Three Essays on Hedge Funds

Dissertation
zur Erlangung des akademischen Grades
des Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)
an der Universität Konstanz
Fachbereich Wirtschaftswissenschaften

vorgelegt von:
Anna Slavutskaya
Radolfzeller Str. 56
78467 Konstanz

Tag der mündlichen Prüfung: 9. Juli 2014

1. Referent: Prof. Dr. Jens Carsten Jackwerth
2. Referent: Prof. Dr. Axel Kind

Моим любимым и самым дорогим родителям

Acknowledgements

First and foremost, I would like to thank Prof. Jens Jackwerth for his advice and guidance during my PhD. Working under his supervision was a challenge and it was absolutely worth it. His enthusiasm and high standards for research were the greatest motivation for me. I appreciate a lot his sense of humor which he always kept even when I lost mine.

I am also very thankful to Prof. Günther Franke, Prof. Winfried Pohlmeier, and Dr. Olga Kolokolova for their valuable comments and suggestions which helped me to improve my thesis. Special thanks to Prof. Axel Kind who kindly agreed to appraise this thesis.

I want to thank my colleagues from the chair Achim Mattes, Marco Menner, Zhen Guo, Horatio Cuesdeanu, and Marc Gerritzen who have contributed a lot to my personal and professional time at the University of Konstanz. I am especially grateful to my dearest friends Fady Barsoum, Sandra Stankiewicz, and Liliya Sattarova who were always there for me.

My warmest thanks to Monika Fischer who helped me a lot with the administrative work. This allowed me to concentrate fully on my thesis. She was also very understanding and supportive in times of down moments.

Last but not least, I would like to express my gratitude to my family for their unconditional love and dedication. Many thanks to Orkan Şahin who also became a part of my family. He always believed in me, provided fruitful discussion on my ideas and constantly encouraged to continue my research work. Without him I would not have been able to complete this thesis.

Contents

List of Figures

List of Tables

Summary

The industry of hedge funds has experienced a dramatic growth over the last 15 years. A number of characteristics make hedge funds particularly attractive for an investor: option-like returns and low correlation with various asset classes are among them.

The estimates of the hedge fund industry size vary depending on the database. According to the report of AIMA [2012], during the period 1999-2007, hedge fund assets quadrupled from $456 billion to $1,868 billion. Even though year 2008 was marked by major outflows[1], by the third quarter of 2012 total assets under management achieved the number of $2,192 billion.

Naturally, interest in hedge funds from academics has increased a lot too. There are numerous studies on the hedge fund performance. Nevertheless, many issues have not been fully investigated yet. This is not surprising as there are several features which allow to categorize hedge funds as very complex investment vehicles. Whereas hedge fund strategies per se can be quite sophisticated and highly dynamic, the absence of the strict governmental regulation and requirements to reveal underlying positions do not contribute to our understanding of what hedge funds actually do.

Richard Bookstaber summarizes the concept of hedge funds in a interesting way, which enlightens why at times it may be challenging to analyze hedge

[1]Partially due to the Financial crisis and subsequent losses which caused redemptions, partially due to the Bernie Madoff scandal which significantly deteriorated the reputation of hedge funds.

funds[1]:

> I believe there is no such thing as a hedge fund. Hedge funds are not a homogeneous class that can be analyzed in a consistent way. The hedge funds/alternative investment moniker is a description of what an investment fund is not rather than what it is. The universe of alternative investments is just that the universe. It encompasses all possible investment vehicles and all possible investment strategies minus the ”traditional” investment funds and vehicles.

In general, a hedge fund is very flexible in its investments. It operates with any investment tool, any strategy and in any market where it sees substantial gains at a lower risk.

The thesis is organized as follows: the first chapter investigates the short-term performance of hedge funds. The second chapter concentrates on the major group of hedge fund institutional investors, pension funds, and tests if hedge fund are promising investment tools. The third chapter introduces a novel measure of the hedge fund performance, relative alpha.

Academically, benefits of hedge funds are often evaluated in terms of alpha within a Fung and Hsieh [2001] (or Agarwal and Naik [2004]) factor model[2]. Due to their simplicity and straightforward implementation, linear models have become a well established methodology for the hedge fund return analysis. At the same time, these models suffer from diverse estimation problems. First of all, since most hedge funds live only for 3 years, linear regressions are subject to over-parameterization and small sample bias. Second, the presence of structural breaks followed by instable estimates can be a serious obstacle for the hedge fund analysis. Third, low persistence in alpha may be forced by the multicollinearity in factors. Finally, seven or eight factors of Fung and Hsieh [2001] (or Agarwal and Naik [2004]) model are not able to capture all possible strategies of hedge funds which leaves us with omitted variable bias. Each of

[1] See Bookstaber [2013].

[2] Alpha can be also interpreted as a manager skill.

the chapters is dealing with the hedge fund performance and, therefore, is concerned with some (or all) of these problems but in different ways.

In the first chapter, I propose a shrinkage methodology which handles the short sample of hedge fund returns. This methodology improves the out-of-sample accuracy of the linear factor model by combining cross-sectional and time series information for groups of hedge funds with similar investment strategies. The additional cross-sectional information allows more precise estimates of risk exposures. I also suggest a trading strategy based on this methodology for extracting substantially larger risk-adjusted returns.

The second chapter (joint work with Prof. Jens Jackwerth) is evaluating potential benefits of investing in hedge funds for pension fund managers. While alpha estimate mentioned above is a wide spread indicator of a fund profitability for most of the investors, there are other important dimensions which may be cherished too: risk diversification, addition of positive skewness, or elimination of left tails in the return distribution. These dimensions are to some extent considered in our measure - total benefit. We measure total benefit via certainty equivalent values (CEVs) which perform better than or on par with alphas or Sharpe ratios. We show that adding a random hedge fund to the typical pension fund portfolio with 10% weight increases diversification benefits. We also strongly reject that adding conventional assets like real estate, commodities, foreign equities, or mutual funds is just as beneficial as adding a random hedge fund. We further demonstrate that conditioning on past changes in CEV significantly outperforms conditioning on past alpha in times of economic stress and when returns are non-normally distributed. Finally, we show that total benefit is persistent during the sample period 1994-2009.

The last chapter (joint work with Prof. Jens Jackwerth) advocates a new measure to evaluate hedge funds - relative alpha. It links each hedge fund to a group of its peers in a straightforward, semi-parametric way. We do not require knowledge of the true factor structure. The alpha within a factor model (or "absolute" alpha) could measure current outperformance over risk-adjusted returns; it could be used to identify funds which generate high future performance and thus determine fund flow; and it could be used to find persistence in alpha. We show that relative alpha outperforms absolute alpha (e.g. based on Fung and Hsieh [2001]) along all three dimensions.

Zusammenfassung

Die Hedgefonds-Industrie hat in den letzten 15 Jahren eine rasante Entwicklung erfahren. Für einen Investor sind Hedgefonds aufgrund verschiedener Eigenschaften besonders attraktiv: Unter anderem sind optionsähnliche Renditen und eine niedrige Korrelation mit diversen Vermögensarten zu erwähnen.

Eine Einschätzung der Größe der Hedgefonds-Industrie variiert von Datenbank zu Datenbank. Ausweislich des Berichts der AIMA [2012] über Hedgefonds, hat sich das verwaltete Vermögen der Hedgefonds im Zeitraum von 1999-2007 von $456 Milliarden auf $1,868 Milliarden vervierfacht. Obwohl das Jahr 2008 von wesentlichem Finanzabfluss geprägt war[1] wurde im dritten Quartal ein Gesamtvermögen in Höhe von $2,192 Milliarden erreicht.

Zwischenzeitlich ist auch in der Wissenschaft ein gesteigertes Interesse an Hedgefonds zu erkennen. Ungeachtet bereits vorhandener, zahlreicher Untersuchungen über die Hedgefonds-Leistung, sind einige Aspekte immer noch vollkommen unerforscht. Dies scheint nicht sonderlich überraschend, da es uns ihre etlichen Besonderheiten erlauben, die Hedgefonds als sehr komplexe Investitionsvehikel zu kategorisieren.

Trotz des Umstands, dass Hedgefonds-Strategien per se äußerst komplex und hochdynamisch sind, trägt das Versäumnis einer staatlichen Reglementierung und die damit verbundene fehlende Möglichkeit die zugrunde liegen-

[1] Teilweise beruhte dies auf der Finanzkrise und dem einhergehenden Verlust, was Tilgungen verursacht hat und teilweise wegen des Bernie Madoff-Skandals, welcher den Ruf der Hedgefonds signifikant verschlechtert hat.

den Positionen aufzudecken, nichts zu unserem Verständnis bei, was Hedgefonds genau genommen bewirken.

Richard Bookstaber fasst das Konzept der Hedgefonds in einer interessanten Art und Weise zusammen und zeigt auf, dass es sehr herausfordernd sein kann, Hedgefonds zu analysieren[1].

> I believe there is no such thing as a hedge fund. Hedge funds are not a homogeneous class that can be analyzed in a consistent way. The hedge funds/alternative investment moniker is a description of what an investment fund is not rather than what it is. The universe of alternative investments is just that the universe. It encompasses all possible investment vehicles and all possible investment strategies minus the "traditional" investment funds and vehicles.

Prinzipiell sind Hedgefonds in Bezug auf Investition sehr flexibel. Sie agieren mit jedem möglichen Investitionsinstrument, mit jeder Strategie und auf jedem Markt, auf dem sie auf beträchtliche Profite unter niedrigem Risiko stoßen.

Der Gang der Dissertation soll wie folgt beschrieben werden: Das erste Kapitel analysiert die Kurzzeitleistung der Hedgefonds. Das zweite Kapitel konzentriert sich auf die Hauptgruppe der institutionellen Investoren der Hedgefonds, die Pensionsfonds, und untersucht, ob Hedgefonds tatsächlich vielversprechende Investitionsinstrumente sind. Das dritte Kapitel befasst sich mit einem neuen Maßder Hedgefonds-Leistung, dem "Relative Alpha".

Akademisch wird die Leistung der Hedgefonds grundsätzlich in Form von Alpha geschätzt im Rahmen eines Fung and Hsieh [2001] (oder Agarwal and Naik [2004]) Faktorenmodells. Angesichts der Simplizität und unkomplizierter Anwendung des Alpha entwickelten sich lineare Modelle zu einer etablierten Methodologie für die Analyse von Hedgefonds-Erträgen. Gleichzeitig, muss erwähnt werden, dass solche Modelle an diversen Schätzungsproblemen leiden. Aufgrund dessen, dass die meisten Hedgefonds nur über die Laufzeit von

[1] Bookstaber [2013].

drei Jahren existieren, neigen lineare Faktorenmodelle zum überparametrieren, zu kleinen Stichproben-Verzerrungen, zur Instabilität der Schätzungen in Gegenwart von strukturellen Brüchen und zu einer niedrigen Persistenz wegen der Mulltikollinearität von Faktoren und Verzerrungen aufgrund von ausgelassenen Variablen. Jedes Paper befasst sich mit der Hedgefonds-Leistung und ist deswegen von dieser Problematik in jeglicher Art und Weise betroffen.

In dem ersten Kapitel stelle ich eine Shrinkage-Methode vor, die sich mit der Kurzeitleistung der Hedgefonds befasst. Ich verbessere die Vorausschätzungen der linearen Modelle, indem ich Querschnittsinformationen mit Zeitreiheninformationen für die Gruppen der Hedgefonds mit ähnlichen Investitionsstrategien kombiniere. Zusätzliche Querschnittsinformationen erlauben es mir präzise Schätzungen der Regressionskoeffizienten vorzunehmen. Auf Basis dieser Methode bringe ich auch eine Handelsstrategie für die Förderung von deutlich größeren risiko-adjüstierten Erträgen ein.

Das zweite Kapitel wurde in Zusammenarbeit mit Prof. Jens Jackwerth entwickelt. In diesem wird die Einschätzung des potentiellen Nutzens der Investition in Hedgefonds für die Hauptgruppe der institutionellen Investoren Pensionsfonds behandelt. Obwohl das obengenannte Alpha im Rahmen diverser linearer Faktorenmodelle ein wichtiger Indikator der Hedgefonds Profitabilität für die meisten Investoren ist, existieren auch andere bedeutende Dimensionen die auch geschätzt werden: Risiko Diversifikation, Addition von positiver Asymetrie, Behebung von Leptokurtosis in der Verteilung. Diese Dimensionen sind zu einem gewissen Ausmaßin Total Benefit berücksichtigt. Wir messen den Total Benefit mithilfe von Sicherheitsäquivalenten (CEVs), welche besser oder auf einer Stufe mit Alpha oder Sharpe Ratio sind. Wir zeigen, dass das Ergänzen eines zufälligen Hedgefonds zu einem typischem Pensionsfond Portfolio mit einem Gewicht von 10%, den Total Benefit erhöht. Wir verwerfen die Hypothese, dass traditionelle Vermögenswerte wie Grundstücke, Rohstoffe, Auslandsbeteiligungen, oder Investmentfonds genauso vorteilhaft sind wie ein zufälliger Hedgefond. Wir illustrieren ferner, dass das CEV das Alpha in Zeiten der Wirtschaftskrise und wenn die Erträge nicht normal verteilt sind, signifikant überbietet. Schließlich zeigen wir, dass der Total Benefit in der Zeitspanne von 1994-2000 persistent ist.

Das letzte Kapitel, in Zusammenarbeit mit Prof. Jens Jackwerth fördert ein neues Maß für die Evaluation der Hedgefonds - "Relative Alpha". Es verbindet jedes Hedgefond mit einer Gruppe von seinesgleichen auf eine direkte, semi-parametrische Art. Wir setzen kein Wissen der wahren Faktorenmodellestruktur voraus. Das Alpha im Rahmen des Faktorenmodells, auch "Absolute Alpha" genannt, könnte die laufende risiko-adjustierte Erträge messen; Es könnte auch Fonds identifizieren, die die Hohe zukünftiger Erträge generieren und deswegen Kapitalfluss bestimmen. Ferner könnte es verwendet werden, um die Persistenz in der Lesitung zu untersuchen. Wir zeigen, dass das Relative Alpha das Absolute Alpha (e.g. basierend auf Fung and Hsieh [2001]) auf allen dieser drei Dimensionen überbietet.

Short-term hedge fund performance

1.1 Introduction

The hedge fund industry has grown quickly over the last two decades. It is not surprising that both practitioners and academic researchers are interested in understanding hedge funds. Despite various attempts to explain hedge fund returns, simple linear factor models are still the most commonly used. Formerly, factors from the CAPM, Fama and French [1992], and Carhart [1997] models were applied to hedge funds[1]. This ad-hoc approach cannot take into account all the peculiarities of hedge funds. Factors extracted directly from hedge fund returns are more specific and, as empirical evidence suggests, have stronger explanatory power[2]. Straightforward implementation of linear models has made them popular tools in the investigation of hedge fund performance. However, risk exposures estimated by these models are unstable in small samples. Since most hedge funds have a life span of only 30-40 months, linear models become inappropriate. This leads to poor forecasting power and a low probability of picking the best hedge fund performers for an investor.

To overcome the problem of a short sample and the resultant unreliability of estimates, it is possible to combine the cross-section with the time series, i.e. to operate with panel data. I group hedge funds by their investment strategy. This approach yields reliable estimates even when the time series are not lengthy. The main idea behind this approach is that the hedge funds following the same investment style are comparable in terms of the magnitude of risk exposures. I

[1]See Hasanhodzic and Lo [2006].

[2]See Agarwal and Naik [2004], and Fung and Hsieh [2001].

examine several panel data methods and show empirically their superior forecasting abilities over conventional linear factor models. Root mean squared prediction error in panel data models is monthly 10-15% smaller than in linear regressions, and the rate of diminishing is significant.

Forecasting power is directly related to the persistence concept. Although the problem existed for years, there is no clear-cut answer as to whether fund returns persist over time. For example, Brown et al. [1999] use raw as well as risk-adjusted returns from the CAPM, and excess returns over the style benchmarks to show little performance persistence in hedge funds. On the contrary, Agarwal and Naik [2000a] reveal substantial persistence in quarter returns using excess returns over the average style-return and (non-)parametric tests. They also find that ”winners” are more persistent than ”losers”. Significant persistence was found by Edwards and Caglayan [2001] for both ”winners” and ”losers”. Capocci and Hübner [2004] apply the four-factor model of Carhart [1997] and find no persistence among either ”winners” or among ”losers”, but limited evidence of persistence in returns of the middle decile funds. More recently, Kosowski et al. [2007] applied Bayesian econometrics and a bootstrap procedure to evaluate hedge fund performance. They find that hedge fund returns persist over a one-year horizon. While some empirical studies reveal a ”reasonable degree” of persistence, the others do not find any predictability in returns (Kat and Menexe [2003]) or find the time-varying persistence (Capocci [2002]). Though being not well explained yet, the persistence is crucial for investors hoping to include ”winner” funds into their portfolios. Persistent returns make funds with good past performance particularly attractive and can be regarded as a motivation for a fund manager.

Controversial results on the performance persistence can be explained by a number of biases existing in the data. The panel data methods solve some of these biases and provide better estimates of an alpha-parameter comparing to the traditional linear factor models which are used in the majority of persistence studies. As a result, the methodology proposed in this study allows estimating alpha which demonstrates higher persistence.

I also suggest investment strategy which gives above-average compensation: the Sharpe ratio is 18% higher than the average Sharpe ratio in the sample.

The remainder of the paper is organized in the following way: section 1.2 presents econometric methods based on the panel data. Section 1.3 describes the data used in the analysis. As the core of the paper, section 1.4 applies proposed methods and tests their performance out-of-sample. Section 1.5 summarizes the main findings of the paper and discusses the remaining drawbacks as well as their possible solutions.

1.2 Model Specification

Small sample biases inevitably affect results of the linear regressions applied to hedge fund returns. Indeed, the history of hedge funds is rather short. From Table 1.B.1 it can be seen that an average life span in the data is around 4.18 years (global macro hedge funds) - 5.24 years (distressed hedge funds). However, as the mode of the Kernel distribution of hedge fund ages suggests, most funds do not live longer than 3.5 years (Figure 1.C.1).

[Table 1.B.1 about here]

[Figure 1.C.1 about here]

A short sample leads to unstable estimates, i.e. adding a few more observations may change the estimates substantially. In order to acquire additional information on returns and "increase" the sample, I use panel data methods. I consider hedge funds within one investment style as panel data with a large number of cross-section elements and not lengthy time series. Here, the classic question - whether it is better to pool the slope estimates, or retain heterogeneous ones - also becomes relevant. Traditionally, this problem does not have a unique answer. Pesaran and Smith [1995] and Pesaran et al. [1996] find that the aggregation of the slope coefficients does not necessarily provide better results; instead they recommend using heterogeneous slopes. Contrarily, Maddala et al. [1997] show that heterogeneous slope estimates can be severely biased and, therefore, be incorrectly interpreted. A compromise between heterogeneous and homogeneous estimates - the shrinkage - becomes especially useful when the number of cross-section elements (hedge funds) is large while the time series cannot be considered either too short (to make the homogeneous model the only one possible alternative) or too long (to be sure of the appropriateness of the heterogeneous estimates).

Here, I would like to analyze and compare the following models: homogeneous panel, heterogeneous panel, and shrinkage panel. In the homogeneous panel (hereafter "common mean") all hedge funds pursuing similar investment strategies have risk exposures of the same magnitudes. The coefficients are relatively stable due to the strict restrictions. However, the restrictions are not necessarily supported by the data: it is doubtful that all hedge funds within the same style demonstrate equal risk exposures. On the contrary, in the heterogeneous panel (hereafter "individual") all hedge funds have risk exposures of different magnitudes. No restrictions are imposed, but coefficients in small samples are highly unstable. Finally, the shrinkage panel combines the advantages of the individual and the common mean estimates and overcomes some of their disadvantages. This method is based on the idea that all hedge funds with the same investment style have risk exposures of different magnitudes which fluctuate around some common mean. The shrinkage can be considered a convex combination of the individual and the common mean estimates. Restrictions imposed by the shrinkage panel are not as heavy as in the case of the common mean model. Moreover, hedge funds within the same style follow similar allocation and, hence, tend to have similar but not necessarily identical exposures which make shrinkage restrictions plausible. Indeed, if one calculates Euclidean distances between the hedge fund risk exposures[1], one can no-

[1]If one fund has a vector of risk exposures $\{\beta_{11}, \beta_{21}, \cdots, \beta_{K1}\}$ and the other fund has exposures denoted as $\{\beta_{12}, \beta_{22}, \cdots, \beta_{K2}\}$ then Euclidean distance between two funds is computed as

tice that the distances between hedge funds which belong to the same style on average are smaller than the distances between the hedge funds from different styles (see Table 1.B.2). Estimated in this way, coefficients are generally more stable than in the case of the individual estimates[1]. Depending on the calculation procedure for the weighting matrix (coefficients of a convex combination), estimates are further called either "naive shrinkage" or "advanced shrinkage".

[Table 1.B.2 about here]

1.2.1 Common mean

The common mean model is based on the extreme assumption that risk exposures of hedge funds from the same cross-sectional group are equal between each other.

Let R_i be a $T \times 1$ vector of monthly hedge fund returns ($i = 1, \ldots, N$), F_i is a $T \times k$ matrix of risk factors (including the constant), ε_i is a $T \times 1$ vector and i.i.d. normally distributed. Then the common mean model is defined in the following way:

$$R_i = F_i \beta_{mean} + \varepsilon_i, \tag{1.1}$$

i.e. the vector of risk exposures ($k \times 1$ β_{mean}) is the same for all hedge funds.

To establish whether the data support the model[2], I test the following hypothesis[3]:

$\sqrt{(\beta_{11}-\beta_{12})^2+(\beta_{21}-\beta_{22})^2+\cdots+(\beta_{K1}-\beta_{K2})^2} = \sqrt{\sum_{k=1}^{K}(\beta_{k1}-\beta_{k2})^2}$.

[1] In general, the shrinkage estimator has superior properties which is shown theoretically in Maddala et al. [2001].

[2] i.e. to check if all funds with the same investment style have exactly the same risk exposures to the factors.

[3] For details on the test see Appendix 1.A.

$$H_0 : \beta_i = \beta_{mean}, \forall i$$

Imposing the restrictions on the risk parameters (under H_0) reduces the variance of the estimates. However, if the restrictions are not valid, computed estimates can be severely biased. That is the reason the issue of pooling the estimates or not becomes crucial.

As alternatives to the common mean estimates described, one could use simple ($\beta_{mean} = 1/N \sum_{i=1}^{N} \beta_i$) or weighted ($\beta_{mean} = \sum_{i=1}^{N} \omega_i \beta_i$) averages of the individual estimates proposed by Swamy [1970][1].

1.2.2 Individual estimates (the benchmark)

Rejecting the hypothesis ($H_0 : \beta_i = \beta_{mean}$) leads to the problem of the appropriate individual estimates. The model with absolutely heterogeneous estimates does not put any restrictions on the slope coefficients; neither does it use any information from the cross-section. Relative to the common mean model, the heterogeneous panel can be placed exactly at the opposite end of the spectrum of panel models examined in this paper.

Using the notation introduced above, the model can be written as follows[2]:

$$R_i = F_i \beta_i + \varepsilon_i, \quad \forall i = 1, \ldots, N \tag{1.2}$$

[1] For details of the derivation see Appendix 1.A.

[2] This is the specification generally used in the hedge fund literature.

1.2.3 Shrinkage

It is not always obvious whether it is better to pool the data and obtain a single estimator for all hedge funds within the same investment style (common mean estimate), or to estimate each series separately (individual estimates). Either way, major assumptions are implied: the risk parameters are the same for all hedge funds in the case of the pooling, or different for each fund with equation-by-equation estimation. A compromise between the fully heterogeneous model and the pooled one is the shrinkage estimator. As a common mean model, it also takes advantage of the panel data information across the cross-section and the time series. However, it does not ignore the individual estimates. It is naturally to think that there is certain similarity between funds which belong to the same investment style. This is where the idea of shrinkage comes from. It is a weighted average between a common mean estimate and the individual estimate of a parameter. Since both parts of this estimator include relevant information, this may decrease the estimation error and, therefore, improve the accuracy of the individual estimates. Generally, a shrinkage estimator is a biased estimator but has a smaller variance (see Figure 1.C.2). The naive shrinkage estimator used in this study does not only have a smaller deviation, but it is also unbiased[1].

[Figure 1.C.2 about here]

Copas [1983], Rao [1975], Smith [1973], and Rubin [1980] claim the out-of-sample superiority of the shrinkage estimator over the single or pooled estimators. The shrinkage methodology allows the incorporation of additional information in the form of a hedge fund investment style in order to improve the out-of-sample properties of the model.

Shrinkage has been used in finance quite often, mainly in portfolio analysis to improve the estimates of the second moments (see Jorion [1986] and Ledoit

[1] For the theoretical proof see Maddala et al. [2001].

and Wolf). Much less is done in the area of returns prediction. The pioneers in this field are Blume [1971] and Vasicek [1973]. The methods they advocate are far away from the methodology used in this study. Nevertheless, there is some similarity as Blume [1971] and Vasicek [1973] also try to improve the quality of the beta estimates from the CAPM model using the information from the cross-section dimension.

The shrinkage estimator is derived within a Bayesian framework. This part of the study is also related to the works which apply Bayesian methodology to investigate hedge fund performance[1]. The crucial aspect of the model applied here is the way in which hedge fund risk exposures fluctuate around some common mean. The same assumption also refers to the fund alpha. This implies that significantly positive alphas are presumably overestimated, whereas significantly negative alphas are more probably underestimated. Thus, shrinking the estimators to the common mean decreases the estimation error.

Simplified, the shrinkage can be presented as:

$$\hat{\beta}_i^* = (I - P_i)\hat{\beta}_i + P_i\hat{\beta}_{mean},$$

where I is identity matrix, P_i is some weighting matrix, $\hat{\beta}_i$ is individual estimate, and $\hat{\beta}_{mean}$ is the estimate of the common mean.

To evaluate P_i, there exist two ways: naive shrinkage estimator which weights individual estimates according to their accuracy and advanced shrinkage which uses cross-validation as the method for obtaining the weights.

In the naive shrinkage $\hat{\beta}_i$ and $\hat{\beta}_{mean}$ are weighted according to uncertainty, i.e. the higher the variance of the individual estimate, the higher the weighting matrix (P_i) and, therefore, the less is the weight of the individual estimate ($\hat{\beta}_i$):

[1] See Baks et al. [1999], Pastor and Stambaugh [2002], and Kosowski et al. [2007].

$$\hat{\beta}_i^* = (\frac{1}{\hat{\sigma}^2{}_i} F_i' F_i + \hat{\Theta}^{-1})^{-1} (\frac{1}{\hat{\sigma}^2{}_i} F_i' F_i \hat{\beta}_i + \hat{\Theta}^{-1} \hat{\mu})$$
$$\hat{\beta}_i^* = (I - P_i)\hat{\beta}_i + P_i \hat{\beta}_{mean},$$

where $P_i = (\hat{V}_i^{-1} + \hat{\Theta}^{-1})^{-1} \hat{\Theta}^{-1}$, $\hat{V}_i^{-1} = \frac{1}{\hat{\sigma}^2{}_i} F_i' F_i$.

Instead of weighting by uncertainty, one could use cross-validation. To obtain cross-validation weights, it is necessary to fit the model to the data $R_{i1}, \cdots, R_{it}$ ($\forall i = 1, \cdots, N$), forecast the next observation ($\hat{R}_{it+1}$) and compute the prediction error ($e_{it+1} = R_{it+1} - \hat{R}_{it+1}$). This procedure is repeated for $t = T_1, \ldots, T_2$ where T_1 is the minimum number of observations needed for fitting the model. The weighting which minimizes the errors is chosen for the further analysis. Estimates obtained by cross-validation are further called ”advanced shrinkage”.

The remaining issue is the computation of standard errors. I apply wild bootstrap procedure (Liu [1988]). The main difference of the wild bootstrap from the standard resampling approach is that each residual is multiplied by a random variable. This methodology allows the relaxing of the homoscedasticity assumption in the original residuals and the replication of the pattern of heteroscedasticity, which makes it more favorable over simple residual sampling for smaller sample sizes.

1.3 Data

The data is merged from several databases (Barclayhedge, Eurekahedge, TASS, HFR, CISDM, and Altvest (as used in Hodder et al. [2013]). It involves funds with monthly returns evaluated over the period from January 1994 to May 2009. Summary statistics for this sample are in Table 1.B.3. For the main runs all funds with history shorter than 36 months are deleted from the sample[1]. Funds are chosen from nine investment styles: equity long/short, fixed income, global macro, event driven, distressed, managed futures, emerging markets, multi-strategy, and arbitrage. Summary statistics of the remaining from the full sample funds after deletion (Table 1.B.4) reveal that fund returns are more skewed and have fatter tails than normally distributed returns.

[Table 1.B.3 about here]

[Table 1.B.4 about here]

While this is typical to use 36 months in the hedge fund analysis, it is important to emphasize that shrinkage models can also be applied to smaller sample sizes. I also provide results for the funds with performance history shorter than 36 months.

There are several important features of hedge funds which make them substantially different from any other investment tool. In evaluating the performance of hedge funds I try to control for most of these features.

To begin with, the absence of strict governmental regulation makes the decision of a fund manager to enter the database completely voluntary. It is natural to suppose that the ultimate ”winners” and ”losers” are not included in the database. For the funds with extremely high returns there is no incentive to advertise themselves additionally, while the funds with very poor performance may not be interested in revealing their returns. Apart from that, the databases commonly suffer from the survivorship bias and instant-history bias. The presence of these biases suggests that the future returns will be lower than the historical ones. The survivorship bias is avoided since the database contains

[1]Further, for each hedge fund first 12 observations are deleted due to instant-history bias considerations.

information on dead funds, and instant-history bias is accounted for by deleting the first 12 months of returns.

It is also known that hedge funds employ rather complex investment strategies. Standard factors do not explain non-linear payoffs (option-like features) of fund returns [1]. The usual procedure in this case is to include the seven factors extracted by Fung and Hsieh [2001]. The factors were downloaded from David A. Hsieh's Hedge Fund Data Library [2] and include equity market factor (S&P 500 index), the size spread factor (Wilshire 1750 Small Cap-Wilshire 750 Large Cap indexes), the bond market factor (change in ten-year treasury yields), the credit spread factor (yield spread between ten-year treasury and Moody's Baa bonds), and three trend-following factors: portfolio of bond options, portfolio of commodity options, and portfolio of currency options.

The vast usage of illiquid assets by hedge funds is another source of biases in the data. The current value of these assets is approximated by the past prices. This results in the significant serial correlation in fund returns. Getmansky et al. [2004] show that true (unobserved) returns ($R^{un}_{i,t-l}$, $l = 1, \cdots, s$) are uncorrelated and related to the reported returns ($R_{i,t}$) through the following adjustment mechanism:

$$
\begin{aligned}
R_{i,t} &= \theta^i_0 R^{un}_{i,t} + \theta^i_1 R^{un}_{i,t-1} + \cdots + \theta^i_s R^{un}_{i,t-s} \\
1 &= \theta^i_0 + \theta^i_1 + \cdots + \theta^i_s \\
\theta_l &\in [0,1], \quad l = 1, \cdots, s
\end{aligned}
$$

It was shown by Getmansky et al. [2004] that de-meaned observed returns follow an MA(s) process and empirically MA of order $s = 2$ explains the correlation in returns best. I estimate θ_0, θ_1, and θ_2 by maximum likelihood to receive

[1] See Fung and Hsieh [2001].

[2] https://faculty.fuqua.duke.edu/ dah7/HFData.htm

de-smoothed returns and include them into the further performance models. Indexes composed from equally weighted returns of various hedge fund styles demonstrate significant auto-correlation up to order 2 (Figure 1.C.3). After de-smoothing, there is almost no autocorrelation in returns (Figure 1.C.4).

[Figure 1.C.3 about here]

[Figure 1.C.4 about here]

Since shrinkage method proposed here requires a decent panel data with enough information about the investment strategy within style, it is impossible to use the whole sample (as only very few funds from each investment style survive through 1994 - 2009). Further results are based on the sample July 2003 - June 2007.

The number of observations for each style is the same for all competing models. The funds which do not survive through the sample period are deleted. Perhaps, the panel data models suffer more from this deletion as they utilize the information from the cross-section. Even despite this fact, panel data models perform better than the model with individual estimates. To test how robust the performance to the chosen sample, I consider sequences of samples starting in January 1996 with a 2-year gap between the starting dates, for the in-sample sizes of 36 months and 12 months. I also present results for different forecasting horizons.

1.4 Empirical Results

Here, I apply the methods introduced in section 1.2. To compare these methods I use out-of-sample forecasting as well as analyze alpha persistence.

1.4.1 Return prediction

Firstly, I test the poolability of the data. Since the p-value of the common mean test described in the previous section is below 0.1, the null hypothesis $H_0 : \beta_i = \beta_{mean}$ $(\forall i = 1, \dots, k)$ is rejected. The data does not support the joint restrictions on the slope coefficients, and, therefore, instead of pooling the coefficients, I use weighted average of individual exposures in the further analysis[1].

Second, I compute the average distance between the true return and its prediction, the root mean squared prediction error (RMSPE), in order to obtain the preliminary comparison of the forecasting power of the models:

$$RMSPE = \sqrt{\frac{1}{T'} \sum_{t=1}^{T'} (\hat{R}_{it+1|t} - R_{it+1})^2} = \sqrt{\frac{1}{T'} \sum_{t=1}^{T'} e^2_{it+1|t}},$$

where T' is the size of forecast sample, $\hat{R}_{it+1|t}$ is predicted return in $t+1$ based on the information set in t, R_{it+1} is actual return in $t+1$, and $e_{it+1|t} = \hat{R}_{it+1|t} - R_{it+1}$ is the prediction error. In the following example $T' = 12$ months, the in-sample size $T = 36$ months[2].

Table 1.B.5 compares models on the basis of their predictive power. It can be seen that for 80-90% of hedge funds the naive shrinkage estimator provides a better forecast compared to the traditional individual estimates (see Figure 1.C.5). Even though for 10-20% of funds individual estimates show better results, the overall improvement is still significant. However, the overall decrease in prediction error differs for different fund styles. In absolute values, this improvement equals to 0.03% (for equity long/short style) - 0.18% (for managed futures style) per month or 0.36% and 2.16% per annum respectively. Taking into consideration that the average return in the full sample is around 0.9-1%,

[1]Simple average estimates provide slightly worse results.
[2]See discussion on other out-of-/in-sample sizes section 1.5.

the enhancement in predictive power can also be regarded as being economically significant. Only for the distressed and arbitrage strategies are the results in favor of individual estimates. This is probably due to the small number of hedge funds available in the cross-section of distressed and arbitrage funds[1]. Advanced shrinkage delivers lower average RMSPE: improvement comparing to the individual estimates is 0.02% (for arbitrage strategy) - 0.39% (for equity long/short strategy) but the percentage of funds with improved is also slightly lower (60-70%). Lowest improvement in forecasting power is shown by distressed and arbitrage funds. For a more detailed comparison, it is useful to refer to Figure 1.C.6, which plots the difference between prediction error of the advanced shrinkage model and the model with individual estimates in the example of global macro strategy ($\Delta RMSPE = RMSPE_{advanced \quad shrinkage} - RMSPE_{individual}$)[2]. The bar-plot demonstrates that almost for each fund within global macro style naive shrinkage provides a better forecast. The Kernel density also shows that the mode of the distribution of $\Delta RMSPE$ is negative. Moreover, most of the distribution mass lies to the left of zero. The comparison between all models is graphically depicted on Figure 1.C.7. Even though the advanced shrinkage model has a smaller prediction error compared to all other models, it also has smaller percentage of funds with an improved forecast (69% instead of 87% provided by the naive shrinkage).

[Table 1.B.5 about here]

[Figure 1.C.5 about here]

[Figure 1.C.6 about here]

[Figure 1.C.7 about here]

As it was already mentioned in section 1.3, in the hedge fund literature the commonly used number of observations is equal to 36 months. However, the

[1]Only 53 funds for distressed strategy and 45 funds for arbitrage strategy which is comparable to the length of time series.

[2]Other styles provide similiar results.

main advantage of the shrinkage is more pronounced for shorter samples. Table 1.B.6 and 1.B.7 demonstrate the root mean square prediction error for the models estimated on 24 and 12 monthly observations respectively. I use the same out-of-sample period as before (July 2006-June 2007). The shrinkage models with the in-sample size of 24 months tend to outperform the model with individual estimates more frequently (i.e. for larger percentage of hedge funds). Further decrease in the in-sample size to 12 months gives even larger improvement in the forecasting power. For almost all investment styles the improvement is observed for 94-98% of funds. This result is two-fold. On the one hand, reducing the sample size leads to the increase in the cross-sectional number of funds. On the other hand, smaller sample deteriorates more the individual estimates. Moreover, the deterioration of the individual estimates seems to be more important than the increased cross-sectional dimension because even for small cross-sections (e.g. arbitrage style) shrinkage outperforms individual estimates.

