CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Hedge Fund Regulation, Characteristics, and Performance

Dissertation thesis

Author: Fan Yang

Study program: Economics and Finance

Supervisor: doc. Bc. Jiří Novák, M.Sc., Ph.D.

Year of defense: 2024

Declaration of Authorship

I hereby declare that I compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

I grant to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agree with the thesis being used for study and scientific purposes.

Prague, August 4, 2024

Abstract

The dissertation thesis consists of three interrelated parts that all analyze the determinants of hedge fund performance. The complexity of investment strategies implemented by hedge funds, their limited regulation, the mostly voluntary nature of the disclosure they provide, and the fragmentation of the performance data in commercial databases all constitute important challenges for hedge fund research. This doctoral thesis revisits prior empirical evidence on hedge funds and examines how various biases affect the inferences about their performance.

The first paper included in this thesis performs a meta-analysis of 1,019 estimates of hedge fund "alphas" collected from 74 research papers and examines whether the reported empirical results are affected by selective publication. We find that even though scholars have considerable discretion in choosing data samples and research methodologies to analyze hedge fund performance, there is little evidence of the presence of publication selection bias in this stream of literature. Our results demonstrate that the publication of empirical results may not be selective when there is no strong a priori theoretical prediction about the sign of the estimated coefficients.

The second paper builds on the first paper and analyzes in greater detail the individual determinants of hedge fund performance. We use the Bayesian Model Averaging (BMA) technique to examine the impact of various hedge fund strategies and research design choices for the reported results. We observe that after correcting for potential biases and adjusting for the variation in research designs, the empirical evidence indicates a steep decline in hedge fund performance over the past decades. This declining trend in hedge fund performance may result from more intensive competition in the hedge fund market or driven by stricter regulatory requirements for greater transparency that facilitates hedge fund oversight and invites imitation of successful investment strategies.

In the third paper, we examine specifically this potential determinant of hedge fund performance. We study the impact of the European Union's (EU) hedge fund regulation adopted in 2013 that substantially increased the transparency of hedge funds marketed in the EU. To isolate the impact of this regulatory change, we use the difference-in-differences and the propensity score matching methods. We observe that the requirement for greater transparency was indeed associated with the decline in hedge fund performance in the EU. The three papers included in this dissertation thesis provide novel insights into the relevance of various performance determinants that are relevant for investors, regulators, as well as for researchers who will analyze hedge fund performance in the future.

Acknowledgments

First of all, I would like to express my thanks to my supervisor doc. Bc. Jiří Novák, M.Sc., Ph.D. for inspiring me, sharing ideas with me, and guiding me on grant application, teaching skills, research methods, writing text, and so on. Thanks for your willingness to hear from me and provide unreserved feedback. I enjoyed working with you and all things learned are expected to benefit my career and future life.

Also, I am extremely grateful to my co-authors prof. PhDr. Tomáš Havránek, Ph.D. and doc. Zuzana Havránková, Ph.D. for their kind guidance and cooperation, I feel lucky to get to know you and collaborate with you. My thesis would not make it without your contribution and help.

Last but not least, thanks to my husband, my parents, and my friends for supporting me all the time, comforting me during tough times, and encouraging me whenever I need. The love, patience, and help from you have made me grow and strong. Love you all.

Typeset in LaTeXusing the IES Thesis Template.

Bibliographic Record

Yang, Fan: *Hedge Fund Regulation, Characteristics, and Performance.* Dissertation thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2024, pages 213. Advisor: doc. Bc. Jiří Novák, M.Sc., Ph.D.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Over the past thirty years, hedge funds experienced a steep increase in popularity. The assets under management (AUA) invested in hedge funds increased about 100 times between 1990 and 2020 (Barth *et al.* 2020; Stulz 2007). Investment devices that initially targeted high net-worth individuals now raise capital from a much broader investment base that includes substantially more conservative investors such as pension funds. The increased importance of hedge funds as a capital allocation device in the economy, as well as the repercussions of some of the notorious hedge fund failures, prompted closer scrutiny of their operations and raised questions about how they should be regulated. The increased interest in hedge fund operations and the social discourse on optimal regulation motivated extensive academic research aimed at measuring hedge fund performance. This research faces numerous challenges stemming from the fundamental characteristics of the industry. Hedge funds frequently engage in rather complex investment strategies that may dynamically evolve over time. Adequately adjusting for relevant risk exposures of dynamically evolving hedge fund investment strategies is far from trivial. Furthermore, hedge funds are typically not obliged to provide systematic disclosures of their holdings and performance to the regulators. Instead, they selectively report their performance to commercial databases that are primarily intended for large investors. The limitations in data coverage of hedge fund performance create a risk of biases in the data samples used in the empirical research. Perhaps due to the above challenges, the empirical findings in prior research are not always consistent, which complicates drawing generalizable conclusions and useful policy implications.

This dissertation thesis addresses the above challenge by synthesizing prior empirical evidence on hedge fund performance and analyzing how the reported estimates are affected by publication and data biases. The thesis consists of three papers that analyze the various aspects of hedge fund performance and its determinants. We first examine whether prior empirical estimates of hedge fund performance tend to be published selectively. Then, we examine the major determinants that explain the variation in the reported estimates in prior research. We identify a strong decline over time in reported estimates of hedge fund performance, and we analyze additional explanatory variables that impact the reported estimates. Finally, we investigate the impact of new rules imposed in the European Union (EU) on hedge funds. The thesis provides important novel insights that are valuable for researchers analyzing hedge fund performance, for investors who consider placing some of their capital in hedge funds, and for regulators who seek optimal ways of regulating the hedge fund industry.

1.1 Can hedge funds beat the market?

The dissertation is inspired by the puzzle involving the tremendous rise of the hedge fund industry, which seems to challenge the efficient market hypothesis (EMH) (Fama 1970). The EMH suggests that in a well-functioning capital market, any opportunities to earn abnormal returns are quickly competed away. Thus, investors can expect to only earn a return that is adequate to the level of systematic risk they bear. As the proportion of passive investments increased from close to zero in 1990 to over 15% in 2017 (Gârleanu & Pedersen 2022), the proportion of actively managed funds has shrunk over time (Stambaugh 2014). Interestingly, the amount of capital invested in hedge funds dramatically increased between 1990 and 2020 (Barth *et al.* 2020; Stulz 2007). This suggests that many investors believe that hedge funds are capable of generating higher returns than other investment vehicles. However, hedge fund managers usually require rather high management and performance fees (Guasoni & Obłój 2016) that diminish the value hedge funds generate for their investors. Even though high-performance fees may incentivize hedge fund managers, they may also be rather problematic for several reasons. The hurdle rate, equivalent to risk-free return, is often set for minimum performance qualified for incentive fees (Connor & Woo 2004). This measure may not constitute an adequate benchmark, especially considering the risks involved in complex and sometimes rather aggressive hedge fund strategies.

Furthermore, the payoff structure for hedge fund managers tends to be rather asymmetric, which incentivizes them to take risks. This asymmetric payoff structure may be partially offset by the so-called "high-watermark provision", which ensures that managers are compensated for their performance only after recovering any potential prior losses. Nevertheless, hedge fund managers may sometimes avoid fee reductions due to prior losses (Fung & Hsieh 1999). Prior research suggests that about 64 percent of the gross excess return is paid as the management fee and incentive fee while only 36 percent is earned by the investors (Ben-David *et al.* 2020). Given the high proportion of generated value paid to hedge fund managers and the fact that some of the less successful hedge funds may be terminated, it is questionable whether hedge funds, on average, create value for investors above the fees they charge. Furthermore, it is not obvious what factors are likely to affect hedge funds' ability to generate value for investors.

1.2 Can hedge funds beat the current market?

It is reasonable to expect the level of market efficiency to evolve over time as the markets become more structured and better regulated. Most market-wide efficiency measures in Rösch *et al.* (2017) show some overall improvements in market efficiency. This increase in market efficiency may have prompted the rise in the proportion of passive investment mentioned above. In contrast, opportunities to exploit market mispricing may have become rarer. As the amount of wealth invested in hedge funds dramatically increased over time, it is conceivable that an increasing number of hedge fund managers chase a diminishing number of investment opportunities to earn abnormal returns. In that situation it is natural to ask whether any superior hedge fund performance may nowadays persist with fewer investment opportunities and fierce competition in the industry.

Investigating these questions is also interesting because the current decade has witnessed many changes in the hedge fund industry. On the one hand, traditionally, hedge funds used to target high-net-worth individuals or institutions who are typically considered to have the knowledge and skills to make qualified investment decisions and be sufficiently aware of the risks various investment strategies involve. Thus, these investors are typically regarded as different from normal investors who require better protection. Therefore, it is commonly argued that due to the sophistication of the investor base, strict hedge fund regulation is not necessary and, in fact, not even desirable. However, as hedge funds' investment base has broadened over time and it starts to involve even retail investors, the above-mentioned argument becomes more questionable. Less sophisticated investors might be more likely harmed by the limited disclosures provided by some hedge funds, and they may fail to fully appreciate the implications that the limited data coverage may have for evaluating hedge funds' prior performance. These investors may also underestimate the implications of the changes in the hedge fund industry for their performance. If some investors have a distorted view of hedge funds, active investments will not mitigate the market deviation as intended, which may impair the efficiency of capital allocation in the economy.

Investigating these questions is crucial for enhancing our understanding of the increasingly prominent hedge fund industry and for contributing to the long-lasting debate on hedge fund performance. Prior studies provide some empirical evidence about hedge fund performance. However, there is substantial heterogeneity in the general conclusions reached by prior studies. Several studies indicate that hedge funds generate value for investors, including Brown *et al.* (1999), Ackermann *et al.* (1999), Liang (1999), Agarwal & Naik (2000), Fung & Hsieh (2004a), Kosowski *et al.* (2007), and Ibbotson *et al.* (2011). Other studies show the hedge funds underperform the benchmarks (Malkiel & Saha 2005; Getmansky *et al.* 2015). Yet another set of studies suggest that hedge fund performance is unstable over time and varies substantially across investment strategy categories (Fung *et al.* 2008; Billio *et al.* 2014; Capocci & Hubner 2004; Ding & Shawky 2007; Griffin & Xu 2009).

Several hedge fund studies analyze their performance over different time periods. For instance, Sullivan (2021) compares the pre-crisis period (1/1994 to 12/2008) and the post-crisis period (1/2009 to 12/2019) and finds a decrease of alpha in the second period. Similarly, Fung *et al.* (2008) analyze the results of subperiods classified by Long Term Capital Management (LTCM) crisis and NASDAQ crash (1995-1998, 1998-2000, 2000-2004) and find that hedge funds have a significantly positive alpha only in the second period. In the remaining time periods, hedge funds do not seem to generate abnormal performance that would significantly differ from zero. Based on the classification of subperiods Fung *et al.* (2008), Naik *et al.* (2007) further discuss the trend of different strategies, and they find some similar results to Fung *et al.* (2008). Nevertheless, They conclude that the performance of different strategies varies, which complicates drawing more general conclusions.

1.3 Challenges for hedge fund research

Investigating hedge fund performance involves multiple challenges. First, data biases resulting from the voluntary nature of reporting hedge fund performance in databases may influence the estimated results. Common biases identified by hedge fund studies are selection bias, survivorship bias (liquidation bias), and backfilling bias (instant-history bias). These biases create challenges for researchers in collecting representative samples and conducting empirical analyses that can be generalized to the overall population of hedge funds. It is important for researchers conducting hedge fund research to be aware of these biases and take measures to mitigate their effect on the estimated results. The self-selection bias may arise when different studies utilize different databases that have various coverages of hedge funds. Due to the voluntary nature of hedge fund performance reporting, fund managers can select whether to report their performance and to what databases (Fung & Hsieh 2004a; Baquero *et al.* 2005). Aiken *et al.* (2013) use the mandatory regulatory filings by registered FoFs, and they observe that only about one-half of these fund-level returns are reported to one of the five major hedge funds databases.

The survivorship bias occurs when some records of the inactive funds are not included in the data pool. Researchers test the impact of survivorship bias by comparing the performance (average return) of surviving funds and the performance of both surviving and inactive funds. For example, Liang (2000) uses data from TASS and HFR and finds that the survivorship bias in TASS exceeds 2% per year while in HFR, it is only .39% per year. Fung *et al.* (2006) find that survivorship bias in TASS, HFR, and CISDM are 2.4%, 1.8%, and 2.4%, respectively. To avoid the effect of survivorship bias, many studies analyzing the performance of hedge funds focus on data after January 1994, when most database vendors distribute data of both live and inactive funds (Aggarwal & Jorion 2010; Kosowski *et al.* 2007).

Biases may also occur when new funds enter the databases. Newly set up hedge funds typically undergo an incubation period during which they build up their performance track record. If the initial performance of a given hedge fund is good, it is likely included in the database so that it can attract additional investors. In such a case, the hedge fund's historical performance tends to be backfilled in the database. In contrast, hedge funds with poor initial performance typically cease their operations, and their weak performance track record never reaches the database. The selective inclusion of initially better-performing funds makes average returns in databases biased upward (Fung & Hsieh 2004a), which can affect inferences about hedge fund overall performance. Some studies use methods to mitigate backfilling bias. A common practice to address this issue is to eliminate the first 12 months of returns from the MAR database. Using this approach, Edwards & Caglayan (2001) find that the average annual returns of hedge funds in the first year are 1.17% higher than in subsequent years. Fung *et al.* (2006) eliminate the first 14 months of returns, and their results show that backfilling bias is around 1.5% in TASS, HFR, and CISDM. Given that some databases, such as TASS, Eurekahedge, and HFR, provide dates when funds start to report to databases, several studies use a more direct method of eliminating returns for each fund between the fund's inception date and the date to start reporting to the database (Agarwal *et al.* 2018; Posthuma & Van der Sluis 2003).

1.4 Measuring hedge fund performance

There are several approaches to measuring hedge fund performance. Two broad categories of methodological approaches are commonly used: (i) the reward-to-risk ratios, e.g., the Sharpe (1966) ratio, and (ii) the intercept terms from regressions of realized returns on risk factors, e.g., the Jensen (1968) alpha. Fung & Hsieh (2004a) develops a seven-factor model that comprises risk factors that mirror various risk exposures common in popular hedge fund investment strategies. After regressing the excess returns on risk factors, the intercept (alpha) is considered the abnormal return earned by the hedge funds. However, there is little consensus about the appropriate combination of various risk factors that should be used to estimate the value that hedge funds generate for investors. The use of different risk models may be one of the important reasons for the divergence in empirical results presented in prior research literature. With the existence of different methods, recent papers tend to use multiple methods to compare or combine several factors to analyze based on the topic. For instance, Capocci & Hubner (2004) simultaneously utilizes the capital asset pricing model, three-factor model, and four-factor model to investigate hedge fund performance. They also develop a multi-factor model combining factors from Carhart (1997), Fama & French (1998), and Agarwal & Naik (2002) and the factors representing emerging bond markets. Dichev & Yu (2011) use the three factor model (Fama & French 1992) as well as the eight factor model (Fung *et al.* 2006) to complement analysis. The multi-developed models foster various results even in the same setting of topics. The empirical results based on these various models are not directly comparable, which complicates their aggregation and the formation of general conclusions about the overall performance of hedge funds.

1.5 Motivation of a meta-analysis

Some of the challenges discussed above may be alleviated by performing a meta-analysis, which aggregates available estimates from prior studies. Aggregating empirical results published in numerous studies increases the time and geographic coverage of the empirical analysis, which increases the reliability of the empirical results. Furthermore, conducting a meta-analysis allows us to control for variation in data samples and method choices used in the primary studies, from which we collect the empirical estimates of hedge fund performance. We can then analyze whether these measures systematically vary with the reported empirical estimates. Performing a meta-analysis also allows us to adjust for potential selectivity in publishing the empirical results. Empirical researchers can exercise considerable discretion over the choice of their data samples and research methodologies. This gives them the possibility to choose results that best conform with their a priori expectations or the results that they consider most likely to be well-received by the journal editors. Hence, it is conceivable that the universe of published empirical estimates of hedge fund performance is biased due to this selectivity in publication. As a part of our research, we investigate whether the publication selection bias is present in this stream of empirical research literature and if so how big impact it has on the conclusions we can draw from these results.

In Chapter 2 of this dissertation thesis, we conduct a meta-analysis to aggregate the empirical results and adjust them for the publication selection bias. This approach allows us to draw more reliable inferences about how much value the funds create for their investors. In Chapter 3, we extend the above-mentioned analysis by investigating the determinants related to the data sample choices, research method choices, and the publication quality of the primary studies on the reported performance estimates. This approach allows us to combine results obtained using a multitude of approaches and to identify measures that best explain the cross-sectional variation in the reported empirical estimates. It also allows us to identify potential time trends in the reported estimates.

In order to conduct the empirical analysis for the studies included in Chapters 2 and 3 of this dissertation thesis, we collect 1019 intercept terms (alphas) from a regression of hedge fund performance on various risk factors. We collect these alpha estimates from 74 published empirical studies. Even though the risk factors used for different estimates vary, the alphas show additional returns generated by hedge funds that are likely due to managerial skills instead of systematic risk rewards.

Following prior research literature, we use several approaches to adjust for the publication selection bias. This allows us to evaluate how robust our results are to the specific approach we use. First, we use the conventional approach and regress the alpha coefficients collected from the primary studies on their standard errors. Absent any selectivity in publishing the empirical results in the primary studies, there should be no systematic relationship between the coefficients and their standard errors. In contrast, the positive relationship between the alpha estimates and their standard error implies that imprecise positive estimates (with a larger standard error) are more likely to be published than equally imprecise low or negative estimates, suggesting selectivity in publishing these results. We complement this conventional approach with the OLS estimates, including study-level fixed effects and between effects to address the idiosyncratic study-level variation and the differences in study size. We also test the weighted least squares estimates considering the standard errors that represent precision and the number of estimates in a given study. Finally, we also use the nonlinear models to relax the assumptions about the underlying distribution of the estimated coefficients and exploit discontinuities in these distributions (Stanley *et al.* 2010; Ioannidis *et al.* 2017; Bom & Rachinger 2019; Furukawa 2020; Andrews & Kasy 2019; van Aert & van Assen 2020).

Using the above-discussed approaches to detect the publication selection bias, we observe little evidence suggesting that the empirical estimates of hedge fund performance are published selectively. Our aggregated estimates adjusted for potential selectivity in publication range between 0.301 and 0.369, which is comparable to the unconditional mean of monthly alpha estimates of 0.36. We thus conclude that despite the considerable discretion researchers have in conducting empirical studies on hedge fund performance, this stream of research literature does not seem to be contaminated by publication bias. This conclusion is further supported by similar findings that we obtain with the use of the other detection approaches. Only the Kinked-meta model developed by (Bom & Rachinger 2019) shows a weak positive relationship that is significant at the 10% level. However, the corrected alpha is 0.32, which is also close to the uncorrected mean of 0.36 and falls within the range of corrected estimates in linear models.

In Chapter 2, we also analyze whether our conclusion on the absence of a substantial publication selection bias can also be observed within more homogenous sub-samples of our data set. This allows us to evaluate whether the finding could be affected by the heterogeneity of the reported estimates. In particular, we analyze separately alpha estimates depending on 1) whether the survivorship and backfilling biases are treated; 2) the one-factor model or the seven-factor model used; 3) whether the instrumental variables are employed; 4) the top 3 or top 5 finance journals published. Even within the sub-samples, we find limited evidence in support of the publication bias. However, for some sub-samples we observe some differences in the ranges of corrected estimates. For example, estimates with bias treatment show lower corrected results than the estimates without bias treatment. The only notable exception, for which we observe a substantial impact of publication selection bias is the sub-sample of estimates that are obtained using instrumental variables (IV). That specific sub-sample seems to suffer from publication bias. After Adjusting for this bias, the estimates range from -0.411 to 0.298, which is lower than the whole sample. The results support the proposition that IV-based estimates tend to suffer from publication bias more frequently than estimates based on other techniques (Brodeur *et al.* 2020) since the choice of an IV and measurements may offer additional discretion for researchers.

Unlike prior studies that identified the presence of publication selection bias in a wide range of economic and finance areas, our results in Chapter 2 show that even though the authors have considerable freedom in data choice and model design, the literature results can be unbiased to a large extent, possibly because both positive and negative relationships are plausible in the prediction. The chapter is the first study that investigates the publication selection bias in hedge fund research and provides an aggregation of hedge fund performance estimates. This research approach allows us to merge the results derived from various combinations of datasets and research designs, correct for potential selectivity in publishing empirical results, and obtain a representative hedge fund performance estimate. Search findings are particularly important for investors who lack access to a comprehensive data set of hedge fund performance data.

1.6 Determinants of hedge fund performance estimates

Chapter 3 uses the same dataset as Chapter 2 to address a different research question. We examine the determinants that explain the variation in the published alpha coefficients. We consider variables related to the data sample choices, research method choices, and the publication quality of the primary studies on the reported performance estimates. This analysis allows us to address questions such as whether the alpha estimates tend to be higher if they are based on more recent samples, samples sourced from specific databases, estimated using specific risk models, or published in higher-quality journals. Because the set of relevant determinants is *a priori* unknown, we use the Bayesian Model Averaging (BMA) approach, which considers various combinations selected from a large pool of potentially relevant explanatory variables and evaluates their ability to explain the variation in the reported alpha coefficients.

Our results suggest that about half of the variables we consider maintain a consistently positive or consistently negative association with the magnitude of the alphas across various regression specifications. Among them, nine variables that have the highest posterior inclusion probability (PIP) are examined as having a "strong" effect on hedge fund performance. The alpha is negatively associated with the data year and publication year, which suggests that the ability of hedge funds to generate value for investors has declined over time. In addition, the alpha coefficients are lower if they are measured on the net-of-fee basis when the backfilling bias is treated when estimating the alpha by the 1-factor model when the underlying performance data is sourced from more databases, when the alphas are estimated at the level of funds-of-funds, and when they are estimated for declining (i.e., "bear") markets. We also quantify the aggregate influence of these factors by the posterior mean of the distribution of regression coefficients and standard deviation of the posterior coefficient distribution so that the extent of the influence by individual factors can be identified.

Our empirical findings presented in Chapter 3 provide valuable insights for researchers, investors, and regulators. We provide information for researchers on how their data sample and research design choices likely affect the alpha estimates. For example, we quantify the impact of explicit treatment of primary data biases, extending the data sample by pooling data from several databases and using various risk models in estimating the abnormal returns generated by hedge funds. Fung & Hsieh (2004a) use funds of hedge funds (FOF) because they expect that FOF can reduce selection bias, survivorship bias, and backfilling bias, while Yang *et al.* (2023) points out that using FOF itself causes the selection bias. Our meta-analysis does show that the performance of FOF is significantly different from other fund types, which is a factor that scholars may consider in future research. Our results also informed investors about how the value generated by hedge funds varies over time, across various regions, investment strategies, and market conditions. These considerations might be important for investors in making their capital allocation choices. Finally, our results provide a comprehensive picture of hedge fund research that aggregates findings in various studies to inform the regulator about the general patterns in this research literature that are adjusted for potential selectivity in reporting these empirical results. Regulators may use our results to form better-informed policy decisions and be more aware of hedge fund performance in various contexts.

