)		(0.18)	(0.14)		(0.20)	(0.20)		(0.19)	(0.15)		(0.16)	(0.14)	

Table VII presents average returns, standard deviations, and minimum DDs for the ten Credit Suisse indexes and each of the corresponding clone portfolios. For each clone, the table also reports a wealth relative ($\frac{W_C}{W_I}$) 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 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). These differences in average returns are significant at the 5% level in only three cases. Focusing on the results across the cloning methods, none of the strategies appears to deliver reliably superior performance. The highest average return is earned by the renormalized five-factor clone in six 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 DDs in seven out of ten cases.

The underperformance of clone portfolios shown in Table VII should not be interpreted as a failure of hedge fund replication for several reasons. First, as noted by Jurek and Stafford (2015), 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 Trading Advisor Selection System (TASS) database over the period 1994–2003 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; Getmansky *et al.*, 2010) and the overall industry becomes more competitive.

4.2.c 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 shown in Figure 4. 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 futures contract. 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 renormalized five-factor clone are limited to positions in the five underlying factors. In contrast, the optimal pool, kitchen sink, and stepwise clones can generate non-zero portfolio positions in as many as twelve futures contracts in any given month.

Table VII. 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 four-model pool, the kitchen sink regression model, the unconstrained five-factor regression model, the renormalized 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 DD over the out-of-sample period, January 1998–December 2015. For each clone portfolio, the table presents a *p*-value for the two-sided test that the mean returns for the clone and its corresponding index are equal. For each replication strategy, 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.

Index				Replication method									
	Index			Optimal pool					Kitchen sink				
	μ_I	σ_I	DD_I	μ_C	*p*-value	σ_C	DD_C	$\frac{W_C}{W_I}$	μ_C	*p*-value	σ_C	DD_C	$\frac{W_C}{W_I}$
Credit Suisse Hedge Fund Index													
Hedge Fund	0.54	1.86	−19.7	0.36	0.10	1.49	−18.0	0.69	0.42	0.43	2.22	−23.4	0.76
Directional strategies													
Long/Short Equity	0.71	2.74	−22.0	0.42	0.02	2.36	−27.1	0.55	0.43	0.06	3.01	−32.4	0.54
Dedicated Short Bias	−0.43	4.81	−77.9	−0.26	0.43	4.15	−61.4	1.55	−0.33	0.68	4.45	−72.6	1.30
Global Macro	0.69	2.13	−26.8	0.29	0.02	1.21	−9.7	0.44	0.37	0.15	2.46	−17.3	0.50
Emerging Markets	0.53	3.61	−39.2	0.40	0.60	2.76	−27.9	0.81	0.34	0.45	3.37	−36.7	0.68
Managed Futures	0.51	3.35	−17.4	0.12	0.05	1.93	−25.9	0.47	0.18	0.16	3.09	−22.9	0.51
Non-directional/arbitrage strategies													
Equity Market Neutral	0.34	3.00	−45.1	0.28	0.77	1.06	−10.0	0.99	0.27	0.70	1.40	−10.4	0.94
Event Driven	0.58	1.86	−19.1	0.42	0.25	1.50	−15.8	0.72	0.40	0.25	2.15	−21.9	0.67
Convertible Arbitrage	0.51	2.00	−32.9	0.27	0.12	1.26	−18.5	0.61	0.28	0.20	1.77	−16.2	0.62
Fixed Income Arbitrage	0.36	1.63	−29.0	0.25	0.41	1.03	−13.3	0.81	0.37	0.98	1.74	−17.4	1.00

(continued)