[Table 1.B.6 about here]

[Table 1.B.7 about here]

To understand better the mechanism of shrinkage, it is helpful to see how the prediction power evolves with different numbers of funds within style (Figure 1.C.8) as well as varying lengths of time series (Figure 1.C.9). Figure 1.C.8 shows that an increase in the number of hedge funds in the cross-section leads to a decrease in prediction error. Indeed, with an increased number of cross-section elements, more information is available in the shrinkage model compared to the traditional linear factor model and therefore more precise estimates can be obtained. As it was mentioned in the section 1.2, shrinkage is a combination of two estimates - the individual and the common mean. Individual estimates cannot be changed by using a larger cross-section. However, this information is useful for the estimation of the common mean and, as a consequence, for shrinkage. On the contrary, Figure 1.C.9 demonstrates that with increasing length of the time series the improvement in the forecasting power of shrinkage diminishes.

This is not surprising, as with more data on time series, individual estimates are more stable.

[Figure 1.C.8 about here]

[Figure 1.C.9 about here]

1.4.2 Alpha persistence

A common measure which allows the evaluation of the performance of hedge funds is the intercept term (alpha) in a factor model. To judge its persistence, generally, two periods are observed: a formation period and an evaluation period. To test for persistence in hedge fund returns, I estimate alphas in the formation period (α_{0i}) and in the evaluation period (α_{1i}) for the individual estimate and the naive shrinkage and calculate Kendall tau rank correlation coefficient (Kendall [1938]):

$$\tau = \frac{N_c - N_d}{1/2N(N-1)}, \tag{1.3}$$

where N_c - number of concordant pairs, N_d - number of discordant pairs, N - total number of ranked funds.

For persistence to be valid, the coefficient τ must be significantly different from zero. The results of the persistence analysis (1.3) for both naive shrinkage and individual estimates are summarized in Table 1.B.23. Kendall rank correlation for the shrinkage model is higher than for individual estimate of alpha. Rank coefficient is much smaller for 3-year evaluation period than for 2-year period.

[Table 1.B.23 about here]

1.4.3 Investment Strategy

The main goal of investors is to identify talented funds with persistent positive returns. To extract these funds from the hedge fund universe is challenging. Sun et al. [2012] reveal that the funds whose returns differ substantially from returns of other funds are able to deliver significantly better future performance. These funds are considered as more innovative. In their paper, Sun et al. [2012] use Strategy Distinctive Index (SDI) to quantify hedge fund uniqueness. Following this idea, I select funds based on their deviations from the common mean. Funds with the largest deviations are of particular interest. Large deviation would mean that these funds follow the allocation differently from their peers. Table 1.B.25 exhibits average decile alphas of funds sorted by deviation from the common mean[1]. The funds with largest deviations (top decile) have significantly larger spread of alpha (large positive alpha as well as large negative alpha). However, average alpha for these funds is positive and equals 0.018. To compare, average alpha in the whole sample equals 0.009[2].

[Table 1.B.24 about here]

[Table 1.B.25 about here]

To test if the funds with largest deviations are indeed skillful, it is important to look at their out-of-sample performance Table 1.B.24 demonstrates preliminary results on rank coefficients in the top and bottom funds. It shows rank correlation in the funds with largest and smallest deviations. I compare these correlations with the persistence in funds sorted on largest/smallest alpha. For evaluation period of 24 months top/bottom funds ranked on their deviation from the common mean are more persistent than top/bottom top/bottom alpha funds. This result holds also for the bottom funds in the evaluation period of 36 months. However, neither top funds with the largest deviation not top

[1] To compute the deviation I use the Euclidean distance between vectors of risk exposures excluding alpha.

[2] Here, I use again shorter sample and the same starting date as before June 2003, 36 months in-sample and 36 months out-of-sample.

alpha-funds show persistence in 36-months evaluation period.

[Table 1.B.24 about here]

To see how large the performance numbers are, I compare the Sharpe ratios of hedge funds sorted on largest deviations with the performance of the funds in the whole sample.

- Sharpe ratios in the funds sorted out according to the deviation strategy are on average by 18% larger than Sharpe ratios of the funds in the whole sample.
- Sharpe ratios of funds within largest-deviations strategy are higher than in the full sample both in funds with largest and smallest deviations from the common mean.
- Threshold separating hedge fund into top and bottom areas do not play a role in the results stated above.

Even though funds implementing investment strategies different from their peers outperform average fund in the sample, they show worse performance than funds with strategies similar to common mean. Despite of the fact that these funds include funds with the highest alpha estimates, they do not necessarily show these alphas out-of-sample. One of the possible explanation here is that alphas of the funds with specific strategies do not capture the skills but rather biased due to the omitted variables. Obviously, seven Fung-Hsieh factors can not represent the strategy of each single fund. Instead, these factors are more suitable for a an average fund. Therefore, outstanding in terms of their investment styles funds may not possess true above-average alpha. To extract skilled funds out of the funds with outstanding strategies, the sample of these funds should be restricted further but on the base of other than alpha-measure[1].

[1] The measure which is less biased.

1.5 Robustness

Here I propose several robustness checks of the main findings. First, it is worth mentioning that the affiliation of the fund to a particular style is crucial for the modeling. Meanwhile, the fact that funds may diverge from the announced style is well-known in the literature and occurs due to several reasons (e.g. chasing superior performance by managers or change of the fund's management). Since funds are grouped according to their styles, it is necessary to check for alternative grouping measures. Instead of the self-reported investment strategy, I apply cluster analysis to identify the hedge funds with the similar risk exposures. More precisely, I use Spearman rank-order correlation coefficients as the measure of the similarity between the funds[1]. The funds with the smallest deviation from each other in terms of the Spearman rank-order correlation coefficient are grouped into one cluster. Styles do not match perfectly. However, most styles are concentrated in one or two clusters. The results are presented in Table 1.B.8 and stay consistent with the results reported in the section 1.4.

[Table 1.B.8 about here]

Second, I use different out-of-sample sizes[2]. Instead of 12 months out-of-sample, I implement the competing models with 24, 18, and 6 out-of-sample months (Tables 1.B.9-1.B.11). The results stay mainly consistent.

[Table 1.B.9 about here]

[Table 1.B.10 about here]

[Table 1.B.11 about here]

[1] I choose this measure as it is considered to be robust against the outliers (Bickel [2003]).
[2] Different in-sample sizes are discussed previously in section 1.4.1.

The results stay robust with respect to the choice of the sample period.[1]. I provide the following samples: January 1996-December 1999, January 1998-December 2001, January 2000-December 2003, January 2002-December 2005, January 2004-December 2007 (Tables 1.B.12-1.B.16). The last sample (January 2004-December 2007) is of a particular interest since out-of-sample period includes the recent financial crisis. The presence of the structural break worsens the forecasting performance of each model. The root mean squared prediction error is approximately by 1-2% larger comparing to no-crisis period (e.g. July 2003-June 2007 in the section 1.4). Also the difference between the performance of the shrinkage models and the model with individual estimates is tightening. Nevertheless, this difference is still significant and the shrinkage model allows to improve the forecasts in 70-80% of all hedge funds within the style[2].

[Table 1.B.12 about here]

[Table 1.B.13 about here]

[Table 1.B.14 about here]

[Table 1.B.15 about here]

[Table 1.B.16 about here]

Applying different periods of a smaller in-sample size also gives significant outperformance of the shrinkage over the model with individual estimates. I again provide the sequence of the time periods starting in January 1996 with a 2-year gap between the starting dates and the estimation is based on 12 monthly observations: January 1996-December 1997, January 1998-December 1999, January 2000-December 2001, January 2002-December 2003, January 2004-December 2005, January 2006-December 2007 (see Tables 1.B.17-1.B.22). Shorter

[1]Recall that for the shrinkage to be valid substantial number of cross-section elements must be available in the sample. For the investigated nine styles, the peak of fund numbers is reached towards middle/end of the full sample.

[2]Except for distressed and arbitrage styles with small numbers of hedge funds.

in-sample period demonstrates better results for the shrinkage model. The number of funds for which the shrinkage delivers superior out-of-sample performance is constantly around 90%.

[Table 1.B.17 about here]

[Table 1.B.18 about here]

[Table 1.B.19 about here]

[Table 1.B.20 about here]

[Table 1.B.21 about here]

[Table 1.B.22 about here]

Finally, apart from Fung and Hsieh [2001] model, I also consider the CAPM, Agarwal and Naik [2004], and Fama and French [1992] models. All of these are linear factor models and work on the same principle as the model of Fung and Hsieh [2001] with the difference that Agarwal and Naik [2004] use factors extracted from fund returns while Fama and French [1992] and the CAPM do not account for non-linearities and illiquidity in fund returns and are more suitable for conventional investment instruments such as stocks, bonds, and equities. The results do not change significantly.

1.6 Conclusion

This paper evaluates the short-term performance of hedge funds. The problem of the short historical data is particularly important for hedge fund area where average life expectancy of a fund is not long. Oft-used factor models appear to be over-parameterized which leads to unstable estimates. To overcome this problem, I suggest complementing the time series with the cross-section, i.e.

considering funds with similar strategies together. This gives additional information which stabilizes - at least partly - the estimates of the factor models.

I use shrinkage as the alternative to traditional factor models. The shrinkage estimate is a trade-off between the individual estimates and the so-called common mean estimate, which is the estimate of the average risk exposure of a particular hedge fund style. Shrinkage provides more accurate estimates of the risk exposures and therefore better forecasting power and more insight into the persistence of hedge funds. The improvement depends on a number of factors. First, for the shrinkage to work well, there must be substantial number of funds with the same investment style. As long as the length of time series is comparable with the number of cross-section elements, the advantages of the shrinkage vanish. Second, the funds within the same style must possess some degree of homogeneity. In other words, by and large funds should follow similar allocations. Groups of funds with totally different allocations do not deliver meaningful estimates of the common mean.

An estimate of the common mean shows average risk exposures within the investment style. A small share of funds with large deviations demonstrate an average alpha which is significantly larger than an average alpha in the whole sample. These funds also demonstrate good future performance, i.e. they show above-average Sharpe ratios. However, funds with smallest deviations show even better performance characteristics. Above-average alpha of funds deviating from the common mean can be result of a bias. It is important to find the measure which is less biased than alpha and which allows extracting truly good funds out of those who deviate.

A remaining drawback of the methods examined in this paper is that they assume constant parameters over time. There is a large body of the literature dealing with the dynamic features of hedge funds. Meanwhile, it has become clear that hedge fund risk exposures are also subject to change (see Racicot and Théoret [2009], Chan et al. [2006], Billio et al. [2010], Blazsek and Downarowitz [2008], and Bollen and Whaley [2009]). Since already constant coefficient models such as the Fung and Hsieh [2001] seven factor model often appear to be over-parameterized; adding time-varying parameters makes this problem even worse. A possible solution to this problem is the combination of so-called state-space models which allow for continuous time-variation in the risk exposures and the shrinkage estimator.

1.A Appendix: Proofs

1A.1 Common mean: testing the hypothesis on joint significance

Under H_0, the pooled model (restricted one) can be rewritten:

$$R = F\beta_{mean} + \varepsilon = \begin{pmatrix} F_1 & 0 & \dots & 0 \\ 0 & F_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & F_N \end{pmatrix} \begin{pmatrix} \beta_{mean} \\ \beta_{mean} \\ \vdots \\ \beta_{mean} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{pmatrix}$$

whereas unrestricted model is given by:

$$R = F\tilde{\beta} + \varepsilon = \begin{pmatrix} F_1 & 0 & \dots & 0 \\ 0 & F_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & F_N \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{pmatrix}$$

Under the assumption that $\varepsilon \sim N(0, \Omega)$, $\Omega = \sigma_\varepsilon^2 \Sigma$, generalized least squared estimator (GLS) for the unrestricted beta-parameters is:

$$\hat{\hat{\beta}} = (F'\Omega^{-1}F)^{-1}F'\Omega^{-1}R. \tag{1.A.1}$$

Restricted GLS is then derived as:

$$\hat{\beta} = \hat{\hat{\beta}} + (F'\Omega^{-1}F)^{-1}S'(S(F'\Omega^{-1}F)S')^{-1}(s - S\hat{\hat{\beta}}), \tag{1.A.2}$$

where S is given matrix of restrictions so that the restriction under H_0 can be rewritten as $S\beta = s$.

The feasible GLS estimator is obtained by substituting the consistent estimate of Ω in Equations (1.A.1) and (1.A.2).

Finally, the standard Chow test statistic is calculated as:

$$F_{test} = \frac{R'(\Sigma^{-1}(\tilde{F}(\tilde{F}'\Sigma^{-1}\tilde{F})^{-1}\tilde{F}' - F(F'\Sigma^{-1}F)^{-1}F')\Sigma^{-1})R/(N-1)k}{(R'\Sigma^{-1}R - R'\Sigma^{-1}\tilde{F}(\tilde{F}'\Sigma^{-1}\tilde{F})^{-1}\tilde{F}'\sigma^{-1}R)/N(T-k)},$$

and follows F-distribution with $(N-1)k$ and $N(T-k)$ degrees of freedom. Since Σ is unknown, it can be replaced with its consistent estimator $\hat{\Sigma}$.

1A.2 Shrinkage

The specification of the model remains as:

$$R_i = F_i\beta_i^* + \varepsilon_i, \quad \forall i = 1, \ldots, N \tag{1.A.3}$$

Furthermore, several assumptions are imposed:

1. $R \sim N(F\beta, \Omega)$, where $\Omega = E[\varepsilon\varepsilon']$
2. $\beta_i^* \sim N(\mu, \Theta)$. This assumption can also be rewritten: $\beta_i^* = \mu + v_i$, where $v_i \sim N(0, \Theta)$

Plugging Assumption 2 into Equation (1.A.3), the following regression is obtained:

$$R_i = F_i\mu + \xi_i, \tag{1.A.4}$$

where $\xi_i \sim N(0, \Omega_i)$, $\Omega_i = (F_i \Theta F_i' + \sigma_i^2 I)$.

The GLS estimator of μ is derived in the following way (Swamy [1970]):

$$\hat{\mu} = (\sum_{i=1}^{N} F_i' \Omega_i^{-1} F_i)^{-1} (\sum_{i=1}^{N} F_i' \Omega_i^{-1} R_i).$$

Swamy [1970] proved that $\hat{\mu}$ can be represented as the sum of weighted individual estimates, $\hat{\beta}_i$:

$$\hat{\mu} = \hat{\beta}_{mean} = \sum_{i=1}^{N} \omega_i \hat{\beta}_i,$$

where $\hat{\beta}_i$ is the standard OLS estimator of the i^{th} fund's risk exposures, and the weight is given by:

$$\omega_i = [\sum_{l=1}^{N} (\sigma_l^2 (F_l' F_l)^{-1} + \Theta)^{-1}]^{-1} (\sigma_i^2 (F_i' F_i)^{-1} + \Theta)^{-1}.$$

In general, μ, σ^2, and Θ are unknown and can be replaced with their sample estimates (see Smith [1973] for details):

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_i \text{ or the same estimator as in Swamy [1970] } \hat{\mu} = \sum_{i=1}^{N}\omega_i\hat{\beta}_i$$
$$\hat{\sigma}_i^2 = \frac{1}{T-k}(R_i - F_i\hat{\beta}_i)'(R_i - F_i\hat{\beta}_i) = \frac{1}{T-k}R_i'M_iR_i$$
$$\hat{\Theta} = \frac{1}{N-1}\sum_i(\hat{\beta}_i - \frac{1}{N}\sum_i\hat{\beta}_i)(\hat{\beta}_i - \frac{1}{N}\sum_i\hat{\beta}_i)'$$
$$-\frac{1}{N}\sum_i(F_i'F_i)^{-1}\hat{\sigma}_i^2.$$

All estimates are based on the OLS estimates of beta parameters, $\hat{\beta}_i$. Finally, naive shrinkage estimator is derived as:

$$\hat{\beta}_i^* = (\frac{1}{\hat{\sigma}^2{}_i}F_i'F_i + \hat{\Theta}^{-1})^{-1}(\frac{1}{\hat{\sigma}^2{}_i}F_i'F_i\hat{\beta}_i + \hat{\Theta}^{-1}\hat{\mu})$$
$$\hat{\beta}_i^* = (I - P_i)\hat{\beta}_i + P_i\hat{\beta}_{mean},$$

where $P_i = (\hat{V}_i^{-1} + \hat{\Theta}^{-1})^{-1}\hat{\Theta}^{-1}$, $\hat{V}_i^{-1} = \frac{1}{\hat{\sigma}^2{}_i}F_i'F_i$.

Even though the estimator for the weight ω_i looks a bit cumbersome, it has an intuitive interpretation if one rewrites it in the following way:

$$\omega_i = [\sum_{l=1}^{N}(V_l + \Theta)^{-1}]^{-1}(V_i + \Theta)^{-1},$$

where $V_i = \sigma_i^2(F_i'F_i)^{-1}$ can be interpreted as the variance-covariance matrix of parameters in each single regression for R_i, while Θ is the variance-covariance matrix of parameters across the regressions.

Since, Θ and σ_i are unknown, Swamy [1970] offered the following estimators[1]:

$$\hat{\sigma}_i^2 = \frac{1}{T-k}(R_i - F_i\hat{\beta}_i)'(R_i - F_i\hat{\beta}_i) = \frac{1}{T-k}R_i'M_iR_i$$
$$\hat{\Theta} = \frac{1}{N-1}\sum_i(\hat{\beta}_i - \frac{1}{N}\sum_i\hat{\beta}_i)(\hat{\beta}_i - \frac{1}{N}\sum_i\hat{\beta}_i)' - \frac{1}{N}\sum_i(F_i'F_i)^{-1}\hat{\sigma}_i^2,$$

To incorporate stochastic prior information about the beta coefficients into the linear model, Lee and Griffiths [1979] apply prior likelihood approach introduced by Edwards [1969]. In this case inferences are based on a likelihood function which incorporates information from the "experimental likelihood" and the "prior likelihood". If additionally ν and ε are multivariate normal, the final likelihood function is the following:

$$\begin{aligned}\mathcal{L}(\mu_i, \sigma_i^2, \beta_i, \Theta|R, F) = const &- \frac{T}{2}\sum_{i=1}^{N}\ln\sigma_i^2 \\ &- \frac{1}{2}\sum_{i=1}^{N}\frac{1}{\sigma_i^2}(R_i - F_i\beta_i)'(R_i - F_i\beta_i) \\ &- \frac{N}{2}\ln|\Theta| - \frac{1}{2}\sum_{i=1}^{N}(\beta_i - \mu)'\Theta^{-1}(\beta_i - \mu).\end{aligned}$$

The prior likelihood estimators are related to the Bayesian ones. If μ, σ^2, and Θ are known, and the prior distributions of β and R are normal, then the posterior distribution for β is also normal with the mean:

[1] Similar estimators were used by Rao [1975].

$$\beta_i^* = (\frac{1}{\sigma^2}_i F_i' F_i + \Theta^{-1})^{-1} (\frac{1}{\sigma^2}_i F_i' F_i \beta_i + \Theta^{-1} \mu)$$
$$\beta_i^* = (I - P_i)\hat{\beta}_i + P_i \beta_{mean},$$

where $P_i = (V_i^{-1} + \Theta^{-1})^{-1} \Theta^{-1}$,

and the variance:

$$Var[\beta_i^*] = (\frac{1}{\sigma^2}_i F_i' F_i + \Theta^{-1})^{-1}.$$

1.B Appendix: Tables

Table 1.B.1: Summary statistics of hedge fund ages

Table demonstrates several statistics on hedge fund ages: mean, minimum, maximum, 25% quantile, median, and 75% quantile. Hedge funds are grouped by seven investment styles: global macro, equity long/short, emerging markets, evetn driven, fixed income, distressed, multi-strategy, managed futures, and arbitrage. ages are presented in months. The sample period is January 1994 - May 2009.

Style	Mean	Min	Max	25%	50%	75%
Global macro	4.18	0.08	15.42	1.50	3.17	5.81
Equity long/short	4.34	0.08	15.42	1.67	3.42	6.00
Emerging markets	5.00	0.08	15.42	2.25	3.92	7.02
Event driven	4.52	0.08	15.42	1.75	3.58	6.15
Fixed income	4.60	0.08	15.42	2.00	3.75	6.17
Distressed	5.24	0.08	15.42	2.00	4.00	7.83
Multi-strategy	4.02	0.08	15.42	1.58	3.17	5.58
Managed futures	5.23	0.08	15.42	2.17	4.08	7.25
Arbitrage	5.14	0.08	15.42	2.00	4.25	7.33

Table 1.B.2: Euclidean distance between hedge fund risk exposures

Table demonstrates the Euclidean distances between hedge fund risk exposures (within and between investment styles). The diagonal of the table shows the average distances between exposures of funds from the same investment style. Off-diagonal elements show the average distances between exposures of funds from different styles. Rank of each distance within a group of distances is in parentheses. The sample period is January 1994 - May 2009.

	Global macro	Equity long/short	Emerging markets	Event driven	Fixed income	Distressed	Multi-strategy	Managed futures	Arbitrage
Global macro	**0.846**	1.459	1.399	0.838	0.722	0.851	0.768	1.017	0.731
	(5)	*(9)*	*(8)*	*(4)*	*(1)*	*(6)*	*(3)*	*(7)*	*(2)*
Equity long/short	1.459	**1.362**	1.804	1.305	1.259	1.309	1.343	1.681	1.324
	(7)	*(6)*	*(9)*	*(2)*	*(1)*	*(3)*	*(5)*	*(8)*	*(4)*
Emerging markets	1.399	1.804	**1.374**	1.393	1.291	1.444	1.327	1.59	1.349
	(6)	*(9)*	*(4)*	*(5)*	*(1)*	*(7)*	*(2)*	*(8)*	*(3)*
Event driven	0.838	1.305	1.393	**0.685**	0.618	0.7	0.701	1.068	0.658
	(6)	*(8)*	*(9)*	*(3)*	*(1)*	*(4)*	*(5)*	*(7)*	*(2)*
Fixed income	0.722	1.259	1.291	0.618	**0.502**	0.639	0.589	0.945	0.544
	(6)	*(8)*	*(9)*	*(4)*	*(1)*	*(5)*	*(3)*	*(7)*	*(2)*
Distressed	0.851	1.309	1.444	0.7	0.639	**0.672**	0.721	1.065	0.667
	(6)	*(8)*	*(9)*	*(4)*	*(1)*	*(3)*	*(5)*	*(7)*	*(2)*
Multi-strategy	0.768	1.343	1.327	0.701	0.589	0.721	**0.643**	0.969	0.608
	(6)	*(9)*	*(8)*	*(4)*	*(1)*	*(5)*	*(3)*	*(7)*	*(2)*
Managed futures	1.017	1.681	1.59	1.068	0.945	1.065	0.969	**1.012**	0.919
	(5)	*(9)*	*(8)*	*(7)*	*(2)*	*(6)*	*(3)*	*(4)*	*(1)*
Arbitrage	0.731	1.324	1.349	0.658	0.544	0.667	0.608	0.919	**0.533**
	(6)	*(8)*	*(9)*	*(4)*	*(2)*	*(5)*	*(3)*	*(7)*	*(1)*

Table 1.B.3: Summary statistics of hedge fund returns

Table demonstrates statistics of hedge fund returns over the whole sample: January 1994 - May 2009. Rows denote the main statistics: mean, standard deviation, minimum, maximum, skewness, kurtosis, and Jarque-Bera p-value; columns describe these statistics with main moments.

Style	Statistics	Mean	Std. dev	Minimum	Maximum
	Mean	0.0072	0.0092	-0.0916	0.0824
	Std. dev	0.0415	0.0290	0.0000	0.2453
Global	Minimum	-0.1000	0.0862	-0.6727	0.0072
	Maximum	0.1279	0.1186	0.0000	1.4074
macro	Skewness	0.1357	0.9274	-3.6900	5.0383
	Kurtosis	4.6651	3.5641	1.5143	37.7708
	Jarque-Bera	0.3002	0.3212	0.0000	0.9994
	Mean	0.0065	0.0091	-0.0543	0.0640
	Std. dev	0.0506	0.0398	0.0000	0.9739
Equity	Minimum	-0.1350	0.1130	-1.3167	0.0048
	Maximum	0.1452	0.1377	0.0000	1.6874
Long/short	Skewness	-0.1019	0.8377	-6.4060	7.0732
	Kurtosis	4.5948	3.4846	1.3684	67.1823
	Jarque-Bera	0.2895	0.3198	0.0000	0.9989
	Mean	0.0071	0.0118	-0.0547	0.0609
	Std. dev	0.0801	0.0532	0.0000	0.3387
Emerging	Minimum	-0.2487	0.1755	-0.8887	0.0025
	Maximum	0.2099	0.1665	0.0000	1.2718
markets	Skewness	-0.5335	1.1581	-6.5445	3.8455
	Kurtosis	6.1422	6.0277	1.7032	56.3408
	Jarque-Bera	0.2237	0.3033	0.0000	0.9980
	Mean	0.0068	0.0080	-0.0292	0.0615
	Std. dev	0.0382	0.0311	0.0000	0.2519
Event	Minimum	-0.1092	0.0905	-0.5381	0.0000
	Maximum	0.1141	0.1252	0.0000	1.5142
driven	Skewness	-0.3179	1.2347	-4.4628	6.0824
	Kurtosis	6.3335	4.8290	1.7479	41.6993
	Jarque-Bera	0.1682	0.2696	0.0000	0.9913
	Mean	0.0024	0.0108	-0.1350	0.0688
	Std. dev	0.0341	0.0372	0.0000	0.3477
Fixed	Minimum	-0.1254	0.1519	-1.1027	0.0079
	Maximum	0.0866	0.0962	0.0000	1.1278
income	Skewness	-0.8266	1.6475	-8.9143	3.1827
	Kurtosis	8.7251	11.0058	1.1227	90.5329
	Jarque-Bera	0.1694	0.2803	0.0000	0.9917
	Mean	0.0064	0.0112	-0.0484	0.0812
	Std. dev	0.0421	0.0486	0.0080	0.4370
	Minimum	-0.1278	0.1152	-0.8713	-0.0007
Distressed	Maximum	0.1366	0.1838	0.0253	1.5031
	Skewness	-0.2729	1.5077	-5.3197	8.4988
	Kurtosis	7.5964	9.6118	1.9707	87.9991
	Jarque-Bera	0.1094	0.2296	0.0000	0.9822
	Mean	0.0071	0.0106	-0.0482	0.0868
	Std. dev	0.0474	0.0425	0.0000	0.4075
Multi-	Minimum	-0.1259	0.1126	-0.8003	0.0102
	Maximum	0.1328	0.1432	-0.0069	1.2137
strategy	Skewness	-0.2477	1.0237	-5.6540	5.9241
	Kurtosis	5.2137	4.7572	1.2763	50.2741
	Jarque-Bera	0.2517	0.2985	0.0000	0.9980
	Mean	0.0077	0.0093	-0.0506	0.0717
	Std. dev	0.0502	0.0315	0.0000	0.2400
Managed	Minimum	-0.1225	0.1020	-0.7423	0.0000
	Maximum	0.1573	0.1346	-0.0130	1.7717
futures	Skewness	0.2082	0.7916	-3.6224	3.8805
	Kurtosis	4.4697	3.2347	1.1635	31.8022
	Jarque-Bera	0.2952	0.3164	0.0000	0.9933
	Mean	0.0040	0.0075	-0.0602	0.0253
	Std. dev	0.0280	0.0353	0.0000	0.3109
	Minimum	-0.0975	0.1659	-1.5710	0.0001
Arbitrage	Maximum	0.0764	0.0783	0.0000	0.4697
	Skewness	-0.4859	1.2030	-6.3359	2.5837
	Kurtosis	6.5840	6.5044	1.9429	55.6783
	Jarque-Bera	0.1797	0.2839	0.0000	0.9918

Table 1.B.4: Summary statistics of hedge fund returns

Table demonstrates statistics of hedge fund returns over the sample: July 2003 - June 2007. Rows denote the main statistics: mean, standard deviation, minimum, maximum, skewness, kurtosis, and Jarque-Bera p-value; columns describe these statistics with main moments.

Style	Statistics	Mean	Std. dev	Minimum	Maximum
	Mean	0.0078	0.0074	-0.0060	0.0557
	Std. dev	0.0339	0.0217	0.0034	0.1494
Global	Minimum	-0.0655	0.0404	-0.2123	-0.0060
	Maximum	0.0951	0.0828	0.0083	0.8966
macro	Skewness	0.1231	0.6904	-2.5096	3.7099
	Kurtosis	3.5600	2.3010	1.9264	20.8394
	Jarque-Bera	0.4396	0.2976	0.0000	0.9972
	Mean	0.0119	0.0075	-0.0222	0.0603
	Std. dev	0.0345	0.0203	0.0004	0.1535
Equity	Minimum	-0.0704	0.0480	-0.5354	-0.0008
	Maximum	0.0964	0.0683	0.0012	0.7389
Long/short	Skewness	0.0166	0.5981	-2.2631	4.8783
	Kurtosis	3.4595	2.1977	1.7094	31.9862
	Jarque-Bera	0.4543	0.3045	0.0000	0.9987
	Mean	0.0212	0.0122	-0.0009	0.0683
	Std. dev	0.0540	0.0340	0.0094	0.2264
Emerging	Minimum	-0.1115	0.0669	-0.3143	-0.0060
	Maximum	0.1423	0.1036	0.0258	0.6869
markets	Skewness	-0.1774	0.7378	-5.6588	2.8411
	Kurtosis	3.6362	2.8999	1.7213	37.1276
	Jarque-Bera	0.3931	0.3215	0.0000	0.9906
	Mean	0.0112	0.0067	0.0016	0.0467
	Std. dev	0.0248	0.0202	0.0018	0.1407
Event	Minimum	-0.0461	0.0388	-0.2613	-0.0034
	Maximum	0.0762	0.0696	0.0080	0.4873
driven	Skewness	0.1135	0.6780	-1.3648	3.1986
	Kurtosis	3.6933	2.0770	1.9708	18.5001
	Jarque-Bera	0.3981	0.3097	0.0000	0.9881
	Mean	0.0065	0.0043	0.0000	0.0242
	Std. dev	0.0176	0.0120	0.0014	0.0605
Fixed	Minimum	-0.0396	0.0331	-0.2260	0.0003
	Maximum	0.0514	0.0408	0.0068	0.2344
income	Skewness	-0.1339	1.0455	-5.5261	3.8653
	Kurtosis	5.0718	4.5381	1.9674	36.0189
	Jarque-Bera	0.3011	0.3020	0.0000	0.9879
	Mean	0.0130	0.0052	0.0047	0.0276
	Std. dev	0.0261	0.0144	0.0084	0.0820
	Minimum	-0.0429	0.0281	-0.1288	-0.0030
Distressed	Maximum	0.0917	0.0810	0.0255	0.5675
	Skewness	0.5650	1.1465	-1.0232	5.9455
	Kurtosis	5.0060	5.4856	1.8932	39.7936
	Jarque-Bera	0.3636	0.3248	0.0000	0.9758
	Mean	0.0083	0.0058	-0.0198	0.0347
	Std. dev	0.0233	0.0169	0.0014	0.1045
Multi-	Minimum	-0.0487	0.0404	-0.2506	0.0066
	Maximum	0.0667	0.0578	0.0054	0.5210
strategy	Skewness	0.0091	0.8590	-4.9108	5.5256
	Kurtosis	3.9695	3.9812	1.5136	36.3498
	Jarque-Bera	0.4155	0.3179	0.0000	0.9896
	Mean	0.0059	0.0084	-0.0104	0.0632
	Std. dev	0.0432	0.0240	0.0064	0.2166
Managed	Minimum	-0.0886	0.0644	-0.7423	-0.0117
	Maximum	0.1153	0.0616	0.0135	0.4438
futures	Skewness	0.2404	0.6532	-1.7001	4.7842
	Kurtosis	3.5388	2.3795	1.8566	29.8536
	Jarque-Bera	0.3880	0.2889	0.0000	0.9884
	Mean	0.0076	0.0047	0.0011	0.0245
	Std. dev	0.0172	0.0120	0.0002	0.0571
	Minimum	-0.0319	0.0224	-0.0842	0.0004
Arbitrage	Maximum	0.0604	0.0539	0.0016	0.2716
	Skewness	0.3349	0.7313	-1.1141	2.3944
	Kurtosis	4.5475	2.5947	2.0743	12.2269
	Jarque-Bera	0.3119	0.2880	0.0000	0.9755

Table 1.B.5: Root mean squared prediction error of models, in-sample size 36 months

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample.

Style	Model	$\Delta RMSPE$				$RMSPE$	% of funds	Nr of funds
		min	max	mean	p-val			
	Individual (benchmark)					3.46%		239
Global macro	Naive shrinkage	-2.04%	0.38%	-0.11%	0.00	3.35%	87.44	239
	Advanced shrinkage	-3.01%	1.11%	-0.34%	0.00	3.12%	68.61	239
	Common mean	-3.09%	1.24%	-0.33%	0.00	3.13%	67.71	239
	Individual (benchmark)					3.45%		1083
Equity long/short	Naive shrinkage	-1.50%	1.05%	-0.03%	0.00	3.42%	89.84	1083
	Advanced shrinkage	-4.23%	3.65%	-0.39%	0.00	3.06%	67.87	1083
	Common mean	-4.37%	4.03%	-0.38%	0.00	3.07%	65.84	1083
	Individual (benchmark)					5.36%		172
Emerging markets	Naive shrinkage	-4.59%	3.93%	-0.16%	0.00	5.20%	65.12	172
	Advanced shrinkage	-4.09%	2.66%	-0.33%	0.00	5.03%	58.14	172
	Common mean	-5.00%	4.18%	-0.25%	0.00	5.11%	53.49	172
	Individual (benchmark)					2.57%		129
Event driven	Naive shrinkage	-3.75%	0.14%	-0.16%	0.00	2.41%	75.19	129
	Advanced shrinkage	-5.02%	0.66%	-0.27%	0.00	2.30%	70.54	129
	Common mean	-5.06%	0.78%	-0.24%	0.00	2.33%	65.12	129
	Individual (benchmark)					1.62%		139
Fixed income	Naive shrinkage	-1.29%	0.18%	-0.10%	0.00	1.52%	92.81	139
	Advanced shrinkage	-1.96%	0.42%	-0.21%	0.00	1.41%	69.78	139
	Common mean	-1.98%	0.45%	-0.20%	0.00	1.42%	65.47	139
	Individual (benchmark)					2.73%		53
Distressed	Naive shrinkage	-0.51%	55.51%	6.77%	0.00	9.50%	11.32	53
	Advanced shrinkage	-0.94%	0.53%	-0.08%	0.00	2.65%	50.94	53
	Common mean	-1.37%	0.48%	-0.28%	0.00	2.45%	67.92	53
	Individual (benchmark)					2.43%		188
Multi-strategy	Naive shrinkage	-2.91%	0.24%	-0.11%	0.00	2.32%	87.77	188
	Advanced shrinkage	-4.49%	0.64%	-0.26%	0.00	2.17%	67.55	188
	Common mean	-4.52%	0.71%	-0.24%	0.00	2.19%	62.77	188
	Individual (benchmark)					4.56%		206
Managed futures	Naive shrinkage	-7.16%	0.31%	-0.18%	0.00	4.38%	77.67	206
	Advanced shrinkage	-10.98%	1.12%	-0.43%	0.00	4.13%	66.99	206
	Common mean	-11.54%	1.44%	-0.39%	0.00	4.17%	62.14	206
	Individual (benchmark)					1.92%		45
Arbitrage	Naive shrinkage	-1.08%	72.84%	11.41%	0.00	13.33%	6.67	45
	Advanced shrinkage	-0.06%	0.02%	-0.02%	0.00	1.90%	84.44	45
	Common mean	-1.51%	0.66%	-0.17%	0.00	1.75%	77.78	45

Table 1.B.6: Root mean squared prediction error of models, in-sample size 24 months

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample.