1.7 How does EU regulation affect the hedge funds?

Chapter 4 analyses the impact of a regulatory change in the European Union (EU) on hedge fund performance. Theoretically, the relationship between the tightness of hedge fund regulation and hedge fund performance is uncertain because there are arguments to support expectations for both the positive and negative relationship (Cumming & Dai 2010). Strict regulation may improve the industry's quality by eliminating weak funds that cannot meet regulatory requirements. The monitoring system could restrict managers from unethical and compensation-oriented behavior and possibly benefit the hedge funds' performance (Cumming & Dai 2010; Frumkin & Vandegrift 2009). The greater transparency that arises from stricter regulation may also allow investors to better monitor the managers, which may, in turn, discipline the managers and prevent them from exploiting the agency problem.

However, there are also several reasons to expect that tighter regulation may also hamper hedge fund performance. Managers may lose the freedom to contract and organize resources in the most efficient way. Common types of regulation, such as restrictions on minimum hedge fund size, restrictions on the location of key service providers, and market channels for hedge fund distributions, set barriers to entry, participation, and choice of efficient human resources. The regulations may lead to worse performance and less efficient hedge fund structures (Cumming & Dai 2010). Furthermore, greater transparency may reveal the "secret ingredient" of hedge funds' investment strategies to competitors and invite imitation. This may disincentivize hedge funds to see new profitable investment strategies as they may understand that they would be unable to exploit their benefits.

Historically, hedge fund regulation was rather weak. The main reason for such an arrangement was that hedge funds target high-net-worth individuals and institutional investors who should be sufficiently qualified to evaluate their investments. It could be argued that, in contrast to small retail investors, due to their size and sophistication, hedge fund investors do not require protection by the regulator. Nevertheless, over time, hedge fund regulation has become tighter in numerous jurisdictions. The regulators attempt to restrict aggressive managers from overusing the leverage that may threaten the stable market. They also require hedge funds to disclose some information about techniques, strategies, asset structures, and performance. The advantage for small investors who need such information to make decisions is obvious. To a large extent, the regulation could restrict the manipulation of managers whose misconduct is motivated by benefits from incentive fees. However, for a very long time, the opponents of the debate on whether regulation should be imposed insist that the restriction on leverage influences managers to take advantage of their skills, especially for arbitrage-type hedge funds. The disclosure encourages free-riders who imitate others' techniques, which is harmful to healthy competition.

The long-lasting debate about the beneficial and detrimental effects of tighter regulation may have motivated a stream of studies investigating the influence of regulations. These studies are mainly based on US data. Frumkin & Vandegrift (2009) find that the improved oversight from the authorities benefits hedge funds and investors because managers are more likely to refrain from unethical behavior. Some rules, like disclosure of portfolio holdings, are prone to reveal hedge funds' secret strategies and lower their competitive advantages and incentives of managers, which reduces hedge fund performance

(Cumming & Dai 2010; Shi 2017; Agarwal *et al.* 2013b). Several studies focus on the overall effect of the US Dodd-Frank Act, which involves the enhancement of registration and disclosure, but the results vary. For example, Cumming *et al.* (2020) compares the pre-Dodd-Frank period and the after-Dodd-Frank period. They show that US funds have worse performance after the effectiveness of the Dodd-Frank Act. On the contrary, Kaal *et al.* (2014) show a positive relationship between the Dodd-Frank Act and hedge fund performance.

Overall, a consistent conclusion can be drawn from prior studies that disclosure of portfolios negatively influences the performance of certain funds that have their special strategies. However, it is uncertain whether the results on the comparison of benefits (reducing managers' manipulation and increasing market efficiency) and costs (compliance costs and competitive strategies disclosed lowering incentives of managers) can be generalized across various settings and for different requirements. Therefore, there are still numerous open questions related to the relationship between the level of strictness of regulations and hedge fund performance.

In Chapter 4, we analyze the impact of the EU's new regulation adopted in 2011 – the Directive 2011/61/EU on Alternative Investment Fund Managers. This piece of regulation concerns all managers that manage any alternative investment funds, such as private equity funds, hedge funds, real estate funds, and infrastructure funds (Kamal 2012). Unlike the US area, which has been extensively investigated for different stages of regulation, research on the impact of the EU is much less common. This may possibly be because of the lower market size compared to the US hedge fund market. However, the EU market is the second market that follows the US and has expanded continuously. Until 2009, the EU accounted for 21%. Chapter 4 thus provides empirical evidence of the regulatory impact in that segment of the hedge fund industry.

The methodology we use to estimate hedge fund performance in Chapter 4 is informed by the systematic review of the impact of research design choices that we discuss in Chapter 3. Some prior studies use returns instead of risk-adjusted performance in estimating the impact of the new regulation. These studies usually find a positive relationship between the hedge fund performance and enforcement of regulations (Frumkin & Vandegrift 2009; Kaal *et al.* 2014). Other studies show a negative relationship. We argue that simply using the returns may not represent the hedge funds' ability to generate value for investors since a proportion of returns may result from market risk factors.

In Chapter 4, we also exploit the differences between the EU and US regulations. First, the coverage of EU regulation is wider since it applies to all hedge funds that market in the EU, while the off-shoring hedge funds marketing in the US are exempted from the US requirements. Second, the EU utilizes depositories for the safekeeping of the assets and monitoring of cash flows. That means the supervision in the EU is not limited to the authority. Besides, the hedge funds should have independent valuation periodically, and failure may result in sanctions. Third, disclosure by EU requirements tends to leak more information to competitors since annual reports are available to both investors and the authorities, while the US only requires disclosure to the authorities to analyze risks. Overall, the EU regulations seem stricter and require more compliance costs than the US rules. Considering that stricter rules imposed by the EU regions may result in a large amount of relocation costs by off-shoring hedge funds, and prior studies using risk-adjusted returns for measuring the performance show a negative relationship between hedge fund performance and strict regulations, we hypothesize that the EU regulations would lead to lower performance than before the enforcement of the law.

To investigate the relationship, we begin with the common method used in prior studies – the difference-in-differences method (DID) to compare the performance changes of EU and non-EU hedge funds. Then, we compare the two subgroups in the EU – the large hedge funds that meet the applicable standards of EU regulation and the smaller ones that are exempt from strict regulations. Furthermore, we combine the two difference-in-difference equations to form the difference-in-difference-in-difference method (DDD) to address the concern that some factors, such as regional sensitivity to market development, may cause group differences and lead to similar results. In this case, the coefficient of the double-interaction represents the relationship between the hedge fund performance and the EU regulation. In addition, we use the propensity score matching (PSM) to address another concern that the dates of authorization of different entities differ because of the time difference in the effectiveness of the AIFMD in national laws and the application date. Our hypothesis is supported by empirical results. The coefficient of triple interaction is significant and negative (-0.24%) and shows that hedge fund performance is negatively affected by the AIFMD. Our PSM analysis has similar results to reflect a negative change in performance by -0.2% in alpha.

The analysis based on the EU area provides additional evidence for the regulatory effect on hedge funds. On the one hand, the pool of EU data represents another important market for the hedge fund industry. On the other hand, the chapter distinguishes itself from other studies based on analyzing potential factors that may influence the empirical results and formulating the DDD to improve the prior methods in investigation.

Chapter 4 provides important insights into the impact of the regulatory change that affected European hedge funds. Both the US rules and EU regulations show the trend of hedge fund regulations, but they are still at an early stage. The analysis provides information for regulators in the EU and other countries that are considering adopting similar rules. They should be aware of the extent of hedge fund performance influenced by the enactment of similar laws but at the same time, they should be able to control the extent and keep the market stable by adjusting the strictness of the rules. Chapter 4 may also provide useful insights for investors by allowing them to compare the implications of the regulation in the US with the regulation in the EU.

1.8 Conclusion

Chapters 2, 3, and 4 of this dissertation thesis analyze various aspects of hedge fund performance. Chapters 2 and 3 offer a broader view of hedge fund performance by aggregating and synthesizing evidence provided in prior empirical research to draw conclusions about general factors affecting hedge fund research. With the overall picture of hedge funds, the thesis also provides information about the main influential factors for measuring performance for further research and regulatory enactment or amendment. The dissertation thesis provides strong evidence that prior hedge fund research literature mostly does not suffer from a publication selection bias, which is rather remarkable given the substantial scope researchers have in research design choices of their studies. We also identify several relevant characteristics that systematically influence the measurement of hedge fund performance. In comparison, Chapter 4 builds on the conclusions presented in the former two chapters to analyze the impact of a specific regulatory change on hedge fund performance. The study shows how European regulation of hedge funds affected the value they generate for investors. Taken together, these studies enrich our understanding of the performance of the hedge fund industry and its determinants.

Chapter 2

Is research on hedge fund performance published selectively? A quantitative survey

Fan Yang, Tomas Havranek, Zuzana Irsova, Jiri Novak

Abstract We examine whether estimates of hedge fund performance reported in prior empirical research are affected by publication bias. Using a sample of 1,019 intercept terms from regressions of hedge fund returns on risk factors (the "alphas") collected from 74 studies published between 2001 and 2021, we show that the selective publication of empirical results does not significantly contaminate inferences about hedge fund returns. Most of our monthly alpha estimates adjusted for the (small) bias fall within a relatively narrow range of 30 to 40 basis points, indicating positive abnormal returns of hedge funds: hedge funds generate money for investors. Studies that explicitly control for potential biases in the underlying data (e.g., backfilling and survivorship biases) report lower but still positive alphas. Our results demonstrate that despite the prevalence of publication selection bias in many other research settings, publication may not be selective when there is no strong *a priori* theoretical prediction about the sign of the estimated coefficients.

This paper was published in *Journal of Economic Surveys, DOI: 10.1111/joes.12574].* This paper is part of a project GEOCEP that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 870245. This work was supported by the Cooperatio Program at Charles University, research area Economics. Novak acknowledges support from the Czech Science Foundation (GA CR, project 24-11583S). Data and code are available in an online appendix at `meta-analysis.cz/hedge`

2.1 Introduction

> "*I can't figure out why anyone invests in active management, so asking me about hedge funds is just an extreme version of the same question. Since I think everything is appropriately priced, my advice would be to avoid high fees. So you can forget about hedge funds.*"

Eugene F. Fama [1]

Over the past three decades, hedge funds have experienced a spectacular increase in popularity. The value of assets under management (AUM) increased about 100 times between 1990 and 2020 (Stulz 2007; Barth *et al.* 2020). This trend is difficult to reconcile with the efficient market hypotheses (EMH) (Fama 1970), which suggests that financial markets rationally process available information and establish unbiased pricing of traded assets. Efficient financial markets should quickly eliminate any opportunities to earn abnormal returns. In efficient markets, an optimal investment strategy involves passive holding of a broad portfolio of assets (i.e., "indexing"). Consistent with this proposition, there has been a sharp rise in passive investment in recent decades. Gârleanu & Pedersen (2022) report that the proportion of wealth invested in passive funds increased from close to zero in 1990 to over 15% in 2017. Easley *et al.* (2021) argue that this trend has been further accelerated by the arrival of exchange-traded funds (ETFs). As passive funds do not participate in information processing and price discovery, the rise of passive investment raised concerns about its potentially detrimental impact on market efficiency. However, as argued by Stambaugh (2014), any deviations from efficient pricing create new opportunities for active investment. Thus, even though the proportion of actively managed funds has decreased over time (Stambaugh 2014), active mutual funds and hedge funds still manage more than 30% of invested wealth, which is more than twice as much as passive mutual funds and ETFs (Gârleanu & Pedersen 2022).

As aptly expressed in the opening quote by Eugene F. Fama, the 2013 Nobel laureate in economics, hedge funds represent an extreme form of active investment management. Consequently, they charge investors high management and performance fees. Ben-David *et al.* (2020) estimate that for every dollar of gross excess return earned by a hedge fund, on average, 64 cents are paid in management and incentive fees and only 36 cents are collected by the investors. Given the magnitude of these fees, it is surprising that hedge funds keep attracting growing amounts of capital.

A potential explanation for this puzzling trend is that investors have distorted views of the value hedge funds actually generate. Hedge funds are relatively lightly regulated and so they remain rather opaque in terms of their investment strategies, asset holdings, and realized returns. The paucity of information constrains systematic analysis of hedge

[1] Source: https://www.azquotes.com/quotes/topics/hedge-fund.html.

fund performance, and it may bias investors' expectations about hedge funds' value-generating potential. Hedge funds are not obliged to periodically publish information on their performance. Some, but not all, voluntarily report their performance data to commercial data providers. This implies that the data available for research is fragmented and may suffer from numerous biases. Furthermore, hedge funds tend to engage in a wide range of unconventional investment strategies, so it is not trivial to adequately adjust for the risks they bear. It is not clear to what extent these constraints bias reported performance estimates.

Prior empirical literature includes numerous conflicting results, which make it difficult to draw clear conclusions. Some studies analyze the potential reasons that include various data sources, data biases and measurement approaches. For instance, Liang (2000) and Fung *et al.* (2006) show that survivorship bias vary from 0.39% to 2.4% across different databases. The selection bias is round 1.17% in Edwards & Caglayan (2001) and 1.5% in Fung *et al.* (2006). Several papers point out that the risks in hedge funds are complicated, and they develop multiple measurement methods to capture the hedge fund performance (e.g., Sadka 2010; Bollen & Pool 2009). However, these examples show their concentration on individual aspects or constraints to evaluate the hedge fund performance but lack a study synthesizing this pool of diverse empirical results. Besides, we find that the fragmentation of hedge fund performance data and the wide range of alternative approaches to controlling for risk give researchers considerable discretion in research design.

Various data sources and estimation techniques may produce different results, some of which may be more attractive for publication than others. This potentially creates opportunities for selective publication that could bias the pool of coefficients reported in research articles. Brodeur *et al.* (2020) argue that research methods that offer researchers more degrees of freedom are more likely to suffer from selective publication as researchers may choose research designs and data samples to generate results that are attractive for publication. Hence, we believe that research literature on hedge fund performance merits a systematic investigation of the prevalence of this potential bias and of its potential impact on the reported estimates. Nevertheless, to the best of our knowledge, no prior study estimates and corrects for publication bias in this stream of literature. We provide the first quantification of the impact of potential selective publication and data biases on the hedge fund performance estimates.

In this paper, we perform the first quantitative survey of research literature on hedge fund performance. We aim to review and integrate published empirical findings and examine how they are affected by publication selection and data biases. We review empirical results in 74 studies on hedge fund performance published between 2001 and 2021. Our analysis is based on a sample of 1,019 estimates of intercept terms (i.e. the "alphas") from regressions of hedge fund returns on risk factors. The risk factors on the right-hand side of the regression equation represent various risk dimensions to which hedge fund investments may be exposed to. The slope coefficients (i.e. the "betas") capture hedge funds' exposure to the individual risk dimensions. The intercept terms, the alphas, represent the portion of realized returns that is not attributable to the fund's exposure to the systematic risk factors. In other words, the alphas represent the abnormal return earned by the hedge fund, which represents the difference between the actual realized return and the return that would be expected as fair compensation for the investment's systematic risk.

We examine the extent to which the alpha estimates are affected by selective publication and data biases. Publication selection bias represents a tendency to publish empirical results that are consistent with the underlying theoretical predictions or with prior empirical findings. Selective publication may result from both conscious and subconscious decisions made by authors, editors, and referees who discard results that look implausible in the light of their *a priori* expectations (Ioannidis *et al.* 2017). Publication selection bias and its implications are extensively discussed in prior literature, including Stanley (2001; 2005); Stanley *et al.* (2010); Havranek (2015); Brodeur *et al.* (2016); Bruns & Ioannidis (2016); Stanley & Doucouliagos (2017); Christensen & Miguel (2018); Brodeur *et al.* (2020); Blanco-Perez & Brodeur (2020); Zigraiova *et al.* (2021). These studies document that publication bias is widespread in a wide range of economic settings, and it substantially impacts the mean value of reported estimates. Given the discretion in research design due to data fragmentation and the variety of risk-adjustment methods, it is worthwhile to examine whether a similar bias is present also in the empirical literature on hedge fund performance.

We use several approaches to test for publication selection bias. First, we exploit the property that tests of statistical significance typically assume that publication bias is a linear function of the standard error. Hence, documenting a correlation between the two can be used as evidence on biased reporting of results in primary studies (Egger *et al.* 1997). We complement this conventional approach with several other recently developed techniques that use different combinations of fixed effects and weighting that relax the assumptions about the underlying distribution of the estimated coefficients and exploit discontinuities in these distributions (Stanley *et al.* 2010; Ioannidis *et al.* 2017; Bom & Rachinger 2019; Furukawa 2020; Andrews & Kasy 2019; van Aert & van Assen 2020). Using these modern techniques allows us to evaluate the robustness of our findings to assumptions that underlie various methodological approaches.

We find that despite the multitude of data sources and methodological approaches, empirical research on hedge fund performance is not substantially contaminated by publication selection bias. The result is inconsistent with the suggestion from Brodeur *et al.* (2020) that publication selection bias is related to the researchers' discretion possibly because less tendency exists regarding the favorable hedge fund results. In our regressions, most of the slope coefficients that capture the impact of publication bias are statistically insignificant. These results also mostly hold when we consider more homogeneous subsamples of alpha estimates that either adjust or do not adjust for survivorship and backfilling biases, subsamples that use a specific asset pricing model to compute the alphas, and subsamples of alphas sourced from the three and five leading journals in finance. A notable exception is the group of empirical estimates based on IV. Such a conclusion is consistent with a recent paper by Brodeur *et al.* (2020) who find that IV-based estimates are more likely to suffer from publication bias than estimates based on other techniques.

Unsurprisingly, the monthly alpha estimates adjusted for this (small) bias are fairly close to our unconditional sample mean of 36 basis points (i.e. 0.36%, which corresponds to 4.3% *per annum*). Our estimates suggest that the "representative" alphas corrected for the publication selection bias range from 0.274 and 0.386.[2] We observe fairly similar ranges when only using observations published in the Top 3 (0.265, 0.358) and Top 5 (0.263, 0.355) finance journals, which suggests that publication quality does not materially affect the reported estimates. In comparison, we document a slightly higher range for estimates based on the one-factor model (0.349, 0.707) and those that are not treated for the survivorship and backfilling biases (0.282, 0.521). In contrast, we observe a lower and wider range for the "corrected" alpha coefficients based on IV techniques (-0.411, 0.298).

We make several contributions to the literature. First, using several recently developed methodological approaches, we aggregate and synthesize fragmented empirical evidence on hedge fund performance. Prior research has long acknowledged that the absence of a comprehensive database may lead to distortion and misleading findings (Fung & Hsieh 2004a; Agarwal *et al.* 2009b). Fung & Hsieh (2004a) point out that differences in coverage across various hedge fund data providers may lead to rather different average returns for a given hedge fund type. To illustrate this observation, they state that two data providers specified two substantially different estimates for equity market-neutral hedge funds for the month of January 2001 (-1.57% vs. 2.13%). Such discrepancies across data sources imply that the choice of the database may have a substantial impact on the estimated hedge fund performance. In a recent working paper, Joenvaara *et al.* (2019a) underscore the importance of combining data from various databases, and they propose a new way of doing so. We offer an alternative approach to overcome the data fragmentation problem. Our approach is based on aggregating the alpha coefficients estimated in prior studies that are themselves based on various data sources. Relative to Joenvaara *et al.* (2019a), the advantage of our approach is that it allows us to include even estimates based on private or hand-collected data and to control for potential data and publication biases simultaneously.

Second, we provide a comprehensive battery of tests to evaluate the impact of pub-

[2] Most of the alphas in our data set are computed on the net-of-fee basis (984 out of 1,019, not tabulated), which implies that the average value we report mostly represents abnormal returns net of management and performance fees.

lication and data biases on hedge fund performance estimates. This paper is the first study that systematically analyzes the impact of selective publication on the reported hedge fund performance results. Recent research suggests that research settings that offer researchers considerable discretion are particularly prone to suffer from selective publication (Brodeur *et al.* 2020). We evaluate this effect in a research field that is characterized by fragmented data and a plurality of methodological approaches to estimate abnormal returns. Furthermore, hedge fund literature frequently mentions a concern that survivorship and backfilling biases may distort estimates of hedge fund performance (Fung & Hsieh 2004a). Prior studies typically address this issue by measuring the returns of funds of funds as their reported performance is less likely to be affected by backfilling historical information for successfully incubated funds and by omitting data for dead funds. However, relying on the data on funds of funds has several shortcomings. First, the inclusion of a hedge fund in a fund of funds is in itself an endogenous decision that has an impact on the reported returns. There is no guarantee that the funds that are actually included in the fund of funds' portfolios are representative of the entire hedge fund population and that the individual funds are treated in these portfolios with appropriate weights relative to the general population. Second, funds of funds charge investors an additional layer of management and performance fees (Stulz 2007), which may distort the quantification of the abnormal return generated by individual hedge funds (Amin & Kat 2003). Due to these additional costs, funds of funds may not represent an attractive investment opportunity for many hedge fund investors. We offer a different approach to adjust for these issues that is based on the aggregation of estimates reported in prior studies.

Third, by documenting a publication selection bias for the subset of estimates based on IV our study provides out-of-sample evidence in support of the recent finding by Brodeur *et al.* (2020), who argue that IV-based estimates are more likely to suffer from publication bias than estimates based on other techniques. When exploring the potential underlying reasons for this finding, Brodeur *et al.* (2020) suggest that it may arise due to the considerable discretion IV estimation gives researchers in designing their empirical tests. In line with this conjecture the authors observe that when the instruments are relatively weak, the second-stage results are likely to be close to the conventional thresholds for statistical significance. Our evidence is consistent with this proposed explanation. We observe that IV-based estimates in our sample seem to be more likely to suffer from selective publication.

Our analysis is relevant to investors who consider investing in hedge funds, to regulators who seek the optimal design of the regulatory framework, as well as to researchers in economics and finance. Our evidence on the absence of a significant publication bias and the fairly narrow range of 30 to 40 basis points that we document for the corrected monthly alpha estimates allow investors to calibrate their expectations of hedge fund performance. Our study also informs regulators that even though hedge funds are not obliged to systematically publish their performance and the data is fragmented in numerous private databases, prior empirical research does not suffer from selectivity in reporting hedge fund performance. Finally, our study demonstrates that despite the prevalence of publication selection bias in numerous other research settings in economics and finance, publication tends not to be selective when there is no strong *a priori* theoretical prediction about the sign of the estimated coefficient and when journals may be more open to publishing statistically insignificant estimates. This may help researchers identify areas where publication bias can be expected and where, in contrast, it is less likely.