Table VII. Continued

Index	Replication method														
	Unconstrained five-factor					Renormalized five-factor					Stepwise				
	μ_C	p-value	σ_C	DD_C	$\frac{W_C}{W_I}$	μ_C	p-value	σ_C	DD_C	$\frac{W_C}{W_I}$	μ_C	p-value	σ_C	DD_C	$\frac{W_C}{W_I}$
Credit Suisse Hedge Fund Index															
Hedge Fund	0.44	0.43	1.85	−23.5	0.82	0.48	0.67	2.32	−27.3	0.86	0.35	0.12	1.87	−20.8	0.67
Directional strategies															
Long/Short Equity	0.47	0.07	2.67	−33.8	0.60	0.47	0.14	3.21	−37.5	0.59	0.43	0.02	2.65	−29.5	0.56
Dedicated Short Bias	−0.29	0.56	4.37	−64.3	1.42	−0.35	0.76	5.28	−71.8	1.15	−0.17	0.26	4.36	−58.8	1.84
Global Macro	0.39	0.15	1.79	−15.2	0.54	0.39	0.26	2.64	−25.4	0.52	0.26	0.01	2.12	−13.4	0.40
Emerging Markets	0.66	0.64	3.65	−34.6	1.31	0.63	0.72	4.38	−43.3	1.17	0.58	0.87	3.65	−35.6	1.10
Managed Futures	0.13	0.10	2.69	−30.1	0.47	0.08	0.13	4.02	−47.8	0.38	0.11	0.08	2.67	−29.8	0.44
Non-directional/arbitrage strategies															
Equity Market Neutral	0.11	0.28	2.04	−35.0	0.66	0.11	0.28	2.22	−37.3	0.65	0.28	0.81	2.03	−21.2	0.95
Event Driven	0.54	0.76	1.81	−21.7	0.91	0.58	0.99	2.19	−26.4	0.99	0.41	0.23	1.77	−19.1	0.70
Convertible Arbitrage	0.46	0.74	1.71	−23.8	0.91	0.57	0.72	2.39	−29.5	1.11	0.25	0.11	1.51	−21.7	0.59
Fixed Income Arbitrage	0.46	0.39	1.40	−16.4	1.26	0.56	0.12	1.92	−21.4	1.51	0.25	0.36	1.62	−19.3	0.79

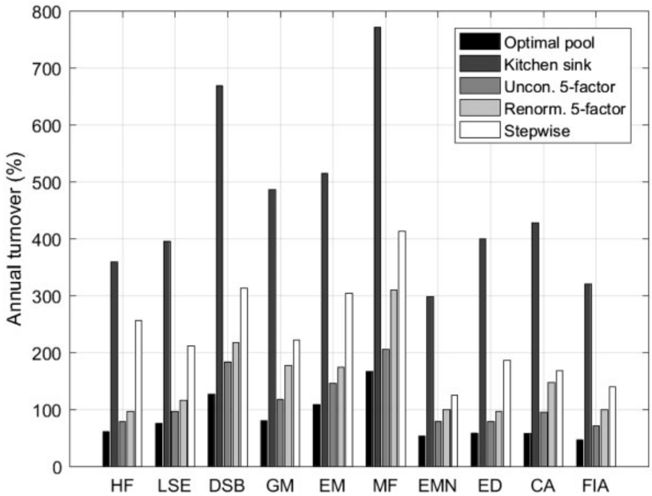

Figure 4. Comparison of annual turnover. For each of the Credit Suisse hedge fund indexes, the figure shows the average annual portfolio turnover for each of the following replication strategies: the optimal four-model pool, the kitchen sink regression model, the unconstrained five-factor regression model, the renormalized five-factor model, and the stepwise regression model. The indexes considered are the broad Hedge Fund Index (HF), the Long/Short Equity Index (LSE), the Dedicated Short Bias Index (DSB), the Global Macro Index (GM), the Emerging Markets Index (EM), the Managed Futures Index (MF), the Equity Market Neutral Index (EMN), the Event Driven Index (ED), the Convertible Arbitrage Index (CA), and the Fixed Income Arbitrage Index (FIA). Monthly portfolio turnover is computed as 0.5 times the sum of the absolute values of the change in portfolio weights in each underlying futures contract. The annual turnover figures are computed by multiplying the monthly turnover values by 12. The sample period is January 1998–December 2015.

From Figure 4, the highest levels of turnover are associated with the kitchen sink clone in all ten cases. The required levels of trading are economically large, with turnover ranging from 298% per year for the Equity Market Neutral Index to 772% for the Managed Futures Index. The stepwise replication procedure also results in substantial levels of turnover, ranging from 126% to 414%. 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 4 shows that the unconstrained and renormalized five-factor clones require less trading than the kitchen sink and stepwise procedures in all cases. Focusing only on the five-factor clones, the unconstrained version tends to require slightly less turnover, with values ranging from 71% for the Fixed Income Arbitrage Index to 206% for the Managed Futures Index.