Style	Model	$\Delta RMSPE$ min	max	mean	p-val	$RMSPE$	% of funds	Nr of funds
Global macro	Individual (benchmark)					3.47%		263
	Naive shrinkage	-2.04%	0.21%	-0.10%	0.00	3.37%	88.97%	263
	Advanced shrinkage	-3.95%	1.60%	-0.46%	0.00	3.01%	68.82%	263
	Common mean	-3.99%	1.69%	-0.46%	0.00	3.01%	68.06%	263
Equity long/short	Individual (benchmark)					3.45%		1367
	Naive shrinkage	-1.31%	0.18%	-0.02%	0.00	3.43%	91.22%	1367
	Advanced shrinkage	-6.02%	3.43%	-0.48%	0.00	2.97%	67.45%	1367
	Common mean	-6.09%	3.97%	-0.47%	0.00	2.98%	66.79%	1367
Emerging markets	Individual (benchmark)					5.71%		210
	Naive shrinkage	-3.08%	0.58%	-0.19%	0.00	5.52%	83.81%	210
	Advanced shrinkage	-7.05%	2.76%	-0.53%	0.00	5.18%	59.52%	210
	Common mean	-7.87%	4.07%	-0.49%	0.02	5.22%	54.29%	210
Event driven	Individual (benchmark)					2.74%		177
	Naive shrinkage	-4.48%	0.28%	-0.14%	0.00	2.60%	80.23%	177
	Advanced shrinkage	-7.84%	0.79%	-0.43%	0.00	2.31%	72.32%	177
	Common mean	-8.08%	0.84%	-0.43%	0.00	2.31%	71.75%	177
Fixed income	Individual (benchmark)					1.72%		193
	Naive shrinkage	-1.66%	0.01%	-0.08%	0.00	1.64%	92.75%	193
	Advanced shrinkage	-3.81%	0.40%	-0.28%	0.00	1.44%	69.95%	193
	Common mean	-3.88%	0.45%	-0.27%	0.00	1.45%	66.84%	193
Distressed	Individual (benchmark)					2.88%		60
	Naive shrinkage	-0.53%	172.29%	14.78%	0.00	17.66%	6.67%	60
	Advanced shrinkage	-2.02%	1.64%	0.29%	0.00	3.17%	31.67%	60
	Common mean	-2.09%	2.05%	0.54%	0.00	3.42%	30.00%	60
Multi-strategy	Individual (benchmark)					2.45%		243
	Naive shrinkage	-3.50%	0.31%	-0.09%	0.00	2.36%	86.42%	243
	Advanced shrinkage	-5.81%	1.09%	-0.36%	0.00	2.09%	74.90%	243
	Common mean	-5.89%	1.16%	-0.35%	0.00	2.10%	72.43%	243
Managed futures	Individual (benchmark)					4.45%		231
	Naive shrinkage	-6.70%	0.08%	-0.14%	0.00	4.31%	80.95%	231
	Advanced shrinkage	-12.50%	1.40%	-0.45%	0.00	4.00%	58.44%	231
	Common mean	-12.51%	1.49%	-0.44%	0.00	4.01%	57.14%	231
Arbitrage	Individual (benchmark)					1.98%		53
	Naive shrinkage	-1.31%	5.83%	1.07%	0.00	3.05%	18.87%	53
	Advanced shrinkage	-1.76%	0.95%	-0.14%	0.16	1.84%	52.83%	53
	Common mean	-1.75%	0.97%	-0.14%	0.20	1.84%	52.83%	53

Table 1.B.7: Root mean squared prediction error of models, in-sample size 12 months

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample.

Style	Model	$\Delta RMSPE$				*RMSPE*	% of	Nr of
		min	max	mean	p-val		funds	funds
	Individual (benchmark)					5.25%		330
Global macro	Naive shrinkage	-3.29%	0.00%	-0.10%	0.00	5.15%	96.97%	330
	Advanced shrinkage	-12.52%	1.50%	-2.22%	0.00	3.03%	90.91%	330
	Common mean	-12.62%	1.53%	-2.22%	0.00	3.03%	90.30%	330
	Individual (benchmark)					5.81%		1706
Equity long/short	Naive shrinkage	-2.45%	0.01%	-0.02%	0.00	5.79%	97.71%	1706
	Advanced shrinkage	-21.73%	2.68%	-2.10%	0.00	3.71%	82.12%	1706
	Common mean	-21.69%	2.73%	-2.09%	0.00	3.72%	81.65%	1706
	Individual (benchmark)					9.01%		264
Emerging markets	Naive shrinkage	-4.27%	0.13%	-0.15%	0.00	8.86%	93.18%	264
	Advanced shrinkage	-32.37%	4.02%	-2.54%	0.00	6.47%	67.05%	264
	Common mean	-31.25%	4.53%	-2.52%	0.00	6.49%	64.39%	264
	Individual (benchmark)					4.38%		228
Event driven	Naive shrinkage	-3.85%	0.01%	-0.11%	0.00	4.27%	93.86%	228
	Advanced shrinkage	-16.19%	1.46%	-1.60%	0.00	2.78%	76.32%	228
	Common mean	-16.34%	1.50%	-1.60%	0.00	2.78%	75.88%	228
	Individual (benchmark)					2.79%		248
Fixed income	Naive shrinkage	-2.59%	0.00%	-0.07%	0.00	2.72%	98.79%	248
	Advanced shrinkage	-9.88%	0.53%	-1.25%	0.00	1.54%	89.11%	248
	Common mean	-9.95%	0.54%	-1.25%	0.00	1.54%	89.11%	248
	Individual (benchmark)					4.16%		64
Distressed	Naive shrinkage	-2.02%	0.10%	-0.29%	0.00	3.87%	93.75%	64
	Advanced shrinkage	-9.18%	0.20%	-1.55%	0.00	2.61%	87.50%	64
	Common mean	-9.15%	0.29%	-1.55%	0.00	2.61%	84.38%	64
	Individual (benchmark)					4.23%		333
Multi-strategy	Naive shrinkage	-2.61%	0.01%	-0.09%	0.00	4.14%	96.70%	333
	Advanced shrinkage	-12.43%	1.79%	-1.54%	0.00	2.69%	85.29%	333
	Common mean	-12.60%	1.82%	-1.54%	0.00	2.69%	84.99%	333
	Individual (benchmark)					6.81%		286
Managed futures	Naive shrinkage	-3.02%	0.00%	-0.14%	0.00	6.67%	98.60%	286
	Advanced shrinkage	-16.87%	1.89%	-2.64%	0.00	4.17%	86.71%	286
	Common mean	-16.79%	1.93%	-2.64%	0.00	4.17%	86.71%	286
	Individual (benchmark)					3.95%		63
Arbitrage	Naive shrinkage	-1.56%	6.91%	0.03%	0.01	3.98%	66.67%	63
	Advanced shrinkage	-4.82%	1.58%	-1.12%	0.00	2.83%	74.60%	63
	Common mean	-4.86%	1.61%	-1.12%	0.00	2.83%	74.60%	63

Table 1.B.8: Root mean squared prediction error of models, cluster analysis

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The fund style breakdown is determined by the cluster analysis (the distance is calculated as the one minus the spearman coefficient)

Style	Model	$\Delta RMSPE$			p-val	*RMSPE*	% of funds	Nr of funds
		min	max	mean				
Cluster 1	Individual (benchmark)					3.38%		410
	Naive shrinkage	-3.10%	0.89%	-0.08%	0.00	3.30%	71.46%	410
	Advanced shrinkage	-3.81%	2.01%	-0.20%	0.00	3.18%	53.17%	410
	Common mean	-4.05%	2.67%	-0.18%	0.05	3.20%	50.00%	410
Cluster 2	Individual (benchmark)					3.84%		904
	Naive shrinkage	-1.26%	0.20%	-0.05%	0.00	3.79%	77.98%	904
	Advanced shrinkage	-4.43%	1.26%	-0.27%	0.00	3.57%	54.09%	904
	Common mean	-4.79%	1.51%	-0.25%	0.00	3.59%	50.77%	904
Cluster 3	Individual (benchmark)					4.71%		77
	Naive shrinkage	-1.96%	23.75%	1.74%	0.11	6.45%	44.16%	77
	Advanced shrinkage	-2.20%	1.35%	-0.28%	0.00	4.43%	58.44%	77
	Common mean	-3.12%	4.57%	-0.15%	0.47	4.56%	48.05%	77
Cluster 4	Individual (benchmark)					2.46%		112
	Naive shrinkage	-0.68%	0.01%	-0.01%	0.00	2.45%	91.96%	112
	Advanced shrinkage	-3.82%	0.51%	-0.20%	0.00	2.26%	66.07%	112
	Common mean	-3.85%	0.53%	-0.20%	0.00	2.26%	65.18%	112
Cluster 5	Individual (benchmark)					2.08%		68
	Naive shrinkage	-0.08%	0.03%	-0.01%	0.00	2.07%	94.12%	68
	Advanced shrinkage	-0.53%	0.14%	-0.12%	0.00	1.96%	92.65%	68
	Common mean	-2.22%	0.84%	-0.19%	0.00	1.89%	70.59%	68
Cluster 6	Individual (benchmark)					3.07%		82
	minimum rspe indit	0.36%	195.31%	16.83%	0.00	19.90%	0.00%	82
	Advanced shrinkage	-0.44%	5.64%	3.86%	0.00	6.93%	2.44%	82
	Common mean	-0.12%	6.46%	4.58%	0.00	7.65%	1.22%	82
Cluster 7	Individual (benchmark)					2.28%		256
	Naive shrinkage	-0.83%	0.39%	-0.06%	0.00	2.22%	71.88%	256
	Advanced shrinkage	-0.83%	0.47%	-0.07%	0.00	2.21%	64.06%	256
	Common mean	-1.50%	1.14%	-0.02%	0.59	2.26%	50.39%	256
Cluster 8	Individual (benchmark)					3.56%		300
	Naive shrinkage	-6.95%	0.07%	-0.13%	0.00	3.43%	80.67%	300
	Advanced shrinkage	-10.69%	1.49%	-0.26%	0.00	3.30%	61.00%	300
	Common mean	-11.59%	2.12%	-0.24%	0.00	3.32%	56.33%	300
Cluster 9	Individual (benchmark)					3.01%		45
	Naive shrinkage	-1.13%	42.66%	6.38%	0.00	9.39%	6.67%	45
	Advanced shrinkage	-0.17%	0.08%	-0.03%	0.03	2.98%	60.00%	45
	Common mean	-1.58%	1.75%	0.52%	0.00	3.53%	17.78%	45

Table 1.B.9: Root mean squared prediction error of models, out-of-sample size 24 months

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample.

Style	Model	$\Delta RMSPE$				*RMSPE*	% of funds	Nr of funds
		min	max	mean	p-val			
Global macro	Individual (benchmark)					4.17%		177
	Naive shrinkage	-1.04%	0.18%	-0.10%	0.00	4.07%	85.88%	177
	Advanced shrinkage	-2.11%	1.05%	-0.22%	0.00	3.95%	68.93%	177
	Common mean	-2.22%	1.14%	-0.20%	0.00	3.97%	65.54%	177
Equity long/short	Individual (benchmark)					4.43%		923
	Naive shrinkage	-1.26%	0.14%	-0.04%	0.00	4.39%	93.28%	923
	Advanced shrinkage	-5.99%	1.55%	-0.33%	0.00	4.10%	70.64%	923
	Common mean	-6.15%	1.69%	-0.32%	0.00	4.11%	67.82%	923
Emerging markets	Individual (benchmark)					6.98%		153
	Naive shrinkage	-0.66%	0.01%	-0.01%	0.00	6.97%	79.09%	153
	Advanced shrinkage	-4.09%	1.13%	-0.30%	0.00	6.68%	60.13%	153
	Common mean	-5.56%	2.09%	-0.26%	0.04	6.72%	54.90%	153
Event driven	Individual (benchmark)					3.29%		112
	Naive shrinkage	-0.23%	0.02%	-0.01%	0.00	3.28%	78.57%	112
	Advanced shrinkage	-3.02%	0.82%	-0.17%	0.00	3.12%	61.61%	112
	Common mean	-3.04%	0.84%	-0.16%	0.00	3.13%	61.61%	112
Fixed income	Individual (benchmark)					2.66%		116
	Naive shrinkage	-0.58%	2.73%	-0.02%	0.00	2.64%	67.24%	116
	Advanced shrinkage	-0.91%	0.87%	-0.13%	0.00	2.53%	71.55%	116
	Common mean	-0.94%	0.92%	-0.13%	0.00	2.53%	68.10%	116
Distressed	Individual (benchmark)					3.57%		46
	Naive shrinkage	2.15%	379.04%	35.98%	0.00	39.55%	0.00%	46
	Advanced shrinkage	-0.33%	0.21%	-0.05%	0.06	3.52%	67.39%	46
	Common mean	-0.14%	4.99%	3.26%	0.00	6.83%	4.35%	46
Multi-strategy	Individual (benchmark)					3.18%		167
	Naive shrinkage	-1.27%	0.17%	-0.09%	0.00	3.09%	82.04%	167
	Advanced shrinkage	-1.80%	0.79%	-0.15%	0.00	3.03%	61.08%	167
	Common mean	-1.81%	0.82%	-0.15%	0.00	3.03%	61.08%	167
Managed futures	Individual (benchmark)					5.20%		180
	Naive shrinkage	-5.29%	0.23%	-0.17%	0.00	5.03%	83.33%	180
	Advanced shrinkage	-7.44%	1.33%	-0.31%	0.00	4.89%	68.89%	180
	Common mean	-7.84%	1.62%	-0.27%	0.00	4.93%	67.22%	180
Arbitrage	Individual (benchmark)					2.54%		40
	Naive shrinkage	-0.04%	118.44%	11.62%	0.00	14.16%	2.50%	40
	Advanced shrinkage	-0.48%	0.89%	0.11%	0.01	2.65%	32.50%	40
	Common mean	-0.13%	4.19%	2.16%	0.00	4.70%	2.50%	40

Table 1.B.10: Root mean squared prediction error of models, out-of-sample size 18 months

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample.

Style	Model	$\Delta RMSPE$				*RMSPE*	% of	Nr of
		min	max	mean	p-val		funds	funds
	Individual (benchmark)					3.99%		200
Global	Naive shrinkage	-1.60%	0.39%	-0.10%	0.00	3.89%	89.50%	200
macro	Advanced shrinkage	-2.22%	1.51%	-0.31%	0.00	3.68%	77.50%	200
	Common mean	-2.21%	1.72%	-0.30%	0.00	3.69%	75.50%	200
	Individual (benchmark)					3.93%		1019
Equity	Naive shrinkage	-1.32%	0.13%	-0.04%	0.00	3.89%	90.48%	1019
long/short	Advanced shrinkage	-6.59%	1.46%	-0.31%	0.00	3.62%	66.83%	1019
	Common mean	-6.77%	1.59%	-0.30%	0.00	3.63%	64.57%	1019
	Individual (benchmark)					6.48%		164
Emerging	Naive shrinkage	-0.82%	0.01%	0.19%	0.00	6.67%	79.27%	164
markets	Advanced shrinkage	-4.66%	1.51%	-0.22%	0.24	6.26%	48.78%	164
	Common mean	-5.93%	2.39%	-0.13%	0.89	6.35%	45.73%	164
	Individual (benchmark)					3.03%		125
Event	Naive shrinkage	-3.79%	2.40%	-0.12%	0.00	2.91%	64.80%	125
driven	Advanced shrinkage	-5.10%	0.71%	-0.20%	0.00	2.83%	63.20%	125
	Common mean	-5.16%	0.75%	-0.20%	0.00	2.83%	63.20%	125
	Individual (benchmark)					2.31%		131
Fixed	Naive shrinkage	-0.71%	0.41%	-0.07%	0.00	2.24%	84.73%	131
income	Advanced shrinkage	-1.06%	0.94%	-0.16%	0.00	2.15%	69.47%	131
	Common mean	-1.10%	1.03%	-0.16%	0.00	2.15%	68.70%	131
	Individual (benchmark)					3.20%		52
Distressed	Naive shrinkage	4.27%	68.62%	48.02%	0.00	51.22%	0.00%	52
	Advanced shrinkage	-0.02%	2.41%	1.47%	0.00	4.67%	1.92%	52
	Common mean	6.68%	16.32%	14.49%	0.00	17.69%	0.00%	52
	Individual (benchmark)					3.04%		176
Multi-	Naive shrinkage	-3.21%	0.29%	-0.10%	0.00	2.94%	85.80%	176
strategy	Advanced shrinkage	-4.83%	0.59%	-0.20%	0.00	2.84%	59.66%	176
	Common mean	-4.86%	0.61%	-0.20%	0.00	2.84%	59.66%	176
	Individual (benchmark)					5.01%		194
Managed	Naive shrinkage	-5.94%	0.41%	-0.18%	0.00	4.83%	85.57%	194
futures	Advanced shrinkage	-8.87%	1.41%	-0.42%	0.00	4.59%	75.26%	194
	Common mean	-9.50%	1.70%	-0.41%	0.00	4.60%	71.65%	194
	Individual (benchmark)					2.24%		44
Arbitrage	Naive shrinkage	-0.23%	26.58%	18.67%	0.00	20.91%	2.27%	44
	Advanced shrinkage	-0.19%	0.22%	-0.03%	0.01	2.21%	72.73%	44
	Common mean	-0.52%	4.18%	2.54%	0.00	4.78%	2.27%	44

Table 1.B.11: Root mean squared prediction error of models, out-of-sample size 6 months

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample.

Style	Model	$\Delta RMSPE$				$RMSPE$	% of funds	Nr of funds
		min	max	mean	p-val			
	Individual (benchmark)					3.65%		253
Global macro	Naive shrinkage	-1.71%	0.77%	-0.10%	0.00	3.55%	74.75%	253
	Advanced shrinkage	-4.11%	2.61%	-0.38%	0.00	3.27%	60.08%	253
	Common mean	-4.35%	2.87%	-0.38%	0.00	3.27%	59.29%	253
	Individual (benchmark)					3.76%		1154
Equity long/short	Naive shrinkage	-2.76%	0.55%	-0.03%	0.00	3.73%	78.08%	1154
	Advanced shrinkage	-6.73%	4.19%	-0.49%	0.00	3.27%	60.05%	1154
	Common mean	-7.03%	4.52%	-0.49%	0.00	3.27%	59.27%	1154
	Individual (benchmark)					5.47%		181
Emerging markets	Naive shrinkage	-5.63%	7.10%	-0.12%	0.00	5.35%	60.22%	181
	Advanced shrinkage	-7.20%	3.82%	-0.35%	0.00	5.12%	54.14%	181
	Common mean	-9.31%	5.36%	-0.30%	0.08	5.17%	53.04%	181
	Individual (benchmark)					2.74%		137
Event driven	Naive shrinkage	-3.73%	0.75%	-0.15%	0.00	2.59%	70.80%	137
	Advanced shrinkage	-5.40%	2.34%	-0.30%	0.00	2.44%	64.96%	137
	Common mean	-5.46%	2.45%	-0.30%	0.00	2.44%	64.23%	137
	Individual (benchmark)					1.60%		143
Fixed income	Naive shrinkage	-1.39%	0.50%	-0.11%	0.00	1.49%	83.22%	143
	Advanced shrinkage	-3.20%	1.00%	-0.33%	0.00	1.27%	71.33%	143
	Common mean	-3.31%	1.06%	-0.33%	0.00	1.27%	69.93%	143
	Individual (benchmark)					2.79%		54
Distressed	Naive shrinkage	-1.11%	216.90%	9.90%	0.00	12.69%	12.96%	54
	Advanced shrinkage	-0.71%	0.41%	0.07%	0.00	2.86%	31.48%	54
	Common mean	-1.30%	2.50%	1.12%	0.00	3.91%	18.52%	54
	Individual (benchmark)					2.49%		199
Multi-strategy	Naive shrinkage	-5.17%	0.42%	-0.12%	0.00	2.37%	75.88%	199
	Advanced shrinkage	-9.10%	1.37%	-0.36%	0.00	2.13%	63.32%	199
	Common mean	-9.30%	1.43%	-0.36%	0.00	2.13%	62.81%	199
	Individual (benchmark)					4.87%		214
Managed futures	Naive shrinkage	-9.59%	0.58%	-0.22%	0.00	4.65%	71.96%	214
	Advanced shrinkage	-16.70%	2.50%	-0.59%	0.00	4.28%	63.55%	214
	Common mean	-17.69%	2.92%	-0.61%	0.00	4.26%	63.08%	214
	Individual (benchmark)					1.73%		48
Arbitrage	Naive shrinkage	-1.62%	100.00%	5.50%	0.00	7.23%	14.58%	48
	Advanced shrinkage	-1.52%	0.86%	-0.20%	0.01	1.53%	60.42%	48
	Common mean	-2.46%	1.62%	0.06%	0.37	1.79%	37.50%	48

Table 1.B.12: Root mean squared prediction error of models, sample: January 1996 - December 1999

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 36 months.

Style	Model	$\Delta RMSPE$				$RMSPE$	% of	Nr of
		min	max	mean	p-val		funds	funds
	Individual (benchmark)					6.49%		118
Global	Naive shrinkage	-0.49%	0.01%	-0.02%	0.00	6.47%	84.75%	118
macro	Advanced shrinkage	-7.42%	1.65%	-0.81%	0.00	5.68%	69.49%	118
	Common mean	-7.54%	1.68%	-0.81%	0.00	5.68%	69.49%	118
	Individual (benchmark)					8.23%		422
Equity	Naive shrinkage	-5.92%	0.40%	-0.16%	0.00	8.07%	84.83%	422
long/short	Advanced shrinkage	-10.32%	1.75%	-0.68%	0.00	7.55%	81.28%	422
	Common mean	-11.39%	2.35%	-0.71%	0.00	7.52%	79.15%	422
	Individual (benchmark)					13.15%		69
Emerging	Naive shrinkage	-7.36%	489.22%	66.56%	0.00	79.71%	1.45%	69
markets	Advanced shrinkage	-2.71%	2.18%	-0.04%	0.71	13.11%	49.28%	69
	Common mean	-8.18%	13.36%	4.29%	0.00	17.44%	10.14%	69
	Individual (benchmark)					3.79%		84
Event	Naive shrinkage	-5.45%	54.90%	6.26%	0.00	10.05%	9.52%	84
driven	Advanced shrinkage	-6.31%	2.16%	0.70%	0.00	4.49%	15.48%	84
	Common mean	-5.87%	2.74%	1.15%	0.00	4.94%	9.52%	84
	Individual (benchmark)					3.92%		39
Fixed	Naive shrinkage	-0.12%	85.95%	11.78%	0.00	15.70%	5.13%	39
income	Advanced shrinkage	-1.89%	2.72%	1.11%	0.00	5.03%	12.82%	39
	Common mean	-0.66%	7.70%	5.18%	0.00	9.10%	2.56%	39
	Individual (benchmark)					3.74%		12
Distressed	Naive shrinkage	0.10%	25.23%	6.28%	0.00	10.02%	0.00%	12
	Advanced shrinkage	0.00%	0.00%	0.00%	1.00	3.74%	0.00%	12
	Common mean	-1.53%	0.63%	-0.03%	0.60	3.71%	33.33%	12
	Individual (benchmark)					6.61%		139
Multi-	Naive shrinkage	-0.18%	0.12%	-0.02%	0.00	6.59%	87.77%	139
strategy	Advanced shrinkage	-3.93%	0.92%	-0.63%	0.00	5.98%	76.98%	139
	Common mean	-4.18%	1.27%	-0.64%	0.00	5.97%	69.78%	139
	Individual (benchmark)					6.76%		113
Managed	Naive shrinkage	-0.71%	0.04%	-0.02%	0.00	6.74%	76.99%	113
futures	Advanced shrinkage	-7.11%	3.84%	1.14%	0.00	7.90%	19.47%	113
	Common mean	-7.63%	4.46%	1.52%	0.00	8.28%	16.81%	113
	Individual (benchmark)					1.83%		16
Arbitrage	Naive shrinkage	-1.76%	16.57%	3.86%	0.00	5.69%	18.75%	16
	Advanced shrinkage	-1.03%	0.55%	-0.03%	0.54	1.80%	31.25%	16
	Common mean	-1.80%	1.34%	0.27%	0.04	2.10%	12.50%	16

Table 1.B.13: Root mean squared prediction error of models, sample: January 1998 - December 2001

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 36 months.

Style	Model	$\Delta RMSPE$				$RMSPE$	% of	Nr of
		min	max	mean	p-val		funds	funds
	Individual (benchmark)					5.42%		120
Global macro	Naive shrinkage	-0.45%	0.13%	-0.01%	0.00	5.41%	74.17%	120
	Advanced shrinkage	-5.76%	1.06%	-0.37%	0.00	5.05%	56.67%	120
	Common mean	-6.07%	1.30%	-0.35%	0.02	5.07%	55.00%	120
	Individual (benchmark)					6.87%		602
Equity long/short	Naive shrinkage	-3.63%	0.32%	-0.17%	0.00	6.70%	88.04%	602
	Advanced shrinkage	-11.09%	2.97%	-0.82%	0.00	6.05%	62.62%	602
	Common mean	-11.30%	3.27%	-0.81%	0.00	6.06%	60.80%	602
	Individual (benchmark)					9.25%		98
Emerging markets	Naive shrinkage	-0.39%	0.01%	-0.03%	0.00	9.22%	81.63%	98
	Advanced shrinkage	-2.04%	0.48%	-0.34%	0.00	8.91%	67.35%	98
	Common mean	-6.46%	3.57%	-0.29%	0.61	8.96%	42.86%	98
	Individual (benchmark)					4.20%		105
Event driven	Naive shrinkage	-0.66%	0.03%	-0.02%	0.00	4.18%	76.19%	105
	Advanced shrinkage	-3.79%	0.54%	-0.29%	0.00	3.91%	56.19%	105
	Common mean	-4.53%	0.98%	-0.23%	0.59	3.97%	44.76%	105
	Individual (benchmark)					3.31%		57
Fixed income	Naive shrinkage	-1.18%	284.57%	12.37%	0.00	15.68%	15.79%	57
	Advanced shrinkage	-0.10%	0.06%	-0.01%	0.05	3.30%	64.91%	57
	Common mean	-2.06%	4.29%	2.19%	0.00	5.50%	10.53%	57
	Individual (benchmark)					4.41%		22
Distressed	Naive shrinkage	-2.59%	15.36%	0.86%	0.58	5.27%	50.00%	22
	Advanced shrinkage	-2.50%	0.62%	-0.21%	0.38	4.20%	54.55%	22
	Common mean	-2.82%	1.11%	-0.04%	0.86	4.37%	50.00%	22
	Individual (benchmark)					5.10%		143
Multi-strategy	Naive shrinkage	-0.36%	0.03%	-0.01%	0.00	5.09%	81.82%	143
	Advanced shrinkage	-7.07%	1.42%	-0.50%	0.00	4.60%	60.84%	143
	Common mean	-8.36%	2.04%	-0.44%	0.02	4.66%	53.85%	143
	Individual (benchmark)					6.25%		138
Managed futures	Naive shrinkage	-2.37%	0.74%	-0.23%	0.00	6.02%	78.26%	138
	Advanced shrinkage	-4.36%	1.58%	-0.41%	0.00	5.84%	60.87%	138
	Common mean	-4.33%	2.19%	-0.38%	0.00	5.87%	57.97%	138
	Individual (benchmark)					2.22%		26
Arbitrage	Naive shrinkage	-0.26%	69.26%	12.15%	0.00	14.37%	15.38%	26
	Advanced shrinkage	-0.84%	0.83%	0.39%	0.00	2.61%	11.54%	26
	Common mean	-0.66%	1.84%	1.16%	0.00	3.38%	7.69%	26

Table 1.B.14: Root mean squared prediction error of models, sample: January 2000 - December 2003

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 36 months.

Style	Model	$\Delta RMSPE$				*RMSPE*	% of	Nr of
		min	max	mean	p-val		funds	funds
Global macro	Individual (benchmark)					4.79%		160
	Naive shrinkage	-4.23%	0.37%	-0.25%	0.00	4.54%	88.75%	160
	Advanced shrinkage	-6.02%	1.25%	-0.67%	0.00	4.12%	78.75%	160
	Common mean	-6.63%	1.47%	-0.69%	0.00	4.10%	76.25%	160
Equity long/short	Individual (benchmark)					6.07%		791
	Naive shrinkage	-2.38%	0.51%	-0.08%	0.00	5.99%	92.04%	791
	Advanced shrinkage	-7.12%	1.73%	-0.79%	0.00	5.28%	70.16%	791
	Common mean	-7.91%	2.17%	-0.81%	0.00	5.26%	67.26%	791
Emerging markets	Individual (benchmark)					7.44%		126
	Naive shrinkage	-0.60%	0.04%	-0.02%	0.00	7.42%	80.95%	126
	Advanced shrinkage	-5.50%	1.86%	-0.55%	0.00	6.89%	59.52%	126
	Common mean	-6.02%	2.15%	-0.52%	0.00	6.92%	55.56%	126
Event driven	Individual (benchmark)					3.31%		111
	Naive shrinkage	-0.80%	0.01%	-0.02%	0.00	3.29%	77.48%	111
	Advanced shrinkage	-3.04%	0.97%	-0.31%	0.00	3.00%	70.27%	111
	Common mean	-3.31%	1.18%	-0.27%	0.00	3.04%	63.96%	111
Fixed income	Individual (benchmark)					2.39%		96
	Naive shrinkage	-4.92%	22.03%	0.77%	0.01	3.16%	37.50%	96
	Advanced shrinkage	-5.68%	0.59%	-0.40%	0.00	1.99%	59.38%	96
	Common mean	-6.47%	0.72%	-0.38%	0.00	2.01%	55.21%	96
Distressed	Individual (benchmark)					4.01%		24
	Naive shrinkage	-3.30%	9.40%	0.59%	0.20	4.60%	25.00%	24
	Advanced shrinkage	-0.33%	-0.02%	-0.10%	0.00	3.91%	100.00%	24
	Common mean	-3.54%	0.18%	-0.78%	0.00	3.23%	83.33%	24
Multi-strategy	Individual (benchmark)					3.35%		140
	Naive shrinkage	-0.54%	0.02%	-0.02%	0.00	3.33%	95.71%	140
	Advanced shrinkage	-6.36%	1.20%	-0.49%	0.00	2.86%	63.57%	140
	Common mean	-6.75%	1.35%	-0.48%	0.00	2.87%	59.29%	140
Managed futures	Individual (benchmark)					5.47%		167
	Naive shrinkage	-7.55%	0.36%	-0.30%	0.00	5.17%	81.44%	167
	Advanced shrinkage	-8.77%	0.98%	-0.53%	0.00	4.94%	70.06%	167
	Common mean	-9.09%	1.19%	-0.51%	0.00	4.96%	67.66%	167
Arbitrage	Individual (benchmark)					2.62%		34
	Naive shrinkage	-1.85%	82.17%	5.47%	0.00	8.09%	5.88%	34
	Advanced shrinkage	-1.19%	0.55%	0.04%	0.19	2.66%	35.29%	34
	Common mean	-2.74%	2.48%	1.21%	0.00	3.83%	5.88%	34

Table 1.B.15: Root mean squared prediction error of models, sample: January 2002 - December 2005

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 36 months.

Style	Model	$\Delta RMSPE$				*RMSPE*	% of funds	Nr of funds
		min	max	mean	p-val			
	Individual (benchmark)					3.79%		194
Global macro	Naive shrinkage	-7.17%	0.28%	-0.14%	0.00	3.65%	87.11%	194
	Advanced shrinkage	-9.88%	0.79%	-0.34%	0.00	3.45%	75.77%	194
	Common mean	-9.81%	0.89%	-0.33%	0.00	3.46%	73.20%	194
	Individual (benchmark)					4.26%		970
Equity long/short	Naive shrinkage	-2.35%	0.03%	-0.05%	0.00	4.21%	93.09%	970
	Advanced shrinkage	-7.62%	1.79%	-0.30%	0.00	3.96%	72.37%	970
	Common mean	-7.84%	2.06%	-0.29%	0.00	3.97%	70.72%	970
	Individual (benchmark)					6.01%		159
Emerging markets	Naive shrinkage	-2.18%	0.40%	-0.16%	0.00	5.85%	72.33%	159
	Advanced shrinkage	-2.90%	0.84%	-0.25%	0.00	5.76%	65.41%	159
	Common mean	-3.55%	1.38%	-0.19%	0.02	5.82%	55.35%	159
	Individual (benchmark)					2.86%		124
Event driven	Naive shrinkage	-0.21%	0.00%	-0.02%	0.00	2.84%	83.87%	124
	Advanced shrinkage	-2.32%	0.61%	-0.17%	0.00	2.69%	61.29%	124
	Common mean	-2.79%	1.08%	-0.14%	0.01	2.72%	54.03%	124
	Individual (benchmark)					1.98%		118
Fixed income	Naive shrinkage	-0.30%	0.01%	-0.01%	0.00	1.97%	88.14%	118
	Advanced shrinkage	-2.52%	0.57%	-0.17%	0.03	1.81%	55.93%	118
	Common mean	-2.56%	0.63%	-0.16%	0.15	1.82%	53.39%	118
	Individual (benchmark)					2.96%		35
Distressed	Naive shrinkage	-1.47%	51.26%	8.85%	0.00	11.81%	14.29%	35
	Advanced shrinkage	-0.10%	0.05%	-0.02%	0.00	2.94%	77.14%	35
	Common mean	-1.45%	1.04%	0.16%	0.02	3.12%	31.43%	35
	Individual (benchmark)					2.79%		154
Multi-strategy	Naive shrinkage	-2.39%	0.94%	-0.10%	0.00	2.69%	86.36%	154
	Advanced shrinkage	-3.03%	1.66%	-0.18%	0.00	2.61%	73.38%	154
	Common mean	-3.07%	1.71%	-0.18%	0.00	2.61%	73.38%	154
	Individual (benchmark)					4.52%		189
Managed futures	Naive shrinkage	-1.36%	1.86%	-0.15%	0.00	4.37%	85.71%	189
	Advanced shrinkage	-2.85%	2.06%	-0.40%	0.00	4.12%	77.78%	189
	Common mean	-3.27%	3.83%	-0.39%	0.00	4.13%	72.49%	189
	Individual (benchmark)					1.93%		380
Arbitrage	Naive shrinkage	-0.65%	311.57%	16.53%	0.00	18.46%	7.89%	380
	Advanced shrinkage	0.00%	0.00%	0.00%	1.00	1.93%	0.00%	380
	Common mean	-1.29%	4.74%	2.96%	0.00	0.0489	5.26%	380

Table 1.B.16: Root mean squared prediction error of models, sample: January 2004 - December 2007

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 36 months.

Style	Model	$\Delta RMSPE$				*RMSPE*	% of	Nr of
		min	max	mean	p-val		funds	funds
Global macro	Individual (benchmark)					4.39%		208
	Naive shrinkage	-1.39%	0.23%	-0.09%	0.00	4.30%	77.40%	208
	Advanced shrinkage	-2.53%	1.07%	-0.26%	0.00	4.13%	70.67%	208
	Common mean	-2.71%	1.32%	-0.23%	0.00	4.16%	69.23%	208
Equity long/short	Individual (benchmark)					4.66%		1142
	Naive shrinkage	-1.32%	0.17%	-0.03%	0.00	4.63%	82.49%	1142
	Advanced shrinkage	-8.03%	1.47%	-0.22%	0.00	4.44%	61.21%	1142
	Common mean	-8.64%	1.83%	-0.20%	0.00	4.46%	58.06%	1142
Emerging markets	Individual (benchmark)					7.84%		182
	Naive shrinkage	-3.35%	0.35%	-0.22%	0.00	7.62%	70.88%	182
	Advanced shrinkage	-3.48%	1.61%	-0.28%	0.00	7.56%	60.99%	182
	Common mean	-4.60%	2.72%	-0.26%	0.04	7.58%	59.34%	182
Event driven	Individual (benchmark)					3.69%		146
	Naive shrinkage	-0.17%	0.01%	-0.01%	0.00	3.68%	71.92%	146
	Advanced shrinkage	-4.16%	0.73%	-0.27%	0.00	3.42%	63.01%	146
	Common mean	-4.36%	0.81%	-0.27%	0.00	3.42%	61.64%	146
Fixed income	Individual (benchmark)					2.77%		160
	Naive shrinkage	-1.51%	0.53%	-0.06%	0.00	2.71%	84.38%	160
	Advanced shrinkage	-1.97%	1.10%	-0.16%	0.00	2.61%	68.13%	160
	Common mean	-2.21%	1.30%	-0.16%	0.00	2.61%	62.50%	160
Distressed	Individual (benchmark)					3.73%		54
	Naive shrinkage	0.38%	110.39%	17.31%	0.00	21.04%	0.00%	54
	Advanced shrinkage	-0.14%	0.06%	-0.03%	0.00	3.70%	72.22%	54
	Common mean	-0.51%	4.55%	2.73%	0.00	6.46%	3.70%	54
Multi-strategy	Individual (benchmark)					3.55%		212
	Naive shrinkage	-1.93%	0.27%	-0.08%	0.00	3.47%	76.42%	212
	Advanced shrinkage	-3.31%	1.70%	-0.13%	0.00	3.42%	57.08%	212
	Common mean	-3.45%	1.89%	-0.12%	0.02	3.43%	55.66%	212
Managed futures	Individual (benchmark)					5.04%		200
	Naive shrinkage	-5.38%	0.35%	-0.14%	0.00	4.90%	83.00%	200
	Advanced shrinkage	-7.35%	1.66%	-0.25%	0.00	4.79%	73.50%	200
	Common mean	-8.13%	2.05%	-0.21%	0.00	4.83%	65.50%	200
Arbitrage	Individual (benchmark)					2.64%		48
	Naive shrinkage	-0.54%	44.65%	4.33%	0.00	6.97%	14.58%	48
	Advanced shrinkage	-0.08%	0.11%	0.01%	0.04	2.65%	37.50%	48
	Common mean	-0.72%	1.42%	0.40%	0.00	3.04%	22.92%	48

Table 1.B.17: Root mean squared prediction error of models, sample: January 1996 - December 1997

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 12 months.