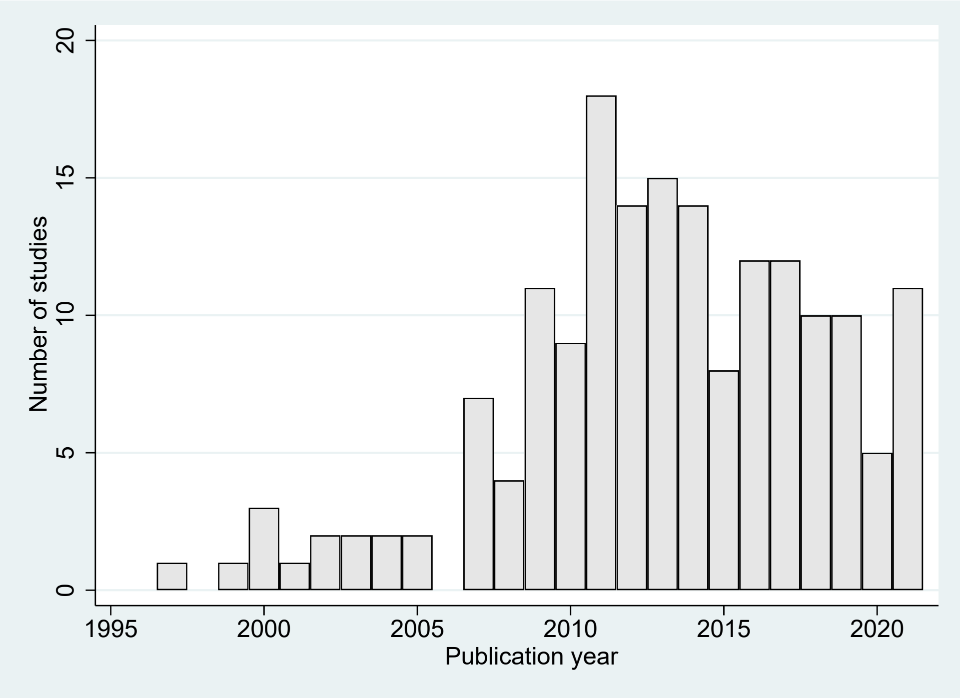
The remainder of the paper is organized as follows. In Section 2.2, we review prior research literature. In Section 3.3, we describe our data collection procedure. In Section 2.4, we present our main empirical results based on the full sample of alpha estimates collected from the primary studies. In Section 2.5, we report the results based on more homogeneous subsamples of alpha estimates. Section 2.6 concludes.

2.2 Previous research on the assessment of hedge fund performance

The increasing prominence of hedge funds as an investment device and the increasing role they play in the economy prompted extensive empirical research aimed at evaluating how well they perform. Over the past decade, numerous studies on hedge fund performance have been published. Figure 2.1 shows the surging number of studies on hedge funds published in leading finance journals – the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*, the *Journal of Financial and Quantitative Analysis*, and the *Review of Finance*.

Connor & Woo (2004) and Agarwal *et al.* (2015) provide narrative reviews of the hedge fund literature. Connor & Woo (2004) give an overview of the history of hedge funds, they discuss the key characteristics that distinguish hedge funds from other investment vehicles, outline their typical investment strategies, and discuss issues in measuring hedge fund risk and performance. Agarwal *et al.* (2015) concentrate more specifically on reviewing research on hedge fund performance and on factors that affect it, e.g., hedge fund characteristics and the risks hedge funds take. They also discuss the role hedge funds play in the economy and their impact on asset prices, market liquidity, quality of corporate governance, and the propagation of financial crises. Furthermore, they discuss a number of issues related to the sources of data used for hedge fund research. We complement these studies by performing a quantitative analysis of whether estimates of hedge fund performance reported in prior empirical research are affected by publication selection bias.

Figure 2.1: Articles on hedge fund performance



Note: The figure shows the number of hedge fund-related articles in top 5 journals in finance (*Journal of Finance*, *Journal of Financial Economics*, *Review of Financial Studies*, *Review of Finance*, *Journal of Financial and Quantitative Analysis*) published in a given year excluding articles that are only published online without a print version.

2.2.1 Estimating Performance

A standard challenge addressed in empirical research analyzing the performance of investment strategies (including those followed by hedge funds) is to adjust for the systematic risk these strategies involve properly. To address this issue, most modern studies on hedge fund performance report the intercept terms (the "alphas") from regressions of hedge fund returns on various combinations of risk factors, as shown in Equation 3.1.

$$(R_p - R_f) = \alpha_p + \sum_{n=1}^{N} \beta_{n,p} \cdot F_n + \epsilon_p \tag{2.1}$$

where R_p denotes the realized return on portfolio p, R_f denotes the risk-free rate of return, α_p represents the intercept term, F_n represents the n-th risk factor, $\beta_{n,p}$ denotes the sensitivity of portfolio p to the n-th risk factor, and ϵ_p represents the error term. Loadings on the risk factors (the "betas") represent a "normal" compensation for the risk that the investment entails. The alphas capture the portion of realized returns unexplained by the set of risk factors. The alphas can thus be interpreted as "abnormal" returns that the fund generates for the investors over and above (or below if negative) what would be expected for a given level of risk. This approach explicitly models an investment's exposure to various risk dimensions. However, the set of relevant risk dimensions is open to question. Thus prior literature provides estimates based on various risk models.

The Jensen (1968) alpha is the simplest of the intercept-based approaches. It was initially designed to measure the investment performance of mutual funds. Returns are measured relative to a benchmark that is relevant for a well-diversified investor. Building on the portfolio theory (Markowitz 1952) and the Capital Asset Pricing Model (CAPM) (Sharpe 1966; Lintner 1965; Mossin 1966; Black 1972), this approach uses the equity market excess return ($R_m - R_f$) as the sole risk factor. It maintains that well-diversified investors only require compensation for an investment's contribution to the volatility of returns on the market portfolio, i.e. for an investment's systematic risk, which is in turn determined by its returns' sensitivity to the variation in market returns. The slope coefficient beta in a regression of an investment's excess return on the market portfolio excess returns captures this sensitivity. In contrast, the intercept term alpha represents the portion of realized excess return that cannot be explained by an investment's contribution to the portfolio risk, i.e. the value generated for investors.

The simplicity of modeling systematic risk with the use of a single risk dimension constitutes a limitation that may be particularly relevant for hedge funds that engage in complex and dynamic investment strategies that are likely to exhibit various forms of exposure to systematic risk. Due to this complexity, prior research develops risk models that are specifically designed to capture risk dimensions relevant to hedge fund strategies. Most prominently, Fung & Hsieh (2004a); Fung *et al.* (2008) propose a seven-factor model that comprises risk factors that mirror various risk exposures common in popular hedge fund investment strategies. Specifically, the model comprises the following risk factors: (i) the stock market excess return, (ii) the spread between the small-capitalization and large-capitalization stock returns, the excess return pairs of look-back call and put options (iii) on currency futures, (iv) on commodity futures, and (v) on bond futures, (vi) the duration-adjusted change in the yield spread of the U.S. 10-year Treasury bond over the 3-month T-bill, and (vii) the duration-adjusted change in the credit spread of Moody's BAA bond over the 10-year Treasury bond. These risk factors are intended to capture risk exposures of a broad set of hedge fund types ranging from equity long-short funds to managed futures funds.

The use of various pricing models has some advantages and disadvantages. The Jensen alpha is well-rooted in financial theory and universally applicable in a wide range of research settings. Hence, empirical results based on the Jensen alpha are easily comparable across various research strategies and data samples. Furthermore, the Jensen alpha is not directly affected by the "model uncertainty problem", which results from the uncertainty over which of the multitude of variables identified in prior research to be associated with realized stock returns actually constitute the "true" returns determinants. Harvey *et al.*

(2016) and Harvey (2017) suggest that the pool of candidate risk factors documented in prior research is overblown by the tendency to publish statistically significant rather than insignificant results, i.e., "p-hacking"[3]. Thus, the authors argue that the variables associated with realized returns may not be true risk factors. Instead, they may represent "false positives" that will likely fail to explain the cross-sectional variation in realized returns in the future. In contrast to the universal measure of the Jensen alpha, the seven-factor model is specifically designed for research in hedge fund performance. Thus, it is likely more effective in filtering away various risk exposures relevant to complex investment strategies followed by hedge funds. Prior research offers alpha based on various risk models. In our robustness checks, we evaluate whether our findings are sensitive to limiting our analysis to a subsample of alpha coefficients that are based on the one-factor and seven-factor models.

2.2.2 Data Fragmentation

Besides the uncertainty about the appropriate risk model, hedge fund performance research faces the challenge of data fragmentation. Due to the relatively light regulatory oversight, hedge funds are mostly not obliged to periodically disclose audited financial information on their performance. Hence, there is no comprehensive central depository of hedge fund data. Only a subset of funds self-select to voluntarily report information on their performance to private data providers. Thus, prior research mostly relies on data sets obtained from commercial databases or hand-collected. The data sets used in prior research may not be comprehensive, and so they may not be fully representative of the entire hedge fund population (Aggarwal & Jorion 2010; Liang 2000; Posthuma & Van der Sluis 2003). This may complicate the interpretation of these findings and raise questions about the generalizability of these results to the universe of hedge funds.

Fung *et al.* (2006) discuss the level of overlap in hedge fund coverage between various databases. Liang (2000) and Agarwal *et al.* (2009b) show that the information provided is not always consistent across all the databases, which implies that the results reported in prior research may be sensitive to the choice of the source database. Similarly, in a recent working paper Joenvaara *et al.* (2019a) propose a new way of combining data from various databases, and they conclude that using this combined database matters for a conclusion about hedge fund performance. They argue that based on this combined database, hedge fund performance appears to be lower but more persistent. These findings underscore the importance of aggregating results based on different segments of hedge fund data.

Hedge funds are not obliged to independently verify reported data by auditors or established data providers. Liang (2003) finds that surviving funds are more likely to be effectively audited, and funds with more reputable auditors report more consistent

[3]We explicitly address the issue of "p-hacking" in our research setting in Subsection 2.4.2

data. Patton *et al.* (2015) find that data on hedge fund returns change depending on when the database is accessed. They also observe that underperforming funds are more likely to alter their performance histories. Data on hedge fund returns may be unreliable because the valuation of illiquid holdings may be imprecise (Cassar & Gerakos 2011) or because the highly incentivized managers may tamper with the reported information to give an impression of better and more stable performance (Bollen & Pool 2009). These complications may contaminate the results of hedge fund research and affect inferences about overall hedge fund performance.

2.2.3 Empirical Findings

Given the multitude of data sources and the range of methodological approaches used to estimate hedge fund performance, it is not surprising that prior research amassed an extensive body of sometimes conflicting empirical findings. Several studies indicate that hedge funds generate value for investors. Brown *et al.* (1999) document superior risk-adjusted returns in offshore hedge funds, but they find little support for performance persistence. Ackermann *et al.* (1999) and Liang (1999) observe that hedge funds earn higher risk-adjusted returns than mutual funds even though they have a higher overall risk due to which hedge funds do not outperform general stock market indices. Agarwal & Naik (2000) find that combining investments in hedge funds with passive investing generates better reward-risk combinations than a passive investment in various asset classes. Fung & Hsieh (2004a) propose seven risk factors relevant to hedge fund research, and they find that jointly these factors explain about 80% of hedge fund returns. Nevertheless, they also find that even after considering these risk factors, hedge funds generate positive alphas for the full sample period. Kosowski *et al.* (2007) use bootstrapping and Bayesian approaches to address some of the limitations common in hedge fund research. They document significant alphas and also substantial persistence in alphas in hedge funds, which suggests that the superior performance of hedge funds cannot be solely attributed to luck. Similarly, Ibbotson *et al.* (2011) conclude that alphas earned by hedge funds are positive and remarkably stable over time even during a financial crisis.

In contrast, Malkiel & Saha (2005) and Getmansky *et al.* (2015) argue that after adjusting for database biases, hedge funds on average underperform their benchmarks. Fung *et al.* (2008) observe a positive and statistical alpha only for an 18-month long subperiod out of the sample covering 120 months. Billio *et al.* (2014) conclude that the alphas generated by hedge funds change dramatically over time and across categories. Capocci & Hubner (2004) observe positive excess return for 10 out of 13 investment strategies that they analyze, but only for one-quarter of individual hedge funds. They also show that best-performing funds follow momentum strategies and have limited holdings of emerging market bonds. Also Ding & Shawky (2007) suggest that the evaluation of hedge fund performance relative to market indices depends on the level of aggregation of hedge fund data and on the adjustments for skewness in hedge fund returns distribution. They conclude that even though all hedge fund categories outperform the general market index less than half of the individual hedge funds beat it. Griffin & Xu (2009) find limited evidence of superior skills of hedge fund managers in timing the market and in picking individual stocks. The alphas they observe are small on a value-weighted basis and insignificant on an equal-weighted basis.

Some of the divergence in the reported results may be due to the data biases resulting from the voluntary nature of reporting of hedge fund performance in databases. A self-selection bias arises when successful hedge funds are more likely to report their performance to commercial databases. Jorion & Schwarz (2014) find that investment companies act strategically and they list in multiple commercial databases their small, best-performing funds, which helps them raise awareness about the funds and attract new investments (Fung & Hsieh 1997; 2000). Agarwal *et al.* (2013a) examine the impact of self-selection bias by comparing data in five commercial databases with information in Form 13F that are reported quarterly by advisors (rather than funds) with the Securities and Exchange Commission (SEC). They find that even though reporting initiation is more likely after a superior performance it subsequently declines. Similarly, Edelman *et al.* (2013) combine previously unexplored data sources with manual data collection to construct a comprehensive dataset of returns earned by large hedge fund management companies. Based on the sample covering more than half of the industry's AUM they observe little differences between the reporting and non-reporting firms. In contrast, Aiken *et al.* (2013) use the mandatory regulatory filings by registered funds-of-funds (FoFs) and they observe that only about one-half of these fund-level returns are reported to one of the five major hedge funds databases.

The backfilling bias or the "instant-history bias" arises when hedge funds are included in databases together with their performance history only after succeeding during an "incubation period" intended to accumulate a performance track record before offering the fund to investors. Recording performance histories of only the successful funds introduces a positive bias into the database (Fung & Hsieh 2000; Posthuma & Van der Sluis 2003). To quantify its effect prior research compares returns generated in the first years of hedge fund existence in the database with other years. Estimates based on this approach range between 1.0% and 1.5% *per annum* (Fung & Hsieh 2000; Edwards & Caglayan 2001).

The survivorship bias may arise when commercial databases terminate coverage of previously included funds. Providers may wish to purge the database of funds that no longer operate because they are not relevant to their clients anymore. Hodder *et al.* (2014) report that on average 15% of hedge funds exit the database every year. A bias arises when the funds that exit the database on average underperform the "surviving" funds. Brown *et al.* (1999) examine survivorship bias in a database of active and defunct offshore funds and observe positive risk-adjusted returns even after adjusting for the bias. Liang (2000) observes that poor performance is the main reason for a fund's disappearance from the databases and finds that the survivorship bias exceeds 2% per annum and it varies with investment styles. Edwards & Caglayan (2001) compare the performance of defunct funds with those that are still in operation and they estimate the impact of the bias at 1.85% per annum. Agarwal *et al.* (2015) propose a range between 2.0% and 3.6% per annum.

The variability of prior empirical results and the potential impact of various data biases complicate the interpretation of this stream of research. Thus, we consider it worthwhile to conduct a quantitative survey to synthesize this pool of diverse empirical results and to examine how they are affected by publication selection and data biases. Our empirical approach builds on and complements earlier research that uses the meta-analysis methodology to study the performance of other types of funds (Rathner 2012; Revelli & Viviani 2015; Coggin *et al.* 1993). Rathner (2012) performs a meta-analysis of 500 performance estimates collected from 25 empirical studies on the performance of Socially Responsible Investment (SRI) funds and he concludes that most primary studies do not find any significant performance difference between the SRI funds and the conventional funds. Furthermore, similar to our paper, Rathner (2012) also studies the impact of treatment of the survivorship bias for the magnitude of the estimates reported in the primary studies and he concludes that primary studies that adjust for the survivorship bias are more likely to report SRI funds outperformance relative to conventional funds. Revelli & Viviani (2015) perform a meta-analysis of 85 studies and 190 experiments that examine the impact of SRI on financial performance. They conclude that, overall, SRI considerations have neither a positive nor negative impact on performance. Coggin *et al.* (1993) conduct a meta-analysis of the investment performance of U.S. equity pension funds and they examine their stock picking and market timing abilities. They identify substantial differences in the performance of individual funds and they observe that some funds produce substantial abnormal returns while others do not. They also conclude that regardless of the choice of benchmark portfolio or estimation model, equity pension funds exhibit superior stock-picking and inferior market-timing skills. Our paper builds on this prior research by applying these methodological approaches to measure hedge fund performance and it extends them by employing a battery of modern econometric techniques for identifying publication selection bias.

2.3 Dataset

To perform a comprehensive analysis of how published evidence on hedge fund performance is affected by selective publication and data biases we collect a large dataset of alpha estimates from primary studies. Alpha estimates represent abnormal returns adjusted for exposures to risk factors. Individual alpha coefficients reported in primary studies thus aim to capture the same underlying concept of value generated by hedge funds for investors. All the collected alpha coefficients are measured in the same unit (i.e. percentage) and they are normalized to monthly frequency. Hence, they are directly comparable, which makes them suitable for aggregation in a quantitative survey.

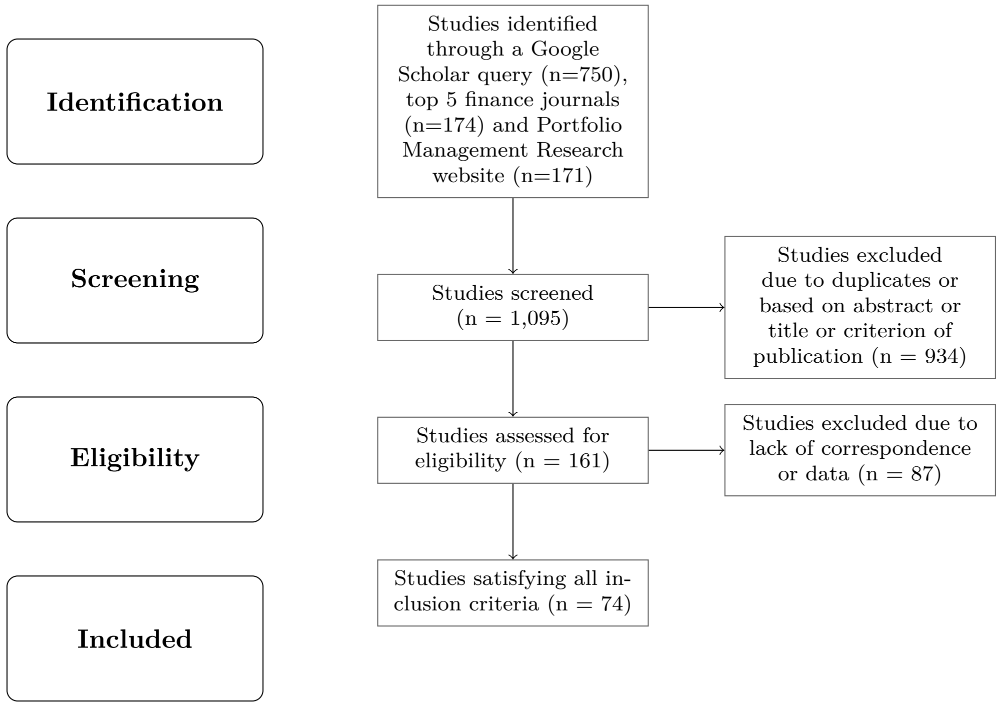
Our data collection process follows the guidelines proposed by Havranek *et al.* (2020). We restrict our analysis to estimates published in peer-reviewed research journals. The peer-review process constitutes an important quality assurance mechanism. Using only estimates that underwent the peer-review process increases the likelihood that the collected alpha coefficients are estimated using established methodological approaches and they are free of error. Furthermore, we expect most researchers and practitioners to form their subjective understanding of typical alpha estimates predominantly based on published articles. Our sample thus likely mirrors the set of studies that shape people's views of hedge fund performance.

Figure 2.2 provides an overview of the individual steps of our data collection process. First, we build a preliminary list of studies based on references included in the sections on hedge fund performance in two comprehensive review articles: Connor & Woo (2004) and Agarwal *et al.* (2015). Second, we perform a systematic Google Scholar search using the following combinations of keywords: "hedge fund returns" OR "hedge fund performance". We search for alpha estimates in the articles as ordered by Google Scholar. We terminate this phase of data collection after having covered the first 750 articles in the Google Scholar list. We observe that after having reached this position at Google Scholar list the articles become less relevant and the likelihood of identifying additional articles with usable alpha estimates drops dramatically. Third, to make sure that our search does not miss any important articles, we perform a slightly broader Google Scholar search using less restrictive keywords: "hedge fund" OR "hedge funds" in the following finance journals: the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*, the *Journal of Financial and Quantitative Analysis*, and the *Review of Finance*. Fourth, to ensure comprehensive coverage of articles in journals aimed primarily at investment professionals and which may not be as highly cited and ranked by Google Scholar, we perform a similar search based on the combination of keywords: "hedge fund" OR "hedge funds" in the journals listed on the Portfolio Management Research website[4]: the *Journal of Portfolio Management*, the *Journal of Financial Data Science*, the *Journal of Impact and ESG Investing*, and the *Journal of Fixed Income*.

To be included in the dataset, a given alpha estimate must be accompanied by a measure of statistical significance, i.e. a t-statistic, a standard error (SE), and/or a p-value. We use these measures to compute the precision of individual alpha estimates. We use the precision variable in our tests of selective publication as well as for our

[4]Source: https://www.pm-research.com/).

Figure 2.2: PRISMA flow diagram



Note: We perform our primary search in Google Scholar based on the following combinations of keywords: "hedge fund returns" OR "hedge fund performance". Furthermore, we perform our secondary search based on slightly broader set of keywords: "hedge fund" OR "hedge funds" in the following finance journals: the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*, the *Journal of Financial and Quantitative Analysis*, and the *Review of Finance*. Finally, we perform our tertiary search based on the combination of keywords: "hedge fund" OR "hedge funds" in the journals listed on the Portfolio Management Research website: the *Journal of Portfolio Management*, the *Journal of Financial Data Science*, the *Journal of Impact and ESG Investing*, and the *Journal of Fixed Income*. We screen for the alpha coefficients the first 750 studies identified by our primary Google Scholar search, as well as 174 studies identified by our secondary search in the top 5 finance journals, and additional 171 studies identified by our tertiary search in journals listed on the Portfolio Management Research website. We are left with 161 studies after the screening. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) is an evidence-based set of items for reporting in systematic reviews and meta-analyses. More details on PRISMA and reporting standard of meta-analysis in general are provided by Havranek *et al.* (2020).

data verification. Before constructing our final sample we attempt to identify alpha coefficients that may have resulted from human error in data hand-collection. To do so we first convert all the measures of statistical significance to a common metric, i.e. t-statistic. Whenever available we collect corresponding t-statistics from primary studies. If the authors report standard errors instead, we compute the implied t-statistic as the ratio of the alpha coefficient and the corresponding standard error. In studies using the Bayesian approach, we divide the alpha coefficient by the reported standard deviation.