More importantly, the optimal pooled clone implies lower turnover than any of the four competing methods for each of the ten index strategies. The pooled replication approach for the broad Hedge Fund Index has an average annual turnover of 61%, compared with values between 79% and 359% for the four alternatives. These improvements are also pronounced within several of the strategy-specific indexes. For example, the pooled clone for the Long/Short Equity Index requires portfolio turnover of only 75% per year, whereas the other four models lead to turnover ranging from 97% to 396%. These results are surprising given that, like the kitchen sink and stepwise methods, the optimal pool allows for allocations to 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 transaction 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 VIII summarizes the levels of gross leverage required by each of the clone portfolios. Consider an investment strategy that combines investments in risky assets (i.e., futures contracts) with positions in cash (i.e., the risk-free asset). Both the risky assets and cash can be held as long or short positions. Gross leverage is simply defined as the sum of the portfolio's long and short positions in all risky assets as a proportion of total assets.[18] Following Equations (4) and (18), the clones are constructed via long and short positions in cash and the available futures contracts. We aggregate the implied long and short futures positions each month and compute gross leverage for each clone as defined above. Also recall that in building the clone portfolios, we limit gross leverage to 4×.

For each clone, Table VIII reports the mean and maximum gross leverage over the sample period. The average leverage tends to be somewhat similar for the optimal pool, unconstrained five-factor, and stepwise clones within each target index. Each of these three approaches routinely requires levels of gross leverage below one. Leverage levels for the kitchen sink and renormalized five-factor clones tend to be higher on average. None of the optimal pooled clones reaches a maximum gross leverage above three during the sample period, suggesting that this approach requires comparatively less extreme investment positions and would be relatively practical to implement.

4.3 Economic benefits of hedge fund clones

The use of a hedge fund clone developed according to the replication methods described in this article 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), the Russell 2000 Index (RTY), the Bank of America (BofA) Merrill Lynch 7–10 Year US Treasury Index (TB), the BofA Merrill Lynch 7–10 Year US Corporate Index (CB), the Morgan Stanley Capital International (MSCI) World ex USA Index (MXWOU), the MSCI Emerging Markets Index (MXEF), and FTSE National Association of Real Estate Investment Trusts (NAREIT) All Equity REITs Index (FNER), in addition to the pooled clone for the broad Hedge Fund Index. The time-series data on index returns are from Bloomberg.

18 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.

Table VIII. Gross leverage

For each Credit Suisse hedge fund index, the table reports the average and maximum gross leverage over the out-of sample period (January 1998–December 2015) for each of the following index replication strategies: the optimal four-model pool, the kitchen sink regression model, the unconstrained five-factor regression model, the renormalized five-factor model, and the stepwise regression model. For a portfolio that invests in risky futures contracts and cash, gross leverage is defined as the proportion of capital deployed as notional futures exposure (i.e., the sum of the portfolio's long and short positions in the futures contracts as a proportion of total assets). Each of the approaches to building clones for the hedge fund indexes is described in the text.

Index	Replication method									
	Optimal pool		Kitchen sink		Unconstrained five-factor		Renormalized five-factor		Stepwise	
	Mean	Max	Mean	Max	Mean	Max	Mean	Max	Mean	Max
Panel A: Credit Suisse Hedge Fund Index										
Hedge Fund Index	0.54	1.12	1.94	4.00	0.69	1.91	0.87	2.24	0.76	3.89
Panel B: Directional strategies										
Long/Short Equity	0.71	1.37	2.16	4.00	0.94	2.30	1.13	3.28	0.85	2.83
Dedicated Short Bias	1.20	2.05	3.46	4.00	1.90	3.57	2.28	4.00	1.22	4.00
Global Macro	0.59	1.89	2.42	4.00	0.80	2.84	1.23	3.68	0.70	3.69
Emerging Markets	0.98	2.76	2.85	4.00	1.38	4.00	1.69	4.00	1.19	4.00
Managed Futures	1.17	2.34	3.66	4.00	1.47	4.00	2.31	4.00	1.17	4.00
Panel C: Non-directional/arbitrage strategies										
Equity Market Neutral	0.40	2.17	1.54	4.00	0.75	4.00	0.88	4.00	0.64	4.00
Event Driven	0.51	1.12	2.39	4.00	0.69	2.00	0.84	2.30	0.56	2.25
Convertible Arbitrage	0.38	1.16	2.36	4.00	0.65	1.63	0.99	2.12	0.47	2.36
Fixed Income Arbitrage	0.34	1.20	1.80	4.00	0.59	1.49	0.81	2.12	0.39	2.11