Style	Model	$\Delta RMSPE$				*RMSPE*	% of	Nr of
		min	max	mean	p-val		funds	funds
	Individual (benchmark)					7.87%		168
Global	Naive shrinkage	-4.83%	0.01%	-0.23%	0.00	7.64%	94.64%	168
macro	Advanced shrinkage	-25.13%	4.07%	-2.33%	0.00	5.54%	75.00%	168
	Common mean	-25.42%	4.16%	-2.32%	0.00	5.55%	74.40%	168
	Individual (benchmark)					9.41%		510
Equity	Naive shrinkage	-2.98%	0.02%	-0.08%	0.00	9.33%	97.06%	510
long/short	Advanced shrinkage	-24.54%	3.56%	-2.16%	0.00	7.25%	76.67%	510
	Common mean	-24.29%	3.73%	-2.16%	0.00	7.25%	75.49%	510
	Individual (benchmark)					12.74%		84
Emerging	Naive shrinkage	-4.22%	0.11%	-0.32%	0.00	12.42%	89.29%	84
markets	Advanced shrinkage	-28.85%	6.47%	-2.31%	0.00	10.43%	70.24%	84
	Common mean	-31.36%	7.76%	-2.22%	0.01	10.52%	60.71%	84
	Individual (benchmark)					5.12%		101
Event	Naive shrinkage	-3.55%	0.02%	-0.23%	0.00	4.89%	93.07%	101
driven	Advanced shrinkage	-15.88%	1.30%	-1.45%	0.00	3.67%	66.34%	101
	Common mean	-16.06%	1.40%	-1.44%	0.00	3.68%	66.34%	101
	Individual (benchmark)					4.01%		43
Fixed	Naive shrinkage	-2.28%	7.11%	0.47%	0.65	4.48%	46.51%	43
income	Advanced shrinkage	-4.91%	1.97%	-0.51%	0.16	3.50%	55.81%	43
	Common mean	-4.71%	2.01%	-0.49%	0.19	3.52%	51.16%	43
	Individual (benchmark)					3.24%		12
Distressed	Naive shrinkage	0.07%	61.53%	21.29%	0.00	24.53%	0.00%	12
	Advanced shrinkage	-0.28%	1.56%	0.68%	0.01	3.92%	16.67%	12
	Common mean	1.17%	4.85%	3.11%	0.00	6.35%	0.00%	12
	Individual (benchmark)					8.04%		186
Multi-	Naive shrinkage	-5.35%	0.08%	-0.20%	0.00	7.84%	86.56%	186
strategy	Advanced shrinkage	-16.26%	4.81%	-1.67%	0.00	6.37%	63.98%	186
	Common mean	-16.40%	5.06%	-1.66%	0.00	6.38%	63.98%	186
	Individual (benchmark)					7.83%		151
Managed	Naive shrinkage	-5.79%	0.32%	-0.28%	0.00	7.55%	94.04%	151
futures	Advanced shrinkage	-24.19%	4.11%	-2.01%	0.00	5.82%	78.81%	151
	Common mean	-24.83%	4.57%	-1.97%	0.00	5.86%	76.16%	151
	Individual (benchmark)					2.37%		19
Arbitrage	Naive shrinkage	-0.91%	32.68%	5.54%	0.00	7.91%	15.79%	19
	Advanced shrinkage	-1.91%	1.39%	0.00%	0.85	2.37%	47.37%	19
	Common mean	-1.92%	1.44%	0.02%	0.79	2.39%	42.11%	19

Table 1.B.18: Root mean squared prediction error of models, sample: January 1998 - December 1999

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 12 months.

Style	Model	$\Delta RMSPE$				$RMSPE$	% of funds	Nr of funds
		min	max	mean	p-val			
	Individual (benchmark)					13.66%		165
Global macro	Naive shrinkage	-12.38%	0.12%	-0.46%	0.00	13.20%	86.67%	165
	Advanced shrinkage	-51.02%	6.11%	-4.97%	0.00	8.69%	67.27%	165
	Common mean	-51.43%	6.26%	-4.95%	0.00	8.71%	66.67%	165
	Individual (benchmark)					19.29%		720
Equity long/short	Naive shrinkage	-9.89%	0.54%	-0.13%	0.00	19.16%	88.33%	720
	Advanced shrinkage	-120.90%	11.40%	-5.03%	0.00	14.26%	61.39%	720
	Common mean	-122.31%	11.63%	-4.99%	0.00	14.30%	61.25%	720
	Individual (benchmark)					32.00%		123
Emerging markets	Naive shrinkage	-16.71%	0.38%	-0.88%	0.00	31.12%	84.55%	123
	Advanced shrinkage	-105.11%	22.47%	-6.74%	0.00	25.26%	60.16%	123
	Common mean	-115.30%	25.54%	-6.56%	0.00	25.44%	59.35%	123
	Individual (benchmark)					10.62%		144
Event driven	Naive shrinkage	-10.54%	1.38%	-0.37%	0.00	10.25%	80.56%	144
	Advanced shrinkage	-37.85%	3.87%	-2.66%	0.00	7.96%	65.28%	144
	Common mean	-35.54%	5.26%	-2.62%	0.00	8.00%	59.72%	144
	Individual (benchmark)					9.29%		67
Fixed income	Naive shrinkage	-7.50%	1.06%	-0.37%	0.00	8.92%	70.15%	67
	Advanced shrinkage	-28.43%	3.61%	-3.85%	0.00	5.44%	64.18%	67
	Common mean	-30.33%	4.06%	-3.84%	0.00	5.45%	64.18%	67
	Individual (benchmark)					13.18%		22
Distressed	Naive shrinkage	-10.01%	56.99%	8.73%	0.01	21.91%	18.18%	22
	Advanced shrinkage	-12.77%	1.74%	-1.76%	0.15	11.42%	50.00%	22
	Common mean	-30.01%	6.79%	-2.46%	0.89	10.72%	45.45%	22
	Individual (benchmark)					14.73%		231
Multi-strategy	Naive shrinkage	-6.63%	1.63%	-0.27%	0.00	14.46%	78.36%	231
	Advanced shrinkage	-49.98%	10.24%	-3.35%	0.00	11.38%	54.11%	231
	Common mean	-50.51%	10.46%	-3.30%	0.00	11.43%	54.11%	231
	Individual (benchmark)					14.30%		172
Managed futures	Naive shrinkage	-7.44%	0.09%	-0.41%	0.00	13.89%	91.86%	172
	Advanced shrinkage	-37.87%	5.81%	-4.99%	0.00	9.31%	71.51%	172
	Common mean	-38.28%	5.96%	-4.99%	0.00	9.31%	70.93%	172
	Individual (benchmark)					5.07%		28
Arbitrage	Naive shrinkage	-3.26%	16.52%	2.66%	0.01	7.73%	32.14%	28
	Advanced shrinkage	-9.76%	1.95%	-1.79%	0.00	3.28%	64.29%	28
	Common mean	-9.18%	2.36%	-1.71%	0.01	3.36%	64.29%	28

Table 1.B.19: Root mean squared prediction error of models, sample: January 2000 - December 2001

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 12 months.

Style	Model	$\Delta RMSPE$				$RMSPE$	% of	Nr of
		min	max	mean	p-val		funds	funds
	Individual (benchmark)					9.41%		185
Global macro	Naive shrinkage	-7.19%	0.02%	-0.27%	0.00	9.14%	91.35%	185
	Advanced shrinkage	-27.78%	2.56%	-3.42%	0.00	5.99%	82.16%	185
	Common mean	-28.22%	2.63%	-3.41%	0.00	6.00%	82.16%	185
	Individual (benchmark)					13.76%		948
Equity long/short	Naive shrinkage	-2.38%	0.03%	-0.06%	0.00	13.70%	92.30%	948
	Advanced shrinkage	-53.83%	6.52%	-4.66%	0.00	9.10%	68.88%	948
	Common mean	-55.82%	6.94%	-4.59%	0.00	9.17%	66.67%	948
	Individual (benchmark)					16.21%		153
Emerging markets	Naive shrinkage	-4.88%	0.08%	-0.37%	0.00	15.84%	94.12%	153
	Advanced shrinkage	-37.26%	9.63%	-3.89%	0.00	12.32%	66.01%	153
	Common mean	-38.80%	10.35%	-3.83%	0.00	12.38%	65.36%	153
	Individual (benchmark)					7.19%		155
Event driven	Naive shrinkage	-5.64%	0.05%	-0.24%	0.00	6.95%	93.55%	155
	Advanced shrinkage	-36.80%	3.38%	-2.17%	0.00	5.02%	61.29%	155
	Common mean	-37.09%	3.45%	-2.16%	0.00	5.03%	61.29%	155
	Individual (benchmark)					6.20%		116
Fixed income	Naive shrinkage	-3.96%	0.00%	-0.29%	0.00	5.91%	93.10%	116
	Advanced shrinkage	-46.63%	2.56%	-2.27%	0.00	3.93%	55.17%	116
	Common mean	-47.07%	2.62%	-2.25%	0.00	3.95%	54.31%	116
	Individual (benchmark)					10.35%		27
Distressed	Naive shrinkage	-5.75%	55.94%	5.54%	0.01	15.89%	22.22%	27
	Advanced shrinkage	-21.82%	4.88%	-2.83%	0.51	7.52%	48.15%	27
	Common mean	-21.86%	5.01%	-2.74%	0.58	7.61%	44.44%	27
	Individual (benchmark)					10.05%		236
Multi-strategy	Naive shrinkage	-17.29%	0.04%	-0.28%	0.00	9.77%	91.10%	236
	Advanced shrinkage	-30.80%	4.00%	-3.23%	0.00	6.82%	58.90%	236
	Common mean	-33.75%	4.69%	-3.14%	0.00	6.91%	57.63%	236
	Individual (benchmark)					10.96%		198
Managed futures	Naive shrinkage	-3.35%	0.11%	-0.26%	0.00	10.70%	92.42%	198
	Advanced shrinkage	-23.56%	4.11%	-3.32%	0.00	7.64%	74.24%	198
	Common mean	-23.79%	4.21%	-3.32%	0.00	7.64%	73.74%	198
	Individual (benchmark)					5.76%		38
Arbitrage	Naive shrinkage	-8.24%	50.35%	4.32%	0.00	10.08%	10.53%	38
	Advanced shrinkage	-39.27%	2.55%	-1.55%	0.90	4.21%	44.74%	38
	Common mean	-39.52%	2.61%	-1.52%	0.97	4.24%	42.11%	38

Table 1.B.20: Root mean squared prediction error of models, sample: January 2002 - December 2003

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 12 months.

Style	Model	$\Delta RMSPE$ min	max	mean	p-val	*RMSPE*	% of funds	Nr of funds
	Individual (benchmark)					11.37%		239
Global	Naive shrinkage	-13.00%	0.23%	-0.38%	0.00	10.99%	92.89%	239
macro	Advanced shrinkage	-61.19%	1.98%	-6.10%	0.00	5.27%	88.28%	239
	Common mean	-69.19%	2.44%	-6.30%	0.00	5.07%	86.61%	239
	Individual (benchmark)					13.00%		1271
Equity	Naive shrinkage	-6.41%	0.03%	-0.10%	0.00	12.90%	91.42%	1271
long/short	Advanced shrinkage	-62.24%	4.69%	-5.83%	0.00	7.17%	78.36%	1271
	Common mean	-68.80%	5.33%	-5.93%	0.00	7.07%	75.61%	1271
	Individual (benchmark)					20.63%		176
Emerging	Naive shrinkage	-11.31%	0.17%	-0.75%	0.00	19.88%	82.39%	176
markets	Advanced shrinkage	-69.85%	9.65%	-7.56%	0.00	13.07%	64.20%	176
	Common mean	-69.79%	9.86%	-7.56%	0.00	13.07%	64.20%	176
	Individual (benchmark)					8.13%		164
Event	Naive shrinkage	-6.50%	0.27%	-0.34%	0.00	7.79%	87.80%	164
driven	Advanced shrinkage	-28.70%	3.66%	-3.09%	0.00	5.04%	68.29%	164
	Common mean	-29.55%	3.92%	-3.09%	0.00	5.04%	67.07%	164
	Individual (benchmark)					7.57%		143
Fixed	Naive shrinkage	-11.14%	1.81%	-0.32%	0.00	7.25%	85.31%	143
income	Advanced shrinkage	-60.56%	2.97%	-3.46%	0.00	4.11%	61.54%	143
	Common mean	-64.72%	3.23%	-3.46%	0.00	4.11%	60.14%	143
	Individual (benchmark)					11.87%		39
Distressed	Naive shrinkage	-9.32%	4.22%	-0.19%	0.93	11.68%	51.28%	39
	Advanced shrinkage	-17.90%	6.20%	-2.75%	0.04	9.12%	58.97%	39
	Common mean	-18.58%	6.48%	-2.73%	0.05	9.14%	58.97%	39
	Individual (benchmark)					8.50%		199
Multi-	Naive shrinkage	-22.23%	2.32%	-0.54%	0.00	7.96%	92.96%	199
strategy	Advanced shrinkage	-56.69%	1.53%	-4.91%	0.00	3.59%	92.96%	199
	Common mean	-63.48%	1.90%	-5.18%	0.00	3.32%	89.95%	199
	Individual (benchmark)					12.45%		219
Managed	Naive shrinkage	-20.47%	0.57%	-0.45%	0.00	12.00%	93.61%	219
futures	Advanced shrinkage	-44.18%	1.63%	-5.28%	0.00	7.17%	90.87%	219
	Common mean	-60.01%	3.05%	-6.47%	0.00	5.98%	85.39%	219
	Individual (benchmark)					5.57%		60
Arbitrage	Naive shrinkage	-15.91%	14.05%	0.55%	0.15	6.12%	48.33%	60
	Advanced shrinkage	-41.66%	2.39%	-2.47%	0.03	3.10%	55.00%	60
	Common mean	-42.52%	2.65%	-2.37%	0.08	3.20%	50.00%	60

Table 1.B.21: Root mean squared prediction error of models, sample: January 2004 - December 2005

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 12 months.

Style	Model	$\Delta RMSPE$				$RMSPE$	% of	Nr of
		min	max	mean	p-val		funds	funds
	Individual (benchmark)					5.49%		291
Global	Naive shrinkage	-4.03%	0.08%	-0.10%	0.00	5.39%	96.22%	291
macro	Advanced shrinkage	-15.35%	1.19%	-1.78%	0.00	3.71%	84.19%	291
	Common mean	-15.48%	1.41%	-1.77%	0.00	3.72%	81.44%	291
	Individual (benchmark)					5.62%		1592
Equity	Naive shrinkage	-11.30%	0.08%	-0.04%	0.00	5.58%	97.86%	1592
long/short	Advanced shrinkage	-46.23%	1.78%	-1.65%	0.00	3.97%	85.05%	1592
	Common mean	-48.77%	2.36%	-1.63%	0.00	3.99%	79.46%	1592
	Individual (benchmark)					8.93%		211
Emerging	Naive shrinkage	-8.17%	0.28%	-0.20%	0.00	8.73%	83.41%	211
markets	Advanced shrinkage	-20.78%	2.54%	-1.41%	0.00	7.52%	57.82%	211
	Common mean	-28.32%	4.28%	-1.38%	0.01	7.55%	49.29%	211
	Individual (benchmark)					4.16%		208
Event	Naive shrinkage	-6.98%	0.02%	-0.11%	0.00	4.05%	95.67%	208
driven	Advanced shrinkage	-11.88%	1.29%	-0.91%	0.00	3.25%	74.52%	208
	Common mean	-13.60%	1.71%	-0.90%	0.00	3.26%	68.75%	208
	Individual (benchmark)					2.81%		202
Fixed	Naive shrinkage	-2.96%	0.01%	-0.09%	0.00	2.72%	96.53%	202
income	Advanced shrinkage	-10.16%	0.79%	-0.90%	0.00	1.91%	73.27%	202
	Common mean	-10.17%	0.96%	-0.87%	0.00	1.94%	65.84%	202
	Individual (benchmark)					3.98%		66
Distressed	Naive shrinkage	-3.13%	0.37%	-0.25%	0.00	3.73%	81.82%	66
	Advanced shrinkage	-6.79%	0.77%	-0.84%	0.00	3.14%	74.24%	66
	Common mean	-7.14%	0.95%	-0.82%	0.00	3.16%	68.18%	66
	Individual (benchmark)					3.68%		270
Multi-	Naive shrinkage	-4.65%	0.02%	-0.08%	0.00	3.60%	92.59%	270
strategy	Advanced shrinkage	-14.58%	0.92%	-0.87%	0.00	2.81%	82.22%	270
	Common mean	-19.67%	1.31%	-0.91%	0.00	2.77%	70.74%	270
	Individual (benchmark)					6.92%		252
Managed	Naive shrinkage	-5.80%	0.00%	-0.13%	0.00	6.79%	96.83%	252
futures	Advanced shrinkage	-20.31%	1.31%	-2.15%	0.00	4.77%	81.75%	252
	Common mean	-17.76%	1.81%	-2.25%	0.00	4.67%	75.79%	252
	Individual (benchmark)					2.73%		59
Arbitrage	Naive shrinkage	-1.53%	0.09%	-0.15%	0.00	2.58%	81.36%	59
	Advanced shrinkage	-6.41%	0.47%	-0.83%	0.00	1.90%	88.14%	59
	Common mean	-6.38%	0.62%	-0.85%	0.00	1.88%	84.75%	59

Table 1.B.22: Root mean squared prediction error of models, sample: January 2006 - December 2007

Table summarizes root mean squared prediction error (RMSPE) statistics of four estimators: individual (the benchmark), naive shrinkage, advanced shrinkage, and common mean. The columns list the descriptive statistics of the $\Delta RMSPE = RMSPE_{model} - RMSPE_{benchmark}$: minimum, maximum, and mean. It also lists the p-value of the Wilcoxon test with null hypothesis that there is no significance difference in RMSPE of the benchmark and the corresponding model. The last three columns show average RMSPE across the funds from the particular style, the percentage of the funds with improved forecasts comparing to the benchmark, and the total number of funds within the style over the sample. The in-sample size is equal to 12 months.

Style	Model	$\Delta RMSPE$ min	max	mean	p-val	$RMSPE$	% of funds	Nr of funds
	Individual (benchmark)					5.77%		312
Global macro	Naive shrinkage	-4.19%	0.10%	-0.07%	0.00	5.70%	91.67%	312
	Advanced shrinkage	-8.12%	1.92%	-1.46%	0.00	4.31%	79.81%	312
	Common mean	-8.15%	2.05%	-1.44%	0.00	4.33%	78.21%	312
	Individual (benchmark)					6.40%		1787
Equity long/short	Naive shrinkage	-2.37%	0.03%	-0.02%	0.00	6.38%	92.00%	1787
	Advanced shrinkage	-29.62%	2.64%	-1.24%	0.00	5.16%	69.73%	1787
	Common mean	-30.03%	2.71%	-1.23%	0.00	5.17%	69.05%	1787
	Individual (benchmark)					9.98%		307
Emerging markets	Naive shrinkage	-7.46%	0.05%	-0.11%	0.00	9.87%	85.67%	307
	Advanced shrinkage	-25.99%	3.66%	-1.69%	0.00	8.29%	64.82%	307
	Common mean	-25.50%	3.83%	-1.67%	0.00	8.31%	64.17%	307
	Individual (benchmark)					5.05%		233
Event driven	Naive shrinkage	-1.81%	0.20%	-0.08%	0.00	4.97%	85.41%	233
	Advanced shrinkage	-10.71%	1.26%	-0.94%	0.00	4.11%	71.67%	233
	Common mean	-10.83%	1.29%	-0.93%	0.00	4.12%	70.82%	233
	Individual (benchmark)					3.78%		258
Fixed income	Naive shrinkage	-1.53%	0.31%	-0.06%	0.00	3.72%	93.80%	258
	Advanced shrinkage	-11.37%	2.42%	-0.82%	0.00	2.96%	81.78%	258
	Common mean	-11.48%	2.52%	-0.82%	0.00	2.96%	81.40%	258
	Individual (benchmark)					4.86%		65
Distressed	Naive shrinkage	-8.58%	0.10%	-0.32%	0.00	4.54%	87.69%	65
	Advanced shrinkage	-11.78%	2.13%	-0.85%	0.00	4.01%	80.00%	65
	Common mean	-11.78%	2.20%	-0.85%	0.00	4.01%	78.46%	65
	Individual (benchmark)					4.86%		372
Multi-strategy	Naive shrinkage	-1.93%	0.01%	-0.07%	0.00	4.79%	90.05%	372
	Advanced shrinkage	-11.35%	1.99%	-0.95%	0.00	3.91%	69.09%	372
	Common mean	-11.45%	2.03%	-0.95%	0.00	3.91%	68.55%	372
	Individual (benchmark)					6.77%		312
Managed futures	Naive shrinkage	-2.17%	0.02%	-0.09%	0.00	6.68%	94.55%	312
	Advanced shrinkage	-9.45%	2.58%	-1.57%	0.00	5.20%	80.77%	312
	Common mean	-9.43%	2.87%	-1.51%	0.00	5.26%	79.17%	312
	Individual (benchmark)					3.68%		60
Arbitrage	Naive shrinkage	-0.62%	947.85%	32.33%	0.00	36.01%	15.00%	60
	Advanced shrinkage	-3.43%	3.80%	1.46%	0.00	5.14%	16.67%	60
	Common mean	-3.41%	3.87%	1.52%	0.00	5.20%	16.67%	60

Table 1.B.23: Alpha persistence

Table demonstrates Kendall rank correlation coefficient (τ) between the ranks of alpha in formation period (α_{0i}) and the ranks of alpha in evaluation period (α_{1i}). Formation period includes 36 months, evaluation period is 24 or 36 months. P-values are given in parentheses.

	36 months formation period 24 months evaluation period		36 months formation period 36 months evaluation period	
Individual alpha	0.328	(0.000)	-0.010	(0.247)
Naive shrinkage alpha	0.360	(0.000)	0.180	(0.080)

Table 1.B.24: Alpha persistence

Table demonstrates Kendall rank correlation coefficient (τ) between the ranks of alpha in formation period (α_{0i}) and the ranks of alpha in evaluation period (α_{1i}). Two strategies are presented: 1) funds are sorted on their deviations from the common mean 2) funds are sorted on their alpha-estimate. Formation period includes 36 months, evaluation period is 24 or 36 months. P-values are given in parentheses.

	Percentage of hedge funds in top/bottom portfolio			
	5%	10%	20%	33%
	36 months formation period 24 months evaluation period			
Alpha-sorted (top)	-0.024	-0.042	-0.132	-0.111
	(0.741)	(0.422)	(0.000)	(0.000)
Largest-deviation (top)	0.268	0.244	0.165	0.133
	(0.000)	(0.000)	(0.000)	(0.000)
Alpha-sorted (bottom)	0.204	0.225	0.161	0.178
	(0.006)	(0.000)	(0.000)	(0.000)
Largest-deviation (bottom)	0.083	0.103	0.078	0.103
	(0.260)	(0.048)	(0.033)	(0.000)
	36 months formation period 36 months evaluation period			
Alpha-sorted (top)	0.005	-0.031	-0.027	0.002
	(0.950)	(0.584)	(0.504)	(0.942)
Largest-deviation (top)	-0.059	-0.122	-0.029	-0.014
	(0.468)	(0.033)	(0.464)	(0.650)
Alpha-sorted (bottom)	0.035	-0.121	-0.025	-0.049
	(0.667)	(0.033)	(0.539)	(0.112)
Largest-deviation (bottom)	0.286	0.217	0.147	0.138
	(0.000)	(0.000)	(0.000)	(0.000)

Table 1.B.25: Alpha of funds sorted by deviations from the common mean

Table shows alpha of the funds in dependence from its deviations from the common mean starting from the funds with the largest deviations (Top 10%) and ending with the funds with smallest deviations (Bottom 10%).

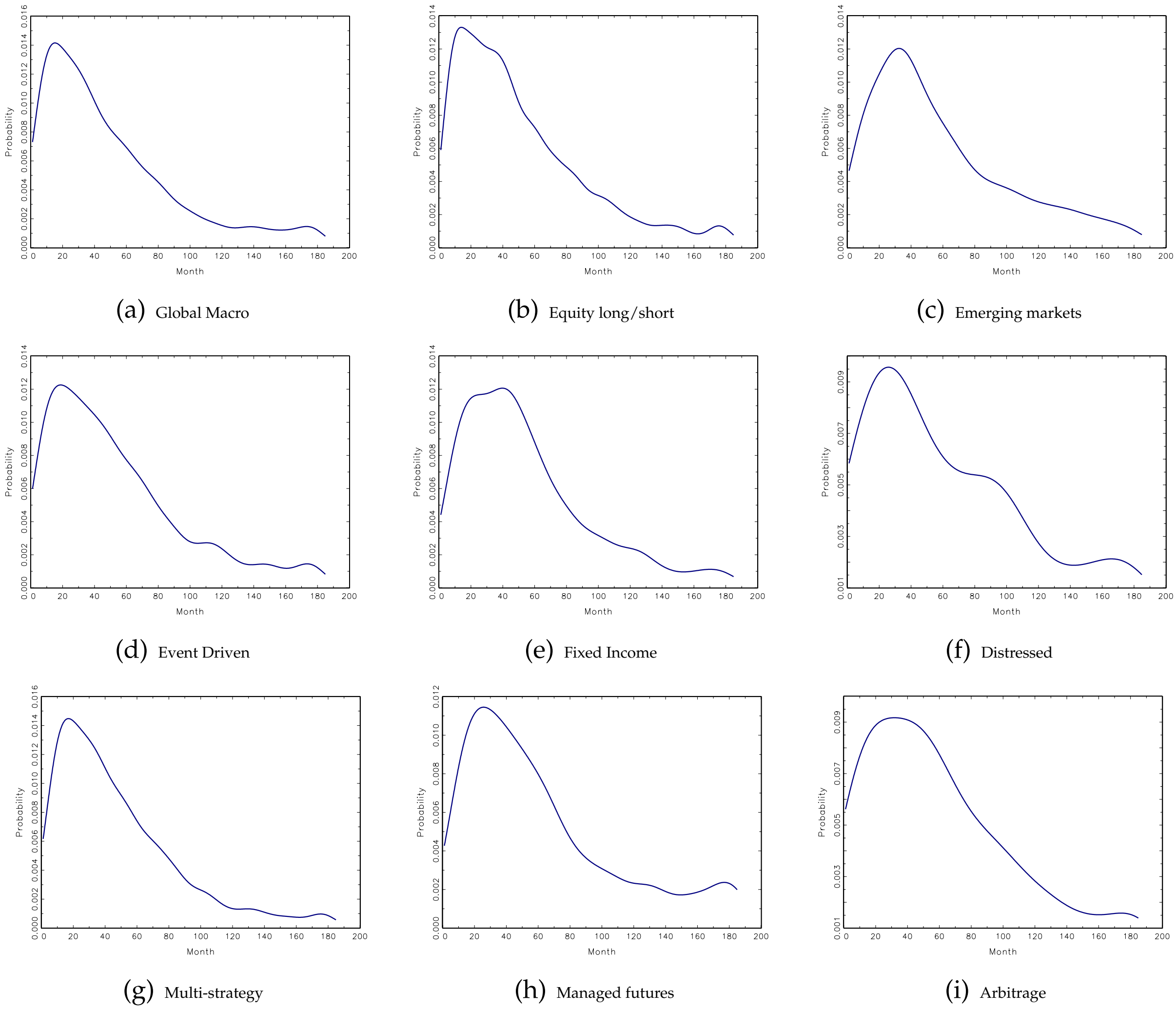
Decile	Average alpha	
	24 months sample	36 months sample
Top 10%	0.018	0.019
2.	0.014	0.016
3.	0.012	0.011
4.	0.010	0.011
5.	0.009	0.008
6.	0.008	0.010
7.	0.007	0.007
8.	0.005	0.008
9.	0.006	0.009
Bottom 10%	0.005	0.004
Average	0.009	0.010

Table 1.B.26: Investment strategy

Table demonstrates the Sharpe ratios of the top decile funds: sorted based on their deviation from the estimated common mean. In this strategy top 5%, 10%, 20%, and 33% funds are selected for the comparison, i.e. 5%, 10%, 20%, and 33% funds with the largest/smallest deviations from the common mean. The table demonstrates the average Sharpe ratios over selected funds. The table provides also average Sharpe ratio for the full sample.

36 months in-sample/12 months out-of-sample								
	5%				10%			
	Top		Bottom		Top		Bottom	
	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample
Strategy	0.318	0.259	0.295	0.427	0.306	0.317	0.321	0.416
Full sample	0.276	0.260	0.276	0.260	0.276	0.260	0.276	0.260
Difference	0.043	-0.002	0.019	0.166	0.030	0.057	0.045	0.156
	20%				33%			
	Top		Bottom		Top		Bottom	
	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample
Strategy	0.301	0.313	0.309	0.354	0.294	0.329	0.303	0.315
Full sample	0.276	0.260	0.276	0.260	0.276	0.260	0.276	0.260
Difference	0.025	0.052	0.033	0.093	0.019	0.069	0.028	0.054
36 months in-sample/24 months out-of-sample								
	5%				10%			
	Top		Bottom		Top		Bottom	
	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample
Strategy	0.223	0.324	0.252	0.289	0.222	0.336	0.269	0.274
Full sample	0.221	0.254	0.221	0.254	0.221	0.254	0.221	0.254
Difference	0.002	0.069	0.031	0.035	0.001	0.082	0.047	0.019
	20%				33%			
	Top		Bottom		Top		Bottom	
	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample
Strategy	0.225	0.316	0.268	0.260	0.244	0.311	0.277	0.250
Full sample	0.221	0.254	0.221	0.254	0.221	0.254	0.221	0.254
Difference	0.004	0.062	0.047	0.005	0.022	0.056	0.056	-0.004
36 months in-sample/36months out-of-sample								
	5%				10%			
	Top		Bottom		Top		Bottom	
Strategy	0.173	0.233	0.359	0.241	0.235	0.221	0.340	0.262
Full sample	0.161	0.233	0.201	0.233	0.201	0.233	0.201	0.233
Difference	0.012	0.000	0.158	0.008	0.034	-0.012	0.140	0.029
	20%				33%			
	Top		Bottom		Top		Bottom	
	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample	in-sample	out-of-sample
Strategy	0.238	0.211	0.296	0.278	0.250	0.217	0.266	0.265
Full sample	0.201	0.233	0.201	0.233	0.201	0.233	0.201	0.233
Difference	0.037	-0.022	0.095	0.045	0.050	-0.016	0.065	0.032

1.C Appendix: Figures



(a) Global Macro
(b) Equity long/short
(c) Emerging markets
(d) Event Driven
(e) Fixed Income
(f) Distressed
(g) Multi-strategy
(h) Managed futures
(i) Arbitrage

Figure 1.C.1: Kernel densities of hedge fund ages

The graphs illustrate Kernel densities of the hedge fund ages; x-axes denote months, y-axes probability density.

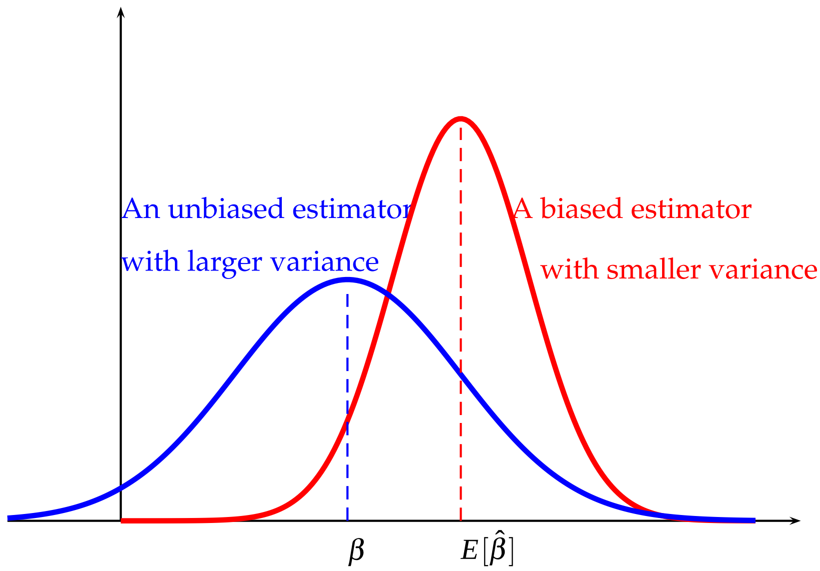
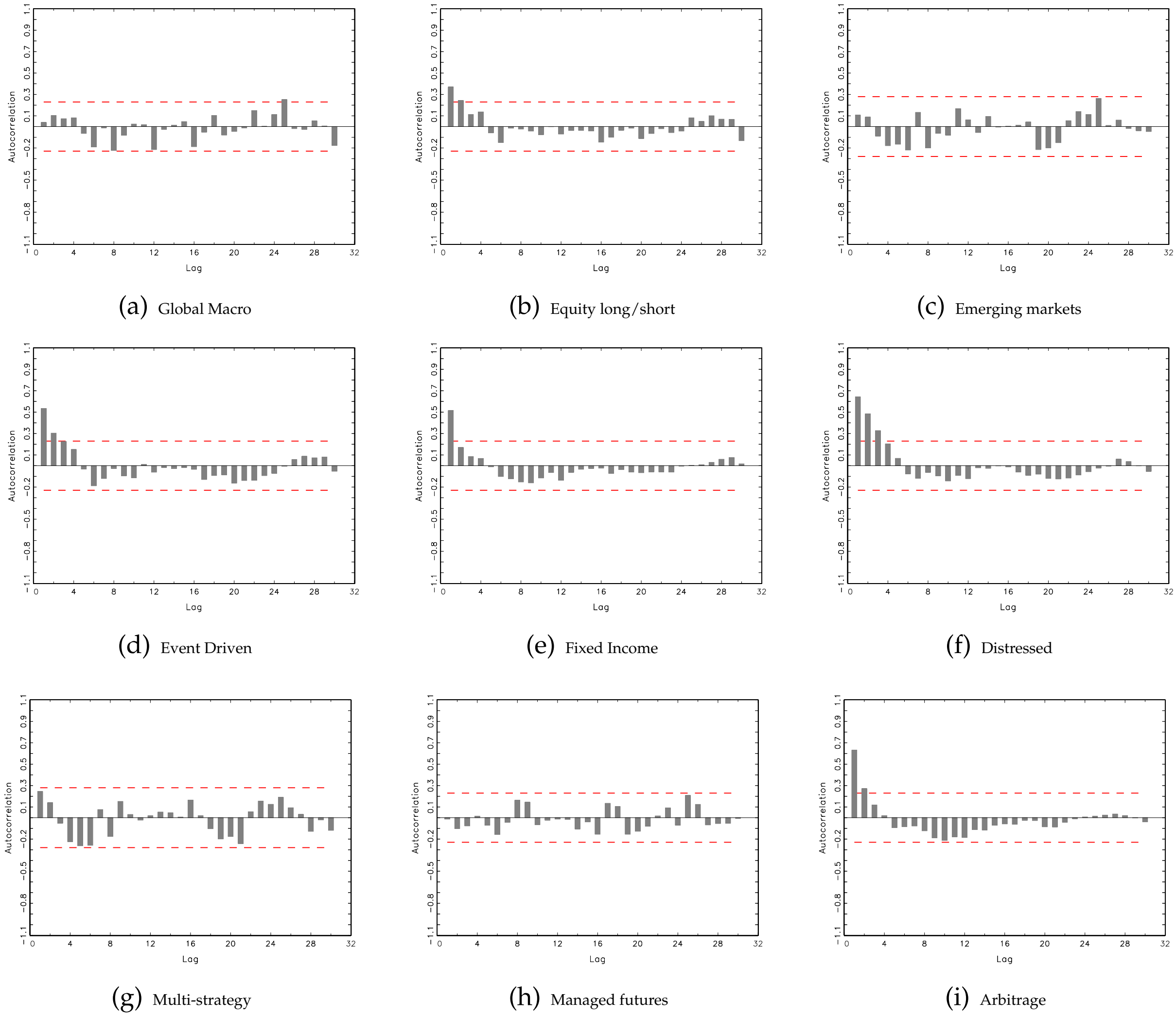
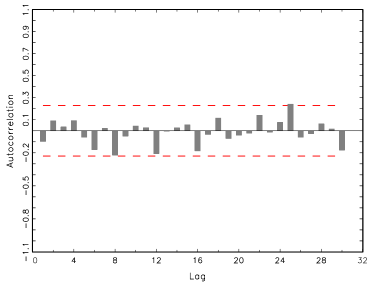


Figure 1.C.2: Idea of the shrinkage estimator
Figure demonstrates the difference between conventional estimator (blue line) and the shrunk one (red line). It plots the distribution of the estimates obtained by two different methods.