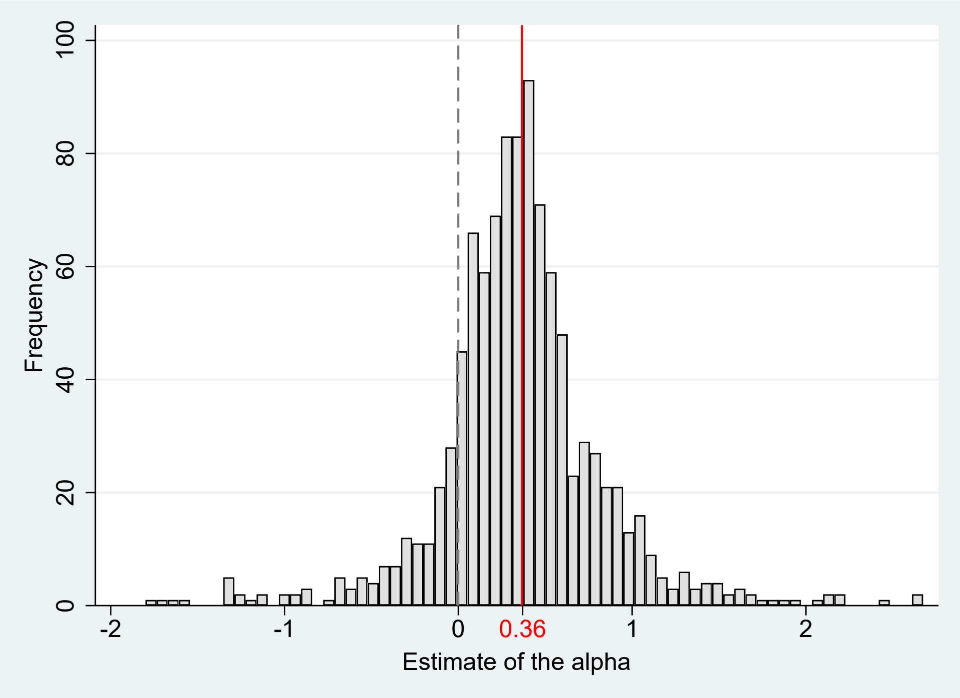
If the authors report p-values we check whether they explicitly state that these are based on a one-tailed or a two-tailed tests. If the type of the test is not explicitly stated in the article we try to infer it from the discussion of the level of statistical significance of results tabulated in the primary studies. If the type of the test cannot be ascertained from the interpretation of the results we assume a two-tailed test (1 study). We then use the inverse t-distribution to convert the reported p-value to a t-statistic. If the authors report the total number of observations based on which a given alpha coefficient is estimated we use that number for the degrees of freedom. If the authors report both the number of observations in the cross-section and in the time series we use the product of the two numbers. If the information on the number of observations is only provided for the cross-section or the time series, we use that number instead. If none of the above is provided we assume 168 observations, which is equal to the sample median for the subsample where the number of observations is explicitly stated. We then check all observations with the implied t-statistic greater than 10 for potential errors in hand-collecting the data. We ensure that such results are presented as highly significant in the main text of the primary study. We discard the 1 observation where the authors report a t-statistic greater than 50.

Our data collection procedure yields 1,019 alpha estimates obtained from 74 primary studies. The data sample size makes our study one of the largest quantitative surveys of prior studies in financial economics. The first alpha estimates that meet the sample collection criteria specified above were published in 2001. We end our data collection on September 1, 2021. The long time span exceeding 20 years ensures that our sample of alpha estimates is representative of the accumulated pool of evidence in this stream of research literature. Figure 2.3 shows the histogram of the alpha estimates in our sample. The figure suggests that the distribution is fairly normal and quite symmetric. Furthermore, we do not observe any significant kinks in the distribution, which indicates that no levels of alpha estimates are significantly underrepresented or over-represented. Figure 2.3 thus offers some preliminary indication that the distribution of our dataset has the expected characteristics and it is free from dramatic discontinuities.

The vertical line in Figure 2.3 denotes the unconditional sample mean of monthly alphas of 0.36%, which corresponds to an annual abnormal return of 4.32%. This result is broadly consistent with values proposed in prominent studies on hedge fund performance. For example, Getmansky *et al.* (2015) report monthly alphas based on the Fung & Hsieh (2001) seven-factor model for various hedge fund strategies between 0.18% and 0.56%. This suggests that our dataset does not dramatically differ from what would be expected based on prior literature. At the same time, the histogram shows that the individual alpha estimates are relatively dispersed. This suggests that there are substantial differences across various studies and estimation approaches.

We report further information about the characteristics of alpha coefficients in the

Figure 2.3: Distribution of alphas



Note: The figure depicts a histogram of the alphas reported by individual studies. The solid red vertical line denotes sample mean. The gray dashed line denotes 0.

individual primary studies in Table 2.1.[5] The table illustrates substantial differences between the individual primary studies. The number of alphas collected from a study ranges from 1 to 61. The median primary study contributes nine alpha estimates to our sample. The number of databases the alpha estimates are based on also substantially varies across the individual primary studies. The median value of one indicates that a typical study uses only one database as a source of data. This observation again highlights the importance of aggregating and synthesizing the hedge fund performance estimates. Nevertheless, the most comprehensive studies include up to seven databases. The sample period of returns data in a typical study spans 171 months, which corresponds to more than 14 years. However, some studies use data sets covering only 31 months (about 2.5 years), while others comprise 475 months, i.e., almost 40 years. Most studies use only one risk model for estimating abnormal returns. However, some studies use up to seven risk models. We collect about 30% of our alpha estimates from studies published in the Top 5 finance journals (i.e., the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*, the *Journal of Financial and Quantitative Analysis*, and the *Review of Finance*).

Table 2.1 also reports the mean value, the standard deviation, the minimum, the

[5]We are grateful to an anonymous referee for proposing this overview table.

median value, and the maximum of the alpha coefficients reported in a given primary study. The mean alpha estimate reported in a study ranges from -0.51 to 1.22. This suggests that the individual primary studies reach very different conclusions about the abnormal returns generated (or destroyed) by hedge funds. Our study aggregates these diverse findings and draws inferences about the overall value reported in this stream of research. The following section presents these aggregated results.

Table 2.1: Primary studies

Study	#Alphas	#Sources	#Months	#Models	Top5	Mean	StDev	Min	Md	Max
Agarwal *et al.* (2017)	61	1	240	3	1	0.27	0.77	-1.80	0.20	3.00
Ahoniemi & Jylha (2014)	1	1	180	1	0	0.41		0.41	0.41	0.41
Aiken *et al.* (2013)	6	6	72	3	1	0.22	0.12	0.09	0.20	0.40
Ammann & Moerth (2005)	3	1	136	2	0	0.03	0.02	0.02	0.03	0.06
Ammann & Moerth (2008a)	4	2	136	2	0	0.25	0.10	0.20	0.20	0.40
Ammann & Moerth (2008b)	2	1	136	2	0	0.00	0.01	-0.01	0.00	0.00
Aragon (2007)	16	1	96	2	1	0.40	0.30	-0.05	0.37	1.02
Asness *et al.* (2001)	20	1	81	2	0	0.00	0.51	-1.35	0.12	0.97
Bali *et al.* (2013)	11	1	216	1	0	0.51	0.16	0.25	0.48	0.87
Bhardwaj *et al.* (2014)	6	1	223	1	1	0.66	0.18	0.36	0.71	0.84
Blitz (2018)	19	2	204	1	0	0.09	0.17	-0.37	0.12	0.33
Bollen & Whaley (2009)	4	1	144	1	1	0.58	0.28	0.25	0.57	0.93
Brown (2012)	2	1	168	1	0	0.83	0.04	0.80	0.83	0.85
Buraschi *et al.* (2014)	45	1	198	3	1	0.30	0.16	0.01	0.28	0.75
Cao *et al.* (2016)	1	1	228	1	1	1.18		1.18	1.18	1.18
Chen & Liang (2007)	10	3	138	7	1	0.37	0.20	-0.01	0.41	0.62
Chen *et al.* (2017)	4	2	216	1	1	0.34	0.17	0.11	0.36	0.52
Chincarini & Nakao (2011)	7	1	475	3	0	0.47	0.07	0.37	0.47	0.55
Clark & Winkelmann (2004)	5	1	119	1	0	0.36	0.11	0.16	0.41	0.43
Dichev & Yu (2011)	8	2	348	2	1	0.36	0.15	0.11	0.38	0.53
Ding & Shawky (2007)	17	1	168	5	0	0.65	0.21	0.40	0.62	1.10
Ding *et al.* (2009)	9	1	144	1	0	0.73	0.45	0.07	0.66	1.58
Do *et al.* (2005)	26	1	31	1	0	0.34	0.51	-0.46	0.18	1.35
Duarte *et al.* (2007)	4	3	194	1	1	0.35	0.14	0.15	0.39	0.48
Edelman *et al.* (2012)	6	4	72	3	0	0.08	0.26	-0.26	0.07	0.44
Edelman *et al.* (2013)	6	4	108	1	1	0.06	0.04	-0.02	0.07	0.09
Edwards & Caglayan (2001)	9	1	104	1	0	0.82	0.27	0.47	0.77	1.27
Eling & Faust (2010)	30	1	152	6	0	0.47	0.91	-1.35	0.48	2.20
Frydenberg *et al.* (2017)	33	1	254	1	0	0.14	0.24	-0.51	0.19	0.72
Fung *et al.* (2002)	9	1	84	1	0	0.36	0.08	0.19	0.37	0.46
Fung & Hsieh (2004a)	22	4	108	1	0	0.59	0.23	0.19	0.58	0.97
Fung & Hsieh (2004b)	8	2	108	4	0	0.97	0.15	0.73	1.02	1.13
Fung *et al.* (2008)	3	3	57	1	1	0.36	0.49	0.06	0.09	0.93

(continued on next page)

Table 2.1: Primary studies (continued)

Study	#Alphas	#Sources	#Months	#Models	Top5	Mean	StDev	Min	Md	Max
Gupta *et al.* (2003)	42	1	152	2	0	0.73	0.35	0.06	0.68	1.64
Hong (2014)	6	1	65	3	0	0.15	0.19	-0.10	0.17	0.42
Huang *et al.* (2017)	5	1	174	1	0	-0.07	0.07	-0.16	-0.03	-0.01
Ibbotson *et al.* (2011)	1	1	180	1	0	0.41		0.41	0.41	0.41
Jame (2018)	10	2	212	2	0	0.20	0.14	0.04	0.14	0.52
Joenvaara *et al.* (2019b)	5	5	228	1	1	0.23	0.12	0.12	0.16	0.39
Joenvaara & Kosowski (2021)	17	7	120	5	1	-0.37	0.26	-0.74	-0.36	0.27
Jordan & Simlai (2011)	12	1	180	1	0	0.57	0.20	0.25	0.58	0.97
Jylha *et al.* (2014)	1	1	216	1	1	0.29		0.29	0.29	0.29
Kanuri (2020)	2	1	225	2	0	0.23	0.01	0.22	0.23	0.24
Klein *et al.* (2015)	30	4	108	1	0	0.34	0.34	-0.30	0.30	1.00
Kooli & Stetsyuk (2021)	2	1	289	2	0	-0.51	0.19	-0.64	-0.51	-0.37
Kosowski *et al.* (2007)	9	5	108	2	1	0.42	0.07	0.27	0.42	0.51
Kotkatvuori-Ornberg *et al.* (2011)	6	1	175	1	0	0.22	0.38	-0.39	0.20	0.64
Liang (2004)	20	1	36	5	0	1.11	0.77	-0.04	0.87	3.50
Ling *et al.* (2015)	3	1	67	2	0	0.79	0.65	0.18	0.71	1.47
Lo (2001)	11	1	47	1	0	1.22	0.70	0.04	1.25	2.64
Malladi (2020)	1	1	155	1	0	0.35		0.35	0.35	0.35
Meligkotsidou & Vrontos (2008)	20	2	143	1	0	0.37	0.34	-0.42	0.39	0.98
Mitchell & Pulvino (2001)	1	1	108	1	1	0.61		0.61	0.61	0.61
Mladina (2015)	10	1	240	1	0	0.19	0.14	-0.06	0.18	0.36
Molyboga & L'Ahelec (2016)	2	1	228	1	0	0.19	0.35	-0.06	0.19	0.43
Mozes (2013)	10	1	234	5	0	0.20	0.16	-0.13	0.27	0.34
Patton & Ramadorai (2013)	23	4	186	4	1	0.26	0.22	-0.39	0.33	0.63
Racicot & Théoret (2009)	47	1	108	1	0	0.57	0.83	0.07	0.36	5.11
Racicot & Théoret (2013)	5	1	183	1	0	0.05	0.20	-0.19	0.03	0.31
Racicot & Théoret (2014)	24	1	183	3	0	-0.07	0.85	-1.65	-0.07	1.22
Ranaldo & Favre (2005)	42	1	152	3	0	0.71	0.26	0.24	0.65	1.14
Diez De Los Rios & Garcia (2011)	11	1	99	1	0	0.35	0.09	0.20	0.36	0.53
Rzakhanov & Jetley (2019)	7	1	264	1	0	0.36	0.21	0.15	0.32	0.71
Sabbaghi (2012)	20	1	203	1	0	0.18	0.35	-0.70	0.25	0.70
Sadka (2010)	11	1	180	1	1	0.42	0.12	0.23	0.46	0.62
Sadka (2012)	12	1	192	1	0	0.33	0.13	0.09	0.35	0.56

(continued on next page)

Table 2.1: Primary studies (continued)

Study	#Alphas	#Sources	#Months	#Models	Top5	Mean	StDev	Min	Md	Max
Sandvik *et al.* (2011)	36	1	182	1	0	0.13	0.25	-0.31	0.09	0.82
Stafylas *et al.* (2018)	40	2	291	1	0	0.39	0.52	-1.18	0.39	1.58
Stafylas & Andrikopoulos (2020)	24	2	291	1	0	0.51	0.43	0.04	0.37	1.71
Stoforos *et al.* (2017)	26	1	237	1	0	-0.22	0.92	-3.64	0.21	0.58
Sullivan (2021)	28	2	312	5	0	0.26	0.18	-0.02	0.23	0.62
Sun *et al.* (2012)	10	1	192	1	1	0.47	0.12	0.31	0.48	0.62
Teo (2009)	5	3	78	2	1	0.59	0.22	0.33	0.50	0.86
Vrontos *et al.* (2008)	5	1	144	5	0	0.50	0.02	0.47	0.51	0.52
Total #Alphas	1019									
Total #Studies	74									
Mean	13.77	1.70	168.99	1.95	0.30	0.37	0.28	-0.08	0.36	0.86
StDev	13.16	1.32	77.41	1.45	0.46	0.31	0.23	0.66	0.29	0.81
Min	1.00	1.00	31.00	1.00	0.00	-0.51	0.01	-3.64	-0.51	-0.37
Md	9.00	1.00	171.00	1.00	0.00	0.35	0.20	0.06	0.36	0.62
Max	61.00	7.00	475.00	7.00	1.00	1.22	0.92	1.18	1.25	5.11

Notes: The table shows descriptive statistics for the alpha coefficients reported in the individual primary studies (sorted alphabetically). *#Alphas* denotes the number of alpha estimates that meet our sample collection criteria reported in a given study. *#Sources* shows the number of databases the alpha estimates in the primary study are based on. *#Months* specifies the number of months of the data sample used in the primary study covers. *#Models* shows the number of risk models used to estimate the alphas in the primary study. *Top5* is the indicator variable equal to 1 if the study is published in the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*, the *Journal of Financial and Quantitative Analysis*, and the *Review of Finance*, and 0 otherwise. The *Mean*, *StDev*, *Min*, *Md*, and *Max* refer to the mean value, the standard deviation, the minimum, the median value, and the maximum of the alpha coefficients reported in the primary study.

2.4 Full Sample Results

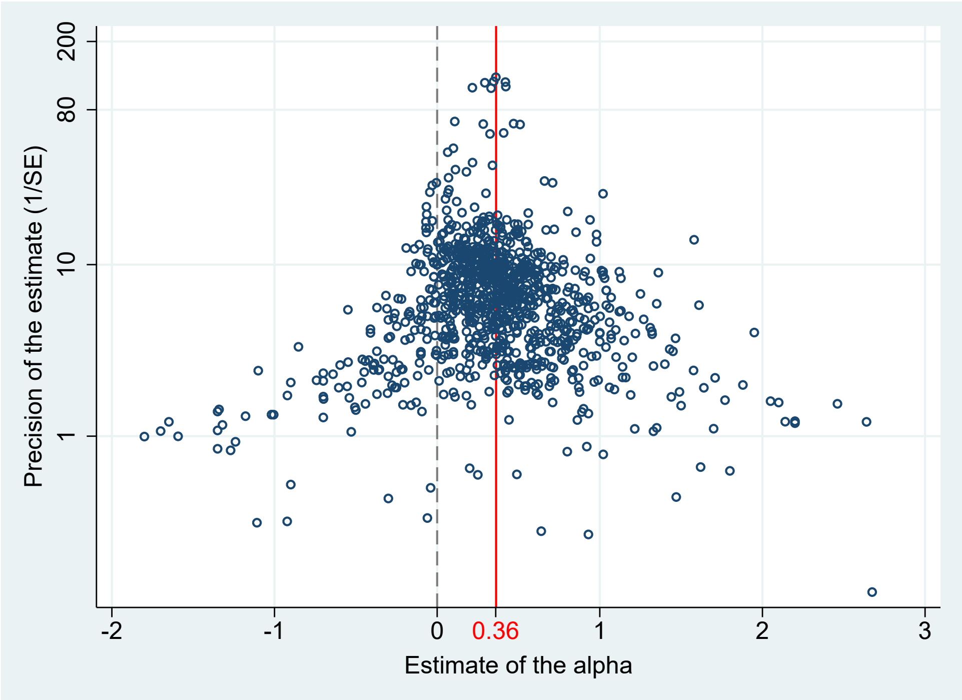
2.4.1 Funnel Plot

Having observed heterogeneity in reported alpha coefficients across various studies we now analyze whether these estimates are affected by publication selection bias. We start this analysis by visualizing how the alpha estimates reported in primary studies depend on their precision, which is defined as one over the estimate's standard error. Tests of statistical significance based on the t-distribution assume that the estimated coefficients and their standard errors are not correlated. Hence, in the absence of publication bias, there should be no systematic relationship between the alpha coefficient and its standard error. In contrast, detecting a positive or a negative association between the coefficients and standard errors suggests selective publication (Stanley 2005; Havranek & Irsova 2010; Havranek 2010; Havranek & Rusnak 2013). The authors of primary studies usually report t-statistics for their estimates, which implies that they assume that the estimates and their standard errors are statistically independent and the ratio of the estimates to their standard errors has a t-distribution. The association between the coefficient and its standard error can thus be used to detect selective publication.

In our setting the association can be depicted with a funnel plot with the alpha coefficients on the x-axis and their precision (i.e. $1/SE$) on the y-axis. We show such a funnel plot in Figure 2.4. In a bias-free world, the graph should resemble a symmetrical inverted funnel. The funnel shape arises because the most precise estimates tend to be concentrated around the underlying mean value, whereas less precise estimates with larger standard errors are more dispersed around the mean. The funnel plot shall be symmetric if for any given level of estimate precision both high and low estimates are equally likely to be published. contrast, if imprecise estimates that are high tend to be reported, while equally imprecise estimates that are low get discarded then the funnel plot shall miss some observations in the left part and consequently, it shall be asymmetric. An asymmetric funnel plot indicates that estimates are reported selectively in primary studies, which implies that their mean value provides a biased estimate of the underlying mean value in the population.

Figure 2.4 exhibits no obvious asymmetry, which is consistent with little or no publication bias. For any given level of precision both high and low estimates seem to be represented in the plot. The funnel plot thus provides initial suggestive evidence indicating that hedge fund alpha estimates reported in primary studies are not significantly contaminated by publication selection bias. Furthermore, a simple visual examination of Figure 2.4 suggests that the funnel plot is slightly "hollow", which might suggest that insignificant estimates (low precision alpha close to zero) are less likely to be published. Below we formally test for the significance of these observed patterns.

Figure 2.4: Funnel plot of alphas



Note: When there is no publication bias, estimates should be symmetrically distributed around the mean denoted by the solid red vertical line. The gray dashed line denotes 0. Outliers are excluded from the figure for ease of exposition but included in all statistical tests.

2.4.2 Formal Tests

Having provided preliminary evidence about the likelihood of a publication selection bias in hedge fund performance literature we proceed with using several approaches to formally test for it. The first set of tests exploits the above-mentioned association between the alpha coefficients reported in primary studies and their standard errors. Since we use the term "alpha" to refer to the intercept term in the regression of returns on risk factors reported in primary studies, we use "kappa" to denote the constant (i.e. the intercept term) in our regressions of alpha coefficients on their standard errors. Furthermore, we use "lambda" to refer to the slope coefficient at the explanatory variable of SE. We estimate the following equation:

$$\alpha_{ij} = \kappa + \lambda \cdot SE(\alpha_{ij}) + \epsilon_{ij}, \tag{2.2}$$

where α_{ij} stands for the i-th estimate of hedge fund alpha reported in the j-th study, $SE(\alpha_{ij})$ denotes its standard error, and ϵ_{ij} is the error term.

In the absence of any publication bias the slope coefficient λ is expected to be zero, which implies no association between the alpha estimates (α_{ij}) and their standard errors ($SE(\alpha_{ij})$). In contrast, if the publication of estimated alpha coefficients is selective and low alpha estimates are more likely to remain unreported in primary studies then imprecise estimates (i.e. those with a large SE) should be more likely to be high rather than low leading to a positive λ coefficient. Conversely, a tendency to discard high rather than low alpha coefficients would lead to a negative λ coefficient. Hence, the slope coefficient λ reflects the effect of publication selection bias and the intercept term κ captures the true mean alpha estimate corrected for the bias.