Investors' expected utility is defined over the expected return, $E(R_t^P)$, and variance, σ_P^2, of their portfolio:

$$E(U) = E(R_t^P) - \frac{1}{2}\lambda\sigma_P^2, \tag{22}$$

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–December 2015.

Panels A and B of Table IX present 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.[19] The restrictions are

19 Source: 2013 NACUBO-Commonfund Study of Endowments.

Table IX. Optimal asset allocation with hedge fund clone investments

The table reports the ex-post optimal portfolio allocations and the corresponding CER in percentage per year for mean-variance optimizing investors during the period January 1998–December 2015. Investors' expected utility is defined over the expected return, $E(R_t^P)$, and variance, σ_P^2, of their portfolio: $E(U) = E(R_t^P) - \frac{1}{2}\lambda\sigma_P^2$, 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 RIY, RTY, TB, CB, MXWOU, MXEF, and 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 percentage per year if investors are forced to exclude the hedge fund clone from their portfolios.

Risk aversion	Portfolio weights								Certainty equivalent returns	
	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	0.000	0.100	0.342	0.308	5.36	0.17
5	0.150	0.000	0.100	0.000	0.000	0.084	0.166	0.500	3.96	2.09
10	0.150	0.000	0.100	0.000	0.110	0.000	0.140	0.500	2.08	5.33
Excluding clone investments										
2	0.058	0.092	0.100	0.000	0.199	0.100	0.451	n/a	5.19	n/a
5	0.150	0.000	0.100	0.000	0.390	0.043	0.317	n/a	1.87	n/a
10	0.150	0.000	0.100	0.000	0.481	0.000	0.269	n/a	−3.25	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	0.043	0.366	0.400	0.000	0.000	0.100	0.091	0.000	6.34	0.00
5	0.163	0.114	0.400	0.000	0.000	0.030	0.044	0.248	5.11	0.19
10	0.048	0.062	0.400	0.000	0.000	0.000	0.000	0.490	4.41	1.48
Excluding clone investments										
2	0.043	0.366	0.400	0.000	0.000	0.100	0.091	n/a	6.34	n/a
5	0.419	0.069	0.400	0.000	0.000	0.016	0.096	n/a	4.92	n/a
10	0.509	0.000	0.400	0.000	0.000	0.000	0.091	n/a	2.93	n/a

in the form of upper bounds on the weights for the following asset classes: US stocks (RIY + RTY $\leq$ 15%), US 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 IX, 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 CER, when the investors are unable to invest in the hedge fund clone. The utility losses for investors with moderate (CER loss of 2.09% per year) and high risk aversion (CER loss of 5.33% 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: US stocks (60%), US 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 IX shows that for investors with moderate risk aversion, the implied allocation to the hedge fund clone is considerable (24.8%). The allocation to the hedge fund clone is even more substantial for investors with high risk aversion (49.0%). The corresponding utility loss from excluding the hedge fund clone in the optimal asset mix is 1.48% 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 article, 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 find that hedge fund replication is relatively more reliable for indexes following directional strategies. Although all of the replication products tend to underperform their target indexes over the 1998–2015 sample period, 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. 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 article has potentially valuable applications in hedging systematic risk inherent in fund strategies and benchmarking managerial performance (e.g., O'Doherty, Savin, and Tiwari, 2016).