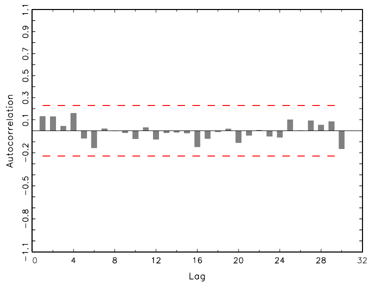


(a) Global Macro
(b) Equity long/short
(c) Emerging markets
(d) Event Driven
(e) Fixed Income
(f) Distressed
(g) Multi-strategy
(h) Managed futures
(i) Arbitrage

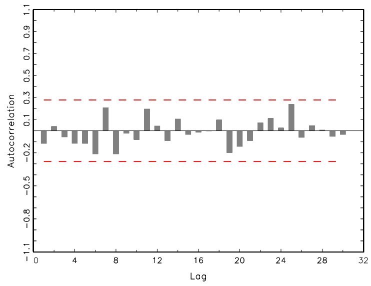
Figure 1.C.3: Autocorrelogram of smoothed returns of hedge fund style indexes
Figure demonstrates autocorrelogram of smoothed (observed) returns of hedge fund style indexes. Style indexes are computed as simple averages across all funds within the style. Styles include global macro, equity long/short, emerging markets, event driven, fixed income, distressed, multi-strategy, managed futures, and arbitrage.



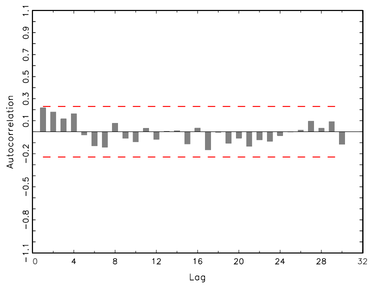
(a) Global Macro



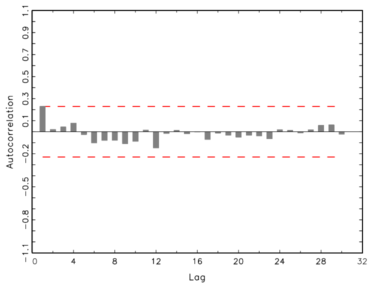
(b) Equity long/short



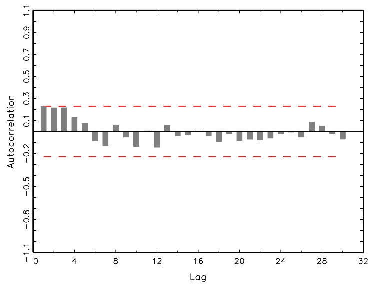
(c) Emerging markets



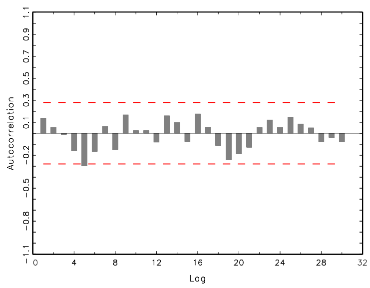
(d) Event Driven



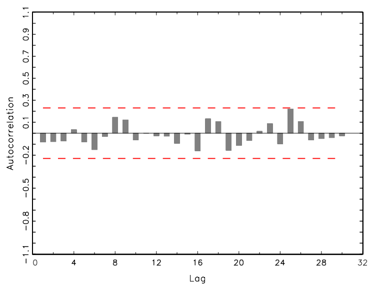
(e) Fixed Income



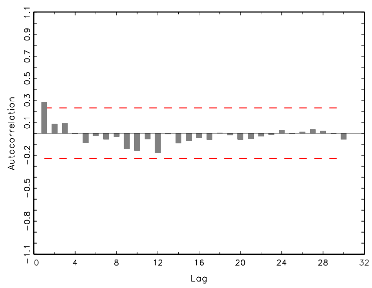
(f) Distressed



(g) Multi-strategy



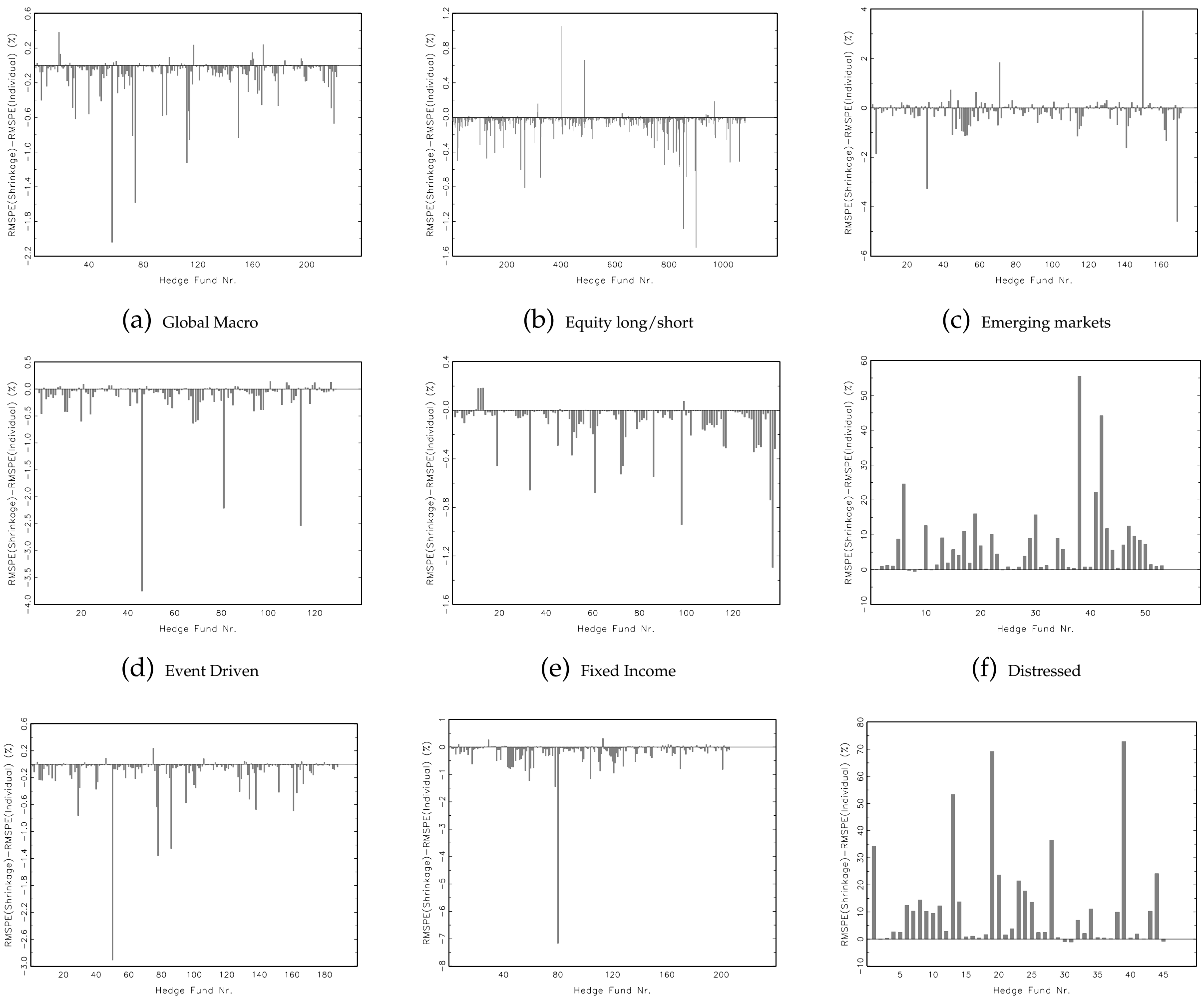
(h) Managed futures



(i) Arbitrage

Figure 1.C.4: Autocorrelogram of de-smoothed returns of hedge fund style indexes

Figure demonstrates autocorrelogram of returns of hedge fund style indexes after desmoothing. Desmoothing is proceeded as in Getmansky et al. [2004]: in order to calculate desmooting weights, it is assumed that unobserved returns can be represented as moving average process of order 2. Style indexes are computed as simple averages across all funds within the style. Styles include global macro, equity long/short, emerging markets, event driven, fixed income, distressed, multi-strategy, managed futures, and arbitrage.

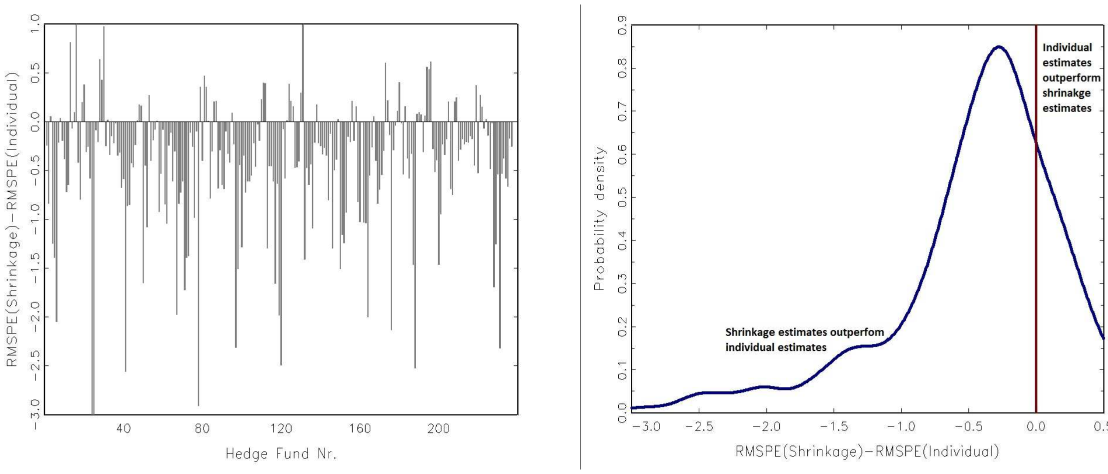


(a) Global Macro (b) Equity long/short (c) Emerging markets

(d) Event Driven (e) Fixed Income (f) Distressed

(g) Multi-strategy (h) Managed futures (i) Arbitrage

Figure 1.C.5: Barplot of the difference in the average root mean squared prediction error between the naive shrinkage and individual estimate
Figure plots the difference in average root mean squared prediction error (RMSPE) over the whole out-of-sample between the naive shrinkage and individual estimates, fund-by-fund. For each fund within the style, RMSPE is calculated using naive shrinkage and individual estimates. Negative bar mean that RMSPE for the naive shrinkage is smaller than for individual estimates and vice versa. Styles include equity long/short, global macro, fixed income, event driven, managed futures, and arbitrage.



(a) Barplot of ΔRMSPE (b) Kernel density of ΔRMSPE

Figure 1.C.6: Difference in the root mean squared prediction error between the advanced shrinkage and individual estimates
Figure (a) illustrates the barplot of the difference in the root mean squared prediction error between the naive shrinkage and the individual estimates ($\Delta RMSPE$). Figure (b) plots the kernel density of this difference. The style used for this illustration is global macro.

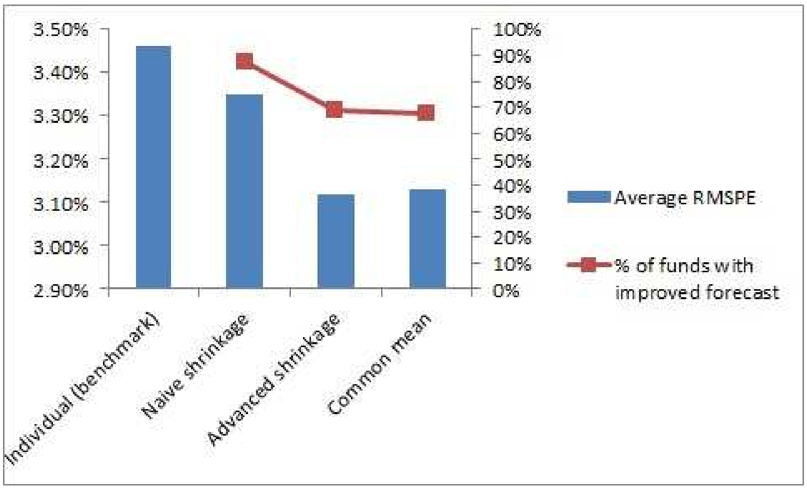


Figure 1.C.7: Barplot of the difference in root mean squared prediction error between the corresponding model and the individual estimate

The figure demonstrates average root mean squared prediction error (RMSPE, on the left y-axe) across the funds in the global macro strategy for the following models: common mean, naive shrinkage, advanced shrinkage, and individual estimates. The figure also provides the percentage of funds with improved forecast for each model comparing to the benchmark (i.e. individual estimates) (right y-axe).

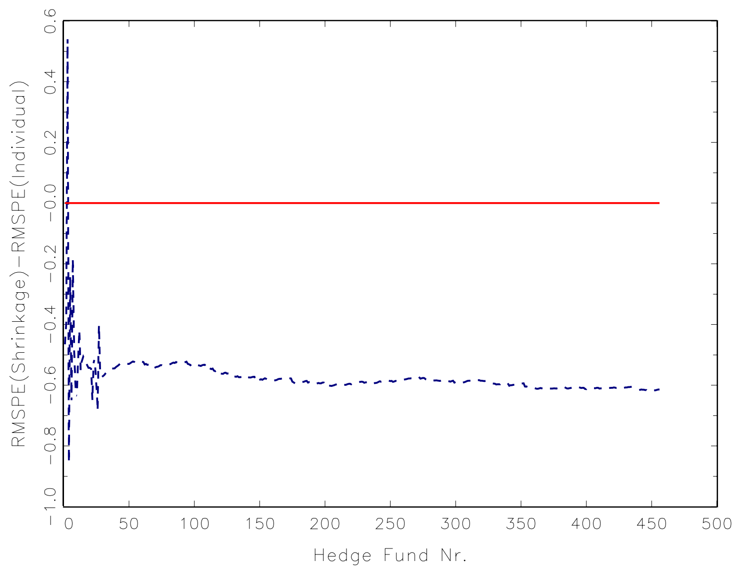


Figure 1.C.8: Change in the root mean squared prediction error difference between the naive shrinkage and individual estimate in dependence from the number of cross-section elements

Figure demonstrates how the difference in root mean squared prediction error between the naive shrinkage and the individual estimate changes with the increasing number of cross-section elements (example of equity long/short strategy).

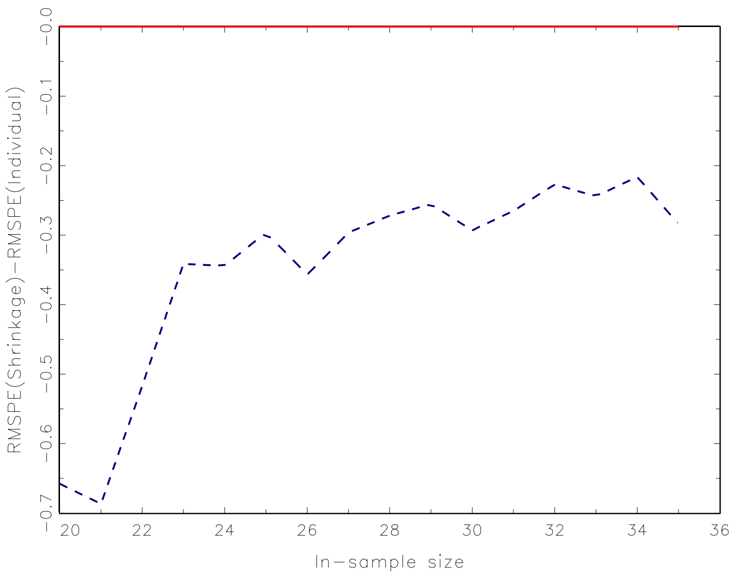


Figure 1.C.9: Change in the root mean squared prediction error difference between the naive shrinkage and individual estimate in dependence from time series length

Figure demonstrates how the difference in root mean squared prediction error between the naive shrinkage and the individual estimate changes with the increasing length of time series (example of equity long/short strategy).

Total Benefit

2.1 Introduction

Hedge funds with their diverse strategies have become an attractive investment tool for many institutional investors, including pension funds looking for enhanced portfolio performance. Pension funds even increased their hedge fund holdings during the financial crisis when a large number of hedge funds failed[1]. In this paper, we analyze the attractiveness of adding hedge funds to a typical pension fund portfolio. While doing so, we take into account that pension fund exposure to hedge funds is only around 5 to 15%.

Choosing the right hedge fund for addition to a pension fund portfolio can be challenging. The complexity of hedge fund strategies and their dynamic nature hinder consistent performance assessment. The de facto standard of hedge fund performance measurement is the alpha within the Fung and Hsieh [2001] seven factor model[2]. The resulting estimated intercept, alpha, of the model is then used to rank hedge funds. Investors chase alpha by investing more into hedge funds with high historical alpha, see Fung et al. [2008].

Alas, one has to wonder about the quality of estimated alpha due to short-samples (typically 36 returns from which 8 parameters need to be estimated in the Fung and Hsieh [2001] model), omitted variables, and the poorly specified

[1] According to the National Association of Pension Funds, UK pension funds substantionally increased their allocation in hedge funds from 1.8% in 2009 to 4.1% in 2011.

[2] Some 439 citations solely in Google Scholar attest to its widespread use. Similarly, Agarwal and Naik [2004], who propose another factor model for hedge funds, are cited 772 times.

linear factor model (the R-squared is typically only some 40%). Further, persistence of alpha is low which means that historically high alpha funds will not necessarily exhibit such high alpha again in the future, see Capocci and Hübner [2004]).

Even though alpha remains an important target of investors, there are other dimensions of portfolio performance about which investors care, such as diversification benefits, the addition of positive skewness, or the reduction of fat (left) tails of the portfolio returns distribution. Less tangible benefits such as having a hedge fund manager of high reputation or personal tailoring of products to the investor's taste should also matter to the investor but are harder to quantify.

In order to measure the total benefit to investors, we suggest using changes in certainty equivalent values (ΔCEV) of returns. We start out with a standard portfolio of 60% in stocks, 35% in bonds, and 5% in cash as our benchmark. This benchmark represents typical pension fund holdings but results do not depend on this exact choice. We sell 10% of the total portfolio (in proportion) and we invest these 10% into an alternative asset. While there is a trend for pension funds to increase allocations in hedge funds, very few pension funds allocate more than 5-15% of total assets to hedge funds (for some countries, there are restrictions on such risky investments), see International Monetary Fund [2004] and Stewart [2007].

We consider seven alternative assets: a portfolio of random hedge funds, a portfolio of mutual funds, the risk-free asset, a real estate index, a commodity index, a foreign equity index, and a random fund of funds. Our sample is from February 1994 to May 2009.

We contribute to the literature in three ways. First, we document that adding a portfolio of random hedge funds to the benchmark portfolio (hereafter, "the hedge fund portfolio strategy") significantly increases the certainty equivalent of the pension fund (ΔCEV = 0.61% annualized) during 1994-2009[1]. Such ΔCEV

[1]These results are insensitive to changes in the risk aversion coefficient used to compute the

of 0.61% in a marginal move (only 10% of the benchmark are replaced) corresponds to a ΔCEV of 4.93% for a 100% investment into a random portfolio of hedge funds, more than 8 times larger than 0.61%. The hedge fund portfolio strategy turns out to be significantly more beneficial than adding any of the alternative assets, namely, real estate, commodity, foreign equity, mutual funds, or fund of funds. Measuring the out-of-sample performance in terms of the Sharpe ratio or alpha gives similar rankings of the alternative assets.

Our second contribution details the conditions under which ΔCEV works best as a performance measure. Results turn out to be stronger once we condition on in-sample information and then measure the out-of-sample ΔCEV over and beyond the hedge fund portfolio strategy. Sorting based on popular performance measures (Sharpe ratio, alpha, ΔCEV, moments of the return distribution, plus others) yields ΔCEVs of up to 0.33%. While the sorts based on mean and ΔCEV perform best (0.33% and 0.32%), show the Sharpe ratio (0.30%) and alpha (0.28%) less improvement, even though they are economically close. Value at risk (0.01%) and standard deviation (-0.06%) perform worst.

The situation changes markedly once we turn to sorts conditional on either the economic conditions or on the statistical properties of the hedge fund returns. In terms of economic conditions, we incorporate information on market volatility and funding liquidity. We dynamically adjust the proportion of the benchmark allocated to the hedge fund portfolio strategy (10% in the base case) within the interval from 5% to 15% according to relative changes in market volatility (VIX) or funding liquidity (NOISE as measured in Hu et al. [2013][1]). In terms of statistical properties of the hedge fund returns, we assess the nonnormality of the returns (NONNORMAL) and the lack of explanatory power of the Fung and Hsieh [2001] seven factor model in explaining the returns (LOW R-SQUARED)[2]. We find in times of economic distress and when faced with

certainty equivalent (γ= 2 for power utility) or to the exact composition of the benchmark; see the robustness section for details.

[1]We calculate the new proportion as 10% times the ratio of VIX_t over $VIX_t - 1$ (and similarly for NOISE). Next, we winsorize the new proportion to lie between 5% and 15%.

[2]NONNORMAL we call a hedge fund is its returns have absolute skewness and kurtosis in

hedge fund returns which are econometrically problematic, that sorts based on ΔCEV perform best, narrowly beating sorts based on mean (by some 0.01%), but substantially outperforming sorts based on Sharpe ratio (by some 0.07%) and alpha (by some 0.19%). Further, managed futures deliver the highest ΔCEV (0.10%) of all hedge fund styles.

Our third contribution is to show that persistence of ΔCEV is higher from one two-year interval to the next than for Fung and Hsieh [2001] alpha.

Our results are robust to a battery of challenges. We add a single random hedge fund instead of a portfolio of hedge funds, use de-smoothed hedge fund returns, change the risk aversion coefficient of the CEV calculation up and down, change the benchmark, plus conduct a number of additional minor changes.

Based on a literature review in Section 2.2, the paper develops the hypotheses and introduces the econometric methodology for testing our performance measure in Section 2.3. All data are presented in Section 2.4. Results follow in Section 2.5 while robustness checks are presented in Section 2.6. Section 2.7 concludes.

2.2 Literature

We relate broadly to the vast literature on performance measurement which can be split into three main groups. In the first group, we have the mean-variance inspired ratios such as the Sharpe ratio. The second group encompasses factor models. The final group consists of utility based performance measurements to which the CEV belongs and where we also subsume stochastic dominance based measures. We relate our work to these three groups and discuss their merits and problems next.

the top tercile for both measures. LOW R-SQUARED we call a hedge fund its R-squared from the Fung and Hsieh [2001] model is in the lowest tercile.

Various ratios have been suggested as performance measures, often much earlier than the advent of hedge funds, and they are still in widespread use. We use the measures of Sharpe [1966], Sortino et al. [1999]. Even though ratios are widely used in the hedge fund literature (Ackermann et al. [1999], Liang [1999]), they were criticized by e.g. Kao [2002] and Amin and Kat [2003]. Most of the ratios assume normally distributed returns. Use of derivatives and dynamic strategies within hedge funds can create non-normally distributed returns with non-zero skewness and fat tails. Plus, hedge fund returns can exhibit serial correlation which could lead to over-estimated Sharpe ratios, see Lo [2002].

Due to the problems associated with ratios, performance assessment for hedge funds is firmly centered around the alpha within the factor models of Fung and Hsieh [2001] or Agarwal and Naik [2004]:

$$r_t - r_{f,t} = \alpha + \sum_{k=1}^{K} \beta_k X_{k,t} + \varepsilon_t, \qquad (2.2.1)$$

where $(r_t - r_{f,t})$ is a hedge fund return at time t in excess of the risk-free rate, α is the abnormal return, $X_{k,t}$ $(k = 1, \ldots, K)$ are the K relevant risk factors at time t, β_k $(k = 1, \ldots, K)$ are the risk exposures to be estimated, and ε_t is an error term.

However, even if a hedge fund delivered high alpha in the past, it is questionable whether the hedge fund can also deliver comparable alpha in the future other than by chance. While some authors find significant persistence in alpha (Agarwal and Naik [2000b] and Edwards and Caglayan [2001]), others find limited evidence for such persistence (Capocci and Hübner [2004]). Berk and Green [2004] provide a theoretical argument for alpha being driven down to zero in a mutual fund setting caused by diminishing mutual fund performance due to fund flows. Still, without such diminishing returns, managerial skill could persist over time and should then lead to persistence in alpha. Berk and van Binsbergen [2013], Kosowski et al. [2007], Jagannathan et al. [2010], and Naik et al. [2007] use factor models to find out if alpha does exist[1]. Within measures based on factor models we also consider timing skills estimators of Henriksson and Merton [1981] and Treynor and Mazuy [1966]. Alas, none of these papers moves beyond the factor model structure. Thus, other dimensions of total benefit, such as diversification, skewness, and tail risk, might be omitted.

Out of these additional dimensions of total benefit, the potential for diversification benefits of managed futures have been voiced by Kat [2004], Amin and Kat [2002], Lhabitant and Learned [2002], and Rollinger [2012]. However, those authors look only at diversification and do not introduce a performance measure, such as ΔCEV, capable of measuring the other dimensions total benefits.

In general, most of the traditional performance measures (ratios, alpha) can be subject to manipulation by the fund manager. Thus, Goetzmann et al. [2007] investigate manipulation-proof performance measures and suggest a metric which is almost identical to our ΔCEV measure. They show their measure to be robust to manipulations of the underlying return distribution, to the introduction of time variation in the underlying distribution, and to the excessive use of dynamic trading strategies by fund managers. Brown et al. apply their measure to hedge funds but solely assess its quality as a manipulation-proof measure; confirming the theoretical predictions of Goetzmann et al. [2007]. We on the other hand advocate the ΔCEV measure as a much broader tool to incorporate a number of other dimensions of total benefit that accrue to investors holding hedge funds (or other alternative assets) in addition to some benchmark portfolio. Also, conditioning on past ΔCEV significantly outperforms conditioning on past alpha. We define our total benefit (CEV) measure as follows:

$$CEV = U^{-1}(E[U(\frac{1+r_t}{1+r_{f,t}})]) - 1, \tag{2.2.2}$$

[1]See also Ferson and Lin [2013] on the use of alpha in a world with heterogeneous investors who differ in their preferences.

where r_t is a hedge fund return at time t, $r_{f,t}$is risk-free rate at time t, and $U(\cdot)$ is a power utility function with risk aversion coefficient $\gamma = 2$. The ΔCEV is then the difference between the augmented portfolio and the original portfolio.

The paper closest to ours is Bali et al., who investigate if hedge funds outperform stock or bond markets. They use the concept of almost stochastic dominance (ASD) of Leshno and Levy [2002] as their performance measure, as well as the manipulation-proof performance measure of Goetzmann et al. [2007]. They look at pure (100%) strategies whereas we look at marginal strategies of allocating only 10% of the benchmark to hedge funds or other alternative assets, which is more realistic. Also, only our marginal analysis informs about the effects on performance due to interactions between the benchmark and the alternative assets.

2.3 Hypotheses and methodology

We test four hypotheses with respect to ΔCEV in its capacity to measure total benefit.

2.3.1 The hedge fund portfolio strategy adds total benefit (ΔCEV)

As a typical pension fund (also referred to as " benchmark") we consider a portfolio which consists of three assets: 60% stocks, 35% bonds, and 5% cash. As a fourth asset we add a random portfolio of hedge funds, a random portfolio of mutual funds, the risk-free asset, a real estate index, a commodity index, a foreign equity index, or a random fund of funds. We shrink the initial pension fund portfolio weights by a factor of 0.9 and invest the remaining 10% into one of the seven alternative assets.

For the hedge fund portfolio strategy, we start with the 25^{th} monthly return, so that we keep a sample of 24 prior returns which, if later needed, we can use for conditioning information. We randomly pick out of all existing hedge funds (with at least prior 24 returns) 20 hedge funds and denote the sum of the equally weighted 25 th returns as r_1. Then, we move one month forward and check if the previously chosen 20 hedge funds have not been delisted from the database. If any of these hedge funds has been delisted from the database, we put a zero return instead of the missing return for this hedge fund and randomly pick another hedge fund with at least 24 past returns as future replacement[1]. We repeat this procedure until the end of our sample which provides us with a time series of T returns of a portfolio of 20 random hedge funds $(r_1, r_2, \ldots, r_T)$. Our choice of 20 hedge funds is driven by two reasons. First, from private conversation with a hedge fund and fund of fund manager, we found out that 20 is a typical value for the number of hedge funds into which a pension fund invests. Second, according to Lhabitant and Learned [2002], 20 hedge funds are enough to provide reasonable diversification for naively constructed portfolios.

We similarly construct the time series for a portfolio of 20 random mutual funds. For a random fund of funds, we only pick a single fund of funds since it already constitutes a portfolio of hedge funds. For the other alternative assets, where only a single index exists, we simply use that index as our time series.

To test our first hypothesis, we calculate ΔCEV as the CEV of the initial benchmark minus the CEV of the hedge fund portfolio strategy. We repeat this calculation 1000 times where each time we randomly pick a set of 20 hedge funds. This generates a distribution of ΔCEV values under the null of no outperformance of the benchmark. We can then use a t-test for the mean of the ΔCEV values being significantly different from zero.

[1] From Hodder et al. [2013] we know that a delisting return is similar to an average hedge fund return. Thus, we are conservative here when we use a zero return instead.

2.3.2 The hedge fund portfolio strategy outperforms adding alternative assets

To test the second hypothesis, we calculate ΔCEV as the CEV of the competing strategy (shrunk benchmark plus 10% weight in one of the alternative assets) minus the CEV of the hedge fund portfolio strategy. We again repeat this calculation 1000 times where each time we randomly pick hedge funds (or mutual funds or fund of funds, as needed). This generates a distribution of ΔCEV and we use a t-test for the mean of the ΔCEV values being significantly different from zero.

2.3.3 Conditioning on in-sample information improves performance

Our third hypothesis states that conditioning the set of hedge funds, from which to randomly pick, by using in-sample information improves investment performance. In particular, conditioning on ΔCEV should yield significantly higher out-of-sample ΔCEV than conditioning on other traditional performance measures which only partially account for the total benefits accruing to the investor.

The third hypothesis is tested in a similar way as the second hypothesis by testing the shrunk benchmark with 10% invested in conditional hedge fund strategies against the hedge fund portfolio strategy. We condition on the extreme 5% in-sample values of the quantity of interest, e.g. the highest ΔCEV, based on the 24 returns before the month for which we want to pick a return. We then randomly pick 20 hedge funds out of that conditional set. We repeat this calculation 1000 times and use a t-test for the mean of the ΔCEV values being significantly different from zero.

For the third hypothesis we use two different groups of conditioning information:

1. First, we use various performance measures: total benefit (CEV), alpha estimates from the seven factor model of Fung and Hsieh (2001), changes in the first four moments of the augmented benchmark over the initial benchmark, and the following traditional performance measurements: Sharpe [1966], Sortino et al. [1999], Information ratio, Henriksson and Merton [1981], and Treynor and Mazuy [1966]. We provide all formulas in the Appendix 2.A on Performance Measures. Here, we also test for the influence of different economic and econometric conditions on the ranks provided by different measures. Regarding the economic conditions, we use option implied volatility (VIX) and funding liquidity (NOISE measure of Hu et al. [2013]) to measure economic distress. We adjust the weight of the random hedge fund portfolio in a pension fund (10% in the base case) by the ratio of VIX_t to VIX_{t-1} (or similarly for NOISE). Next, we winsorize the new weight to lie in the interval from 5% to 15%. As for econometric conditions, we reduce the set of hedge funds to those demonstrating heavily non-normal returns (with absolute skewness and kurtosis to be both in the highest tercile of all hedge funds) or having an R-squared for the Fung and Hsieh [2001] factor model in the lowest tercile of all hedge funds.

2. Second, we condition on the investments styles, namely: equity long/short, multi-strategy, global macro, event driven, technical trading, managed futures, emerging markets, relative value, fixed income, sector trading, merger arbitrage, distressed, convertible arbitrage, stock index, and other.

2.3.4 Total benefit is more persistent than alpha

For the fourth hypothesis, we consider six consecutive 48-month periods, starting with the 1^{st}, the 25^{th}, the 49^{th}, ... observation. Each of these periods is divided into two 24 month sub-periods: a formation period (months 1-24) and an evaluation period (months 25-48). For each hedge fund i, which survives throughout the whole period, we compute the ΔCEV during the formation period (denoted as ΔCEV_{1i}) and the evaluation period (denoted as ΔCEV_{2i}). Finally, we regress ΔCEV_{2i} on ΔCEV_{1i}:

$$\Delta CEV_{2i} = a_{CEV} + b_{CEV}\Delta CEV_{1i} + \omega_i, \tag{2.3.1}$$

where a_{CEV}, b_{CEV} are the parameters to be estimated, and ω_i is an error term. We run this regression on each of our six 48-month periods and also jointly for a stacked regression. For total benefit (CEV) to be persistent, the slope coefficient b_{CEV} should be positive and significant.

We repeat this study for the estimated alpha from a seven factor Fung and Hsieh [2001] model:

$$\alpha_{2i} = a_\alpha + b_\alpha \alpha_{1i} + v_i, \tag{2.3.2}$$

where $\alpha_{1i}(\alpha_{2i})$ are the alpha estimates in the formation (evaluation) period for fund i; a_α, b_α are the parameters to be estimated, and v_i is an error term.

2.4 Data

Our benchmark consists of stocks (Datastream S&P 500 index total return), bonds (Datastream Barclays US aggregate Bond index), and cash (Datastream US T-bill 3 month). The sample runs from February 1994 until May 2009 and consists of monthly returns.

For hedge fund information we use the MOAD database described in Hodder et al. [2013]. MOAD is a merged database of six commercially available databases (CISDM, Barclays, TASS, HFR, Altvest, and Eurekahedge). We use only USD-denominated, net-of-fees returns which leaves us with 19,109 hedge funds. The descriptive statistics of our sample are presented in Table 2.B.1, Panel A. We document excess kurtosis and left-skewness in hedge fund returns, suggesting that returns are often not normally distributed. Also, returns exhibit positive serial correlation.

Mutual fund returns are extracted from the Morningstar database and account for 259,088 individual funds. The descriptive statistics are in Table 2.B.1, Panel B. Compared to the hedge funds, mutual funds demonstrate smaller average returns, kurtosis, and skewness, but larger standard deviation.

[Table 2.B.1 about here]

Hedge funds differ from other asset classes in many aspects. One of them is the absence of strict regulation. This leads to database biases as reporting is voluntary. We address those biases as follows. First, our joint database is free of survivorship bias because it contains both live and dead funds. Second, to control for the instant history bias, we delete the first 12 months of each hedge fund's returns. Our main results we compute on the reported returns as we find them in the database. However, results are robust to using de-smoothed returns; see Getmansky et al. [2004] and robustness results in Section 2.6. For funds of funds, we extract the USD-denominated, net-of-fees returns in a similar fashion to the hedge fund returns from our database and are left with 5,216 funds of funds.

We also use the seven factors of the Fung and Hsieh [2001] model which are available at David A. Hsieh's Hedge Fund Data Library[1].

For our alternative assets, we use a real estate index (Datastream US real estate index), a commodity index (Datastream S&P-GSCI index), and a foreign equity index (Datastream MSCI EAFE). For information on the state of economy, we use the option implied volatility VIX (Datastream) as a proxy for market volatility and the NOISE measure from Hu et al. [2013] as a proxy for funding liquidity.

[1]https://faculty.fuqua.duke.edu/ dah7/HFData.htm

2.5 Results

We are now ready to present our results for each of our four hypotheses in turn.

2.5.1 The hedge fund portfolio strategy adds total benefit

In line with our first hypothesis, we show in the first row of Table 2.B.2 that the hedge fund portfolio strategy generates an annualized improvement in CEV of 0.61% (=ΔCEV)[1]. To appreciate the magnitude a ΔCEV of 0.61%, note that we are conducting marginal analysis. A pure (as opposed to marginal) strategy of going from 100% investment in the benchmark to 100% investment in a random portfolio of hedge funds leads to a ΔCEV of 4.93%, more than 8 times larger than 0.61%.

[Table 2.B.2 about here]

2.5.2 The hedge fund portfolio strategy outperforms adding other alternative assets

We significantly reject that adding the risk-free asset (ΔCEV=-0.42%), a real estate index (-0.08%), a commodity index (-0.21%), a foreign equity index (-0.56%), a random fund of funds (-0.18%), or a portfolio of random mutual funds (-0.29%) is as beneficial as the hedge fund portfolio strategy. All ΔCEV values in Table 2.B.2 are negative with zero p-values. The poor performance of fund of funds compared to portfolios of hedge funds seems to lie in the higher total fees. All hedge funds and fund of funds returns are observed post-fee. Still, funds of funds have two layers of fees, one at the hedge fund level and one at

[1] The weights of the benchmark are: 60% in stocks, 35% in bonds, and 5% in cash. The weights of the augmented benchmark are: 54% in stocks, 31.5% in bonds, 4.5% in cash, and 10% in a random portfolio of hedge funds.

the fund of funds level. See also Brown et al. [2008].

We also provide assessments of the different alternative assets in terms of changes to the Sharpe ratio (ΔSR) and alpha ($\Delta\alpha$). The rankings of the methods do not change at all when using alpha and minimally so when using the Sharpe ratio. We thus refrain from reporting the Sharpe ratios and alphas in the remaining tables.