Panel A of Table 2.2 shows the results for several alternative ways of estimating Equation 2.2. In the first column, we report the conventional ordinary least squares (OLS) estimate. As discussed above the OLS estimate represents the most straightforward way of testing for selective publication that is commonly used in prior research. However, it could yield spurious results in case unobserved features of the primary study design are correlated with the reported alphas. To address this potential problem we complement the OLS estimates with several alternative estimation techniques. The results reported in the second column of Table 2.2 are based on an estimation that includes study-level fixed effects (FE). Including study-level fixed effects filters out idiosyncratic study-level variation. Hence, as long as alpha estimates in a given primary study are estimated using similar methodologies including study-level fixed effects removes the potential confounding effect of these methodological choices on the reported alpha estimates. Identification of the fixed-effect estimator rests on studies that report more than one estimate. Thus we complement the analysis with study-level between-effect estimation (BE) that accounts for the differences in study size. We report these results in the third column of Table 2.2.

To further address the issue of potential endogeneity in the method choices and reported standard errors in the primary studies we follow Stanley (2005), Bajzik *et al.* (2020), Cazachevici *et al.* (2020), Matousek *et al.* (2022), Havranek *et al.* (2023), Ehrenbergerova *et al.* (2023), and Irsova *et al.* (2023) and use the inverse of the square root of the number of observations in primary studies as an IV for the standard error. This measure has the desirable characteristics of a valid instrument. By construction, the number of observations is correlated with the standard error. At the same time, it is plausibly unrelated to the chosen estimation technique. Furthermore, it seems reasonable to assume that the number of observations is quasi-randomly distributed among the primary studies. The results based on this instrument thus constitute an important robustness check. We report these results in the fourth column in Table 2.2.

In the last two columns of Panel A of Table 2.2 we report our weighted least squares estimates of Equation 2.2. In the fifth column, we weigh the observations by the inverse of their squared standard error (WLS). This approach gives less weight to less precise estimates, which helps to adjust for potential heteroskedasticity in our observations. The sixth column shows our results from estimation when the observations are weighted by the inverse of the number of estimates reported in a given study (wNOBS). This approach provides a more comparable basis for larger and smaller studies.

Table 2.2: Full sample results

Panel A: Linear models	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	-0.0152	-0.0265	0.0602	0.178	0.324	0.0497
	(0.188)	(0.215)	(0.131)	(0.353)	(0.320)	(0.127)
	[-0.534, 0.455]			[-0.526, 0.971]	[-0.415, 1.120]	[-0.348, 0.457]
				{-0.626, 0.983}		
Effect beyond bias (κ)	0.366***	0.369***	0.350***	0.316***	0.301***	0.353***
	(0.0426)	(0.0540)	(0.0474)	(0.0854)	(0.0440)	(0.0380)
	[0.277, 0.458]			[0.157, 0.475]	[0.186, 0.412]	[0.270, 0.436]
First-stage robust F-stat				12.71		
Studies	74	74	74	73	74	74
Observations	1,019	1,019	1,019	979	1,019	1,019

Panel B: Non-linear models	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				0.183*	P = 0.631	L = 0.364
				(0.106)	(0.092)	(p = *0.834*)
Effect beyond bias	0.310***	0.325***	0.355***	0.320***	0.274***	0.386***
	(0.026)	(0.009)	(0.093)	(0.008)	(0.03)	(0.045)
Studies	74	74	74	74	74	74
Observations	1,019	1,019	1,019	1,019	1,019	1,019

Notes: The first two panels report the results of the regression $\alpha_{ij} = \kappa + \lambda \cdot SE(\alpha_{ij}) + \epsilon_{ij}$, where α_{ij} denotes the i-th alpha coefficient estimated in the j-th study, and $SE(\alpha_{ij})$ denotes its standard error. FE: study-level fixed effects, BE: study-level between effects, IV: the inverse of the square root of the number of observations is used as an instrument for the standard error, WLS: model is weighted by the inverse of the squared standard error of an estimate, wNOBS: model is weighted by the inverse of the number of estimates per study. In Panel B, Top10 is model by Stanley *et al.* (2010), WAAP stands for Weighted Average of the Adequately Powered model by Ioannidis *et al.* (2017), Kinked-meta is endogenous kink model by Bom & Rachinger (2019), Stem model is by Furukawa (2020), selection is model by Andrews & Kasy (2019) using clustered SEs, P denotes the probability that estimates insignificant at the 5% level are published relative to the probability that significant estimates are published (normalized at 1), p-uni* is by van Aert & van Assen (2020), L denotes test statistic of p-uniform's publication bias test. Standard errors, clustered at the study level, are reported in parentheses. 95% confidence intervals from wild bootstrap in square brackets (Roodman *et al.* 2018). In curly brackets we show the two-step weak-instrument-robust 95% confidence interval based on Andrews (2018) and Sun (2018).* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Considering the above-discussed results reported in Panel A of Table 2.2 we find little evidence of publication selection bias for the alpha estimates reported in the primary studies. The λ coefficients that capture the effect of selective publication are all small in magnitude and statistically insignificant. These findings are remarkably consistent across the alternative ways of estimating Equation 2.2. Thus, consistent with our preliminary findings based on the funnel plot in Figure 2.4, our formal tests provide evidence consistent with the non-selective publication of estimated monthly alphas in the primary studies in our sample.

Panel A of Table 2.2 also shows the κ estimates that reflect the estimated magnitude of the monthly alphas adjusted for the publication selection bias. We observe that these estimates range between 0.301 and 0.369 and they are strongly statistically significant at a better than 1 percent level in all specifications. It is noteworthy that the unconditional mean of monthly alpha estimates of 0.36 highlighted in the histogram in Figure 2.3 falls within this range of estimates corrected for the publication selection bias. Considering perhaps the most conservative estimate reported in Panel A of Table 2.2 we observe that the lower bound of the bootstrapped confidence interval of the IV specification is 0.157. This suggests that the true hedge fund alpha is unlikely to be below 1.9% per annum (0.157*12). These results further strengthen our earlier conclusion that the alpha estimates in our sample are not contaminated by selective publication and hedge funds do earn positive alphas for their investors.

Our estimates of Equation 2.2 reported in Panel A of Table 2.2 are subject to several limitations. First, these tests of selective publication are based on an assumption of a linear relationship between the estimate and its standard error. In reality, this association may not be linear. For example, it may exhibit discontinuities around conventional levels of statistical significance, i.e. when the t-statistics approaches 1.96. Second, the IV specification may not fully remedy the endogeneity problem because it may arise for reasons other than the bias due to omitted variables related to the research design in primary studies. Gechert *et al.* (2022) point out that endogeneity may arise even when deliberately reporting spuriously precise estimates, for example, due to reverse causality. Furthermore, since the standard error is itself an estimate endogeneity can also manifest itself through the measurement error. We address these shortcomings in Panel B of Table 2.2. To address the first issue, we use nonlinear techniques for publication bias correction. To address the second limitation, we use the p-uniform* approach recently developed by van Aert & van Assen (2020) that does not rely on the assumption of exogeneity.

In the first column of Panel B of Table 2.2 we report results based on the Top10 method proposed by Stanley *et al.* (2010). The method is based on a simple proposition that the bias arising from aggregating potentially selectively reported coefficients can be addressed by simply considering only the 10% most precise estimates. The second column of Panel B of Table 2.2 shows results based on the Weighted Average of the Adequately Powered (WAAP) model proposed by Ioannidis *et al.* (2017). Similarly to Top10 also WAAP is based on averaging only a subset of published coefficients. Ioannidis *et al.* (2017) examine the statistical power of the results published in the field of economics and they propose dropping all estimates with statistical power lower than 80% and weighting the remaining estimates by the inverse of their variance. In the third column of Panel B of Table 2.2 we report the results from the Stem-based method recently developed by Furukawa (2020). The stem-based method builds on Stanley *et al.* (2010) but it aims at limiting the loss of sample variation that results from discarding 90% of the less precise estimates. Furukawa (2020) optimizes the trade-off between the bias and variance, discards only the estimates that do not add value in the light of this trade-off, and uses the remaining estimates to compute the average value. The first three columns in Panel B of Table 2.2 show estimates ranging from 0.310 to 0.355, which falls within the range documented earlier for the linear methods reported in Panel A (0.301 to 0.369). Thus, even based on these alternative methods we reach a similar conclusion on the limited impact of publication bias and on the values of alpha estimates corrected for a potential publication bias.

In the fourth column of Panel B of Table 2.2 we report results based on the endogenous kink model (Kinked-meta) proposed by Bom & Rachinger (2019). The model is based on the assumption that the relationship between an estimate and its standard error is only linear to some point because for some levels of reported coefficients, there is no reason to expect the presence of publication bias. Hence, there is an endogenously determined cut-off value (or a "kink") at which the relationship changes. The Kinked-meta model yields some weak evidence (significant at 10% level) on the presence of selective publication (λ 0.183, SE 0.106). Nevertheless, even this approach yields a κ coefficient of 0.320, which is very close to the uncorrected mean of 0.36 and comfortably within the interval of 0.301 and 0.369 shown in Panel A.

The fifth column of Panel B of Table 2.2 shows our results for the selection model recently developed by Andrews & Kasy (2019). The model is based on the assumption that the probability of publishing an estimate depends on its statistical significance. The model identifies how likely it is for an estimate to fall into different intervals determined by the critical values of t-statistics. The model gives more weight to intervals that are underrepresented. Our results from the selection model suggest that statistically insignificant estimates may be somewhat less likely to get published than statistically significant estimates (63% probability vs. 100% probability). However, the corrected mean of alpha estimates decreases only slightly to 0.274. Hence, even this methodological approach does not suggest that inferences about the magnitude of alpha coefficients are greatly affected by selective publication.

The Andrews & Kasy (2019) model relies on several assumptions. It requires the estimates and their standard errors to be statistically independent. It also assumes that the probability of publication is the same for all estimates in a given interval. We test these assumptions in Table A1 using the Kranz & Putz (2022) framework. These tests suggest some of the underlying assumptions of the selection model (especially the independence assumption) may be violated in many of our samples. Therefore, as robustness checks, we also use models that do not rely on the underlying assumption of no correlation between the estimates and their standard errors in the absence of publication bias.

The last column in Panel B of Table 2.2 shows the results of the p-uniform* model by van Aert & van Assen (2020). Harvey *et al.* (2016) and Harvey (2017) suggest that "p-hacking", i.e. a greater tendency to publish statistically significant rather than insignificant results, is a major problem in financial economics research. They argue that "p-hacking" may have led to a number of "false positives" to be reported in published research on factors explaining the cross-sectional variation in realized returns. This introduces important distortions in modeling systematic risk in asset pricing because the association between many of these variables and realized returns is unlikely to persist in the future. Following this argument, we examine potential "p-hacking" in research on hedge fund performance. To do so, we use the p-uniform* model recently proposed by van Aert & van Assen (2020) that is based on evaluating the distribution of p-values around the 5% cut-off level that is conventionally used to assess statistical significance. A tendency to publish statistically significant results implies an over-representation of p-values just below the 5% cut-off and an under-representation of p-values just above it. P-uniform* corrects for this potential bias by assigning different weights to estimates of various degrees of statistical significance based on the estimated publication probability. This selection model is robust to the assumption of zero correlation between estimates and standard errors in the absence of any publication bias. Our results based on this methodological approach are consistent with our previous findings. The test statistic for the publication bias (denoted "L") is statistically insignificant, which again suggests that the publication of alpha estimates in primary studies is not selective. In fact, this method suggests a somewhat higher value of 0.386 for alpha estimates corrected for the publication bias. Thus, we reach similar conclusions about the absence of publication selection bias and a somewhat similar estimate of the true mean value of the alpha coefficient even when using the p-uniform* method, which does not require the exogeneity assumption for the standard errors to be satisfied.

In contrast to the more conventional linear approaches reported in Panel A of Table 2.2, the more sophisticated non-linear approaches shown in Panel B of Table 2.2 do not require the linearity and exogeneity assumptions to be met. Overall, these approaches lead to fairly similar conclusions about the limited impact of selective publication on the alpha estimates reported in primary studies. Only the Kinked-meta model shows some marginally significant evidence of publication selection bias. However, even this approach does not dramatically alter the estimated value of alpha coefficients corrected for the publication bias. Furthermore, the selection model suggests that statistically insignificant estimates may be somewhat less likely to get published. However, the estimate for the mean alpha coefficient does not dramatically change after correcting for this bias. The interval of corrected alpha estimates based on the more sophisticated approaches reported in Panel B is slightly wider and it ranges from 0.274 to 0.386. However, both the upper bound and the lower bound of this interval are fairly close to the unconditional mean of 0.36. These results thus provide further support for our conclusion that inferences about the magnitude of the alpha coefficient in the literature on hedge fund performance are not significantly affected by publication selection bias.

To complement our analysis we examine whether our results may be affected by the inclusion of studies that are coauthored by similar research teams. We acknowledge that different teams of co-authors may plausibly have various preferences over the choice of estimation methodology and they may have access to different data sources. Hence, the alpha coefficients estimated by members of a given research team may be interdependent.[6] To address this issue, we recompute our results and we cluster the standard errors by author teams. We report these results in Table A2. These results show that potential inter-dependencies between the alphas estimated in individual research teams do not materially affect our conclusions. Consistent with our main results, we find little evidence of publication selection bias after clustering the standard errors at the research team level. All the λ coefficients reported in Table A2 are statistically insignificant, which suggests against selective publication. Furthermore, the κ coefficients fall within a fairly narrow interval of (0.301, 0.369) that is fully subsumed by the corresponding interval that we observe for our main results (0.274, 0.386). We thus conclude that clustering of standard errors at the level of a research team leads us to similar conclusions on the absence of selective publication and on the magnitude of the "corrected" alpha estimate as our main results.

Finally, we follow recent advances in econometrics and conduct a test of "p-hacking" based on Elliott *et al.* (2022). We report the results in Figure A1 and Table A3. The "p-hacking" tests are conceptually different from the publication bias tests. Therefore, they constitute a good complement to the results reported earlier in this section. Figure A1 shows no obvious breaks at the value of 1.96 (represented by the vertical solid red line), which represents the most important threshold for statistical significance at 5% level. Similarly, our formal test show no indication of over-reporting of statistically significant estimates in primary studies (see Table A3). Consistent with our earlier conclusions, these results also suggest that the pool of empirical evidence on hedge fund performance is not substantially contaminated by selective reporting of estimates in research journals.

We find these results remarkable, especially when contrasted with the abundant empirical evidence on the prevalence of publication selection bias in a multitude of other settings in economics and finance, e.g. Stanley (2001; 2005); Stanley & Doucouliagos (2010); Havranek (2015); Brodeur *et al.* (2016); Bruns & Ioannidis (2016); Stanley & Doucouliagos (2017); Christensen & Miguel (2018); Brodeur *et al.* (2020); Blanco-Perez & Brodeur (2020), and Zigraiova *et al.* (2021). We can only speculate about the underlying reasons why we do not observe selective publication in research on hedge fund performance. We consider it likely that the presence of two opposing perspectives, both of which may be quite plausible, limits researchers' and editors' incentives to systemically discard either high or low estimates of hedge fund performance. On the one hand, hedge funds likely employ very talented wealth managers who are highly incentivized to generate returns for investors. Hence, it may be reasonable to expect that these bright and highly motivated minds are capable of identifying assets that are temporarily mispriced due to investor irrationality or the impact of passive investment. Hedge funds can possibly earn abnormal returns by investing in these assets, which implies positive alpha

[6] We are grateful to an anonymous referee for pointing this out.

coefficients. On the one hand, following the EMH, hedge funds mostly trade on competitive markets where it may be challenging to systematically earn more than the "normal" rate of return. In addition, the high management and performance fees that hedge funds charge may imply that their net-of-fee performance may be inferior to passive indexing. This would imply either insignificant or negative alphas. We consider it plausible that the lack of a clear *a priori* theoretical prediction about the expected sign of estimated coefficients may limit the incentives for selective reporting and increase the readiness of academic journals to publish both positive and insignificant or negative results on hedge fund performance.

To further strengthen our analysis, we perform several robustness checks intended to ensure that our results are not driven by the heterogeneity in the mix of various alpha coefficients estimated in the primary studies using a wide range of techniques. Heterogeneity in estimation may potentially lead to offsetting biases that would compromise our ability to detect selective publication in the full sample. For example, it is acknowledged that the p-uniform* method tends to overestimate the measured effect when large heterogeneity is present among the estimates collected from the primary studies (Carter *et al.* 2019). To further strengthen the confidence in our findings and to rule out the possibility that our tests are adversely affected by the diversity of the techniques used in estimating the alpha coefficients in the primary studies we proceed by analyzing more homogeneous subsets of alpha estimates to determine whether selective publication can be observed in any of these subsamples.

2.5 Subsample Results

In this section, we report our results for various subsamples of our dataset. Our data contain alphas estimated using different data sources and estimation techniques. We argue that it is important to aggregate these different estimates and report the representative alphas, as all estimates reported in primary studies are likely to help shape researchers' and investors' views of the abnormal returns that hedge funds earn on average. We assume that different estimation approaches are used in the literature to date because opinions differ about their relative appropriateness. We assume that readers of the research literature on hedge fund performance have differing views on these techniques and rely most heavily on the alphas estimated using the methods they believe to be the best. Our analysis of the full sample recognizes that none of the approaches is universally superior and takes into account all of the estimates that are likely to shape researchers' and practitioners' views on the subject. Despite this advantage, the overall results may be affected by the diversity of the underlying data and their impact on the power of our test. To ensure that our results are not affected by the underlying heterogeneity of the data, we re-estimate our regressions for more narrowly defined and thus more homogeneous subsamples. We observe whether the results of our subsample are consistent with the results of the whole sample.

We consider several subsamples of more homogeneous alpha estimates. First, we partition our sample based on whether the survivorship and/or backfilling biases are adjusted for in a given primary study. Since these data biases may potentially have a significant impact on the documented returns estimating our regressions for the two subsamples separately lets us draw stronger inferences from our results. Second, we consider alpha coefficients estimated using two commonly used risk models: (i) the one-factor model, and (ii) the seven-factor model. Hedge funds exhibit unusual risk exposures to various risk dimensions and so the choice of a risk model may have an impact on the estimated abnormal return. Third, we recompute our results for the subsample of alpha coefficients estimated with the use of IV. Prior research shows that IV-based estimates are more likely to suffer from a publication bias because they tend to have larger standard errors (Brodeur *et al.* 2020). We examine whether we detect selective publication in this subset of estimates. Fourth, we report our results for the subsample of alpha estimates published in the Top 3 (the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*), or the Top 5 (plus the *Journal of Financial and Quantitative Analysis*, and the *Review of Finance*) finance journals. These journals are particularly prominent in the field of finance, and so the alpha estimates they publish are likely to be particularly impactful in shaping researchers' views of hedge funds performance. In addition, many researchers are strongly incentivized to publish their research in these leading journals. This implies that if the publication of hedge fund performance is selective, the bias is likely to be particularly strong in these journals.

To pave the way for computing our subsample results, we visualize in the Appendix the distribution of the alpha coefficients in these subsamples. Figure A2 shows the histograms and Figure A3 shows the funnel plots for the individual subsamples we analyze in this section. In the histograms, we do not observe any major deviations from normality. In a similar vein, with the exception of alpha estimates estimated with the use of IV-based methods, the funnel plots are rather symmetric, which points towards little publication selection bias. In the following subsections, we formally test for selective publication in these subsamples.

2.5.1 Survivorship and Backfilling Biases

Prior research has long argued that the survivorship and backfilling biases may have a substantial impact on hedge fund performance estimates (Fung & Hsieh 2004a; Aggarwal & Jorion 2010; Kosowski *et al.* 2007). The survivorship bias arises when a data sample excludes performance results of funds that are no longer in existence. From the perspective of data providers, excluding these funds from their database is sensible because funds that no longer operate are not interesting for investors anymore. Nevertheless, since the performance of funds that stop reporting information on their performance to the database may systematically differ from the performance of surviving funds purging this information biases the research results based on the database. The backfilling bias arises when funds undergo an "incubation period" intended to accumulate performance track record before they are offered to investors. If performance history is backfilled into the database only for those funds that succeed in the incubation period the database overstates the performance of the entire hedge fund population in the early years of their existence.

In the following analysis we consider separately a subsample of alpha estimates that explicitly controls for the survivorship and/or backfilling biases. Then we consider only those alpha estimates that do not adjust for these biases. Given that the survivorship and backfilling biases may have a significant impact on the estimated alpha coefficients considering only one subsample at a time makes the individual alpha estimates more homogenous. We examine whether our main conclusions on the limited publication selection bias are robust to testing these relationships within the two subsamples.

In Table 2.3, we report our results for two subsets of alpha estimates: (i) those adjusted for the survivorship and/or backfilling bias are reported in Part 1, and (ii) those adjusted for neither survival bias nor backfilling bias are reported in Part 2. The results based on the conventional approaches reported in Panel A of Part 1 convey a fairly similar albeit slightly weaker message as the results based on the full sample (reported in Table 2.2). In Panel A of Part 1 all but one λ coefficient are statistically insignificant. Only for the WLS model, which weighs the observations by the inverse of their standard error, we observe a significant λ coefficient of 0.472 (SE 0.233, significant at 5% level). This finding provides limited evidence on some publication bias within the subsets of alpha estimates adjusted for the survivorship and/or backfilling biases weighted by their precision. Furthermore, among the results based on the non-linear approaches reported in Panel B of Part 1 we observe one statistically significant λ for the Kinked-meta model of 0.519 (SE 0.234, significant at 1% level). Similarly, to the WLS model also Kinked-meta attributes different weights to alpha estimates based on their precision. However, overall, we find only limited evidence of selective publication of alpha coefficients within the subsample of estimates that adjust for the survivorship and/or backfilling biases.