References

Agarwal, V. and Naik, N. Y. (2004) Risks and portfolio decisions involving hedge funds, *Review of Financial Studies* **17**, 63–98.

Amenc, N., Gehin, W., Martellini, L., and Meyfredi, J. C. (2008) Passive hedge fund replication: A critical assessment of existing techniques, *Journal of Alternative Investments* **11**, 69–83.

Amenc, N., Martellini, L., Meyfredi, J. C., and Ziemann, V. (2010) Passive hedge fund replication – Beyond the linear case, *European Financial Management* **16**, 191–210.

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

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

Ang, A., Gorovyy, S., and van Inwegen, G. B. (2011) Hedge fund leverage, *Journal of Financial Economics* **102**, 102–126.

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

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

Bollen, N. P. and Whaley, R. E. (2009) Hedge fund risk dynamics: Implications for performance appraisal, *Journal of Finance* **64**, 985–1035.

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

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

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

Fung, W. and Hsieh, D. A. (2001) The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies* **14**, 313–341.

Gelman, A., Hwang, J., and Vehtari, A. (2014) Understanding predictive information criteria for Bayesian models, *Statistics and Computing* **24**, 997–1016.

Getmansky, M., Liang, B., Schwarz, C., and Wermers, R. (2010) Share restrictions and investor flows in the hedge fund industry. Unpublished working paper, University of Massachusetts.

Geweke, J. and Amisano, G. (2011) Optimal prediction pools, *Journal of Econometrics* **164**, 130–141.

Geweke, J. and Amisano, G. (2014) Analysis of variance for Bayesian inference, *Econometric Reviews* **33**, 270–288.

Giacomini, R. and White, H. (2006) Tests of conditional predictive ability, *Econometrica* **74**, 1545–1578.

Giamouridis, D. and Paterlini, S. (2010) Regular(ized) hedge fund clones, *Journal of Financial Research* **33**, 223–247.

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

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

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

Jurek, J. W. and Stafford, E. (2015) The cost of capital for alternative investments, *Journal of Finance* **70**, 2185–2226.

Kat, H. M. and Palaro, H. P. (2005) Who needs hedge funds? A copula-based approach to hedge fund return replication. Unpublished working paper, City University.

Kat, H. M. and Palaro, H. P. (2006) Superstars or average Joes? A replication-based performance evaluation of 1917 individual hedge funds. Unpublished working paper, City University.

Klein, R. W. and Brown, S. J. (1984) Model selection when there is "minimal" prior information, *Econometrica* **52**, 1291–1312.

Koijen, R. S., Moskowitz, T. J., Pedersen, L. H., and Vrugt, E. B. (2015) Carry. Unpublished working paper, London Business School.

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

Malkiel, B. G. and Saha, A. (2005) Hedge funds: Risk and return, *Financial Analysts Journal* **61**, 80–88.

Mitchell, M. and Pulvino, T. (2001) Characteristics of risk and return in risk arbitrage, *Journal of Finance* **56**, 2135–2175.

Naik, N. Y., Ramadorai, T., and Stromqvist, M. (2007) Capacity constraints and hedge fund strategy returns, *European Financial Management* **13**, 239–256.

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

Newey, W. K. and West, K. D. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 703–708.

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

O'Doherty, M., Savin, N., and Tiwari, A. (2016) Evaluating hedge funds with pooled benchmarks, *Management Science* **62**, 69–89.

Payne, B. C. and Tresl, J. (2015) Hedge fund replication with a genetic algorithm: Breeding a usable mousetrap, *Quantitative Finance* **15**, 1705–1726.

Pedersen, L. H. (2015) *Efficiently Inefficient*, Princeton University Press, Princeton, NJ.

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

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

Stock, J. H. and Watson, M. W. (2003) Forecasting output and in inflation: The role of asset prices, *Journal of Economic Literature* **41**, 788–829.

Stock, J. H. and Watson, M. W. (2004) Combination forecasts of output growth in a seven-country data set, *Journal of Forecasting* **23**, 405–430.

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

Zellner, A. (1971) *An Introduction to Bayesian Inference in Econometrics*, Wiley, Hoboken, NJ.