2.5.3 Conditioning on in-sample information improves performance

In our third hypothesis, we want to assess, to what extent we can improve upon the hedge fund portfolio strategy by conditioning on in-sample information during the last 24 months. We use two sets of conditioning information in turn: performance measures and investment styles.

2.5.3.1 Conditioning on performance measures

When conditioning on performance measures, we use the previous 24 months of hedge fund returns to compute the different measures. The other alternative assets we do not investigate any more since the hedge fund portfolio strategy performs much better. Instead of picking random hedge funds for addition to the benchmark, we now pick hedge funds from the set with the highest 5% of measures (lowest 5% for kurtosis and standard deviation).

[Table 2.B.3 about here]

Our comparisons are computed with respect to the hedge fund portfolio strategy. The results are summarized in Table 2.B.3. The largest improvement in ΔCEV is a sort based on the highest in-sample mean with 0.33%. Interestingly, each group of performance measures has a representative close to this value: a sort by ΔCEV yields 0.32%, a sort by the Sharpe ratio 0.30%, and a sort by alpha 0.28%. It emerges that all these performance measurements benefit strongly from a higher mean and thus perform similarly. Sorts by other ratios perform less well (0.13% to 0.25%), except for the ratio of Sortino et al. [1999] with 0.31%. The higher moments achieve solid results (kurtosis 0.25% and skewness 0.14%) but standard deviation performs abysmally at -0.06%. The explanation is that standard deviation of a hedge fund does not account for co-variation with the benchmark portfolio, a problem which also affects skewness and kurtosis to a lesser degree. Value-at-risk also performs poorly at 0.01%.

[Table 2.B.4 about here]

Next, we wonder under what conditions ΔCEV provides an edge over the other measures such as the Sharpe ratio or alpha. Results are collected in Table 2.B.4. For one, ΔCEV should work particularly well when the economy is volatile and under stress. Starting from our base case of the hedge fund portfolio strategy, we increase (decrease) the percentage (initially 10%) allocated to the hedge fund portfolio strategy when the economy is under more (less) stress so that the percentage lies within the interval from 5% to 15%. We measure stress in two ways, namely, as the growth rate of monthly option implied volatility (VIX) and as the growth rate of the monthly funding liquidity measure (NOISE) of Hu et al. [2013]. The results for VIX in Table 2.B.4, Model VIX, show that ΔCEV and mean significantly outperform all other performance measures with 1.06% and 1.05%[1]. The Sharpe ratio comes in with only 0.98% and alpha with only 0.82%. The results for the funding liquidity measure in Table 2.B.4, Model NOISE, suggest a similar picture, although at a slightly lower level. Mean and ΔCEV lead the pack with 0.72% and 0.71%, while Sharpe ratio and alpha come in significantly worse with 0.64% and 0.49%.

Other than under economic stress, ΔCEV should also work well when the historical hedge fund returns are non-normal or when factor models do not

[1] All differences which are larger than 0.01% are statistically significant, see Table 2.B.11.

work well (i.e. funds might be more sophisticated and dynamic in their investment strategies; features which might not be captured by the conventional Fung and Hsieh [2001] risk factors). We restrict the initial sample of all hedge funds to only those hedge funds from the highest terciles of kurtosis and absolute skewness (NONNORMAL) or to funds from the lowest tercile of R-squared according to the Fung and Hsieh (2001) model (LOW R-SQUARED). Here, we keep the investment proportion into the random portfolio of hedge funds at 10% but we restrict the sets of hedge funds to choose from according to our above criteria. The results for NONNORMAL in Table 2.B.4, Model NONNORMAL, show that ΔCEV and mean outperform all other performance measures with 0.58% and 0.57%. The Sharpe ratio comes in with only 0.49% and alpha with only 0.43%. The results for LOW R-SQUARED in Table 2.B.4, Model LOW R-SQUARED, give a somewhat different picture. Kurtosis is the leading measure with 0.45%, followed by ΔCEV with 0.40%, while Sharpe ratio and alpha come in significantly worse with 0.37% and 0.26%. We conclude that ΔCEV is the overall best performance measure, in particular in times of economic stress and for econometrically problematic hedge fund returns.

2.5.3.2 Conditioning on investment style

As a final set of conditioning information, we use the largest 14 self-reported investment styles in the hedge fund database. The remaining styles (labeled "other") account for a total of 9% of all hedge funds. We present the results in Table 2.B.5. The investment style which gives the highest improvement in terms of ΔCEV is the managed futures style (also known as CTA) with a ΔCEV of 0.10% and a zero p-value. This result is consistent with Kat [2004] and Rollinger [2012] who show empirically that investing in managed futures funds reduces overall risk at limited costs.

[Table 2.B.5 about here]

A number of styles (multi-strategy, global macro, and sector trading) exhibit positive ΔCEVs (0.06% to 0.08%) with zero p-values. The least beneficial styles include technical trading, relative value, fixed income, merger arbitrage, convertible arbitrage, and stock index with negative average Δ CEVs(-0.15% to -0.63%).

2.5.4 Total benefit is more persistent than alpha

As there is considerable disagreement about the persistence of alpha, we are keen to show that ΔCEV is more persistent than alpha. We adopt a methodology commonly used in alpha persistence analysis when we estimate the regressions in Equations (2.3.1) and (2.3.2): Cross sectional ΔCEVs (alphas) estimated in one 24 months period are regressed on cross sectional ΔCEVs (alphas) estimated in a subsequent 24 months period.

[Table 2.B.6 about here]

We exhibit in Table 2.B.6 the estimated slope coefficients for six windows of two year estimation periods and two year evaluation periods. Periods always start with the February return of the beginning year and end with the January return of the ending year. The estimate of the slope coefficient for total benefit (b_{CEV} from Equation (2.3.1)) is positive and significant in four out of six sample periods: 1994-1998, 2000-2004, 2002-2006, and 2004-2008. The period 1998-2002 covers the dotcom bubble. Brunnermeier and Nagel [2004] claim that, prior to the dotcom bubble bursting in March 2000, hedge funds allocated substantial shares of their capital to technology stocks, thus gaining during the up-market. Even though they were selling the stocks prior to decline they could not completely avoid the downturn as the bubble burst.

Slope estimates in the alpha regressions are significantly positive in only two out of six periods: 1994-1998 and 1996-2000. Similar to ΔCEV, alpha does not show persistence during the dotcom bubble in the period 1998-2002. After 2000, there is no or even negative persistence in alpha.

We also estimate aggregate slope coefficients over the whole period 1994-2008 by stacking the six periods. The results are provided in the last row of Table 2.B.6. The aggregate slope coefficient on total benefit is 0.081 and has a zero p-value. On the contrary, alpha has an insignificant aggregate slope coefficient of 0.035. We conclude that while alpha is not persistence during our sample, ΔCEV is strongly persistent.

2.6 Robustness

We perform several robustness checks for the main unconditional results reported in Table 2.B.2(hypotheses 1 and 2). Our results are robust to adding single random hedge funds instead of a portfolio of hedge funds, using desmoothed hedge fund returns, changes in the risk aversion coefficient of the CEV calculation, changes to the benchmark, and a number of smaller methodological changes.

2.6.1 Single hedge fund instead of portfolio of hedge funds

Even though institutional investors such as pension funds typically augment their benchmark with a portfolio of some 20 hedge funds, we test if adding a single hedge fund is still beneficial. We measure the ΔCEV of the alternative assets over the addition of a single random hedge fund and report the results in Table 2.B.7. ΔCEVs are about 0.04% lower than in Table 2.B.2 for all alternative assets. This difference quantifies the added benefit of diversification by using a random portfolio of hedge funds as opposed to a single random hedge fund.

[Table 2.B.7 about here]

2.6.2 Return de-smoothing

We de-smooth returns as suggested in Getmansky et al. [2004]. De-smoothed returns are characterized by higher volatility than smoothed (observed) returns. Therefore, it is not surprising that the ΔCEV values in Table 2.B.8 are slightly smaller by about 0.01% than in Table 2.B.2.

[Table 2.B.8 about here]

2.6.3 Different coefficients of risk-aversion

When computing ΔCEVs in our base case, we set the risk aversion coefficient of the power utility function in Equation (2.2.2) to $\gamma = 2$. We take this to be a fairly typical value for a diversified investor. We replicate the main results with $\gamma = 4$ (Table 2.B.9, Panel A) and $\gamma = 1$ (Table 2.B.9, Panel B). Our results remain largely unchanged and strongly significant. In general, the move to less risk aversion ($\gamma = 1$) narrows the gap between the random portfolio of hedge funds and the other alternative assets by around 0.10%, while the move to more risk aversion ($\gamma = 4$) widens this gap by around 0.20%. All Δ CEVs remain negative, except for real estate in the low risk aversion setting ($\gamma = 1$) which turns out to be 0.12%.

[Table 2.B.9 about here]

2.6.4 Different weights for the benchmark portfolio composition

Instead of our typical pension fund portfolio with the largest allocation in stocks (60% in stocks, 35% in bonds, and 5% in cash), we try the following portfolio as a benchmark: 35% in stocks, 60% in bonds, and 5% in cash. The results in Table 2.B.10 move around somewhat but are on average almost unchanged. All

ΔCEVs remain negative but for real estate which turns out to be 0.06%.

[Table 2.B.10 about here]

2.6.5 Other robustness checks

We change a number of methodological choices and find that our main results of Table 2.B.2 do not change much and stay significant. We use

- different in-sample size of 36 returns instead of 24 returns, see Table 2.B.11
- different weights for additions to the benchmark of 5% and 25% (instead of 10% in the base case), see Table 2.B.12
- only hedge funds open to new investment, see Table 2.B.13

Altogether we find that our results are very stable with respect to all changes discussed.

2.7 Conclusion

Pension funds have been exploring alternative assets beyond stock and bond on a limited scale. Possibly driven by the financial crisis, those stakes have been increased to around 10% of the benchmark portfolio. Still, we know little about which alternative assets to add and how to account for their total benefits to the pension fund which might include diversification, the addition of positive skewness, and the avoidance of tail risk. We advocate the use of changes in certainty equivalent values (ΔCEV) for performance measurement as the main traditional measures, Sharpe ratio and Fung and Hsieh [2001] factor model alpha, are problematic. The Sharpe ratio assumes normally distributed returns and factor models suffer from low persistence and explanatory power.

Using a large merged database of hedge funds returns from February 1994 to May 2009, we find that the hedge fund portfolio strategy improves benchmark performance by an annualized ΔCEV of 0.61% and is superior to adding other alternative assets. This advantage can be further improved upon by conditioning on in-sample Δ (yielding an additional Δ CEV of 0.32%). However, some other performance measures exhibit similar results. The advantage of the ΔCEV performance measure comes to light, when we further condition on economic stress (high option implied volatility (VIX) or high funding liquidity (NOISE measure of Hu et al. [2013]) or on econometrically problematic hedge fund returns (NONNORMAL or LOW R-SQUARED). As a result, we can push the additional ΔCEV up as high as 1.06% in case of VIX. In terms of hedge fund styles, the highest performance gain comes from conditioning on the managed futures style.

Persistence of total benefit is significantly positive in our sample while it is insignificant for alpha. Results presented in the paper are robust to a wide-ranging battery of tests.

2.A Appendix: Performance Measures

$\tilde{r}$ - excess return, $\tilde{r}_b$ - return of the benchmark, $\tilde{m}_t$ - excess market return, F_l - cdf.

Measure	Formula
Sharpe [1966]	$\frac{E[\tilde{r}]}{\sqrt{Var[\tilde{r}]}}$
Sortino et al. [1999]	$\frac{E[Max(\tilde{r},0)]}{\sqrt{E[Min^2(\tilde{r},0)]}}$
Information Ratio	$\frac{E[\tilde{r}-\tilde{r}_b]}{\sqrt{Var[\tilde{r}-\tilde{r}_b]}}$
Henriksson and Merton [1981]	$\gamma_2 : \tilde{r}_t = \gamma_0 + \gamma_1 \tilde{m}_t + \gamma_2 \tilde{w}_t + \varepsilon_t ; \tilde{w}_t = Max(-\tilde{m}_t, 0)$
Treynor and Mazuy [1966]	$\gamma_2 : \tilde{r}_t = \gamma_0 + \gamma_1 \tilde{m}_t + \gamma_2 \tilde{w}_t + \varepsilon_t ; \tilde{w}_t = \tilde{m}_t^2$
Value-at-Risk (q-quantile)	$\inf\{l \in \mathbb{R} : F_l \geq q\}$

2.B Appendix: Tables

Table 2.B.1: Descriptive statistics

The summary statistics are the equally weighted cross-sectional averages, standard deviations, and medians of the: mean monthly return, μ; the standard deviation of monthly returns, σ; the skewness, Skewness; the excess kurtosis, Kurtosis. The sample is February 1994 to May 2009. Panel A covers hedge funds, Panel B mutual funds.

(a) Panel A.Hedge fund returns

	Mean	Std.dev	Median
μ	0.0052	0.0335	0.0058
σ	0.0469	0.0439	0.0355
Skewness	-0.1332	1.1059	-0.0713
Kurtosis	4.8708	5.2342	3.5729

(b) Panel B.Mutual fund returns

	Mean	Std.dev	Median
μ	0.0022	0.0155	0.0035
σ	0.0564	0.0409	0.0509
Skewness	-0.4621	0.6968	-0.4527
Kurtosis	4.5979	3.8368	4.0637

Table 2.B.2: Unconditional strategies

The table presents the unconditional results of our CEV analysis for a random hedge fund portfolio. We measure the (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the alternative asset strategy (shrunk benchmark with 10% allocated to one of the alternative assets: risk-free rate, real estate, commodity, foreign equity, random fund of funds, or a random portfolio of mutual funds) minus the CEV of the augmented benchmark with 10% allocated to a random hedge fund portfolio. We also report the change in Sharpe ratio (ΔSR) and the change in alpha ($\Delta\alpha$).

Strategy in-sample	avg ΔCEV over a random hedge fund portfolio	p-val	avg ΔSR over a random hedge fund portfolio	avg $\Delta\alpha$ over a random hedge fund portfolio
real estate index	-0.08%	0.000	0.012	0.04%
random fund of funds	-0.18%	0.000	-0.017	-0.18%
commodity index	-0.21%	0.000	-0.016	-0.15%
random portfolio of mutual funds	-0.29%	0.000	-0.026	-0.29%
risk-free rate	-0.42%	0.000	-0.042	-0.44%
foreign equity index	-0.56%	0.000	-0.039	-0.51%
benchmark	-0.61%	0.000	-0.044	-0.57%

Table 2.B.3: Conditional strategies: performance measures

The table presents the out-of-sample results of the CEV analysis for conditional strategies. The conditional sets include portfolios of hedge funds with the highest 5% in-sample performance measures (lowest 5% for kurtosis and standard deviation). We measure the (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the conditional strategy (shrunk benchmark with 10% allocated to a random hedge fund portfolio from the conditional set) minus the CEV of the augmented benchmark with 10% allocated to a random hedge fund portfolio.

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
highest change in mean	0.33%	0.000
highest ΔCEV	0.32%	0.000
highest change in ratio, Sortino et al. [1999]	0.31%	0.000
highest change in ratio, Sharpe [1966]	0.30%	0.000
highest alpha, Fung and Hsieh [2001]	0.28%	0.000
lowest change in kurtosis	0.25%	0.000
highest change in Information Ratio	0.25%	0.000
highest measure, Treynor and Mazuy [1966]	0.19%	0.000
highest change in skewness	0.14%	0.000
highest measure, Henriksson and Merton [1981]	0.13%	0.000
highest change in Value-at-Risk	0.01%	0.043
lowest change in standard deviation	-0.06%	0.000

Table 2.B.4: Conditional strategies: performance measures

The table presents the out-of-sample results of the CEV analysis for conditional strategies. The conditional sets include portfolios of hedge funds with the highest 5% in-sample performance measures (lowest 5% for kurtosis and standard deviation). We measure the (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the proposed strategy (shrunk benchmark with 5-15% in columns (1)-(2) or strictly 10% in columns (3)-(4) allocated to a random hedge fund portfolio from the conditional set) minus the CEV of the augmented benchmark with 5-15% or strictly 10% allocated to a random hedge fund portfolio. In column (1) is the 5-15% portfolio weight allocated to the random hedge fund portfolio determined by the growth rates of the option implied volatility (VIX). In column (2) is the 5-15% portfolio weight determined by the growth rates of the funding liquidity measure (NOISE) from Hu, Pan, and Wang (2013). In column (3) is the random portfolio of hedge funds selected from the 33% highest absolute skewness and the 33% highest kurtosis hedge funds (NONNORMALITY), where both conditions have to be met. In column (4) is the random portfolio of hedge funds selected from the 33% lowest in-sample R-squared hedge funds (LOW R-SQUARED), using the Fung and Hsieh [2001] seven factor model. The rows are sorted by the average rank.

Strategy in-sample	State of the economy		Econometric characteristics	
	VIX	NOISE	NON-NORMAL	LOW R-SQUARED
	(1)	(2)	(3)	(4)
highest ΔCEV	1.06%	0.71%	0.58%	0.40%
highest change in mean	1.05%	0.72%	0.57%	0.34%
highest change in ratio, Sharpe [1966]	0.98%	0.64%	0.49%	0.37%
highest change in ratio, Sortino et al. [1999]	0.96%	0.57%	0.56%	0.31%
highest measure, Treynor and Mazuy [1966]	0.94%	0.61%	0.35%	0.26%
lowest change in kurtosis	0.70%	0.36%	0.47%	0.45%
highest change in Information Ratio	0.90%	0.46%	0.40%	0.29%
highest alpha, Fung and Hsieh [2001]	0.82%	0.49%	0.43%	0.26%
highest measure, Henriksson and Merton [1981]	0.88%	0.53%	0.26%	0.14%
highest change in skewness	0.83%	0.53%	0.17%	0.16%
lowest change in Value-at-Risk	0.62%	0.27%	0.03%	0.19%
lowest change in standard deviation	0.68%	0.33%	-0.04%	0.06%

Table 2.B.5: Conditional strategies: investment styles

The table presents the out-of-sample results of the CEV analysis for conditional strategies. The conditional sets include the 14 largest reported investment styles of hedge funds. We measure the (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the proposed strategy (shrunk benchmark with 10% allocated to a random hedge fund portfolio from the conditional set) minus the CEV of the augmented benchmark with 10% allocated to a random hedge fund portfolio.

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
equity long/short (34%)	0.00%	0.909
multi-strategy (10%)	0.06%	0.000
global macro (8%)	0.08%	0.000
event driven (6%)	-0.03%	0.000
technical trading (6%)	-0.15%	0.000
managed futures (5%)	0.10%	0.000
emerging markets (5%)	-0.06%	0.000
relative value (5%)	-0.21%	0.000
fixed income (4%)	-0.29%	0.000
sector trading (3%)	0.07%	0.000
merger arbitrage (2%)	-0.17%	0.000
distressed (1%)	-0.06%	0.000
convertible arbitrage (1%)	-0.39%	0.000
stock index (1%)	-0.63%	0.000
other (9%)	0.09%	0.000

Table 2.B.6: Persistence analysis (six 48-month periods)

The table presents estimated slope coefficients (b) from two linear regressions: $\Delta CEV_{2i} = a_{CEV} + b_{CEV}\Delta CEV_{1i} + \omega_i$ and $\alpha_{2i} = a_\alpha + b_\alpha \alpha_{1i} + v_i$. It also provides p-values of the t-test on the significance of the b estimate. Both regressions are run for each of six time periods. $\Delta CEV_{1i}(\Delta CEV_{2i})$ are first calculated for each fund during six consecutive periods (ΔCEV_{1i} during the first two years and ΔCEV_{2i} during the second two years): 1994-1998, 1996-2000, 1998-2002, 2000-2004, 2002-2006, 2004-2008 for all hedge funds which survive during the corresponding period. $\Delta CEV_{1i}(\Delta CEV_{2i})$ are then stacked in order to estimate the aggregated slope coefficient. We proceed similarly for the alphas. By stacking the regressions, we assume that the slope coefficients are constant across periods. Standard errors are Newey-West corrected.

Sample	ΔCEV		Alpha	
	b	p-val	*b*	p-val
Feb. 1994 - Jan. 1998	0.267	0.000	0.223	0.000
Feb. 1996 - Jan. 2000	0.000	0.996	0.550	0.000
Feb. 1998 - Jan. 2002	-0.248	0.000	-0.022	0.073
Feb. 2000 - Jan. 2004	0.039	0.009	-0.043	0.009
Feb. 2002 - Jan. 2006	0.385	0.000	0.042	0.105
Feb. 2004 - Jan. 2008	0.255	0.000	0.009	0.465
Feb. 1994 - Jan. 2008	0.081	0.000	0.035	0.119

Table 2.B.7: Robustness check: single hedge fund

The table presents the unconditional results of the CEV analysis for a random hedge fund as opposed to a portfolio of random hedge funds. We measure the (annualized) avg. ΔCEV over a random hedge fund as the CEV of the proposed strategy (shrunk benchmark with 10% allocated to one of the competing assets: risk-free rate, real estate, commodity, foreign equity, fund of funds, or a random portfolio of mutual funds) minus the CEV of the augmented benchmark with 10% allocated to a random hedge fund.

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
real estate index	-0.03%	0.244
random fund of funds	-0.14%	0.000
commodity index	-0.16%	0.000
random portfolio of mutual funds	-0.21%	0.000
risk-free rate	-0.37%	0.000
foreign equity index	-0.52%	0.000
benchmark	-0.57%	0.000

Table 2.B.8: Robustness check: de-smoothed returns

The table presents the unconditional results of the CEV analysis for a random hedge fund portfolio when returns are de-smoothed as in Getmansky et al. [2004]. We measure (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the proposed strategy (shrunk benchmark with 10% allocated to one of the competing assets: risk-free rate, real estate, commodity, foreign equity, fund of funds, or a random portfolio of mutual funds) over the CEV of the augmented benchmark with 10% allocated to a random hedge fund portfolio.

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
real estate index	-0.07%	0.000
random fund of funds	-0.17%	0.000
commodity index	-0.19%	0.000
random portfolio of mutual funds	-0.28%	0.000
risk-free rate	-0.41%	0.000
foreign equity index	-0.55%	0.000
benchmark	-0.60%	0.000

Table 2.B.9: Robustness check: risk-aversion coefficient $\gamma = 1$ or 4

The table presents the unconditional results of the CEV analysis for a random hedge fund portfolio when the CEV calculation uses a power utility function $\gamma = 4$ or 1 (log-utility). We measure (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the proposed strategy (shrunk benchmark with 10% allocated to one of the competing assets: risk-free asset, real estate, commodity, foreign equity, fund of funds, or a random portfolio of mutual funds) over the CEV of the augmented benchmark with 10% allocated to a random hedge fund portfolio.

(a) Panel A.$\gamma = 4$

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
random fund of funds	-0.18%	0.000
commodity index	-0.28%	0.000
random portfolio of mutual funds	-0.32%	0.000
risk-free rate	-0.36%	0.000
real estate index	-0.49%	0.000
foreign equity index	-0.77%	0.000
benchmark	-0.83%	0.000

(b) Panel B.$\gamma = 1$

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
real estate index	0.12%	0.000
random fund of funds	-0.18%	0.000
commodity index	-0.18%	0.000
random portfolio of mutual funds	-0.28%	0.000
risk-free rate	-0.45%	0.000
foreign equity index	-0.47%	0.000
benchmark	-0.51%	0.000

Table 2.B.10: Robustness check: different benchmark

The table presents the unconditional results of the CEV analysis for a random hedge fund portfolio when the benchmark is a portfolio consisting of 35% stocks, 60% bonds, and 5% cash. We measure (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the proposed strategy (shrunk benchmark with 10% allocated to one of the competing assets: risk-free rate, real estate, commodity, foreign equity, fund of funds, or a random portfolio of mutual funds) over the CEV of the augmented benchmark with 10% allocated to a random hedge fund portfolio.

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
real estate index	0.06%	0.000
random fund of funds	-0.18%	0.000
commodity index	-0.20%	0.000
random portfolio of mutual funds	-0.29%	0.000
risk-free rate	-0.44%	0.000
foreign equity index	-0.48%	0.000
benchmark	-0.65%	0.000

Table 2.B.11: Conditional strategies: performance measures, pairwise comparisons

The table presents the pairwise t-tests of the conditional strategies in Table 4, Model (VIX). The lower left shows the average difference in Δ CEV in % where we subtract the model from the leftmost column from the model on the top row). The upper right shows the corresponding p-values.

		1	2	3	4	5	6	7	8	9	10	11	12	
1	highest total benefit	-	0.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
2	highest change in mean	0.01	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
3	highest change in ratio, Sharpe [1966]	0.08	0.07	-	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	p
4	highest change in ratio, Sortino et al. [1999]	0.10	0.09	0.02	-	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-
5	highest measure, Treynor and Mazuy [1966]	0.12	0.11	0.04	0.02	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	v
6	highest change in Information Ratio	0.16	0.16	0.08	0.06	0.04	-	0.01	0.00	0.00	0.00	0.00	0.00	a
7	highest measure, Henriksson and Merton [1981]	0.18	0.17	0.10	0.08	0.06	0.02	-	0.00	0.00	0.00	0.00	0.00	l
8	highest change in skewness	0.23	0.23	0.15	0.13	0.11	0.07	0.05	-	0.73	0.00	0.00	0.00	u
9	highest alpha, Fung and Hsieh [2001]	0.24	0.23	0.16	0.14	0.12	0.07	0.06	0.00	-	0.00	0.00	0.00	e
10	lowest change in kurtosis	0.36	0.35	0.28	0.26	0.24	0.20	0.18	0.13	0.12	-	0.02	0.00	s
11	lowest change in standard deviation	0.38	0.38	0.30	0.28	0.26	0.22	0.20	0.15	0.14	0.02	-	0.00	
12	lowest change in Value-at-Risk	0.44	0.43	0.36	0.34	0.32	0.28	0.26	0.20	0.20	0.08	0.06	-	
							avg ΔCEV(%)							

Table 2.B.12: Robustness check: in-sample size 36

The table presents the unconditional results of the CEV analysis for a random hedge fund portfolio with the in-sample size of 36 months. We measure the (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the proposed strategy (shrunk benchmark with 10% allocated to one of the competing assets: risk-free asset, real estate, commodity, foreign equity, fund of funds, or a random portfolio of mutual funds) minus the CEV of the augmented benchmark with 10% allocated to a random hedge fund portfolio.

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
random fund of funds	-0.17%	0.000
real estate index	-0.25%	0.000
random portfolio of mutual funds	-0.28%	0.000
commodity index	-0.37%	0.000
risk-free rate	-0.38%	0.000
foreign equity index	-0.60%	0.000
benchmark	-0.64%	0.000

Table 2.B.13: Robustness check: 5% or 25% weight in hedge fund portfolio

The table presents the unconditional results of the CEV analysis for a random hedge fund portfolio added with 5% weight (Panel A) or 25% weight (Panel B) to the benchmark. We measure the (annualized) avg. ΔCEV over a random hedge fund portfolio as the CEV of the proposed strategy (shrunk benchmark with 5% (25%) allocated to one of the competing assets: risk-free asset, real estate, commodity, foreign equity, fund of funds, or a random portfolio of mutual funds) minus the CEV of the augmented benchmark with 5% (25%) allocated to a random hedge fund portfolio.

(a) Panel A.5%

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
real estate index	-0.04%	0.000
commodity index	-0.09%	0.000
random fund of funds	-0.09%	0.000
random portfolio of mutual funds	-0.15%	0.000
risk-free rate	-0.21%	0.000
foreign equity index	-0.28%	0.000
benchmark	-0.31%	0.000

(b) Panel A.25%

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
real estate index	-0.27%	0.000
random fund of funds	-0.48%	0.000
commodity index	-0.74%	0.000
random portfolio of mutual funds	-0.76%	0.000
risk-free rate	-1.05%	0.000
foreign equity index	-1.46%	0.000
benchmark	-1.49%	0.000

Table 2.B.14: Robustness check: only open to new investment funds

The table presents the unconditional results of the CEV analysis for a random hedge fund which is open to new investment. We measure the (annualized) avg. ΔCEV over a random hedge fund as the CEV of the proposed strategy (shrunk benchmark with 10% allocated to one of the competing assets: real estate, commodity, foreign equity, fund of funds, or a random portfolio of mutual funds) minus the CEV of the augmented benchmark with 10% allocated to a random hedge fund.

Strategy in-sample	avg. ΔCEV over a random hedge fund portfolio	p-val
real estate index	-0.16%	0.0000
random fund of funds	-0.26%	0.0000
commodity index	-0.29%	0.0000
random portfolio of mutual funds	-0.38%	0.0000
risk-free rate	-0.51%	0.0000
foreign equity index	-0.65%	0.0000
benchmark	-0.70%	0.0000

Relative Alpha

3.1 Introduction

Investment analysis makes much use of alpha, the intercept in a factor model of returns. First, alpha should measure the current outperformance of a fund over its risk-adjusted return. Here, it is important to account for all relevant factors in order to avoid missing factor bias. Second, alpha should predict future fund performance and thus funds should flow to high alpha funds. Third, alpha should be persistent so that high performance funds today remain high performance funds in the future. While it is not clear that one and the same alpha serves all three goals, in the analysis of hedge funds, the seven factor model of Fung and Hsieh [2001] has become the de facto standard and is used for all three purposes. Our goal is to show that our new measure, relative alpha, performs better than the traditional, absolute alpha measures (e.g. Fung and Hsieh [2001] and others) along all three dimensions.

To argue our case, we first define absolute and relative alpha. Hedge funds are often evaluated in terms of alpha within a factor model:

$$r_t - r_{f,t} = \alpha + \sum_{l=1}^{L} \beta_l X_{l,t} + \varepsilon_t, \tag{3.1.1}$$

where $(r_t - r_{f,t})$ is the excess hedge fund return at time t, $X_{l,t}$ is the factor l at time t, β_l is the risk exposure to factor l, and ε_t is a mean zero error term. Given the relatively short life span of the average hedge fund (typically only 36 to 48 monthly observations), popular sets of factors number just seven or eight (e.g. Fung and Hsieh [2001] or Agarwal and Naik [2004]) so that they can be estimated with some precision.

As an alternative, we propose a new measure (”relative alpha”). Relative alpha links each hedge fund to a group of its peers and averages over the expected differences between returns of the hedge fund and returns of each of its peers. Peers are characterized by the low variance of these differences in returns and receive the more weight in our average the lower their variance of differences in returns. Thus, relative alpha measures the outperformance of a hedge fund over its closest peers without resorting to a particular factor model. The kernel-based estimation technique is simple to implement and uses solely hedge fund returns.

We will now argue that relative alpha performs better than absolute alpha along all three dimensions of interest: measuring current outperformance over risk-adjusted returns; identifying funds which generate high future performance and thus determining fund flow; and finding persistence in alpha.

First, in terms of measuring current outperformance, absolute alpha has the intuitive interpretation as the excess skill added to the hedge fund return by the manager beyond the risky investments into the factors. However, hedge funds are sophisticated investment vehicles and the universe of their strategies might lie beyond the seven or eight factors typically employed in factor models. If there are other relevant, omitted factors, then absolute alpha will erroneously include the returns associated with the omitted factors. The Fung and Hsieh [2001] model only explains with its seven factors 41% of the return variation (adjusted R-squared) in our sample, leaving us worried that some of the missing explanatory power is due to omitted factors. Interpreting the seven closest peers as factors, we achieve an adjusted R-squared of 81%. Also, the average standard error around the alpha estimates is only 0.57 for relative alpha vs. 0.86 for Fung and Hsieh [2001].

Second, we find that relative alpha is better than absolute alpha in identifying funds which generate high future performance and thus determining fund flow. In particular, we use rolling windows of 36 months to estimate both relative and absolute alphas. We then invest in equally weighted portfolios based on hedge funds in the highest deciles, respectively, and measure the out-of-sample monthly returns. On a monthly basis, the relative alpha strategy has higher mean return (1.12%) at lower volatility (2.11%), yielding a Sharpe ratio of 0.49, compared with the absolute alpha strategy with a lower mean return (1.02%) at higher volatility (3.25%), yielding a Sharpe ratio of 0.27 which is about half the Sharpe ratio of the relative alpha strategy. The monthly Sharpe ratio of a top-minus-bottom decile strategy is a surprising 11 times higher for relative over absolute alpha. We wondered if any risk averse investor would prefer the sorts based on relative alpha over the ones based on absolute alpha. Thus, we employ a test if investing into the top relative alpha portfolio stochastically non-dominates (in the second order) investing into the top absolute alpha portfolio. The test of Davidson and Duclos [2013] rejects the null at the 1% level. We conclude that any risk averse investor would rather use a sort based on relative alpha. We obtain the same result when we investigate top-minus-bottom deciles. Relative alpha also dominates the use of the appraisal ratio of Treynor and Black [1973], the manipulation-proof performance measure (MPPM) of Goetzmann et al. [2007], and the strategy distinctiveness index (SDI) of Sun et al. [2012]. Thus, portfolios based on sorting by relative alpha perform significantly better than sorts based on absolute alpha in statistical and economic terms.

Moreover, we find a strong positive relation between fund flows and past relative alpha, i.e. the larger the past relative alpha, the larger the flows into assets under management. We do not find the same pattern for the fund flows and past absolute alpha relation.

Third, alpha is often interpreted as the skill a hedge fund manager possesses. If the managers skill is persistent, then alpha should also be persistent[1]. Tests

[1] For mutual funds, Berk and Green [2004] argue that competition should eliminate such

of alpha persistence show that relative alpha exhibits significant positive coefficients while absolute alpha exhibits insignificant coefficients or even negative coefficients. Our findings for absolute alpha are in line with Capocci and Hübner [2004] who also find very little or no persistence in absolute hedge fund alpha.

Furthermore, we design a simulation study to analyze under which circumstances relative alpha works best. Relative alpha works the better the more omitted variable bias exists. Also, a larger cross-section of hedge funds contributes to the superior performance of relative alpha. Such large cross-section increases the probability for each hedge fund to find a relevant group of peers which spans the investment opportunity set of the hedge fund.

We also show that our results are not qualitatively affected by minor changes in methodology and sample.

Based on a literature review in Section 3.2, the paper develops the hypotheses and introduces the econometric methodology for testing our performance measure in Section 3.3. All data are presented in Section 3.4. Results follow in Section 3.5 while robustness checks are presented in Section 3.6. Simulation study is in Section 3.7. Section 3.8 concludes.

3.2 Literature

There are several related branches of the literature. First, a number of papers directly try to improve the omitted factor bias by adding additional factors: the factor models of Agarwal and Naik [2004], Fung and Hsieh [2001], or the hedge fund index model of Jagannathan et al. [2010]. These models have around 7-

skill related alpha. In Berk and van Binsbergen [2013] they thus argue that therefore mutual funds should not exhibit persistence in alpha. However, for hedge funds Glode and Green [2011] show theoretically that potential information spillovers (associated with innovative trading strategy or emerging sector) could lead to persistent alpha after all.

10 factors and are thus much richer than a single market factor model or the Fama-French three factor model. In particular, Fung and Hsieh [2001] is nowadays the de facto standard factor model for hedge fund research, but even Fung and Hsieh [2001] only accounts in our sample for 41% of the return variation of hedge funds with their seven factors. Thus, there is concern that the remaining 59% of return variation might contain omitted factors.

Hunter et al. [2014] augment a conventional factor model with an active peer benchmark. They determine peer-groups of mutual funds based on their investment objectives. While this is a simple approach that significantly improves the selection of successful funds, there are several issues that hinder us from applying this methodology directly to hedge funds. First, hedge fund strategies are not as well specified as the strategies of mutual funds. Not merely is the description of hedge fund strategies vague, but hedge funds may change their strategies according to market conditions, available financial resources, and current management objectives. Second, a four-factor model of Carhart [1997] used for mutual funds is able to explain 60-70% of the variation in returns, which is almost twice as much as the amount of explained variation by typical hedge fund linear factor models.

A paper by Wilkens et al. [2013] adopts the endogenous benchmark approach of Hunter et al. [2014] to hedge funds. Wilkens et al. [2013] construct weighted endogenous benchmarks to relate individual funds to their style peers in order to reduce the problem of missing variables. Still, Wilkens et al. [2013] require a linear factor model with the associated problem that about half of the hedge fund return variation is left unexplained. We add to this literature by defining the peer group of a hedge fund to be only hedge funds which are very similar in terms of the variance of the return differences. That allows us to reduce the omitted variable bias greatly and to implicitly offset much of the unexplained variation in hedge fund returns. As a result, we can estimate relative alphas of one hedge fund with respect to its peer group more precisely.

Hoberg et al. [2014] use competition between mutual funds to develop a new measure of skill, customized peer alpha (CPA). They show that this new measure predicts alpha for at least four quarters. Competitors are defined based on the data on mutual fund holdings. Funds are placed into a 3-dimensional (also 2- or 4-) space of characteristics (including size, value-growth orientation, momentum, dividend yield) based on the dollar-weighted average of their stock holdings. Two funds are considered competitors if the spatial distance between them is smaller than a given value. CPA is then measured as a fund's outperformance over its spatial peers. A fund is considered skillful if it is able to beat funds with similar strategies. In our relative alpha approach, we also compare funds to their peers. However, we derive relative alpha directly as a way to reduce the omitted variable bias and do not use the competition concept. The approach of Hoberg et al. [2014] does not work for hedge funds as hedge fund holdings are unknown but for the 13F reports which only apply to US equity holding of large funds ($>$ \$100 million).