Results presented in Part 1 of Table 2.3 also indicate that limiting the analysis to the subsample of estimates adjusted for the survivorship and/or backfilling biases does not dramatically affect the conclusions about the magnitude of the alpha coefficients. The κ coefficients reported in Part 1 reflect the average alpha coefficients adjusted for the publication selection bias range between 0.194 and 0.351. This range is only slightly lower than the corresponding interval for κ coefficients based on the full sample between 0.274 and 0.386 reported in Table 2.2. The most significant deviation from this pattern is

Table 2.3: Survivorship and backfiling biases

Part 1: Survivorship and/or Backfiling Bias Treated						
Panel A: linear	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	-0.00148	-0.0846	0.0795	0.493	0.472**	0.0814
	(0.180)	(0.0961)	(0.189)	(0.402)	(0.233)	(0.195)
	[-0.435, 0.568]			[-0.492, 2.534]	[-0.092, 0.948]	[-0.406, 0.561]
				{-0.342, 1.566}		
Effect beyond bias (κ)	0.329***	0.351***	0.300***	0.194*	0.241***	0.300***
	(0.0373)	(0.0252)	(0.0622)	(0.106)	(0.0337)	(0.0487)
	[0.243, 0.411]			[-0.104, 0.392]	[0.133, 0.377]	[0.183, 0.403]
First-stage robust F-stat				11.29		
Studies	50	50	50	49	50	50
Observations	605	605	605	565	605	605
Panel B: non-linear	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				0.519***	P = 0.632	NA
				(0.125)	(0.115)	(NA)
Effect beyond bias	0.267***	0.248***	0.220***	0.234***	0.262***	0.325***
	(0.028)	(0.017)	(0.068)	(0.012)	(0.029)	(0.048)
Studies	50	50	50	50	50	50
Observations	605	605	605	605	605	605
Part 2: Survivorship and Backfiling Biases Untreated						
Panel A: linear	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	-0.0168	0.0491	0.0602	-0.466	0.438	0.0337
	(0.363)	(0.486)	(0.170)	(0.725)	(0.492)	(0.140)
	[-0.941, 0.870]			[NA]	[-0.558, 1.749]	[-0.844, 0.671]
				{-5.850, NA}		
Effect beyond bias (κ)	0.416***	0.400***	0.436***	0.521***	0.334***	0.443***
	(0.0782)	(0.114)	(0.0686)	(0.163)	(0.0559)	(0.0563)
	[0.245, 0.615]			[NA]	[0.112, 0.482]	[0.326, 0.554]
First-stage robust F-stat				3.76		
Studies	29	29	29	29	29	29
Observations	414	414	414	414	414	414
Panel B: non-linear	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				0.343*	P = 0.719	NA
				(0.191)	(0.124)	(NA)
Effect beyond bias	0.301***	0.359***	0.331***	0.351***	0.282***	0.507***
	(0.038)	(0.012)	(0.040)	(0.009)	(0.086)	(0.077)
Studies	29	29	29	29	29	29
Observations	414	414	414	414	414	414

Notes: Part 1: Sample in which both biases are treated for (either the survivorship or the backfiling bias is treated or both biases are treated for). Part 2: Sample in which biases are not treated for (neither the survivorship nor the backfiling bias is treated for). NA = non-convergence to the result. [NA] = confidence interval could not be bounded. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the slightly lower and only marginally significant κ coefficient based on the IV estimate that uses the inverse of the square root of the number of observations as an instrument for the standard error. This κ of 0.194 (SE 0.106, significant at 10% level) is reported in the fourth column in Panel A of Part 1.

In Part 2 of Table 2.3 we report results based on the subsample of alpha estimates that do not explicitly control neither for the survivorship nor for the backfilling biases.

The conclusion based on this subsample is very similar to the main results reported in Table 2.2. In line with the full-sample results, the λ coefficients are statistically insignificant with the exception of the one based on the Kinked-meta model, which is equal to 0.343 and similarly to the full-sample result it is marginally significant at 10% (SE 0.191). Furthermore, the κ coefficients reported in Part 2 of Table 2.3 range between 0.282 and 0.521. Relative to the corresponding range for the κ coefficients based on the full sample this range is slightly wider. The difference is mainly driven by the higher upper bound, which is consistent with the proposition that studies that control for survivorship and/or backfilling biases tend to report lower alpha estimates than those that do not. Overall, these findings suggest that the alpha coefficients that are not adjusted for backfilling and survivorship biases are not reported selectively.

2.5.2 Risk Models

One of the key methodological issues in hedge fund performance research concerns the choice of risk models used to adjust for the systematic risk that a given investment strategy involves. These models define the risk dimensions considered relevant for a given investment strategy. Prior hedge fund performance research uses several risk models. Models that feature fewer risk factors (e.g. the CAPM, the three-factor, and the four-factor model) are well-established in general asset pricing and investment research, which implies that the alpha coefficients based on these models are easily comparable with alpha coefficients estimated to evaluate the performance of other types of investments, e.g. mutual funds. On the other hand, hedge funds commonly employ complex investment strategies that may exhibit unusual risk profiles and exposures to risk dimensions that are not essential for conventional asset classes. Thus, standard risk models may not fully capture the exposure of hedge funds' investment strategies to all relevant risk dimensions. More complex risk models featuring additional risk dimensions designed specifically to measure hedge fund performance may thus be more effective in capturing the plurality of risk exposures that hedge fund strategies involve. The choice of a risk model is thus one of the important drivers for the heterogeneity in the alpha coefficients that we collect from primary studies. In evaluating the robustness of our findings to various ways of reducing heterogeneity in our sample we re-estimate our regressions using two subsamples of alpha coefficients estimated based on two frequently used risk models: (i) the one-factor model, and (ii) the seven-factor model.

Part 1 of Table 2.4 shows the results of our tests of selective publication for the subset of alpha coefficients based on the one-factor model. In these tests we include all the alpha estimates that use a single risk factor based on market portfolio returns, i.e. both the estimates that use raw market returns and those that use market returns in excess of the risk-free rate. These methodological modifications are relatively small and so we do not expect them to have a substantial impact on the reported alpha coefficients. In line with our earlier results, we do not find evidence of a significant publication bias for this narrowly defined subsample of alpha coefficients. The λ coefficients that capture the impact of a potential publication bias are all statistically insignificant. This suggests that the alpha estimates based on a one-factor model are not reported selectively in prior literature.

Table 2.4: Risk models

Part 1: One-Factor Model						
Panel A: linear	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	-0.456	-0.328	0.0326	-1.115	0.453	-0.338
	(0.488)	(0.657)	(0.494)	(0.819)	(0.602)	(0.249)
	[-1.649, 1.468]			[NA]	[-0.776, 2.269]	[-1.247, 0.909]
				{-3.629, 0.102}		
Effect beyond bias (κ)	0.562***	0.534***	0.411***	0.707***	0.404***	0.482***
	(0.0447)	(0.145)	(0.119)	(0.175)	(0.0883)	(0.0931)
	[0.465, 0.642]			[0.324, 1.411]	[0.044, 0.515]	[0.252, 0.702]
First-stage robust F-stat				14.47		
Studies	18	18	18	18	18	18
Observations	167	167	167	167	167	167
Panel B: non-linear	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				0.450	P = 0.613	L = 0.188
				(0.347)	(0.177)	(*p = 0.911*)
Effect beyond bias	0.446***	0.454***	0.349**	0.405***	0.426***	0.427***
	(0.072)	(0.030)	(0.163)	(0.028)	(0.088)	(0.103)
Studies	18	18	18	18	18	18
Observations	167	167	167	167	167	167
Part 2: Seven-Factor Model						
Panel A: linear	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	-0.142	-0.0729	0.305*	0.624	0.0683	0.226
	(0.137)	(0.0547)	(0.155)	(0.557)	(0.296)	(0.265)
	[-0.666, 0.555]			[NA]	[-0.571, 0.948]	[-0.732, 0.644]
				{0.0184, NA}		
Effect beyond bias (κ)	0.326***	0.308***	0.200**	0.128	0.284***	0.222***
	(0.0392)	(0.0141)	(0.0730)	(0.150)	(0.0330)	(0.0641)
	[0.239, 0.413]			[NA]	[0.132, 0.361]	[0.073, 0.375]
First-stage robust F-stat				3.41		
Studies	33	33	33	33	33	33
Observations	298	298	298	298	298	298
Panel B: non-linear	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				0.019	P = 0.900	L = 0.269
				(0.173)	(0.212)	(*p = 0.874*)
Effect beyond bias	0.229***	0.297***	0.325***	0.298***	0.302***	0.305***
	(0.036)	(0.013)	(0.059)	(0.008)	(0.059)	(0.040)
Studies	33	33	33	33	33	33
Observations	298	298	298	298	298	298

Notes: Part 1: Sample in which the one-factor model or its modifications are used to estimate the alpha. Part 2: Sample in which the seven-factor model or its modifications are used to estimate the alpha. [NA] = confidence interval could not be bounded. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Part 1 of Table 2.4 also shows that the κ coefficients that reflect the estimated true magnitude of the alpha estimates corrected for the potential bias range from 0.349 to 0.707. This interval includes the unconditional mean of all the alpha estimates in our sample of 0.36. However, these estimates are somewhat higher than the ones we document for the full sample in Table 2.2. This may suggest that the single market-based risk factor does not fully control for the systematic risk hedge fund strategies involve and so the abnormal return based on the model is higher. Overall, these results provide additional support for the conclusion that the alpha estimates reported in prior literature are not subject to selective publication.

Prior literature acknowledges that the complexity and the dynamic nature of hedge funds' investment strategies may induce exposure to risk dimensions that are not included in conventional risk models. Prior research thus proposes alternative risk models designed specifically for investment strategies common in hedge funds. The most notable example of these models is the seven-factor model (Fung & Hsieh 2004a; Fung *et al.* 2008). The model comprises the following risk factors: (i) the stock market excess return, (ii) the spread between the small capitalization and large capitalization stock returns, the excess return pairs of look-back call and put options (iii) on currency futures, (iv) on commodity futures, and (v) on bond futures, (vi) the duration-adjusted change in the yield spread of the U.S. 10-year Treasury bond over the 3-month T-bill, and (vii) the duration-adjusted change in the credit spread of Moody's BAA bond over the 10-year Treasury bond. These risk factors are intended to capture risk exposures of a broad set of hedge fund types ranging from equity long-short funds to managed futures funds.

In Part 2 of Table 2.4, we report the results of our tests of selective publication for the alpha coefficients based on the seven-factor model. Similarly to the results on the one-factor model reported in Part 1 of Table 2.4, the results in Part 2 of Table 2.4 show little evidence of publication selection bias. The λ coefficients are insignificant with the exception of the between effect estimation (BE) that produces marginally significant λ of 0.305 (SE 0.155, significant at 10% level). Furthermore, the κ coefficients, which reflect the expected value of abnormal returns generated by hedge funds after adjusting for selective publication, range from 0.128 to 0.326, which is lower than the corresponding range in Part 1 of Table 2.4.

The range is also below the unconditional mean of all monthly alpha estimates in our sample of 0.36. This magnitude of the κ coefficients is unlikely to be driven by selective publication.

Taken together, our results do not suggest that the heterogeneity in methodological approaches used for estimating alpha coefficients reported in primary studies is the underlying reason for not detecting any publication bias. We do not observe a significant publication bias even when concentrating on fairly homogeneous subsamples of alphas that are estimated using one of the common risk models.

2.5.3 Instrumental Variables

We further consider a subsample of alphas estimated based on IV, for which selective publication is particularly likely. Prior research shows that IV-based estimates tend to suffer from publication bias more frequently than estimates based on other techniques (Brodeur *et al.* 2020). The authors argue that research methods that offer researchers more degrees of freedom are more likely to suffer from selective publication as researchers may exercise discretion in choosing research designs that help them achieve results that may be viewed as more attractive for publication. The choice of an IV and the specific way of measuring it give researchers considerable leeway. Researchers may choose to report IV-based estimates that are consistent with their prior beliefs or that are otherwise more attractive for publication. Brodeur *et al.* (2020) show that when IVs are relatively weak, the second stage results are likely to be close to the conventional thresholds for statistical significance, which is consistent with selectivity in the process that determines what coefficients eventually get published.

Motivated by this argument recently proposed in research literature we test for selective publication within the subsample of IV-based alpha coefficients. Primary studies typically use higher moments of the distribution of returns, such as skewness and kurtosis, as IV for the excess returns of the mimicking portfolios. This approach follows earlier research that shows that higher moments of the returns distribution are valid instruments and they are effective in removing errors-in-variables problem (Durbin 1954; Pal 1980; Dagenais & Dagenais 1997). We collect 46 IV-based alpha estimates from 3 different studies.

Our results reported in Part 1 of Table 2.5 are consistent with the proposition in prior literature that IV-based estimates tend to exhibit a greater publication selection bias. Five out of seven λ coefficients are positive and statistically significant at 5% level or better. The positive association between reported alphas and their standard errors indicates that highly positive alpha estimates tend to be reported when they are rather imprecise, i.e. they have a large standard error. Such a pattern is characteristic of selective publication. Furthermore, we also observe that for the subsample of IV-based estimates, the magnitude of the κ coefficients that represent the expected value of alpha estimates after adjusting for selective publication is substantially lower than in our main results. The κ coefficients reported in Part 1 of Table 2.5 range from -0.411 to 0.298, many of them are close to 0, and five out of twelve are actually negative. This suggests that after correcting the IV-based alpha estimates for selective publication there is only limited evidence that they actually are positive and statistically significant. In fact, in contrast to our previous results, only two out of twelve κ coefficients are statistically different from zero. Consistent with the *a priori* expectations this evidence suggests that the composition of the pool of published IV-based alpha estimates tends to be affected by selective publication. These findings thus provide one of the first pieces of out-of-sample evidence in support of the recent proposition that IV-based estimates are more likely to suffer from publication bias than estimates based on other techniques Brodeur *et al.* (2020).

Table 2.5: Instrumental variables

Part 1: Methods Using Instrumental Variables						
Panel A: linear	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	1.378***	1.307**	3.514	2.459	2.418***	1.445***
	(0.221)	(0.283)	(6.083)	(2.111)	(0.299)	(0.178)
	[0.760, 1.634]			[NA]	[1.886, 2.945]	[0.776, 1.647]
				{-1.512, 6.430}		
Effect beyond bias (κ)	0.127	0.144	-0.411	-0.137	-0.0601	0.104
	(0.0851)	(0.0691)	(1.525)	(0.438)	(0.0875)	(0.102)
	[-0.102, 0.235]			[NA]	[-0.195, 0.260]	[-0.113, 0.267]
First-stage robust F-stat				155.41		
Studies	3	3	3	3	3	3
Observations	46	46	46	46	46	46
Panel B: non-linear	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				2.418***	P = 0.341	L = 3.551
				(0.431)	(0.103)	(p = *0.169*)
Effect beyond bias	-0.036	0.018	0.078	-0.060	0.231***	0.298***
	(0.027)	(0.067)	(0.091)	(0.048)	(0.088)	(0.071)
Studies	3	3	3	3	3	3
Observations	46	46	46	46	46	46
Part 2: Methods Not Using Instrumental Variables						
Panel A: linear	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	-0.0335	-0.0230	0.0147	0.203	0.284	0.0364
	(0.191)	(0.220)	(0.132)	(0.351)	(0.321)	(0.131)
	[-0.543, 0.442]			[-0.493, 0.999]	[-0.453, 1.080]	[-0.351, 0.458]
				{-0.597, 1.003}		
Effect beyond bias (κ)	0.366***	0.363***	0.358***	0.305***	0.308***	0.353***
	(0.0429)	(0.0553)	(0.0487)	(0.0849)	(0.0440)	(0.0382)
	[0.277, 0.459]			[0.138, 0.469]	[0.192, 0.424]	[0.271, 0.435]
First-stage robust F-stat				12.98		
Studies	74	74	74	73	74	74
Observations	973	973	973	933	973	973
Panel B: non-linear	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				0.166	P = 0.636	L = 0.378
				(0.108)	(0.100)	(p = *0.828*)
Effect beyond bias	0.333***	0.329***	0.355***	0.324***	0.276***	0.388***
	(0.026)	(0.009)	(0.092)	(0.008)	(0.030)	(0.045)
Studies	74	74	74	74	74	74
Observations	973	973	973	973	973	973

Notes: Part 1: Sample where the instrumental variable approach (including 2SLS, GMM, Hasuman) is used for estimation of the alpha. Part 2: Sample where the instrumental variable approach is not used for estimation of the alpha (mostly ordinary least squares). [NA] = confidence interval could not be bounded. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In contrast, our results based on the subsample of the remaining alpha coefficients that are not estimated with the use of IV reported in Part 2 of Table 2.5 are in line with our main results. All λ coefficients are statistically insignificant, which indicates that these alpha estimates are not substantially affected by the publication selection bias. In comparison to the full-sample results, within this subsample, even the λ coefficient based on the Kinked-meta model is statistically insignificant. Furthermore, the κ coefficients fall within a fairly narrow range between 0.276 and 0.388, which is very similar to the full sample result. Taken together, the subsample of alpha coefficients that are not estimated based on IV do not seem to be affected by publication bias and their mean value corrected for any (small) biases are quite close to the unconditional sample mean of 0.36.

2.5.4 Top Journals

In this section we report our results for the subsample of alpha estimates published in the Top 3 (the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*), or the Top 5 (plus the *Journal of Financial and Quantitative Analysis*, and the *Review of Finance*) finance journals. These journals are particularly prominent in the field of finance, and so the alpha estimates they publish are likely to be particularly impactful in shaping researchers' views of hedge funds performance. In addition, many researchers are strongly incentivized to publish their research in these leading journals. This implies that if the publication of hedge fund performance is selective, the bias is likely to be particularly strong in these journals.

Table 2.6 shows our results for the subsamples of alphas collected from the Top 3 (Part 1) and Top 5 (Part 2) finance journals. We observe that these results are mostly consistent with our main findings reported in Table 2.2. In line with our main results, we find little evidence of publication selection bias for the alpha estimates published in top finance journals. Most of the λ coefficients that capture the effect of selective publication are statistically insignificant. Furthermore, some of the λ coefficients are positive and others are negative, which points towards the absence of a systematic tendency to over-report or under-report high estimates of hedge fund performance. Only two λ coefficients are statistically significant at the conventional 5% level. In Part 1, in the model where the observations are weighted by the inverse of the number of estimates per study (wNOBS) the λ coefficient is positive and significant. In contrast, Part 2 shows a significant negative λ coefficient for the model that includes study-level fixed effects (FE). Thus, the only two statistically significant results point in the opposite direction. Consistent with our main results, we conclude that also our results based on subsamples of alpha estimates published in the Top 3 and Top 5 finance journals exhibit little signs of publication selection bias.

We also observe the magnitude of the κ estimates that reflect the estimated magnitude of the monthly alphas adjusted for the publication selection bias. These results are again broadly consistent with our main findings based on the full sample. All the κ coefficients

Table 2.6: Top journals

Part 1: Top 3 journals						
Panel A: linear	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	0.00308	-0.0876*	0.125	0.0711	0.241	0.131***
	(0.104)	(0.0469)	(0.0982)	(0.190)	(0.441)	(0.0363)
	[-0.226, 0.388]			[NA]	[-1.055, 2.069]	[-0.177, 0.594]
				{-0.301, 0.443}		
Effect beyond bias (κ)	0.328***	0.352***	0.358***	0.310***	0.274***	0.356***
	(0.027)	(0.012)	(0.049)	(0.057)	(0.059)	(0.044)
	[0.265, 0.416]			[NA]	[0.086, 0.386]	[0.253, 0.450]
First-stage robust F-stat				2.444		
Studies	16	16	16	16	16	16
Observations	218	218	218	218	218	218
Panel B: non-linear	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				0.113	P = 0.817	L = 0.289
				(0.257)	(0.188)	(*p = 0.866*)
Effect beyond bias	0.265***	0.314***	0.353***	0.314***	0.272***	0.306***
	(0.031)	(0.014)	(0.055)	(0.008)	(0.070)	(0.070)
Studies	16	16	16	16	16	16
Observations	218	218	218	218	218	218
Part 2: Top 5 journals						
Panel A: linear	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	-0.0537	-0.121***	0.0916	-0.00437	0.0795	0.0893
	(0.114)	(0.040)	(0.167)	(0.237)	(0.413)	(0.060)
	[-1.000, 0.182]			[NA]	[-1.003, 1.686]	[-0.517, 0.227]
				{-0.469, 0.460}		
Effect beyond bias (κ)	0.300***	0.317***	0.354***	0.287***	0.279***	0.355***
	(0.042)	(0.011)	(0.074)	(0.075)	(0.053)	(0.060)
	[0.187, 0.397]			[NA]	[0.114, 0.384]	[0.234, 0.481]
First-stage robust F-stat				2.10		
Studies	22	22	22	22	22	22
Observations	256	256	256	256	256	256
Panel B: non-linear	Top10	WAAP	Stem-based	Kinked-meta	Selection model	p-uniform*
Publication bias				-0.071	P = 0.956	L = 0.307
				(0.230)	(0.216)	(*p = 0.858*)
Effect beyond bias	0.263***	0.312***	0.355***	0.313***	0.289***	0.343***
	(0.030)	(0.012)	(0.061)	(0.008)	(0.061)	(0.064)
Studies	22	22	22	22	22	22
Observations	256	256	256	256	256	256

Notes: Part 1: Sample extracted from top 5 journals in finance (*Journal of Finance*, *Journal of Financial Economics*, *Review of Financial Studies*, *Review of Finance*, *Journal of Financial and Quantitative Analysis*). Part 2: Sample extracted from top 3 journals in finance (*Journal of Finance*, *Journal of Financial Economics*, *Review of Financial Studies*). [NA] = confidence interval could not be bounded. $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

fall within a fairly narrow interval of (0.265, 0.358) for the Top 3 journals and (0.263, 0.355) for the Top 5 journals. These intervals for the "corrected" alphas published in the top finance journals greatly overlap with the corresponding interval that we observe for the full sample (0.274, 0.386). In fact, the intervals for the top finance journals are slightly narrower which may be driven by greater consistency in the data sets and estimation methods that may be required by these leading journals. Thus, we again conclude that our results for the Top 3 and Top 5 finance journals do not materially differ from our main findings based on the full sample.

2.5.5 Overview

Table 2.7 provides an overview of our subsample results. The table shows that, for the full sample, our estimates based on various techniques of the "representative" alpha coefficient corrected for the publication selection bias range from 0.274 and 0.386. The mean and median values of 0.335 and 0.338 respectively are both fairly close to the unconditional sample mean of 0.36. Relative to these full sample results, our estimates of the "representative" alpha have a lower minimum of 0.194 for the subsample of alphas treated the survivorship and/or backfilling biases, and a higher upper bound for the subsample of alphas untreated for either of the biases. The range of "corrected" alphas is wider and higher for alphas estimated with the use of the one-factor model (0.349, 0.707) relative to the seven-factor model (0.128, 0.326). While the range of the "representative" alpha coefficients estimated without the use of IV (0.276, 0.388) virtually coincides with our full sample results, the range of the "corrected" alpha estimates based on IV techniques (-0.411, 0.298) is substantially wider and includes negative values. Finally, the ranges of the "corrected" alpha estimates based on observations published in the Top 3 (0.265, 0.358) and Top 5 (0.263, 0.355) are only slightly below our full sample results.

Table 2.7: Results overiew

Table	Part	Note	#Studies	#Alphas	Mean	StDev	Min	Md	Max
Table 2.2		Full sample	74	1019	0.335	0.033	0.274	0.338	0.386
Table 2.3	1	Bias treated	50	605	0.273	0.048	0.194	0.265	0.351
Table 2.3	2	Bias untreated	29	414	0.390	0.077	0.282	0.380	0.521
Table 2.4	1	1F model	18	167	0.467	0.095	0.349	0.437	0.707
Table 2.4	2	7F model	33	298	0.269	0.061	0.128	0.298	0.326
Table 2.5	1	Instruments	3	46	0.025	0.187	-0.411	0.048	0.298
Table 2.5	2	No instruments	74	973	0.338	0.032	0.276	0.343	0.388
Table 2.6	1	Top 3	16	118	0.317	0.034	0.265	0.314	0.358
Table 2.6	2	Top 5	22	256	0.314	0.032	0.263	0.313	0.355

Notes: The table provides an overview of the results presented in this paper. *Table* and *Part* specify the table and its part where the results are reported. *Note* provides a brief description of a given set of results. *#Studies* and *#Alphas* show the number of studies and the number of alpha estimates a given set of results is based on. *Mean*, *StDev*, *Min*, *Md*, and *Max* refer to the mean value, the standard deviation, the minimum, the median value, and the maximum of a given set of results.

Our results make several important contributions to prior literature. First, we synthesize fragmented empirical evidence on hedge fund performance and present estimates that are corrected for any publication selection bias. Second, our results demonstrate that despite the prevalence of the publication selection bias in numerous other research settings, publication may not be selective when there is no strong *a priori* theoretical prediction about the sign of estimated coefficients, which may induce greater readiness to publish statistically insignificant results. Third, we provide one of the first out-of-sample tests of the proposition by (Brodeur *et al.* 2020) who argues that IV-based estimates tend to suffer from publication bias more frequently than estimates based on other techniques.

2.6 Conclusion

We perform a meta-analysis of prior empirical studies evaluating hedge fund performance. We examine whether published estimates of hedge fund alphas (abnormal returns) are affected by publication bias and by data biases. Prior research detects publication selection bias in a wide range of economic and finance settings, e.g. Stanley (2001; 2005); Stanley & Doucouliagos (2010); Rusnak *et al.* (2013); Havranek (2015); Brodeur *et al.* (2016); Bruns & Ioannidis (2016); Havranek & Irsova (2017); Stanley & Doucouliagos (2017); Christensen & Miguel (2018); Irsova *et al.* (2018); Havranek *et al.* (2018a;b); Astakhov *et al.* (2019); Brodeur *et al.* (2020); Blanco-Perez & Brodeur (2020), and Zigraiova *et al.* (2021). In contrast to these findings, using a wide range of techniques and data partitions we do not detect selective publication in hedge fund performance literature with the exception of estimates based on instrumental variables. In contrast, we provide evidence that not controlling for the potential biases in the underlying data (e.g. the backfilling bias and the survivorship bias) affects reported alpha coefficients systematically.

The fragmentation of hedge fund performance data and the wide range of alternative approaches for controlling for risk give researchers considerable discretion over the design of their research. This potentially creates opportunities for selective publication because the use of various estimation techniques based on different data sources may yield diverse results, some of which may be more attractive for publication than others. Our results demonstrate that despite the prevalence of publication selection bias in numerous other research settings, publication may not be selective when there is no strong *a priori* theoretical prediction about the sign of estimated coefficients, which may induce greater readiness to publish statistically insignificant results.

The heterogeneity in methodological approaches and data sources used in estimating hedge funds' alphas opens up additional research opportunities. Future research can examine whether and how the various aspects of methodological choices affect the magnitude of reported alpha coefficients. Our aim in this paper is to propose a representative alpha coefficient that is aggregated across the plurality of these approaches and corrected for publication and data biases. Therefore, in this study we provide robustness checks based on subsamples that narrow down the pool of collected alpha estimates to more homogeneous subsets but we do not explicitly exploit the full sample heterogeneity to analyze and draw conclusions about individual subsets or about the relative magnitude of alpha coefficients for the individual subsets. We leave the analysis of the impact of this heterogeneity on the reported alpha coefficients for future research that can examine the importance of various dimensions of methodological choices on the alpha coefficients reported in primary studies.

Chapter 3

Where Have All the Alphas Gone? A Meta-Analysis of Hedge Fund Performance

Fan Yang, Tomas Havranek, Zuzana Irsova, Jiri Novak

Abstract We conduct a systematic meta-analysis of the factors influencing hedge fund performance estimates published between 2001 and 2021. Using a sample of 1,019 intercept terms from regressions of hedge fund returns on risk factors (the "alphas") collected from 74 studies, we document a strong downward trend in reported alphas that persists even after controlling for heterogeneity in hedge fund characteristics and research design choices in the underlying studies. Our best-practice estimates of current performance are not reliably different from zero for all common hedge fund strategies. In addition, we provide an estimate of the sizeable impact of management and performance fees charged by hedge funds. We also document how reported performance estimates vary with hedge fund and study characteristics. Our results suggest that while hedge funds have generated positive value for investors in the past, on average, they no longer do so.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 823748 (project CoSP). This work was supported by the Cooperatio Program at Charles University, research area Economics. Novak acknowledges support from the Czech Science Foundation (GA CR, project 24-11583S). Data and code are available in an online appendix at `meta-analysis.cz/hedge`

3.1 Introduction

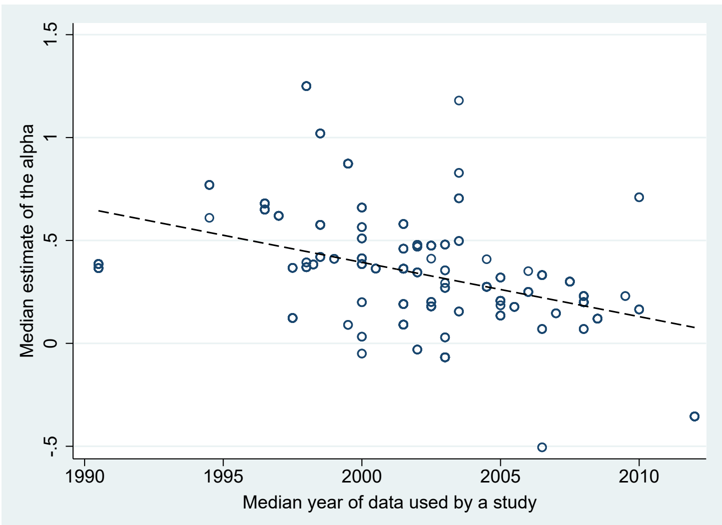
The prominence of hedge funds in the economy and the capital invested in them has dramatically increased over the past years. Stulz (2007) and Barth *et al.* (2020) document a one-hundred-fold increase in the assets under management (AUM) since the early 1990s. Hedge funds often target high-net-worth individuals and institutional investors, which allows them to take advantage of exceptions from regulatory requirements and to be structured to benefit from a favorable tax treatment (Connor & Woo 2004; Stulz 2007). The exemptions from regulatory oversight allow hedge funds to remain rather secretive about their proprietary, often unorthodox, trading strategies, to concentrate their exposure to the idiosyncratic risk components that are inherent to the information trading they perform (Brown *et al.* 2018), and to make extensive use of financial derivatives to modulate their risk exposures (Aragon & Spencer Martin 2012).

The surge in hedge funds' popularity, their limited transparency, and the strength of macroeconomic repercussions of some of their notorious failures (Connor & Woo 2004; Stulz 2007) naturally led to questions on their performance and on how much value they generate for their investors. On the one hand, relative secrecy may allow hedge funds to more effectively exploit their innovative investment strategies. The focus on large investors may allow them to take risks that would otherwise not be tenable in conventional mutual funds. Hedge funds may also benefit from the favorable tax treatment. On the other hand, their limited transparency and the voluntary nature of many performance disclosures make monitoring funds and systematically tracking their performance difficult. Limited oversight may lead to underperformance. Furthermore, hedge funds typically charge rather sizable management and performance fees (Ben-David *et al.* 2020), which may diminish the net-of-fee returns that investors actually get. Hence, *a priori* it is not obvious whether hedge fund performance is, on average, sufficiently strong to generate value for investors.

In this paper, we perform a systematic analysis of prior empirical research on hedge fund performance to evaluate whether and under what circumstances hedge funds generate value for investors. We collect 1,019 intercept terms (i.e., the "alphas") from regressions of returns on risk factors from 74 empirical studies published between 2001 and 2021. The alphas represent the abnormal return earned or value generated for investors after adjusting for the level of returns investors are expected to require for a given level of their investment's systematic risk. We use several meta-analytical techniques that allow us to adjust for potential data and publication biases in this stream of research. Our approach thus allows us to draw inferences about the magnitude of the value generated for investors, taking into consideration various forms of adjusting for risk differences, adjusting for potential data and publication biases, and conditional on specific hedge fund strategies, time periods, etc.

We find the unconditional sample mean of 36 basis points (i.e., 0.36%), which corresponds to the annual abnormal return of 4.3%. This number falls within ranges of alpha estimates reported in several prominent prior studies on hedge fund performance (Fung & Hsieh 2001; Getmansky *et al.* 2015). Nevertheless, we also find that the estimates of the value created by hedge funds have steeply decreased over time. We visualize this trend in Figure 3.1, which shows the median hedge fund alpha reported in academic journals in individual years. The dashed trend line crosses the horizontal axis (that denotes zero abnormal returns) around the year 2015. This suggests that estimates of hedge fund performance published after this year are, on average, negative. Our empirical analysis discussed below establishes the conclusion on the declining trend stands even after taking into consideration various factors that may have affected the reported estimates. We simultaneously control for a host of variables related to hedge fund characteristics and research design choices made in the primary studies. We demonstrate that our best practices estimate of hedge funds' current performance is not reliably different from zero. Furthermore, when classifying hedge funds into common categories based on their investment strategies, we observe that the current performance estimate is not significantly positive for any of these categories. Thus, our findings suggest that even though hedge funds used to generate positive value for investors in the past, on average, they do not do so anymore.

Figure 3.1: Are the markets getting more efficient?



Notes: The vertical axis shows the median estimate of the alpha reported in individual studies. The horizontal axis shows the median year of the data used in the studies. The dashed line denotes a linear trend. Outliers are omitted from the figure for ease of exposition but are included in all tests.

To identify the time trend in hedge fund performance after controlling for factors that may have affected the estimates reported in prior research, we take into consideration a number of factors related to hedge fund characteristics and research design choices. In this analysis, we document that the value generated by hedge funds is substantially diminished by the fees they charge. Prior literature on the management and performance fees charged by hedge funds concludes that these fees may indeed be quite sizable. Ben-David *et al.* (2020) estimates that, on average, hedge funds appropriate in fees almost two-thirds of the excess return they generate. The literature also suggests that these fees are difficult to quantify due to their conditional nature. We offer an alternative way of estimating the effective fees paid by hedge fund investors by exploiting the composition of our sample that includes both alphas estimated using gross returns and alphas estimated net of fees. Our indicator variable captures the effect of hedge fund fee adjustment after controlling for all other relevant characteristics that affect the magnitude of reported alpha estimates. Our regression analysis shows that the indicator variable that captures whether hedge fund performance is estimated on a gross or a net-of-fee basis is the most powerful variable explaining the variation in the reported alpha coefficients. We show that, on average, alphas reported on the net-of-fee basis are -0.439 lower than alphas based on gross returns.

Our methodological approach also allows us to evaluate how the reported alpha estimates vary with the research design choices in the primary studies. Prior research frequently voices concerns that the measurement of hedge fund performance may be distorted by the survivorship and backfilling biases (Fung & Hsieh 2000; 2002; 2004a; Fung *et al.* 2008). The backfilling bias arises when hedge funds are included in databases together with their performance history only after succeeding during an "incubation period" intended to accumulate a performance track record before offering the fund to investors. Backfilling performance histories of successful funds introduces a positive bias into the database since the performance of the funds that perform poorly in the incubation period are never recorded in the database (Fung & Hsieh 2000; Posthuma & Van der Sluis 2003). The survivorship bias may arise when commercial databases terminate coverage of previously included funds. Providers may wish to purge the database of funds that no longer operate because they are no longer relevant to their clients (Edelman *et al.* 2013; Getmansky *et al.* 2015). Hodder *et al.* (2014) report that on average 15% of hedge funds exit the database every year. A bias arises when the funds that exit the database on average underperform the "surviving" funds.

In this paper, we offer an alternative approach to estimating the impact of these biases by comparing estimates that adjust for them with those that do not. Consistent with the concerns that these biases may indeed matter for the performance estimates voiced in the prior research literature, our results show that the reported alpha estimates tend to be significantly lower when the research design of the primary studies explicitly adjusts for the backfilling bias. Furthermore, we also observe lower alpha estimates for the fund of funds. Both of these findings are consistent with the proposition that these biases have a meaningful impact on the reported results, and empirical findings in studies that do not adjust for these biases should be interpreted with caution.

Furthermore, one of the major challenges in measuring hedge fund performance is the choice of the appropriate risk model. Hedge funds frequently engage in complex and dynamically evolving investment strategies. Thus, they may exhibit exposures to fundamental risk factors that differ from those that are typical for more conventional asset classes, such as common equities and fixed-income securities. Fung & Hsieh (2001; 2004a); Fung *et al.* (2008) propose a risk model that is specifically designed to capture risk exposures relevant to hedge funds. Given that this model was explicitly designed for measuring hedge fund alpha, it is plausible to expect that it should best capture the risk factors relevant to investment strategies commonly used by hedge funds. Nevertheless, the specificity of this model for hedge fund research also implies that results based on it are not directly comparable to performance estimates of other investment forms, such as mutual funds. Thus, prior hedge fund research also frequently reports alpha estimates based on several other asset pricing models, such as the Capital Asset Pricing Model (CAPM) (Sharpe 1966; Lintner 1965; Mossin 1966; Black 1972), the three-factor model (Fama & French 1995; 1996), and the four-factor model (Carhart 1997). Our results show that the choice of the risk model indeed matters for estimating how much value hedge funds actually create.

Our results also show that the estimates of the value created by hedge funds depend on the market conditions when they are measured. Hedge funds sometimes aspire to be “market neutral”, i.e., to generate fairly stable returns regardless of the general stock market conditions. Market neutrality should be valued by investors because robust returns during market downturns help investors diversify away risks. Nevertheless, empirical research does not provide strong support for hedge funds’ market neutrality (Capocci *et al.* 2005; Patton 2009). We document that hedge fund alphas tend to be lower when estimated for the declining “bear” markets.

Finally, we observe that reported hedge fund alpha coefficients tend to be lower when more databases are used as a source of data in a given primary study, and they are, on average, higher when the CISDM database is used as one of the data sources. These findings suggest that using more comprehensive datasets typically implies reporting lower hedge fund performance. Furthermore, researchers should be aware that alpha estimates based on the CISDM database are commonly higher than those based on other databases.

We make several important contributions to the prior research literature. First, we document that even though hedge funds used to generate positive value for investors in the past, on average, they do not do so anymore. This finding may potentially be driven by several underlying forces. First, the number of hedge funds has steeply increased over time, which may have intensified the competition among them and diminished abnormal returns that the early hedge ones were able to achieve. This explanation would be consistent with the prediction of a rational model of active portfolio management proposed by Berk & Green (2004). The model suggests that when managers have a differential ability to identify profitable investment opportunities with decreasing returns to scale, the likelihood of generating abnormal returns decreases in the volume of resources that these managers allocate. The increasing volume of resources managed by a given hedge fund manager may have, over time, eroded their ability to identify profitable investment opportunities and earn abnormal returns for their investors. These findings are also consistent with Eugene F. Fama's famous quote[1] that suggests that in efficient markets where everything is appropriately priced, hedge funds (that represent an extreme form of active investment management) cannot be expected to generate abnormal returns.

It is also possible that the decrease in hedge fund abnormal performance is due to progressively tighter hedge fund regulation that requires greater transparency. Much tighter hedge fund regulation was enacted in the aftermath of the 2008 financial crisis, in which some considered hedge funds one of the culprits (Fagetan 2020). Better transparency may reduce monitoring costs to investors and promote better oversight that may discipline hedge fund managers. Some argued that stricter regulation requiring greater transparency may be instrumental in discouraging hedge fund managers from overly aggressive investment strategies based on high leverage and extensive use of financial derivatives that may be particularly damaging in times of market turmoil. In the US, the government proposed the Dodd-Frank Act in 2009, and the registration and greater disclosure requirements became effective in 2012. The EU implemented the Alternative Investment Fund Managers Directive (AIFMD) in 2012. Nevertheless, greater transparency may also reveal some of the funds' proprietary information, make it easier for free riders to imitate successful investment strategies, make it more difficult for hedge fund strategies to reap the benefits of their ideas, and ultimately dilute managerial incentives to innovate (Bianchi & Drew 2010; Shi 2017). Furthermore, the new regulation may entail a significant compliance cost that may further depress hedge fund performance (Kamal 2012; Cumming *et al.* 2020).

Second, we also provide an alternative approach to estimating the impact of the management and performance fees that hedge funds charge for the value they generate for their investors. These fees tend to be difficult to quantify because their magnitude may depend on a fairly complex set of conditions agreed upon by the investors. Our methodological approach allows us to measure the effective impact these fees have on the realized hedge fund performance.

Third, we identify several research design choices in the primary studies that systematically affect the magnitude of the reported hedge fund alphas. In particular, we

[1]Eugene F. Fama: "*I can't figure out why anyone invests in active management, so asking me about hedge funds is just an extreme version of the same question. Since I think everything is appropriately priced, my advice would be to avoid high fees. So you can forget about hedge funds.*" Source: https://www.azquotes.com/quotes/topics/hedge-fund.html.

show that the estimated alpha coefficients depend on the choice of the asset pricing model. They also depend on the general market conditions during which hedge fund performance is measured. Furthermore, we also show that the reported hedge fund performance estimates are sensitive to the choice of data sources. The reported coefficients tend to be lower when more databases are used as a source of data and are, on average, higher when the CISDM database is used as one of the data sources. These findings help researchers and practitioners interpret prior empirical findings and inform them about the likely impact of their methodology and sample choices in future research.

The remainder of this paper is organized as follows. In Section 3.2, we provide an overview of the institutional background, and we discuss the arguments why hedge funds may or may not be expected to generate abnormal returns. In Section 3.3, we describe our data sample and the methodology we use. In Section 3.4 we discuss our empirical results. Section 3.5 concludes.

3.2 Background

3.2.1 Hedge Fund Characteristics

It is commonly believed that the first hedge fund was created in 1949 by a former Fortune magazine writer Alfred Winslow Jones (Connor & Woo 2004; Stulz 2007). Even though the financial industry has undergone a dramatic development over the seven decades that have passed since then many of the features of this first hedge fund resemble common hedge fund characteristics today. First, A. Jones structured the fund to be exempt from the Investment Company Act of 1940, which was the main pillar of the Security and Exchange Commission (SEC) regulations of investment entities at the time (Connor & Woo 2004). This exemption gave the fund greater flexibility in the use of investment techniques. Second, the fund made a relatively concentrated (rather than well-diversified) investment in a limited number of stocks that it considered undervalued and it hedged some of its risks by short selling other stocks. The long-short equity strategy still remains one of the most popular hedge fund strategies. It is also a strategy that gave "hedge" funds their name. Third, to build investors' confidence A. Jones co-financed a substantial portion of the fund's assets (40%) with his own money (Stulz 2007). Fourth, A. Jones used financial leverage to increase risk and simultaneously enhance the fund's ability to earn a higher return on the base capital. Fifth, A. Jones charged the investors a performance fee of 20% of returns earned (Connor & Woo 2004). All of these features are quite common in hedge funds even nowadays.

In the 1960s, news about the high and relatively stable returns earned by A. Jones' hedge fund inspired imitation and many new hedge funds arose. Many of these new funds modified the original investment strategy. First, due to the hedged long-short strategy, hedge funds missed out on the strong bull market of the late 1960s. That prompted many hedge funds to abandon hedging against market downturns and to pursue a leveraged long-bias strategy that keeps the fund exposed to overall market movements (Connor & Woo 2004). In the 1980s, new global macro funds started to appear, e.g. Julian Robertson's Tiger Fund, George Soros' Quantum Fund (Connor & Woo 2004). In contrast to the original hedge fund that aimed at limiting its exposure to overall market conditions, these funds aimed to exploit the impact of general macroeconomic conditions typically in foreign exchange markets. The global macro funds made highly leveraged bets on the appreciation or depreciation of specific currencies. When successful (e.g. the Tiger Fund's bet on U.S. dollar appreciation, the Quantum Fund's bet on U.K. pound depreciation) these strategies generated spectacular returns, which attracted further investors (Connor & Woo 2004). However, betting against currencies also earned hedge funds a reputation as a destabilizing force that profits from financial market turmoil.

Naturally, not all hedge fund bets turned out successful. Especially, the events of the late 1990s with the dot-com equity market bubble and the Russian debt crisis uncovered many vulnerabilities in hedge fund investment strategies. Both the Tiger Fund and the Quantum Fund lost billions on bets against the new economy that they were not able to sustain. The late 1990s also witnessed perhaps the most infamous hedge fund collapse of the Long-Term Capital Management (LTCM). The fund was started in 1993 by John Meriwether (a renowned Wall Street trader) and Myron Scholes and Robert Merton (Nobel Prize-winning economists). Between 1994 and 1998 it was very successful in pursuing the fixed-income arbitrage strategy that exploits small interest rate spreads between various debt securities. Pricing discrepancies in fairly efficient bond markets tend to be relatively small. Thus, the LTCM used very high leverage to earn an acceptable return on the capital provided by investors. This leverage became unsustainable during the Russian debt crisis when debt markets exhibited temporary anomalies. Some large investors "flew to safety" and closed their positions in riskier debt securities (Connor & Woo 2004), which prompted the LTCM's collapse. To avoid wider contagion in financial markets the Federal Reserve Bank (FED) organized a bailout. The cost of this bailout led to further discussions about the potentially destabilizing macroeconomic impact of hedge funds. It became widely acknowledged that notwithstanding their prominent role in promoting financial markets' efficiency hedge funds may also play a more detrimental role. This understanding provided a strong motivation for systematic research in hedge funds.

Despite their growth there is, in fact, no universally accepted definition of a hedge fund (Brav *et al.* 2008). However, hedge funds share several characteristics that distinguish them from other investment facilities. First, hedge funds are structured to take advantage of exceptions from regulatory requirements and to benefit from a favorable tax treatment (Connor & Woo 2004). The legal framework that regulates investment entities, such as the Securities Act of 1933 and the Investment Company Act of 1940, typically allows funds with a number of investors below some threshold (often 100) to be exempted from regulatory requirements that commonly apply to mutual funds (Connor & Woo 2004; Stulz 2007). To qualify for such exceptions hedge funds target a limited number of high-net-worth individuals and institutional investors. From the regulatory perspective, these investors may be considered sufficiently competent to make investment decisions and sufficiently wealthy to sustain potential losses. Hence, regulators may consider it unnecessary to protect these investors from potentially adverse consequences of their investment decisions (Stulz 2007). Furthermore, hedge funds tend to be organized as limited partnerships to benefit from pass-through tax treatment where the returns are only taxed at the individual investors' level rather than at the level of the hedge fund (Connor & Woo 2004).