We also connect to a second literature on fund flow and fund performance (see e.g. Fung et al. [2008] and Ding et al. [2009]. The results are somewhat contradictive which is due to the different samples and different methodologies. Sirri and Tufano [1998] find that mutual fund flows chase good performance. As a performance measure they use fractional rank quantiles which represent fund performance relative to other funds in the same period. Getmansky [2012] adopts similar methodology and records a positive relation between flows and past performance for the middle and bottom terciles of hedge funds. At the same time, top performing funds do not grow proportionally as much as the average fund in the industry. Goetzmann et al. [2003] on the contrary reveal that hedge funds demonstrate a decrease in investments, conditional on the past returns.

A third literature concerns the predictability of alpha. The existing literature (Ammann et al. [2010], Capocci and Hübner [2004]) shows mixed evidence concerning the predictability of absolute alpha. That is, historically measured alpha has little predictive value for future alpha. But that means that allocating investments to past high alpha funds will not lead to high alpha in the future. In contrast, we document strong persistence of relative alpha during our sample.

3.3 Hypotheses and methodology

We assume that there exists a complete factor model which explains hedge funds without omitted factors and with uncorrelated error terms. Without loss of generality, we assume these factors to have been orthogonalized. In particular, we do not limit ourselves to the seven or eight factors of Fung and Hsieh [2001] or Agarwal and Naik [2004]. Thus, our assumed factor model would have perfect explanatory power but for the error term, i.e. an R-squared of close to 1. Obviously, we might not be able to enumerate all these factors, but we do not need to; we argue below that we can still assess our performance measure relative alpha without the explicit knowledge of the full factor model by essentially netting out much of the unknown factor structure. The complete factor models for hedge funds i and j are then as follows:

$$r_{it} - r_{f,t} = \alpha_i + \sum_{k=1}^{K} \beta_{ik} X_{k,t} + \varepsilon_{it}, \tag{3.3.1}$$

$$r_{jt} - r_{f,t} = \alpha_j + \sum_{k=1}^{K} \beta_{jk} X_{k,t} + \varepsilon_{jt}, \tag{3.3.2}$$

where $(r_{i,t} - r_{f,t})$ is the excess hedge fund return of hedge fund i at time t, $X_{k,t}$ is the factor k at time t, β_{ik} is the risk exposure of hedge fund i to factor k, and ε_{it} is a mean zero error term for hedge fund i. Definitions for hedge fund j are similar.

We next take expectations of the differences in returns. If hedge funds i and j implement identical strategies (i.e. they load on the same risk factors $X_{k,t}$ and have $\beta_{ik} = \beta_{jk}$), then their factor loadings cancel, leaving only differences in alphas:

$$E[r_{it} - r_{jt}] = \alpha_i - \alpha_j + \sum_{k=1}^{K} (\beta_{ik} - \beta_{jk}) E[X_{k,t}] + E[\varepsilon_{it} - \varepsilon_{jt}] = \alpha_i - \alpha_j \qquad (3.3.3)$$

We can thus obtain relative alpha, the difference in hedge fund i's alpha from hedge fund j's alpha. Now clearly, hedge funds typically do not have perfectly identical betas. Instead, we allow for small discrepancies in betas. Such discrepancies in beta would not even affect the expectation in Equation (3.3.3) as long as beta differences are random and uncorrelated from one hedge fund to the next. Also, we allow for more than one peer hedge fund with respect to which fund i's relative alpha is being calculated and suggest taking weighted averages.

For these two additional steps (discrepancies in betas and a larger peer group), we first need a distance measure in order to establish the size of beta discrepancies and second an averaging technique in order to work out the relative alpha with respect to a group of peers.

As a distance measure, we calculate the variance of the return differences:

$$Var[r_{it} - r_{jt}] = (\beta_i - \beta_j)' Cov(X_{k,t})(\beta_i - \beta_j) + \sigma_i^2 + \sigma_j^2, \qquad (3.3.4)$$

where β_i (β_j) is the $(K \times 1)$ vector of risk exposures of hedge fund i (j), $Cov(X_{k,t})$ is the $(K \times K)$ variance-covariance matrix of the factors (assumed to be bounded from above), and $\sigma_i^2(\sigma_j^2)$ are the variances of the error terms ε_{jt} and ε_{jt}. We assume that the variances of the error terms (σ^2) are similar in magnitude for all hedge funds. Note that there are no covariances for the error terms as we assumed that all common components are reflected in the factor structure. Using our assumption that the factors are orthogonal to each other, the expression for the variance of return differences becomes:

$$Var[r_{it} - r_{jt}] = \sum_k \left(\beta_{ik} - \beta_{jk}\right)^2 Var\left(X_{k,t}\right) + \sigma_i^2 + \sigma_j^2. \tag{3.3.5}$$

It follows that the variance of return differences is smaller if the funds' betas are closer to each other (i.e. the differences in betas decrease for at least one factor and do not increase for the other factors). We provide more details on this mechanism in our simulation study in Section 3.7. We will thus use the variance of return differences as a measure of how similar two hedge funds are to each other.

As an averaging technique, we compute relative alpha as the sum of kernel weighted expected return differences. The kernel gives more importance to peer funds which are closer to the hedge fund in terms of the variance of return differences as in Equation (3.3.4). Therefore, the relative alpha of hedge fund i is being determined with respect to all other hedge funds by first calculating, one at a time, the expected differences in returns according to Equation (3.3.3). Those pairwise expected differences in returns are then aggregated into one relative alpha via kernel weights which add up to one. Relative alpha is calculated as:

$$\Delta_{i,T} = \frac{\sum_{i \neq j} K\left(Var[r_{it} - r_{jt}]/h\right) E[r_{it} - r_{jt}]}{\sum_{i \neq j} K\left(Var[r_{it} - r_{jt}]/h\right)}, \tag{3.3.6}$$

where $K(\cdot)$ is a Gaussian Kernel, h is the bandwidth according to Silverman [1986] rule of thumb, $Var[\cdot]$ and $E[\cdot]$ are the variance and expectation of the return differences between hedge funds i and j for t=T-35,...,T.

Our results do not change much if we use bandwidths from a fifth to five-times of the value of Silverman [1986] rule of thumb. Kernel estimates are biased on the boundary of the data and we suffer from this problem as we evaluate the kernel estimate at a point where the variance is zero based on variances which are all positive. Thus, we show in our robustness section that results do not change when we use the local regression technique proposed by Hastie and Loader [1993] which better accommodates the boundary bias. Relative alpha is distinct from absolute alpha, but the two alpha measures are related to each other. The correlation between relative alpha and absolute Fung and Hsieh [2001] alpha is on average 0.62 ranging between 0.37 and 0.86, depending on the sample.

A number of advantages emerge which argue in favor of using relative alpha over absolute alpha. First, we do not require knowledge of the true factor model. Thus, we are much less prone to omitted variable bias: if two hedge funds are similar, then they presumably have a rather similar (but possibly partially unknown) factor structure which will cancel out in the relative alpha calculation. Second, implementation is straightforward. Third, relative alpha performs considerably better than absolute alpha along our three dimensions of interest: explaining high current performance, predicting high future performance, and predicting high future alpha. We now detail the methodology for establishing the superior performance of relative alpha in each dimension in turn.

To assess how well alpha measures the current outperformance over risk-adjusted returns, we investigate the adjusted R-squared of the Fung and Hsieh [2001] model for absolute alpha in terms of average and standard deviation. For relative alpha, we use the returns of the seven closest peer funds in lieu of the seven Fung and Hsieh [2001] factors. We can then regress hedge fund returns on those peer fund returns and again obtain values of adjusted R-squared.

To show that relative alpha predicts future high performance, we compare the out-of-sample performance of portfolios based on sorts on relative and absolute alphas. We use 36-month rolling windows to estimate relative and absolute alpha, sort hedge funds into top and bottom deciles, and form equally weighted portfolios. We record returns of top, bottom, and top-minus-bottom portfolios in the 37^{th} month and repeat the procedure by moving one month out. We take care of look-ahead bias by recording a zero return instead of a missing return in case a hedge fund is delisted. Since hedge funds may have a lock-up period, as a robustness check we increase portfolio holding period to 12 months. We report means, standard deviations, and Sharpe ratios for these strategies. Also, we use the test of Davidson and Duclos [2013] for second order stochastic non-dominance. A rejection of this test has the powerful implication that any risk-averse investor would prefer investing into hedge funds sorted by relative alpha as opposed to sorts based on absolute alpha. A related point is that high alpha should then generate high fund flow as investors allocate investments according to past alpha. We simply regress fund flows for fund *i* (measured from time *t-1* to *t*) on its relative (absolute) alpha (measured from time *t-1*,...,*t-36*).

To analyze persistence of alpha, we adopt a methodology commonly used in the hedge fund literature. We consider consecutive 72 (48, 24)-month periods, starting with the 1^{st}, 37^{th}(25^{th}, 13^{th}), 73^{rd}(49^{th}, 25^{th}), ...observations. Each of these periods is divided into two 36 (24, 12)-month sub-periods: a formation period (1-36^{th} (1-24^{th}, 1-12^{th}) months) and an evaluation period (37-72^{nd} (25^{th}-48^{th}, 13^{th}-24^{th}) months). For each hedge fund which survives the whole 72 (48, 24)-month period, we compute relative alpha in the formation ($\triangle_{1i}$) and the evaluation ($\triangle_{2i}$) periods and estimate the following regression:

$$\triangle_{2i} = a_{\Delta} + b_{\Delta}\triangle_{1i} + \omega_i, \tag{3.3.7}$$

where a_{Δ}, b_{Δ} are the parameters to be estimated, and ω_i is an error term. We stack all observations for the different non-overlapping periods and then run the joint regression. Persistence in relative alpha is determined by a significantly positive coefficient b_{Δ}.

The persistence study is repeated for absolute alphas from the seven factor Fung and Hsieh [2001] model:

$$\alpha_{2i} = a_\alpha + b_\alpha \alpha_{1i} + v_i, \tag{3.3.8}$$

where $\alpha_{1i}(\alpha_{2i})$ are the alpha estimates in the formation (evaluation) period for fund i; a_α, b_α are the parameters to be estimated, and v_i is an error term.

3.4 Data

For hedge fund information we use the MOAD database described in Hodder et al. [2013]. MOAD is a merged database of six commercially available databases (CISDM, Barclays, TASS, HFR, Altvest, and Eurekahedge). We use only USD-denominated, net-of-fees returns with at least 36 month historical returns which leaves us with 11,597 hedge funds. Our sample runs from February 1994 until June 2011. The descriptive statistics of our sample are presented in Table 3.C.1. We document excess kurtosis and left-skewness in hedge fund returns, suggesting that returns are often not normally distributed.

[Table 3.C.1 about here]

Hedge funds differ from other asset classes in many respects. One of them is the absence of strict regulation. This leads to database biases as reporting is voluntary. We address those biases as follows. First, our joint database is free of survivorship bias because it contains both live and dead funds. Second, to control for the instant history bias, we delete the first 12 months of each hedge fund's returns. We compute our main results based on the reported returns as we find them in the database.

For funds of funds, we extract the USD-denominated, net-of-fees returns in a similar fashion to the hedge fund returns from our database and are left with 9,314 funds of funds. The descriptive statistics of our sample are presented in

Table 3.C.2. Compared to single hedge funds, fund of funds have smaller net-of-fee average return which is explained by the double layer of fees. As a portfolio of funds, they also demonstrate smaller variance.

[Table 3.C.2 about here]

We further use the seven factors of the Fung and Hsieh [2001] model which are available at David A.Hsieh's Hedge Fund Data Library[1].

For the alternative factor models of Capocci and Hübner [2004], Edelman et al. [2012], Agarwal and Naik [2004] we downloaded the risk factors from Thomson Reuters Datastream. Option-based factors from Agarwal and Naik [2004] were graciously provided by the authors.

3.5 Results

We now present our results where we first investigate the ability of relative and absolute alpha to assess current hedge fund performance. Next, we analyze the ability to predict future performance and the associated fund flow which should result from such ability. We continue with an investigation of the persistence of relative alpha versus absolute alpha. Last, we document drivers of relative alpha.

3.5.1 Current hedge fund performance

When focusing on alpha as a measure of current performance, we are much concerned that the returns associated with omitted factors might show up in the alpha of a misspecified factor model. For absolute alpha, we use the adjusted R-squared of the Fung and Hsieh [2001] model in order to assess the

[1] https://faculty.fuqua.duke.edu/ dah7/HFData.htm

explanatory power of the factor model. With an average adjusted R-squared of 41%, the Fung and Hsieh [2001] model leaves quite some variation in returns unexplained. Also, the average standard error around the alpha estimates is fairly large at 0.86.

For our relative alpha measure, we cannot directly obtain comparable quantities. We thus use the returns of the seven closest peer funds as pseudo factors. Then, we are able to repeat the above exercise of computing average adjusted R-squared which at 81% explains significantly more return variation than the Fung and Hsieh [2001] model at all significance levels. Also, the average standard error around the alpha estimates is much tighter at 0.57 which is significantly different from 0.86 at the 10% level.

3.5.2 Predicting future performance

In assessing the predictive ability of alpha, we turn to out-of-sample returns on decile portfolios of hedge funds, sorted by 36-month in-sample alpha. We roll the sample one month forward and repeat the exercise.

Table 3.C.3 provides monthly results on top decile, bottom decile, and top-minus-bottom decile portfolios. The top relative alpha portfolio delivers a slightly higher mean return (1.12%) at substantially lower standard deviation (2.11%) in comparison to the top absolute alpha portfolio (mean of 1.02% and standard deviation of 3.25% based on Fung and Hsieh [2001]), which is reflected in an almost doubled Sharpe ratio (0.49 for the top relative alpha portfolio vs. 0.27 for the top absolute alpha portfolio). We formally test if investing into the top relative alpha portfolio stochastically non-dominates (in the second order) investing into the top absolute alpha portfolio. The test of Davidson and Duclos [2013] rejects the null at the 1% level.[1] That means that any risk averse investor prefers the top relative alpha portfolio to the top absolute alpha portfolio. The

[1]We apply the Davidson and Duclos [2013] test based on the t statistic version. We simulate the distribution of the test statistics in order to account for small samples.

same result obtains when we turn to the top-minus-bottom portfolios.

[Table 3.C.3 about here]

Another important result is that the top-minus-bottom portfolio constructed by sorting on relative alpha has a significantly positive mean return, while the mean return of the top-minus-bottom portfolio constructed by sorting on absolute alpha is insignificant. The Sharpe ratio of top-minus-bottom portfolio is 11 times higher for relative alpha when compared to absolute alpha. We conclude that relative alpha works better in distinguishing between future winners and losers than absolute alpha.

To strengthen our argument, we repeat the analysis by sorting hedge funds into decile portfolios based on alternative performance measures. Again, all results are based on 36-month rolling windows and on recording the subsequent out-of-sample returns. We use the appraisal ratio of Treynor and Black [1973], the manipulation-proof performance measure (MPPM) of Goetzmann et al. [2007], and the strategy distinctiveness Index (SDI) of Sun et al. [2012]. The relevant formulas are collected in Appendix 3.A. All three performance measures are second-order stochastically dominated by relative alpha, although the manipulation-proof performance measure only with a p-value of 0.07, see Table 3.C.4.

[Table 3.C.4 about here]

There are several alternatives to the Fung and Hsieh [2001] factor model. As alternatives, we evaluate absolute alpha based on the models proposed in Agarwal and Naik [2004][1], Edelman et al. [2012], and Capocci and Hübner [2004]. Please consult Appendix 3.B for details on all the factors used. The results are

[1]Our sample of option-based factors from Agarwal and Naik [2004] runs until August 2009. For the SSD-tests in Table 3.C.5 we use the corresponding shorter sample of relative alpha.

presented in Table 3.C.5.

[Table 3.C.5 about here]

Our previous findings are still valid for the factor models of Edelman et al. [2012], and Capocci and Hübner [2004]. In Agarwal and Naik [2004], absolute alpha works well in picking the top performers. However, it fails to pick bottom performers properly, i.e. the difference between top and bottom portfolios is statistically insignificant (as opposed to relative alpha). All Davidson and Duclos [2013] tests are significant at the 10% level except for the top portfolios of Agarwal and Naik [2004]. The ratios of monthly Sharpe ratios reflect these findings and range from 1.4 to 7.2, reflecting the superiority of relative alpha but for the top portfolios based on Agarwal and Naik [2004] which have the same Sharpe ratio as achieved under relative alpha.

We try to explain the qualitative differences between top and bottom portfolios. In order to do so, we first regress portfolios based on relative alpha on the seven factors of Fung and Hsieh [2001]. Table 3.C.6 summarizes average risk exposures across funds and time. Generally, the estimates from Fung and Hsieh [2001] are quite similar for both portfolios, except for three factors: the size spread factor, the credit spread factor, and the currency trend-following factor. The top portfolio loads positively on the size spread factor (estimated average risk exposure is 0.03), while the bottom portfolio has a negative exposure to this factor (estimated average risk exposure is -0.02). Exposures to the credit spread factor as well as the currency trend-following factor are also significantly higher for the top portfolio (-0.45 and 0.02 respectively) than for bottom portfolio (-0.53 and 0.01). However, there are several issues one should be aware of: a) the exposures are aggregated across two dimensions; b) the estimates are still exposed to the omitted variable bias.

[Table 3.C.6 about here]

A second way to explain the portfolio construction is to repeat the study within investment styles. Table 3.C.7 shows the characteristics of top, bottom, and top-minus-bottom portfolios for the largest 6 investment strategies: equity long/short, fixed income, global macro, CTA/managed futures, event driven, and relative value. Naturally, the styles differ in their characteristics: some styles are considered more risky and yield higher returns (CTA/managed futures), while others are safer but less profitable (Equity long/short). Each portfolio based on relative alpha yields higher mean return and equal or lower standard deviation in comparison to absolute alpha. The only exception is CTA/managed futures where the top portfolio based on relative alpha has a higher standard deviation (8.12 for relative alpha vs. 7.89 for absolute alpha top portfolio). On the contrary, the bottom portfolio based on relative alpha has generally a lower return.

[Table 3.C.7 about here]

Our main results on single hedge funds are also confirmed for funds of funds, see Table 3.C.8: relative alpha works better in differentiating between future winners and losers than absolute alpha.

[Table 3.C.8 about here]

3.5.3 Fund flow and alpha

To analyze the relation between fund flow and alpha, we simply regress fund flows (from t-1 to *t*) and past relative (absolute) alpha (measured from *t-36* to *t-1*). The results of these regressions are presented in Table 3.C.9.

[Table 3.C.9 about here]

Table 9 documents a significantly positive relation between past relative alpha and current fund flows. Even though our relative alpha approach is novel and has not been applied in the literature before, the underlying idea of comparing hedge funds to their peers seems well-understood according to the data: funds with higher past relative alpha attract more investments. We do not find the same pattern for the absolute alpha: the slope coefficient in the regression is insignificantly different from zero.

3.5.4 Persistence of alpha

We know from the hedge fund literature that there is rather mixed evidence on hedge fund absolute alpha persistence. We are therefore interested if relative alpha is more persistent than absolute alpha and present results in Table 3.C.10. Relative alpha is more persistent than absolute alpha and it is robust to different sample sizes: estimates of the b-coefficients for relative alpha (Equation (3.3.7)) are consistently positive and significant for different sample sizes (12, 24, and 36 months) as opposed to the b-coefficients for absolute alpha (Equation (3.3.8)). It is worth noting that the b-coefficients for relative alpha are decreasing as we increase the sample size which is consistent with time-decay in persistence.

[Table 3.C.10 about here]

We further analyze the sensitivity of top portfolios to the business cycle by regressing top portfolio returns on a crisis-dummy during the recent financial crisis. We find that both top portfolios (constructed by relative and absolute alphas) are losing, but the magnitude of the crisis-dummy for the relative alpha portfolio is smaller than for the absolute alpha portfolio.

Table 11 demonstrates persistence results for funds of funds. Slope coefficients for relative alpha stay positive for all sample sizes and are significant for the 24 and 12-month samples. Moreover, the significance of the slope coefficients is decreasing in the out-of-sample size. The sign of slope coefficient for absolute alpha is switching for different sample sizes.

[Table 3.C.11 about here]

3.5.5 Results on drivers of relative alpha

In our attempt to see what drives relative alpha, we run panel data regressions. On the right hand side of our regressions we include the seven factors from Fung and Hsieh [2001], a dummy variable indicating whether the fund is open to new investments, the logarithm of assets under management, a dummy variable indicating a high water mark, a dummy variable indicating whether the fund is using leverage, the management fee, and the performance fee. We divide the sample into non-overlapping 12-month periods and compute relative alpha for each fund within these periods. Time-varying explanatory variables (the seven factors of Fung and Hsieh [2001] and assets under managements) are averaged over the same periods. Table 3.C.12 summarizes several panel data regressions: robust OLS, fixed effect, and random effect. The significance of the coefficients is robust to different methodologies. Apart from some Fung and Hsieh [2001] factors, the following variables have an effect on relative alpha: openness to new investment and performance fee. If a fund is opened to new investment, it has on average lower relative alpha, consistent with the intuition that superior hedge funds (with high relative alpha) are more likely to be closed to new investment. Also, the higher the performance fee, the higher the relative alpha of a fund which is consistent with the role of fees in signaling better performance (see Habib and Johnsen [2012] on mutual funds).

[Table 3.C.12 about here]

3.6 Robustness

As a part of our robustness checks, we split the sample into several subsamples: February 1994 - February 2000 (dotcom bubble), March 2000 - July 2007 (intermediate period), and August 2007 - June 2001 (financial crisis). Table 3.C.13 summarizes performance of top portfolios based on relative alpha and absolute alpha during the subsamples. During the first subsample (the dotcom bubble from February 1994 - February 2000) both measures are performing similarly and yield a Sharpe ratio of 0.61-0.63. Our main results stay largely unaffected during the last two subsamples, the intermediate period and the financial crisis, i.e. relative alpha outperforms absolute alpha in terms of Sharpe ratios and the Davidson and Duclos [2013] test confirms that any risk averse investor would prefer top relative alpha portfolios. During the recent financial crisis the returns overall decreased dramatically which is also reflected in the portfolios (mean return is 0.71% instead of 1.12% during the whole period for relative alpha and 0.53% instead of 1.02% for absolute alpha).

[Table 3.C.13 about here]

- Furthermore, our results are robust to a range of minor methodological and sample changes:
- eliminating the hedge funds which are closed to new investment (Table 3.C.14)
- eliminating small funds with assets under management under $20 million as proposed in Kosowski et al. [2007] (Table 3.C.15)
- changing the rolling window size to 24 months instead of 36 months (Table 3.C.16)
- using portfolios of only 20 hedge funds with the largest alphas instead of larger deciles (Table 3.C.17)

- comparing relative alpha to a random hedge fund portfolio (Table 3.C.18)
- holding top and bottom portfolios for 12 months instead of 1 month (Table 3.C.19)
- correcting for boundary bias of the kernel estimates by using locally weighted regressions as proposed in Hastie and Loader (1993) (Table 3.C.20)

3.7 Simulation

We use a simulation study to investigate why relative alpha outperforms absolute alpha. We are concerned about three aspects, namely the poor estimation of absolute alpha due to short time-series, multi-collinear factors, and omitted factors.

We first simulate hedge fund returns by assuming that the true model is the seven factor model of Fung and Hsieh [2001] with independently normally distributed residuals. In order to preserve the empirical characteristics of our hedge fund returns, we set the assumed true parameters equal to the estimated parameters as observed in the data. That gives us empirical residuals from which we compute the variances of the residuals. We then assume that the true residuals are independently normally distributed with variances equal to the empirical variances and zero mean. Given this structure, we can simulate hedge fund returns where the initial cross-sectional dimension is 2000, the time series length is 36 months, and the number of simulation runs is 100.

We next estimate relative alpha based on the simulated returns using our usual methodology as in Equation (3.3.6). We report the average value of 0.15 in Table 3.C.21 for the true model in the first row (for estimation method Relative alpha) and column two (labeled relative alpha) as we use Equation (3.3.6). An absolute alpha value is not available in this case (n/a in column three, labeled absolute alpha). Next we report the true average alpha of 1.20 in the column absolute alpha. Also, we would like to report a relative alpha version which we base on Equation (3.3.3). Namely, if two hedge funds are identical in terms of their beta exposures, then

$$E[r_{it} - r_{jt}] = \alpha_i - \alpha_j + \sum_{k=1}^{K} (\beta_{ik} - \beta_{jk}) E[X_{k,t}] + E[\varepsilon_{it} - \varepsilon_{jt}] = \alpha_i - \alpha_j. \tag{3.7.1}$$

Thus, we replace in the relative alpha formula of Equation (3.3.6) the expectation $E[r_{it} - r_{jt}]$ with the difference in true alpha:

$$\triangle^{\alpha}_{(i,T)} = \frac{\sum_{i \neq j} K\left(Var[r_{it} - r_{jt}]/h\right)(\alpha_i - \alpha_j)}{\sum_{i \neq j} K\left(Var[r_{it} - r_{jt}]/h\right)}. \tag{3.7.2}$$

We compute the modified relative alpha $\triangle^{\alpha}_{(i,T)}$ from Equation (3.7.2) and obtain an average value of 0.17, insignificantly different from 0.15 for average relative alpha itself. Next, we estimate the Fung and Hsieh [2001] model for each hedge fund and obtain for the true model exactly the same absolute alpha (estimation method FH7) - the variation introduced by simulating the residuals does not affect the average absolute alpha value. Using the differences of those estimated alphas instead of the expectation $E[r_{it} - r_{jt}]$, we modify Equation (3.3.6) yet again:

$$\triangle^{\widehat{\alpha}}_{(i,T)} = \frac{\sum_{i \neq j} K\left(Var[r_{it} - r_{jt}]/h\right)(\widehat{\alpha}_i - \widehat{\alpha}_j)}{\sum_{i \neq j} K\left(Var[r_{it} - r_{jt}]/h\right)}. \tag{3.7.3}$$

Calculating $\triangle^{\widehat{\alpha}}_{(i,T)}$ from Equation (3.7.3) yields 0.17 as the average relative alpha version for the estimation method FH7. Finally, we introduce omitted variable bias in the estimation method Market model with just the market and an intercept. The Market model ignores the remaining six factors of the true model. The average absolute alpha under the Market model is 0.95 while the relative alpha version thereof (calculated according to Equation (3.7.3)) is 0.05. These estimates based on the estimation method with omitted factors (Market model) are significantly different from the values using the true alphas. A big advantage is thus that our Relative alpha estimation method finds a value for relative alpha (0.15) close to the true value of 0.17, whereas the estimation method Market model with omitted factors is far off at 0.05.

[Table 3.C.21 about here]

Next, we vary the length of the time series and show the influence of small sample bias on absolute alpha in comparison to relative alpha. From Table 3.C.21, columns four and five, we see that the small sample (of only 6 instead of 36 monthly returns) affects the estimation methods Relative alpha and Market model but not the estimation method FH7. Using the correct model (FH7) gives - even on very short samples - the correct average alpha (1.20, 0.17 when expressed as relative alpha) which is insignificantly different from the true alpha (1.27, 0.17 when expressed as relative alpha). Relative alpha performs poorly (0.03) on the short samples and the Market model, too (0.05, 0.02 when expressed as relative alpha). The simulation suggests that a minimal sample length is needed for the relative alpha method to work.

Finally, we reduce the cross section from which to pick the peer group from 2000 to 50. Results are in Table 3.C.21, columns six and seven. Using the correct model (FH7) gives identical alphas (0.44, 0.10 when expressed as relative alpha) to the true alphas. The Market model has a hard time yet again due to the omitted factor bias with estimated alphas of -0.01 when expressed as relative alpha. Relative alpha is also performing poorly at 0.06, significantly different from 0.10 for the same value based on the true alphas. It shows that for relative alpha to perform well, a minimal cross sectional dimension is required.

We conclude from our simulation that the good performance of relative alpha is due to its capability of dealing with omitted factor bias. In order to achieve that feat, the method needs reasonably long samples for the estimation (36 months) and a large enough cross section of peer funds (2000) so that the investment opportunity set is being spanned by its peers.

In Section 3.3 we argued that the variance of return differences $V[r_i - r_j]$ is a good measure of funds proximity. It consists of a systematic ($\triangle\beta' Cov(X)\triangle\beta$) and an idiosyncratic ($\sigma_i^2 + \sigma_j^2$) component. A concern is our assumption that the idiosyncratic component is fairly constant across hedge fund pairs while the systematic component measures the similarity of the hedge fund pair. Based on our simulation study and for one sample hedge fund, we know both components and plot them together in Figure 3.D.1 (sorted by increasing total variance, $V[r_i - r_j]$). We find that the idiosyncratic component ($\sigma_i^2 + \sigma_j^2$) is quite similar for different funds, while the systematic component ($\triangle\beta' Cov(X)\triangle\beta$) increases strongly in total variance, $V[r_i - r_j]$. We are thus comfortable with our argument that the variance of return differentials is a good measure for the similarity in the hedge fund risk exposures.

[Table 3.D.1 about here]

According to Equation (3.3.3), for hedge funds with identical strategies, the expectation of return differences is equal to the differences in true alpha. If two hedge funds differ in their strategies (i.e. have different betas), then Equation (3.3.3) will not hold perfectly anymore and $E[r_i - r_j] - (\alpha_i - \alpha_j) = \triangle\beta' E[X] + E[\varepsilon_i - \varepsilon_j]$. We argued above that this discrepancy should be small for similar hedge funds where we interpret similarity as small variances in return differences, $V[r_i - r_j]$. Using our simulation, we can depict that relation in Figure 3.D.2 where we plot (for one sample hedge fund) the discrepancies $E\left[r_i - r_j\right] - (\alpha_i - \alpha_j)$ across variances in return differences, $V[r_i - r_j]$. Figure 2 shows that the smaller the variance, the closer is the expectation of differences, $E\left[r_i - r_j\right]$, to the true alpha differences $(\alpha_i - \alpha_j)$. We are thus comfortable with our above assumption that Equation (3.3.3) holds reasonably well for similar hedge funds.

[Table 3.D.2 about here]

3.8 Conclusion

We propose a novel performance measure, relative alpha, which assesses the out-performance of a hedge fund with respect to a group of peers. It exhibits the intriguing property that omitted factor bias cancels, as the peer group is selected by exhibiting the least variance of return differentials. A nice side effect is that the investor does not even need to know the exact factor structure, nor the omitted factors - simply the similarity of hedge funds according to our distance measure leads to the reduction of omitted factor bias.

Different results obtain for different uses of alpha, but relative alpha tends to beat absolute alpha in all three dimensions. Concerning alpha as a measure of current outperformance of the risk-adjusted return, we find that relative alpha explains more return variation (81%) than the Fung and Hsieh [2001] model (41%). When using alpha in order to predict high future performance, then relative alpha can be used to construct portfolios of hedge funds and the out-of-sample performance of the top decile portfolio second order stochastically dominates sorts based on absolute Fung and Hsieh [2001] alpha, the appraisal ratio, the manipulation-proof performance measure, and the strategic distinctiveness index. Sharpe ratios of relative alpha sorted top portfolios are about twice those of the competitors. Related, we find fund flow is more closely related to relative alpha than it is to absolute alpha. Finally, using past alpha to predict future alpha, relative alpha is strongly persistent in our sample, as opposed to absolute alpha.

Our results stay robust to various methodological changes and sample manipulations. In our simulation study we show that relative alpha works better than absolute alpha when there are omitted variables, if there is a large number of hedge funds in the cross-section, and if there is a reasonable sample length.

Work lies ahead in several directions. Concerning absolute alpha, there is a lively discussion about skill versus luck as the driving force and we would like to study this question for relative alpha, too. Relative alpha might also predict voluntary delisting of hedge funds from the databases.

3.A Appendix: Performance Measures

3A.1 Appraisal Ratio

$$AR = \frac{\widehat{\alpha}}{\widehat{\sigma}_{\varepsilon}}, \tag{3.A.1}$$

where $\widehat{\alpha}$ is the alpha estimate of a hedge fund, $\widehat{\sigma}_{\varepsilon}$ is the residual standard deviation.

3A.2 Manipulation-proof Performance Measure of Goetzmann et al. [2007]

$$\widehat{\Theta} = \frac{1}{(1-\rho)\triangle t} ln\left(\frac{1}{T}\sum_{t=1}^{T}[\frac{1+r_t}{1+r_{ft}}]^{1-\rho}\right), \tag{3.A.2}$$

where T is the total number of observations, $\triangle t$ is the length of time between observations, r_t is the return of a hedge fund in t, r_{ft} is the risk-free rate, ρ is the relative risk-aversion coefficient.

3A.3 Strategy Distinctiveness Index of Sun et al. [2012]

$$SDI = 1 - corr(r_t, \mu), \tag{3.A.3}$$

where r_t is the return of a hedge fund in t, and μ is the average return of all funds belonging to the same style.

3.B Appendix: Factor Models

3B.1 Fung and Hsieh [2001]

1. Bond Trend-Following Factor, lookback straddles
2. Currency Trend-Following Factor, lookback straddles
3. Commodity Trend-Following Factor, lookback straddles
4. Excess return on the S&P 500 index over the risk-free rate
5. Difference in the returns on the Wilshire Small Cap 1750 index and Wilshire Large Cap 750 index
6. The monthly change in the 10-year treasury constant maturity yield
7. The monthly change in the spread between Moody's Baa yield and 10-year treasury constant maturity yield

3B.2 Agarwal and Naik [2004]

1. Returns on Russel 3000 Index
2. Returns on Morgan Stanely Capital International world excluding US Index
3. MSCI emerging market index
4. Salomon Brothers government and corporate bond index
5. Salomon Brothers world government bond
6. Lehman high yield index
7. Federal Reserve Bank competitiveness-weighted dollar index
8. Goldman Sachs commodity index
9. Factor-mimicking portfolio for size
10. Factor-mimicking portfolio for book-to-market equity
11. Factor-mimicking portfolio for the momentum effect
12. The monthly change in the spread between Moody's Baa yield and 10-year treasury constant maturity yield

13. At-the-money European call on the S&P 500 composite index
14. At-the-money European put on the S&P 500 composite index
15. Out-of-the-money European call on the S&P 500 composite index
16. Out-of-the-money European put on the S&P 500 composite index

3B.3 Edelman et al. [2012]

1. Bond Trend-Following Factor, lookback straddles
2. Currency Trend-Following Factor, lookback straddles
3. Commodity Trend-Following Factor, lookback straddles
4. Excess return on the S&P 500 index over the risk-free rate
5. Difference in the returns on the Wilshire Small Cap 1750 index and Wilshire Large Cap 750 index
6. The monthly change in the 10-year treasury constant maturity yield
7. The monthly change in the spread between Moody's Baa yield and 10-year treasury constant maturity yield
8. Excess return on the IFC Emerging Markets Index

3B.4 Capocci and Hübner [2004]

1. Excess return on the Russel 3000 Index
2. Factor-mimicking portfolio for size
3. Factor-mimicking portfolio for book-to-market equity
4. Factor-mimicking portfolio for the momentum effect
5. Excess return of the MSCI World Index excluding US
6. Excess return on the Lehman Aggregate US Bond Index
7. Excess return on the Salomon World Government Bond Index
8. Excess Return of the JP Morgan Emerging Bond Index
9. Excess Return of the Lehman BAA Corporate Bond Index
10. Excess Return of the Goldman Sachs Commodity Index

3.C Appendix: Tables

Table 3.C.1: Descriptive statistics: Hedge Funds

The summary statistics are the equally weighted cross-sectional averages, standard deviations, minimum, and maximum of the: mean monthly return, μ; the standard deviation of monthly returns, σ; the skewness, Skewness; the excess kurtosis, Kurtosis. The sample is February 1994 to June 2011.

	Mean	Std.dev	Minimum	Maximum
μ	0.87	0.79	-6.68	8.56
σ	4.06	3.07	0.00	38.89
Skewness	-0.18	1.31	-10.33	9.93
Kurtosis	6.53	6.84	1.56	111.81

Table 3.C.2: Descriptive statistics: Funds of Funds

The summary statistics are the equally weighted cross-sectional averages, standard deviations, minimum, and maximum of the: mean monthly return, μ; the standard deviation of monthly returns, σ; the skewness, Skewness; the excess kurtosis, Kurtosis. The sample is February 1994 to June 2011.

	Mean	Std.dev	Minimum	Maximum
μ	0.53	0.35	-2.63	5.26
σ	2.17	1.31	0.12	19.50
Skewness	-0.90	1.22	-12.10	8.66
Kurtosis	7.34	6.78	1.63	153.76

Table 3.C.3: Predicting future portfolio performance: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test.

HF deciles	Relative alpha			Absolute alpha			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	1.12	2.11	0.49	1.02	3.25	0.27	<0.01
Bottom	0.40	2.30	0.16	0.85	3.37	0.23	n/a
Top-Bottom	0.72	1.24	0.52	0.17	2.80	0.05	<0.01

Table 3.C.4: Predicting future portfolio performance: alternative hedge fund performance measures

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on Appraisal ratio of Treynor and Black (1973), Manipulation-Proof Performance Measure (MPPM) of Goetzmann et al. (2007), and Strategy Distinctiveness Index (SDI) of Sun, Wang, and Zhang (2012). The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test of each measure against relative alpha portfolios.

HF deciles	Appraisal ratio			DD (2013)	MPPM			DD (2013)	SDI			DD (2013)
	mean	std.dev	Sharpe Ratio	test	mean	std.dev	Sharpe Ratio	test	mean	std.dev	Sharpe Ratio	test
Top	0.75	2.71	0.23	0.01	0.89	3.24	0.23	0.07	0.77	3.70	0.17	0.01
Bottom	0.72	3.80	0.15	n/a	0.80	3.21	0.20	n/a	0.55	1.54	0.29	n/a
Top-Bottom	0.03	3.55	0.00	0.01	0.09	2.96	0.00	0.02	0.22	4.53	0.01	0.00

Table 3.C.5: Predicting future portfolio performance: alternative hedge fund factor models

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on absolute alphas of the Agarwal and Naik (2004), Edelmann et al. (2012), and Capocci and Hübner (2004) factor models. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test of each measure against relative alpha portfolios.

HF deciles	Agarwal and Naik (2004)			DD (2013)	Edelmann et al. (2012)			DD (2013)	Capocci and Hübner (2004)			DD (2013)
	mean	std.dev	Sharpe ratio	test	mean	std.dev	Sharpe ratio	test	mean	std.dev	Sharpe ratio	test
Top	0.94	1.96	0.43	0.30	1.18	3.16	0.33	0.07	1.07	2.91	0.32	0.08
Bottom	0.63	2.82	0.17	n/a	0.45	3.47	0.10	n/a	0.61	3.83	0.12	n/a
Top-Bottom	0.31	1.67	0.14	0.00	0.73	3.10	0.19	0.01	0.46	3.05	0.12	0.06

Table 3.C.6: Risk exposures within top and bottom portfolio deciles

The table demonstrates risk exposure to the factors form Fung and Hsieh (2001) of top and bottom decile portfolios constructed by sorting based on relative alphas. The risk exposures are averaged across time. The last column provides p-values of the mean differences between top and bottom portfolios risk exposures.

Factor	Top	Bottom	Difference	p-value
Alpha	2.60	-0.19	2.79	0.00
Equity Market Factor	0.08	0.07	0.01	0.33
The Size Spread Factor	0.03	-0.02	0.05	0.05
The Bond Market Factor	-0.27	-0.28	0.01	0.50
The Credit Spread Factor	-0.45	-0.53	0.07	0.06
Bond Trend-Following Factor	-0.02	-0.02	0.00	0.22
Currency Trend-Following Factor	0.02	0.01	0.01	0.00
Commodity Trend-Following Factor	0.01	0.01	0.00	0.98

Table 3.C.7: Predicting future portfolio performance within self-reported styles: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha across the largest self-reported styles: equity long/short, fixed income, global macro, CTA/managed futures, event driven, and relative value. The characteristics include monthly mean and standard deviation.

	Relative alpha			Absolute alpha		
	Top	Bottom	Top-Bottom	Top	Bottom	Top-Bottom
Equity long/short						
mean	1.15	0.71	0.43	1.08	0.91	0.16
std.dev	3.30	6.02	5.19	4.75	6.10	6.34
Fixed income						
mean	0.91	0.00	0.91	0.77	-0.30	1.07
std.dev	1.91	2.71	2.85	1.89	3.30	3.38
Global macro						
mean	0.97	0.24	0.72	0.89	0.44	0.45
std.dev	2.16	1.93	1.63	2.87	2.27	2.75
CTA/managed futures						
mean	1.75	0.66	1.09	1.64	1.00	0.64
std.dev	8.12	5.22	9.50	7.89	6.04	9.69
Event driven						
mean	1.20	0.43	0.77	1.16	0.57	0.59
std.dev	2.73	2.96	1.97	3.02	2.92	2.59
Relative value						
mean	1.16	0.49	0.67	0.94	0.43	0.50
std.dev	1.81	2.00	1.87	2.30	2.01	2.27

Table 3.C.8: Predicting future portfolio performance for funds of funds: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test.

HF deciles	Relative alpha			Absolute alpha			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	0.84	2.23	0.35	0.82	2.71	0.26	0.14
Bottom	0.05	1.95	0.01	0.06	2.20	0.01	n/a
Top-Bottom	0.79	1.74	0.42	0.76	2.27	0.30	0.04

Table 3.C.9: Fund flow vs past alpha

We regress fund flows (measured from *t-1* to *t*) on the estimates of relative (absolute) alpha (measured from t-1,...,*t-36*) and a constant. The table summarizes the OLS estimates, t-statistics, and p-values.

	Relative Alpha			Absolute Alpha		
	coeff.	t-stat	p-val	coeff.	t-stat	p-val
constant	0.01	10.21	0.00	0.01	9.84	0.00
alpha	0.00	3.47	0.00	0.00	0.85	0.40

Table 3.C.10: Persistence of alpha: relative alpha vs. absolute alpha

The table presents estimated slope coefficients (b) from stacked, non-overlapping linear regressions: $\triangle_{2i} = a_{\Delta} + b_{\Delta}\triangle_{1i} + \omega_i$ (for relative alpha) and $\alpha_{2i} = a_{\alpha} + b_{\alpha}\alpha_{1i} + v_i$ (for absolute alpha). It also provides t-statistics and p-values on the significance of the b estimates. By stacking the regressions, we assume that the slope coefficients are constant across periods.

Sample size	Relative alpha			Absolute alpha		
Formation/Evaluation	b	t-stat	p-val	b	t-stat	p-val
12/12	0.17	11.49	0.00	-0.08	-4.17	0.00
24/24	0.14	5.10	0.00	-0.01	-0.74	0.46
36/36	0.07	2.90	0.00	0.04	1.90	0.06

Table 3.C.11: Persistence of alpha for funds of funds: relative alpha vs. absolute alpha

The table presents estimated slope coefficients (b) from stacked, non-overlapping linear regressions: $\triangle_{2i} = a_{\Delta} + b_{\Delta}\triangle_{1i} + \omega_i$ (for relative alpha) and $\alpha_{2i} = a_{\alpha} + b_{\alpha}\alpha_{1i} + v_i$ (for absolute alpha). It also provides t-statistics and p-values on the significance of the b estimates. By stacking the regressions, we assume that the slope coefficients are constant across periods.

Sample size	Relative alpha			Absolute alpha		
Formation/Evaluation	b	t-stat	p-val	b	t-stat	p-val
12/12	0.14	7.49	0.00	0.31	17.21	0.00
24/24	0.17	6.56	0.00	-0.09	-3.43	0.00
36/36	0.04	1.00	0.32	-0.05	-1.04	0.30

Table 3.C.12: Explanatory panel regressions for relative alpha

The dependent variable is relative alpha. The independent variables are the seven Fung and Hsieh (2001) factors, the logarithm of assets under management, a dummy variable indicating whether the fund is using leverage, a dummy variable indicating whether the fund is using a high water, a dummy variable indicating whether the fund is open to new investments, the management fee, the performance fee, and a constant. All variables are averages over non-overlapping one-year periods. Below the coefficients we report p-values.

Variable Variable	OLS robust	Fixed effect	Fixed effect robust	Random effect	Random effect robust
equity market factor	0.09	0.12	0.12	0.10	0.10
	0.00	0.00	0.02	0.00	0.00
size spread factor	0.09	0.09	0.09	0.08	0.08
	0.02	0.02	0.07	0.02	0.03
bond market factor	0.03	0.09	0.09	0.05	0.05
	0.23	0.03	0.05	0.19	0.07
credit spread factor	0.00	0.08	0.08	0.02	0.02
	0.91	0.10	0.08	0.59	0.48
bond trend-following factor	0.01	0.00	0.00	0.01	0.01
	0.22	0.73	0.71	0.15	0.14
currency trend-following factor	0.02	0.02	0.02	0.02	0.02
	0.13	0.06	0.13	0.00	0.10
commodity trend-following factor	0.00	0.00	0.00	0.00	0.00
	0.97	0.70	0.64	0.93	0.89
log assets under management	0.01	-0.23	-0.23	0.01	0.01
	0.27	0.00	0.00	0.02	0.37
leverage {0,1}	-0.03	-	-	-0.05	-0.05
	0.45	-	-	0.40	0.31
high water mark {0,1}	0.02	-	-	-0.01	-0.01
	0.81	-	-	0.85	0.90
open to new investment {0,1}	-0.23	-	-	-0.26	-0.26
	0.01	-	-	0.00	0.03
management fee	0.17	-	-	0.22	0.22
	0.29	-	-	0.00	0.31
performance fee	0.01	-	-	0.01	0.01
	0.00	-	-	0.02	0.00
constant	-0.36	3.27	3.27	-0.37	-0.37
	0.14	0.00	0.00	0.00	0.23

Table 3.C.13: Predicting future portfolio performance for different samples: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test.

Period	Relative alpha			Absolute alpha			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Feb. 1994-Feb. 2000	1.70	2.77	0.61	2.36	3.77	0.63	0.56
Mar. 2000-Jul. 2007	1.10	1.87	0.59	0.75	2.76	0.27	0.01
Aug. 2007-Jun. 2011	0.71	1.90	0.38	0.53	2.66	0.20	0.10

Table 3.C.14: Predicting future portfolio performance for funds open to new investment: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test. Here the sample of the hedge funds is restricted to funds which are opened to new investments.

HF deciles	Relative alpha			Absolute alpha			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	1.10	2.19	0.47	1.05	2.97	0.31	0.01
Bottom	0.30	2.02	0.14	0.51	2.17	0.20	n/a
Top-Bottom	0.80	1.35	0.54	0.54	2.47	0.21	0.00

Table 3.C.15: Predicting future portfolio performance for large funds: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test. Here the sample of the hedge funds is restricted to the funds with assets under management larger than $20 million.

HF deciles	Relative alpha			Absolute alpha			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	1.38	2.86	0.44	1.32	3.56	0.34	0.02
Bottom	0.55	2.44	0.20	0.78	2.62	0.27	n/a
Top-Bottom	0.83	2.43	0.31	0.53	3.45	0.13	0.00

Table 3.C.16: Predicting future portfolio performance for shorter rolling window: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test. Here we use 24-month rolling windows instead of 36-month used in the main runs.

HF deciles	Relative alpha			Absolute alpha			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	1.26	2.07	0.59	1.07	3.10	0.31	0.01
Bottom	0.56	2.51	0.20	0.60	2.20	0.25	n/a
Top-Bottom	0.70	1.86	0.35	0.46	3.02	0.13	0.00

Table 3.C.17: Predicting future portfolio performance for smaller portfolios: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom 20 fund portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test. Here we use 20 hedge funds for our portfolios instead of deciles as used in the main runs.

HF 20-fund portfolios	Relative alpha			Absolute alpha			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	1.13	2.11	0.49	1.03	3.02	0.31	0.00
Bottom	0.50	2.85	0.14	0.71	2.87	0.22	n/a
Top-Bottom	0.63	2.08	0.27	0.33	3.09	0.09	0.00

Table 3.C.18: Predicting future portfolio performance: relative alpha vs. random portfolio

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and randomly selected portfolios. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test.

HF deciles	Relative alpha			Random portfolio			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	1.12	2.11	0.49	0.60	1.77	0.31	0.02
Bottom	0.40	2.30	0.16	n/a	n/a	n/a	n/a
Top-Bottom	0.72	1.24	0.55	n/a	n/a	n/a	n/a

Table 3.C.19: Predicting future portfolio performance for longer holding period: relative alpha vs. absolute alpha

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order-stochastic non-dominance test. Portfolios are held for a period of 12 months instead of one month as used in the main runs.

HF deciles	Relative alpha			Random portfolio			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	1.04	2.32	0.34	1.03	3.50	0.20	0.00
Bottom	0.56	3.23	0.12	0.70	3.36	0.15	n/a
Top-Bottom	0.48	2.59	0.14	0.33	3.63	0.04	0.02

Table 3.C.20: Correcting for the boundary bias of the kernel estimates

The table demonstrates out-of-sample performance characteristics of top, bottom, and top-bottom decile portfolios constructed by sorting based on relative alpha and absolute alpha. The characteristics include monthly mean, standard deviation, and Sharpe ratio. Last column provides p-values of the Davidson and Duclos (DD 2013) second-order stochastic non-dominance test. We correct for the boundary bias of the kernel estimates from the Equation (3.3.6) by using the locally weighted regression method proposed by Hastie and Loader (1993).

HF deciles	Relative alpha			Absolute alpha			DD (2013)
	mean	std.dev	Sharpe ratio	mean	std.dev	Sharpe ratio	test
Top	1.11	2.08	0.49	1.02	3.25	0.27	<0.01
Bottom	0.41	2.26	0.16	0.85	3.37	0.23	n/a
Top-Bottom	0.70	1.27	0.51	0.17	2.80	0.05	<0.01

Table 3.C.21: Simulation study

The table demonstrates results of the simulation study as described in Section 7. We simulate hedge fund returns by using a factor model that we assume to be true, using 100 simulation runs. In order to preserve the empirical characteristics of our hedge fund returns, we use the seven factors of Fung and Hsieh (2001) and set assumed true parameters equal to the estimated parameters as observed in the data. For the true model (columns two and three), we report two results according to four estimation methods. In column three ”Absolute alpha”, we report average absolute alphas. In column two ”Relative alpha”, we report relative alpha based on Equation (3.3.6) where we substitute $E[r_{it} - r_{jt}]$ with the expected differences in alpha. In the first row of results, we use our usual estimation method ”Relative alpha” as in Equation (3.3.6). In the second row, we report ”True alpha”. In the third row, we use the estimation method ”FH7” based on the Fung and Hsieh (2001) model. In the fourth row, we use the estimation method ”Market model” where we repeat the calculations using only a single market factor model. In columns four and five, marked ”only 6 returns”, we repeat the study using only 6 months of observations instead of 36 months. In columns six and seven, marked ”Only cross section 50”, we repeat the study using only 50 peer funds instead of 2000.

Estimation	True model, 36 returns, cross section 2000		Only 6 returns		Only cross section 50	
method used	Relative alpha	Absolute alpha	Relative alpha	Absolute alpha	Relative alpha	Absolute alpha
Relative alpha	0.15	n/a	0.03***	n/a	0.06***	n/a
True alpha	0.17	1.20	0.17	1.27	0.10	0.44
FH7	0.17	1.20	0.17	1.20	0.10	0.44
Market model	0.05**	0.95***	0.02***	0.05***	-0.01***	0.39

3.D Appendix: Figures

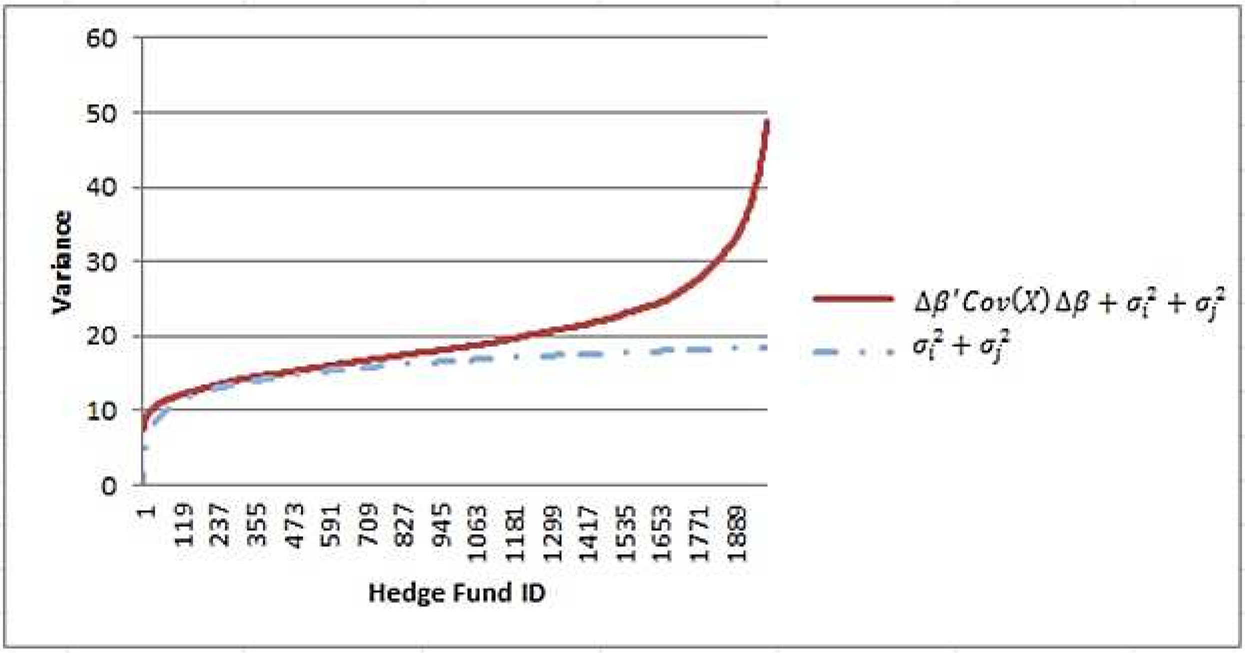


Figure 3.D.1: Simulation: variance decomposition

The figure shows the proportion between the idiosyncratic component ($\sigma_i^2 + \sigma_j^2$) of the variance of return differences $V[r_i - r_j]$ and the systematic component ($\triangle\beta' Cov(X)\triangle\beta$) for the simulation study based on one randomly selected hedge fund and the differences with all other hedge funds. We sorted the hedge fund peers by increasing total variances on the x-axis and numbered them by integers, using a variable called ID. The idiosyncratic components vary somewhat as a function of $V[r_i - r_j]$ and we thus fit a regression function to the data according to logarithmic function to the data of the model $(\sigma_i^2 + \sigma_j^2) = \text{a} + \text{b} \ln(ID) + \varepsilon$. All values are in percent per month.

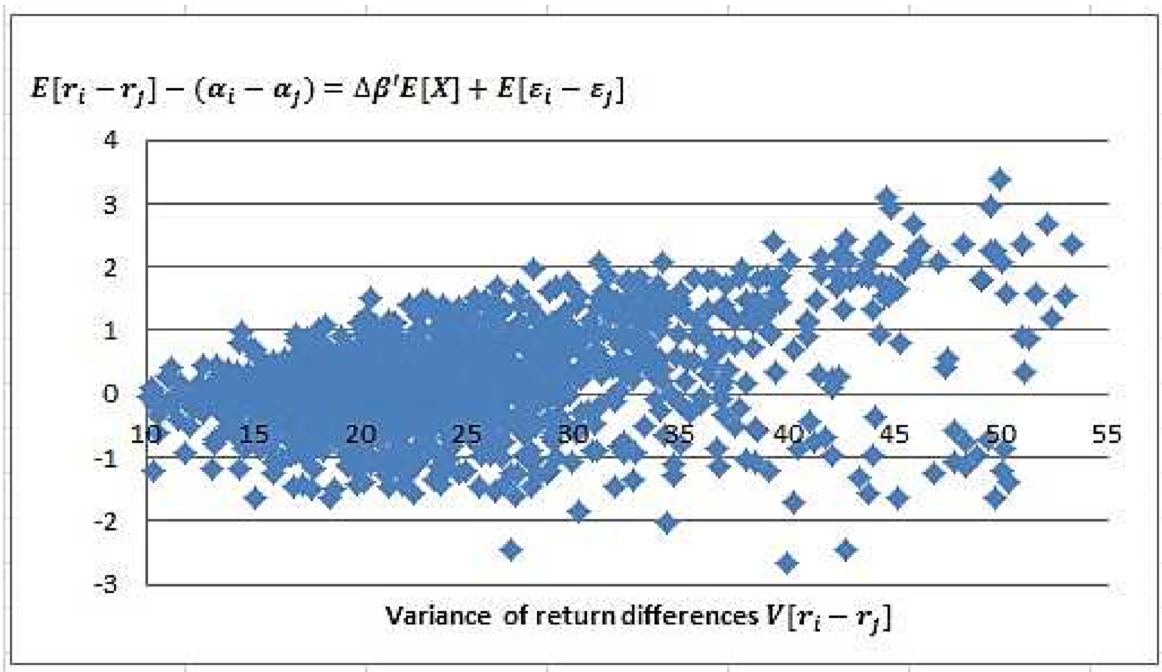


Figure 3.D.2: Simulation: Discrepancy of expectation of return differences and differences in alpha vs. variances of return differences

The figure shows the relation between the variance of the return differences $V[r_i - r_j]$ and the expression: $E[r_i - r_j] - (\alpha_i - \alpha_j) = \sum_{k=1}^{K} (\beta_{ik} - \beta_{jk}) E[X_{k,t}] + E[\varepsilon_{it} - \varepsilon_{jt}]$ for the simulation study based on one randomly selected hedge fund and the differences with all other hedge funds. All values are in percent per month.

References

Carl Ackermann, Richard McEnally, and David Ravenscraft. The Performance of Hedge Funds: Risk, Return, and Incentives. *Journal of Finance*, 54(3):833–874, 1999. 77

Vikas Agarwal and Narayan Y. Naik. On Taking the "Alternative" Route: Risks, Rewards, Style and Performance Persistence of Hedge Funds. *Journal of Alternative Investments*, 2(4):6–23, 2000a. 10

Vikas Agarwal and Narayan Y. Naik. Multi-Period Performance Persistence Analysis of Hedge Funds. *Journal of Financial and Quantitative Analysis*, 35(3): 327–342, 2000b. 77

Vikas Agarwal and Narayan Y Naik. Risks and Portfolio Decisions Involving Hedge Funds. *Review of Financial Studies*, 17(1):63–98, 2004. 2, 6, 9, 29, 73, 77, 109, 111, 114, 120, 122, 123, 135

AIMA. Roadmap to hedge funds. Report, Alternative Investment Management Association (AIMA), Deutsche Bank, 2012. 1, 5

Gaurav S. Amin and Harry M Kat. Diversification and Yield Enhancement with Hedge Funds. *The Journal of Alternative Investments*, 5(3):50–58, 2002. 78

Gaurav S. Amin and Harry M. Kat. Hedge Fund Performance 1990-2000: Do the Money Machines Really Add Value? *Journal of Financial and Quantitative Analysis*, 38(2):251–274, 2003. 77

Manuel Ammann, Otto R. Huber, and Markus M. Schmid. Hedge Fund Characteristics and Performance Persistence. Working Paper, University of St. Gallen, 2010. 113

Klaas Baks, Andrew Metrick, and Jessica Wachter. Bayesian Performance Evaluation. Working Paper, National Bureau of Economic Research, 1999. 16

Turan G. Bali, Stephen J. Brown, and K. Ozgur Demirtas. Do Hedge Funds Outperform Stocks and Bonds? *Management Science*, (59):1887–1903. 79

Jonathan B. Berk and Richard C. Green. Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy*, 112(6):1269–1295, 2004. 77, 110

Jonathan B. Berk and Jules H. van Binsbergen. Measuring Managerial Skill in the Mutual Fund Industry. Working Paper, Stanford University, 2013. 77, 111

David R. Bickel. Robust Cluster Analysis of Microarray Gene Expression Data with the Number of Clusters Determined Biologically. *Bioinformatics*, 19(7): 818–824, 2003. 27

Monika Billio, Mila Getmansky, and Loriana Pelizzon. Crises and Hedge Fund Risk. Working Paper, Yale School of Management, 2010. 30

Szabolcs Blazsek and Anna Downarowitz. Regime Switching Models of Hedge Fund Returns. Working Paper, School of Economics and Business Administration, University of Navarra, 2008. 30

Marshall E. Blume. On the Assessment of Risk. *Journal of Finance*, 26(1):1–10, 1971. 16

Nicolas P.B. Bollen and Robert E. Whaley. Hedge Fund Risk Dynamics: Implications for Performance Appraisal. *Journal of Finance*, 64(2):985–1035, 2009. 30

Richard Bookstaber. Hedge Fund Existential. *Financial Analyst Journal*, 59(5): 19–23, 2013. 2, 6

Stephen J. Brown, MaengSoo Kang, Francis Haeuck In, and Gunhee Lee. Resisting the Manipulation of Performance Metrics: An Empirical Analysis of the Manipulation-Proof Performance Measure. Working Paper, Stern Business School, New York University. 78

Stephen J. Brown, William N. Goetzmann, and Roger G. Ibbotson. Offshore Hedge Funds: Survival & Performance 1989-1995. *The Journal of Business*, 72 (1):91–117, 1999. 10

Stephen J. Brown, Thomas L. Fraser, and Bing Liang. Hedge Fund Due Diligence: A Source of Alpha in a Hedge Fund Portfolio Strategy. *Journal of Investment Management*, (6):23–33, 2008. 86

Markus K. Brunnermeier and Stefan Nagel. Hedge Funds and the Technology Bubble. *The Journal of Finance*, 59(5):2013–2040, 2004. 89

Daniel P.J. Capocci. An Analysis of Hedge Fund Performance 1984-2000. Working Paper, Luxembourg School of Finance, 2002. 10

Daniel P.J. Capocci and Georges Hübner. An Analysis of Hedge Fund Performance. *Journal of Empirical Finance*, 11(1):55–89, 2004. 10, 74, 77, 111, 113, 120, 122, 123, 136

Marc M. Carhart. On Persistence in Mutual Fund Performance. *Journal of Finance*, 52(1):57–82, 1997. 9, 10, 112

Nicholas Chan, Mila Getmansky, Shane M. Haas, and Andrew W. Lo. Systemic Risk and Hedge Funds. In Marc Carey and Réne M. Stulz, editors, *Risks of Financial Institutions*. Chicago and London: University of Chicago Press, 2006. 30

John B. Copas. Regression, Prediction, and Shrinkage. *Journal of the Royal Statistical Society*, 45(3):311–354, 1983. 15

Russel Davidson and Jean-Yves Duclos. Testing for Restricted Stochastic Dominance. *Econometric Reviews*, 32(1):84–125, 2013. 110, 118, 121, 123, 127

Bill Ding, Mila Getmansky, Bing Liang, and Russ Wermers. Share Restrictions and Investor Flows in the Hedge Fund Industry. Working Paper, University of Massachusetts at Amherst, 2009. 113

Daniel Edelman, William Fung, David A. Hsieh, and Narayan Y. Naik. Funds of Hedge Funds: Performance, Risk and Capital Formation 2005 to 2010. *Financial Markets and Portfolio Management*, 26(1):87–108, 2012. 120, 122, 123, 136

Anthony W.F. Edwards. Statistical Methods in Scientific Inference. *Nature*, 222 (June), 1969. 36

Franklin R. Edwards and Mustafa Onur Caglayan. Hedge Fund Performance and Manager Skill. *Journal of Futures Markets*, 21(11):1003–1028, 2001. 10, 77

Eugene F. Fama and Kenneth R. French. The Cross-Section of Expected Stock Returns. *Journal of Finance*, 47(2):427–465, 1992. 9, 29

Wayne E. Ferson and Jerchern Lin. Alpha and Performance Measurement: The Effects of Investor Disagreement and Heterogeneity. *forthcoming, Journal of Finance*, 2013. 78

William Fung and David A. Hsieh. The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers. *Review of Financial Studies*, 14(2):313–341, 2001. 2, 4, 6, 8, 9, 19, 29, 31, 73, 75, 76, 77, 82, 83, 84, 88, 92, 97, 98, 104, 108, 109, 111, 112, 114, 117, 118, 120, 121, 122, 123, 126, 128, 129, 132, 135

William Fung, David A. Hsieh, Narayan Y. Naik, and Tarun Ramadorai. Hedge Funds: Performance, Risk, and Capital Formation. *Journal of Finance*, 63(4): 1777–1803, 2008. 73, 113

Mila Getmansky. The Life Cycle of Hedge Funds: Fund Flows, Size and Performance. Working Paper, MIT Sloan School of Management, 2012. 113

Mila Getmansky, Andrew W. Lo, and Igor Makarov. An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns. *Journal of Financial Economics*, 74(3):529–609, 2004. 19, 67, 84, 91, 101

Vincent Glode and Richard C. Green. Information Spillovers and Performance Persistence for Hedge Funds. *Journal of Financial Economics*, 101(1):1–17, 2011. 111

William N. Goetzmann, Jonathan Ingersoll, and Stephen Ross. High-Water Marks and Hedge Fund Management Contracts. *Journal of Finance*, 58(4): 1685–1718, 2003. 113

William N. Goetzmann, Jonathan Ingersoll, Matthew Spiegel, and Ivo Welch. Portfolio Performance Manipulation and Manipulation-proof Performance Measures. *The Review for Financial Studies*, 20(5):1503–1546, 2007. 78, 79, 110, 122, 134

Michel A. Habib and D. Bruce Johnsen. Moral Hazard in Mutual Fund Management: The Quality-Assuring Role of Fees. Working Paper, School of Law, George Mason University, 2012. 126

Jasmina Hasanhodzic and Andrew W. Lo. Can Hedge-Fund Returns Be Replicated?: The Linear Case. *Journal of Investment Management*, 5(2):5–45, 2006. 9

Trevor Hastie and Clive Loader. Local Regression: Automatic Kernel Carpentry. *Statistical Science*, 8(2):120–143, 1993. 117

Roy D. Henriksson and Robert C. Merton. On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills. *The Journal of Business*, 54(4):513–533, 1981. 78, 82, 94, 97, 98, 104

Gerard Hoberg, Nitin Kumar, and Nagpurnanand Prabhala. Mutual Fund Competition, Managerial Skill, and Alpha Persistence. Working Paper, University of Maryland, 2014. 112, 113

James E. Hodder, Jens C. Jackwerth, and Olga Kolokolova. Recovering Delisting Returns of Hedge Funds. *forthcoming, Journal of Financial and Quantitative Analysis*, 2013. 17, 80, 83, 119

Grace Xing Hu, Jun Pan, and Jiang Wang. Noise as Information for Illiquidity. *Journal of Finance*, 68(6):2223–2772, 2013. 75, 82, 84, 87, 93

David Hunter, Eugene Kandel, Shmuel Kandel, and Russ Wermers. Mutual Fund Performance Evaluation with Active Peer Benchmarks. *Journal of Financial Economics*, 112(1):1–29, 2014. 112

International Monetary Fund. Global Financial Stability Report. 2004. 74

Ravi Jagannathan, Alexey Malakhov, and Dmitry Novikov. Do Hot Hands Exist among Hedge Fund Managers? An Empirical Evaluation. *Journal of Finance*, 65(1):217–255, 2010. 77, 111

Philippe Jorion. Bayes-Stein Estimation for Portfolio Analysis. *Journal of Financial and Quantitative Analysis*, 21(3):279–292, 1986. 15

Duen-Li Kao. Battle for Alphas: Hedge Funds versus Long-Only Portfolios. *Financial Analysts Journal*, 58(2):16–36, 2002. 77

Harry M. Kat. Managed Futures and Hedge Funds: A Match Made in Heaven. *Journal of Investment Management*, 2(1):1–9, 2004. 78, 88

Harry M. Kat and Faye Menexe. Persistence in Hedge Fund Performance: The True Value of a Track Record. *Journal of Alternative Investments*, 5(4):66–72, 2003. 10

Maurice G. Kendall. A New Measure of Rank Correlation. *Biometrika*, 30(1/2): 8189, 1938. 24

Robert Kosowski, Narayan Y. Naik, and Melvyn Teo. Do Hedge Funds Deliver Alpha? A Bayesian and Bootstrap Analysis. *Journal of Financial Economics*, 84 (1):229–264, 2007. 10, 16, 77, 127

Olivier Ledoit and Michael Wolf. Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection. *Journal of Empirical Finance*. 15

Lung-Fei Lee and William E. Griffiths. The Prior Likelihood and Best Linear Unbiased Prediction in Stochastic Coefficient Linear Models. Working Paper, University of New England, 1979. 36

Moshe Leshno and Haim Levy. Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance. *Management Science*, 48(8): 1074–1085, 2002. 79

Francois Serge Lhabitant and Michelle Learned. Hedge Fund Diversification: How Much Is Enough? *Journal of Alternative Investments*, 5(3), 2002. 78, 80

Bing Liang. On the Performance of Hedge Funds. *Financial Analysts Journal*, 55 (4):72–85, 1999. 77

Regina Y. Liu. Bootstrap Procedure under Some Non-i.i.d. Models. *The Annals of Statistics*, 16(4):1696–1708, 1988. 17

Andrew W. Lo. The Statistics of Sharpe Ratios. *Financial Analysts Journal*, 58(4): 36–52, 2002. 77

G.S. Maddala, Robert P. Trost, Hongyi Li, and Frederick Joutz. Estimation of Short-Run and Long-Run Elasticities of Energy Demand from Panel Data Using Shrinkage Estimators. *Journal of Business and Economic Statistics*, 15(1): 90–100, 1997. 12

G.S. Maddala, Hongyi Li, and V.K. Srivastava. A Comparative Study of Different Shrinkage Estimators for Panel Data Models. *Annals of Economics and Finance*, 2(1):1–30, 2001. 13, 15

Narayan Y. Naik, Tarun Ramadorai, and Maria Strömqvist. Capacity Constraints and Hedge Fund Strategy Returns. *European Financial Management*, 13(2):239–256, 2007. 78

Ľuboš Pastor and Robert F. Stambaugh. Mutual Fund Performance and Seemingly Unrelated Assets. *Journal of Financial Economics*, 63(3):315–349, 2002. 16

M. Hashem Pesaran and Ronald Patrick Smith. Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics*, 68(1):79–113, 1995. 12

M. Hashem Pesaran, Ronald Patrick Smith, and Kyung So Im. Dynamic Linear Models for Heterogeneous Panels. In László Mátyás and Patrick Sevestre, editors, *The Econometrics of Panel Data: a Handbook of the Theory with Applications*. Kluwer Academic Publishers, Dordrehct, 1996. 12

François-Éric Racicot and Raymond Théoret. Hedge Fund Returns, Kalman Filter, and Errors-in-Variables. *Atlantic Economic Journal*, 38(3):377–378, 2009. 30

C. Radhakrishna Rao. Simultaneous Estimation of Parameters in Different Linear Models and Applications to Biometric Problems. *Biometrics*, 31(2):545–554, 1975. 15, 36

Thomas N. Rollinger. Revisiting Kat's Managed Futures and Hedge Funds: A Match Made in Heaven, 2012. 78, 88

Donald B. Rubin. Using Empirical Bayes Techniques in the Law School Validity Studies. *Journal of the American Statistical Association*, 75(372):801–816, 1980. 15

William F Sharpe. Mutual Fund Performance. *Journal of Business*, 39(1):119–138, 1966. 77, 82, 94, 97, 98, 104

Bernard W. Silverman. In *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London, 1986. 116

Erik Sirri and Peter Tufano. Costly Search and Mutual Fund Flows. *Journal of Finance*, 53(5):1589–1622, 1998. 113

A.F.M. Smith. A General Bayesian Linear Model. *Journal of the Royal Statistical Society*, 35(1):67–75, 1973. 15, 34

Frank A. Sortino, Robert van der Meer, and Auke Plantinga. The Dutch Triangle. *The Journal of Portfolio Management*, 26(1):50–57, 1999. 77, 82, 87, 94, 97, 98, 104

Fiona Stewart. Pension Fund Investment in Hedge Funds, 2007. 74

Zheng Sun, Ashley Wang, and Lu Zheng. The Road Less Traveled: Strategy Distinctiveness and Hedge Fund Performance. *Review of Financial Studies*, 25 (1):96–143, 2012. 25, 110, 122, 134

P.A.V.B. Swamy. Efficient Inference in a Random Coefficient Regression Model. *Econometrica*, 38(2):311–323, 1970. 14, 34, 35, 36

Jack L. Treynor and Fischer Black. How to Use Security Analysis to Improve Portfolio Selection. *Journal of Business*, 46(1):66–86, 1973. 110, 122

Jack L. Treynor and Kay L. Mazuy. Can Mutual Funds Outguess the Market? *Harvard Business Review*, 44(4):131–136, 1966. 78, 82, 94, 97, 98, 104

Oldrich A. Vasicek. A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas. *Journal of Finance*, 28(5):1233–1239, 1973. 16

Marco Wilkens, Juan Yao, Nagaratnam Jeyasreedharan, and Patrick Oehler. Measuring the Performance of hedge Funds Using Two-Stage Endogeneous Benchmark. Working Paper, University of Augsburg, 2013. 112

Abgrenzung

Ich versichere hiermit, dass ich Kapitel 1 der vorliegenden Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel verfasst habe.

Kapitel 2 und Kapitel 3 entstammen einer gemeinsamen Arbeit mit Prof. Jens Jackwerth von der Universität Konstanz. Meine individuelle Leistung bei der Erstellung des Kapitel 2 und 3 beträgt jeweils 50%.