Second, the exemptions from regulatory oversight allow hedge funds to implement unorthodox and often dynamic investment strategies that may exploit a wide range of diverse investment opportunities. Furthermore, hedge funds typically use limited amounts of base capital and they use substantial leverage to increase the return earned on their investment strategies. Leverage makes hedge fund investments substantially riskier than what is common for mutual funds. Nevertheless, hedge funds frequently engage in short selling and they make a complex use of financial derivatives (Aragon & Spencer Martin 2012) to concentrate their exposure to the idiosyncratic risk components that are inherent to the information trading they perform (Brown *et al.* 2018). Besides the investment strategies already discussed above (long-short equity, global macro, and fixed-income arbitrage) hedge funds also engage in event-driven strategies that are based on investing in anticipation of major corporate events, e.g. mergers and acquisitions (M&As), spin-offs, reorganizations, and bankruptcies (Stulz 2007). The success of event-driven strategies crucially depends on fund managers' ability to predict the outcome and the price impact of these events and on identifying the optimal time to make the investment.

Third, hedge funds often require their investors to commit their investment for a fairly long time (Teo 2009). The "lockup periods" may last for several years. Even after the expiration of the lockup periods investors may be obliged to notify managers several months in advance when they want to redeem their investment (Aragon 2007). These withdrawal restrictions give managers more flexibility in investing in illiquid assets, the value of which may remain depressed for some time. Hedge funds may also exploit opportunities that arise when more conservative investment entities such as pension funds are obliged to divest distressed securities (Connor & Woo 2004). Holding distressed securities is typically associated with higher liquidity risk. Hence, hedge funds may have substantial exposures to macroeconomic liquidity shocks (Boyson *et al.* 2006; Sadka 2010). The lockup period and redemption notice period thus limit the likelihood that hedge funds will be forced to quickly liquidate these assets under unfavorable conditions.

Fourth, being exempted from many regulatory requirements allows hedge funds to remain rather opaque, which helps them protect their proprietary trading strategies from imitation by competitors. Hence, investors can typically barely learn about the rough contours of investment strategies that a given fund aims to pursue. Furthermore, unlike mutual funds, most hedge funds are not obliged to periodically report audited financial statements to regulators. Nevertheless, some funds may provide information on their performance on a voluntary basis (Stulz 2007). Hedge funds are not allowed to publicly advertise and so having their performance record included in commercial databases may help them attract investors (Fung & Hsieh 2004a; Baquero *et al.* 2005). This discretion was constrained by the Dodd-Frank Act of 2010 which mandates investment funds domiciled in the U.S. that manage more than $150 million in aggregate assets to register with the SEC and to provide basic periodic disclosures on asset values, returns, borrowings, strategy classifications, investor composition, and their largest counterparties (Barth *et al.* 2020). The asset value cutoff implies that the regulation applies only to the large hedge funds that may be systemically important.

Fifth, hedge funds typically charge their investors substantial management and performance fees (Malkiel & Saha 2005). A common arrangement consists of a flat management fee of 1% to 2% of AUM and a variable performance fee usually 20% of realized returns above the risk-free rate (Fung & Hsieh 1999; Connor & Woo 2004; Stulz 2007; Kouwenberg & Ziemba 2007; Getmansky *et al.* 2015). The performance fee is usually paid only after reaching the so-called "high water mark", i.e. the minimum level of absolute performance over the entire investment lifetime (Asness *et al.* 2001; Goetzmann *et al.* 2003; Lim *et al.* 2016; Stulz 2007). In other words, in a given year fund managers receive the performance fee only after having recovered any losses incurred in previous years. However, effectively the performance fees constitute even a larger portion of realized returns because investors cannot offset gains and losses across funds, they tend to withdraw capital after a poor past performance, and managers sometimes terminate hedge funds after large losses, which renders the high water mark provision irrelevant. Ben-David *et al.* (2020) find that due to these three reasons the effective performance fees approach one-half of the aggregate gross profits. The high level of hedge fund managers' participation in realized returns strongly incentivizes them to perform and it allows successful managers to earn compensation similar to what they would earn in mutual funds 10 times their hedge fund size (Connor & Woo 2004; Jobman 2002). Furthermore, unlike in mutual funds, the performance fee in hedge funds makes the compensation structure highly asymmetric. Hedge fund managers are compensated for gains, but they are not equivalently penalized for commensurate losses. These option-like payoffs strongly motivate them to take risk. The high-water mark provisions are likely to only partially moderate these risk-taking incentives because managers of unsuccessful hedge funds may opt to close the fund down and open a new one (Stulz 2007). Getmansky *et al.* (2015) report that only about one-half of hedge funds survive for more than five years. Hence, hedge funds are likely to take substantial risks, which should be taken into consideration when measuring their performance.

3.2.2 Hedge Fund Performance

A priori, it is not quite obvious whether hedge funds should be expected to outperform other types of investments. Hedge funds typically make their investments in financial markets that are rather competitive and where investors have strong incentives to quickly eliminate any mispricing. In efficient markets, any quest for mispriced assets that subsequently earn abnormal returns may be elusive. In the past, many famous hedge fund successes were followed by spectacular failures, which suggests that extraordinary performance may be temporary and driven by chance. For example, the once-lauded and abundantly financed investment strategy of the LTCM later failed and necessitated a massive bailout (Stulz 2007). Furthermore, competition is intensive even within the hedge fund industry. Light regulation implies relatively low barriers to entry. Any profitable strategies discovered by hedge funds may invite imitation by competitors and their ability to generate abnormal returns may quickly disappear.

Furthermore, the generous and convex "option-like" compensation packages that reward success but do not commensurately penalize failure may encourage excessive risk-taking (Cao *et al.* 2016). Hedge fund managers may thus take aggressive positions that expose investors to substantial risks. Stulz (2007) argues that hedge fund risk profiles may resemble those of firms selling earthquake insurance. They may exhibit stable profitability for a long time but incur catastrophic losses at rare events when a disaster strikes. The LTCM's arbitrage strategy was ex-post likened to "*picking up pennies in front of a steamroller*" (Stulz 2007). Since most hedge funds are not obliged to systematically report their performance some of these failures may be kept off the radar. If successful hedge funds are more likely to be included in the private databases and become better known to investors than the failed ones (Posthuma & Van der Sluis 2003), investors' view of overall hedge fund performance may be distorted.

In addition, the light regulatory oversight and limited reporting requirements may impair managerial accountability and complicate monitoring by investors. Information on the portfolio composition and periodic performance may not be independently audited and so its reliability may be in question. Hedge fund managers may thus be able to camouflage inferior performance for some time, which may prevent investors from taking timely corrective action. When investors are kept in the dark they may find it difficult to base their investment decisions on a pragmatic economic calculus. Rather, they may fall prey to hedge fund managers' personal charm and keep trusting them for longer than appropriate. The Bernard L. Madoff Investment Securities investors mention the founder's personality as one of the reasons why they remained confident in the fund for so long.[2]

Finally, hedge funds charge very substantial management and performance fees. Thus, it is also conceivable that hedge funds actually beat the benchmark but the return they generate is not sufficient to cover these high fees. Paying these fees may thus leave the investors worse off than they would be by simply tracking the market index at a modest cost.

On the other hand, the flexibility resulting from the regulatory status puts hedge funds in a strong position to exploit opportunities that others cannot. It allows them to adopt a wide range of rather unorthodox investment styles that cannot be pursued by more tightly regulated mutual funds and pension funds. The light regulation allows hedge funds to remain secretive about the nature of their strategies, their holdings, and annual performance, which may allow them to protect their proprietary trading strategies and keep exploiting them longer than conventional mutual funds could. Hedge funds can thus act as investment strategy innovators and benefit from their first-mover advantage. They can also benefit from being a counterparty to transactions when more conventional investment entities are obliged due to regulation to divest distressed assets. Hedge funds may also benefit from introducing competition into previously oligopolistic market segments such as fixed-income arbitrage that used to be the domain of investment banks (Connor & Woo 2004; Schneeweis 1998).

Furthermore, investors typically agree to forgo some of the diversification benefits, which allows hedge funds to keep asset holdings relatively concentrated and to specialize in a fairly narrowly defined niche. Investment concentration may allow hedge funds to realize some gains from their high degree of investment specialization. The lack of aspiration to hold well-diversified portfolios may also give hedge funds an opportunity to act more aggressively in acquiring substantial stakes in firms and to become "activist", i.e. they can actively use their ownership rights to alter how the companies are run. Hedge fund activism can make the companies more valuable by rectifying some of the agency conflicts between the owners and managers, by adopting more suitable business strategies, and by reducing inefficiencies in operations. Besides hedge funds' stock picking and market timing skills, their activism may be another source of generating value for investors.

Hedge funds may also benefit from their flexibility in designing specific contractual arrangements with their investors and their managers. The lockup periods and the withdrawal restrictions may relieve hedge fund managers from potentially myopic short-term performance pressures and allow them to pursue strategies that may temporarily underperform but that are profitable in the long run avoiding the risk of being forced to

[2]Source: https://www.nbcnews.com/business/business-news/madoff-exploited-weak-oversight-did-regulators-learn-their-lesson-n1264094.

liquidate some of their assets at temporarily depressed prices. Hedge funds may also benefit from their flexibility in designing managerial compensation contracts (Agarwal *et al.* 2009b; Cao *et al.* 2016). Mutual fund regulation obliges incentive compensation to be symmetric, i.e. gains and losses of equal size must have an identical opposite effect on managerial compensation (Elton *et al.* 2003; Stulz 2007). Most mutual funds thus make limited use of performance-based compensation and they mostly remunerate managers based on the value of assets under their management (Elton *et al.* 2003; Stulz 2007). In contrast, hedge funds are unconstrained in the design of their compensation packages. They may offer managers highly asymmetric (i.e. "option-like") compensation contracts, which helps them attract managerial talent and keep managers incentivized to perform and to take risks. Exceptionally talented and strongly motivated managers are likely to have a positive impact on hedge fund performance.

Hedge funds may also be more efficient on the cost side. They target a limited number of accredited investors, which allows them to save on marketing and communication costs. Their opacity also implies that they avoid the disclosure and attestation (audit) costs. Thus, whether hedge funds generate abnormal returns to investors is ultimately an empirical question.

3.2.3 Prior Empirical Findings

The studies on hedge fund performance produced rather diverse estimates of the value that hedge funds produce. Some of the variation in the published results likely arises due to different methodological approaches in various studies. A commonly voiced concern related to the measurement of hedge fund performance is related to the deviations from normality in the distribution of hedge fund returns (Malkiel & Saha 2005). Several studies explicitly address this issue. Agarwal & Naik (2004) document a significant left-tail risk in a wide range of hedge fund strategies. To account for this left-tail risk, they develop a conditional value-at-risk framework, which shows that the conventional mean-variance measures may underestimate expected left-tail losses by more than half. Amin & Kat (2003) use an approach that does not require specific characteristics of the underlying returns distribution, and they conclude that the vast majority of individual hedge funds and hedge fund indices are inefficient relative to the general market index. Also Bali *et al.* (2013) use an approach that accommodates the non-normality in returns distribution. Out of eleven hedge fund indices they consider, they find outperformance only for two of them – the long-short equity and emerging markets hedge fund indices. In a similar vein, Agarwal *et al.* (2009a) document that hedge funds are exposed to the risks associated with the higher moments of their returns distribution and that adjusting for this exposure substantially reduces the observed abnormal performance, especially for equity-based hedge fund strategies.

Another research stream investigates the dependence of hedge fund performance on macroeconomic conditions. Bali *et al.* (2011) report that hedge funds with higher exposure to default risk premium and lower exposure to inflation earn higher returns. Avramov *et al.* (2013) consider four variables related to the macroeconomic conditions: the default spread, the dividend yield, the volatility index (VIX), and the aggregate fund flows into hedge funds, and they show that they predict future hedge fund returns. Similarly, Agarwal *et al.* (2017) measure hedge funds' exposure to uncertainty about aggregate volatility and they show that funds with low exposure to this uncertainty outperform those with high exposure. Building on these findings that underscore the relevance of macroeconomic conditions for hedge fund performance Bali *et al.* (2014) include measures of macroeconomic uncertainty directly in the risk model used to measure hedge fund performance, and they demonstrate the relevance of most of the macroeconomic factors in this setting.

Related to the macroeconomic conditions, other papers also examine how hedge fund performance depends on conditions in financial markets. Hedge funds sometimes aspire to be "market neutral", i.e., generate fairly stable returns regardless of the general market conditions. Market neutrality should be valued by investors because robust returns during market downturns help investors diversify away risk. Nevertheless, empirical research does not provide strong support for hedge funds' market neutrality. Capocci *et al.* (2005) examine hedge fund performance in bull and bear markets, and they conclude that hedge fund outperformance is concentrated in periods of rising markets. Patton (2009) considers five different ways of measuring market neutrality and he concludes that hedge fund returns tend to be positively correlated with market returns. The author also finds that about one-quarter of funds classified in the market-neutral style exhibit substantial exposure to market risk.

Another reason for the divergence in the reported results may be the data deficiencies that may arise due to the voluntary nature of reporting of hedge fund performance in hedge fund databases. Fung & Hsieh (2000) and Fung & Hsieh (2002) and Fung *et al.* (2008) argue that the impact of these biases may be mitigated by using data on the funds of hedge funds (FoFs). FoFs' returns should not be affected by backfilled returns and they should appropriately reflect returns of hedge funds that decide not to report returns in commercial databases and that cease to exist (Posthuma & Van der Sluis 2003). However, using FoF returns generates new problems. FoFs endogenously decide on what hedge funds to include in their holdings, which may not be representative of the overall hedge fund population. Furthermore, FoFs charge investors an additional layer of management and performance fees (Stulz 2007) that reduce the realized return, which may distort the quantification of the abnormal return generated by individual hedge funds (Amin & Kat 2003). Brown *et al.* (2004) find that due to the extra layer of fees individual funds actually dominate FoFs in terms of net-of-fee returns, which make FoFs unattractive to investors. Getmansky *et al.* (2015) observe a decline in the number of

FoFs over time, which the authors ascribe to their fee structure, competition from multi-strategy funds, and their limited ability to protect investors from losses during financial downturns. With fewer available FoFs, any analysis of their holdings and performance is less generalizable for the universe of hedge funds.

3.3 Research Design

3.3.1 Measuring Hedge Fund Performance

Hedge fund performance is commonly measured by the intercept terms (the "alphas") from regressions of hedge fund returns on risk factors, see Equation 3.1.

$$(R_p - R_f) = \alpha_p + \sum_{n=1}^{N} \beta_{n,p} \cdot F_n + \epsilon_p \tag{3.1}$$

where R_p denotes the realized return on portfolio p, R_f denotes the risk-free rate of return, α_p represents the intercept term, F_n represents the n-th risk factor, $\beta_{n,p}$ denotes the sensitivity of portfolio p to the n-th risk factor, and ϵ_p represents the error term. The factor models adjust for the portfolio returns exposure to the systematic risk. The alphas that represent the unexplained portion of the realized return may thus be interpreted as the "abnormal" returns that hedge funds earn for their investors.

Various factor models differ in the set of factors they consider. Thus, the alpha estimates obtained based on the different models may also vary. The simplest approach based on the Capital Asset Pricing Model (CAPM) (Sharpe 1966; Lintner 1965; Mossin 1966; Black 1972) uses the difference between the stock market return and the risk-free rate ($R_m - R_f$) as the only risk factor. Notwithstanding the conceptual appeal this approach has, since it models the expected excess return based on an asset's contribution to the overall portfolio risk, which should correctly reflect the relevant risks exposure of well-diversified investors, prior research establishes that the single risk dimension might be too restrictive in capturing all the relevant risk exposures. Thus, the three-factor model (Fama & French 1995; 1996) and the four-factor model (Carhart 1997) are frequently proposed as more comprehensive alternative approaches to capturing the systematic risk. Furthermore, due to the complexity of measuring a risk exposure in Hedge funds that frequently engage in complex and dynamically evolving investment strategies, Fung & Hsieh (2004a) propose a model featuring seven factors that are particularly relevant for risk exposures that common hedge fund strategies involve. These seven dimensions involve (i) the stock market excess return, (ii) the spread between the small-capitalization and large-capitalization stock returns, the excess return pairs of look-back call and put options (iii) on currency futures, (iv) on commodity futures, and (v) on bond futures, (vi) the duration-adjusted change in the yield spread of the U.S. 10-year Treasury bond over the 3-month T-bill, and (vii) the duration-adjusted change in the credit spread of Moody's BAA bond over the 10-year Treasury bond.

3.3.2 Sample Collection

We collect our sample of hedge fund alphas from peer-reviewed research articles published between January 1, 2001, and September 1, 2021. The alpha estimates are the intercept terms from regressions of hedge fund returns on risk factors. The alphas represent risk-adjusted returns generated by hedge funds, which makes them comparable and suitable for aggregation by means of a meta-analysis. We ensure that all the alphas are expressed as a percentage, and we normalize them by dividing annual and quarterly alphas by twelve and three, respectively. We consider only published estimates as these successfully cleared the peer-review process that assures the quality of published findings. This increases the likelihood that the alphas we consider are estimated using established methodologies and free of error. In addition, estimates published in academic journals likely represent empirical evidence that is most influential in shaping the views of investment professionals and academics on hedge fund performance.

Our procedure of identifying primary studies, from which we source the alpha estimates, follows the guidelines proposed by Havranek *et al.* (2020). First, we consider studies cited in two prominent reviews of empirical research on hedge fund performance: Connor & Woo (2004) and Agarwal *et al.* (2015). We then perform a systematic Google Scholar search based on the following combinations of keywords: "hedge fund returns" OR "hedge fund performance". To ensure that our search has a good coverage of relevant articles we verify that the used combination of keywords identifies the vast majority of studies cited in the two above-mentioned review articles. We go through the first 750 articles in the Google Scholar list and we manually collect hedge fund alpha estimates reported in them. We terminate our screening of primary studies after having covered the first 750 articles from the Google Scholar list because we observe that after this point, the relevance of articles substantially decreases and the likelihood of finding additional usable alpha estimates is rather small in articles further down in the list.

We further complement our main keyword search with another search that is more general in the combination of used keywords: "hedge fund" OR "hedge funds", and that is limited to five journals where empirical research on hedge fund performance is likely to be published: the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*, the *Journal of Financial and Quantitative Analysis*, and the *Review of Finance*. Finally, to ensure comprehensive coverage of estimates published in journals aimed primarily at investment professionals, we perform the third - search using the following keywords: "hedge fund" OR "hedge funds" in the journals listed on the Portfolio Management Research website[3]: the *Journal of Portfolio Management*, the *Journal of Financial Data Science*, the *Journal of Impact and ESG Investing*, and the *Journal of Fixed Income*.

In our quantitative analysis, we control for potential publication selection bias in reporting alpha coefficients. This requires a measure of the precision of collected alpha estimates. Consequently, we only collect alpha estimates accompanied by a measure of statistical significance, i.e., a t-statistic, a standard error (SE), and/or a p-value. When more than one measure of statistical significance is provided, we apply the following procedure. We collect corresponding t-statistics directly from the primary studies whenever available. When standard errors are reported, we compute the t-statistic by dividing the alpha by its standard error. Correspondingly, in Bayesian studies, we compute the t-statistic by dividing the alpha by its standard deviation. When a primary study reports p-values, we check if the paper discusses whether these are based on one-tailed or two-tailed tests. When the information on the type of the test is not explicitly stated, we try to infer it from the discussion of the statistical significance of reported results. We assume a two-tailed test whenever the type of the test cannot be ascertained from the discussion of statistical significance (1 study).

We convert the reported p-value to a t-statistic with the use of the inverse t-distribution. Whenever available we use the total number of observations used to estimate the corresponding regression as the number of the degrees of freedom. When the primary study specifies both the number of observations in the cross-section and in the time-series we use the product of the two numbers. If only one of the two numbers is reported we use it instead. If no information about the number of observations is provided we assume 168 observations, which corresponds to the median value for the sub-sample of coefficients, for which the information on number of observations is reported. We manually verify that all the coefficients with the implied t-statistic greater than 10 are referred to as highly significant in the text of primary studies. We discard 1 observation with a reported t-statistic greater than 50.

Table 3.1 provides a list of 74 primary studies identified by our data collection procedure. From these research articles we collected 1,019 alpha estimates that constitute the sample for our empirical analysis. The number of data points makes our study one of the largest meta-analyses in finance. The substantial number of primary studies on this topic and the number of reported alpha coefficients imply that hedge fund performance has been extensively studied in prior research and the alpha coefficients have been estimated in a variety of ways with the use of various data samples. It thus seems worthwhile to aggregate the results from diverse studies by means of a meta-analysis.

[3]Source: https://www.pm-research.com/).

Table 3.1: List of Primary Studies

Agarwal *et al.* (2017)	Edelman *et al.* (2013)	Malladi (2020)
Ahoniemi & Jylha (2014)	Edwards & Caglayan (2001)	Meligkotsidou & Vrontos (2008)
Aiken *et al.* (2013)	Eling & Faust (2010)	Mitchell & Pulvino (2001)
Ammann & Moerth (2005)	Frydenberg *et al.* (2017)	Mladina (2015)
Ammann & Moerth (2008a)	Fung & Hsieh (2004a)	Molyboga & L'Ahelec (2016)
Ammann & Moerth (2008b)	Fung & Hsieh (2004b)	Mozes (2013)
Aragon (2007)	Fung *et al.* (2002)	Patton & Ramadorai (2013)
Asness *et al.* (2001)	Fung *et al.* (2008)	Racicot & Théoret (2009)
Bali *et al.* (2013)	Gupta *et al.* (2003)	Racicot & Théoret (2013)
Bhardwaj *et al.* (2014)	Hong (2014)	Racicot & Théoret (2014)
Blitz (2018)	Huang *et al.* (2017)	Ranaldo & Favre (2005)
Bollen & Whaley (2009)	Ibbotson *et al.* (2011)	Diez De Los Rios & Garcia (2011)
Brown (2012)	Jame (2018)	Rzakhanov & Jetley (2019)
Buraschi *et al.* (2014)	Joenvaara & Kosowski (2021)	Sabbaghi (2012)
Cao *et al.* (2016)	Joenvaara *et al.* (2019b)	Sadka (2010)
Chen & Liang (2007)	Jordan & Simlai (2011)	Sadka (2012)
Chen *et al.* (2017)	Jylha *et al.* (2014)	Sandvik *et al.* (2011)
Chincarini & Nakao (2011)	Kanuri (2020)	Stafylas *et al.* (2018)
Clark & Winkelmann (2004)	Klein *et al.* (2015)	Stafylas & Andrikopoulos (2020)
Dichev & Yu (2011)	Kooli & Stetsyuk (2021)	Stoforos *et al.* (2017)
Ding & Shawky (2007)	Kosowski *et al.* (2007)	Sullivan (2021)
Ding *et al.* (2009)	Kotkatvuori-Ornberg *et al.* (2011)	Sun *et al.* (2012)
Do *et al.* (2005)	Liang (2004)	Teo (2009)
Duarte *et al.* (2007)	Ling *et al.* (2015)	Vrontos *et al.* (2008)
Edelman *et al.* (2012)	Lo (2001)	

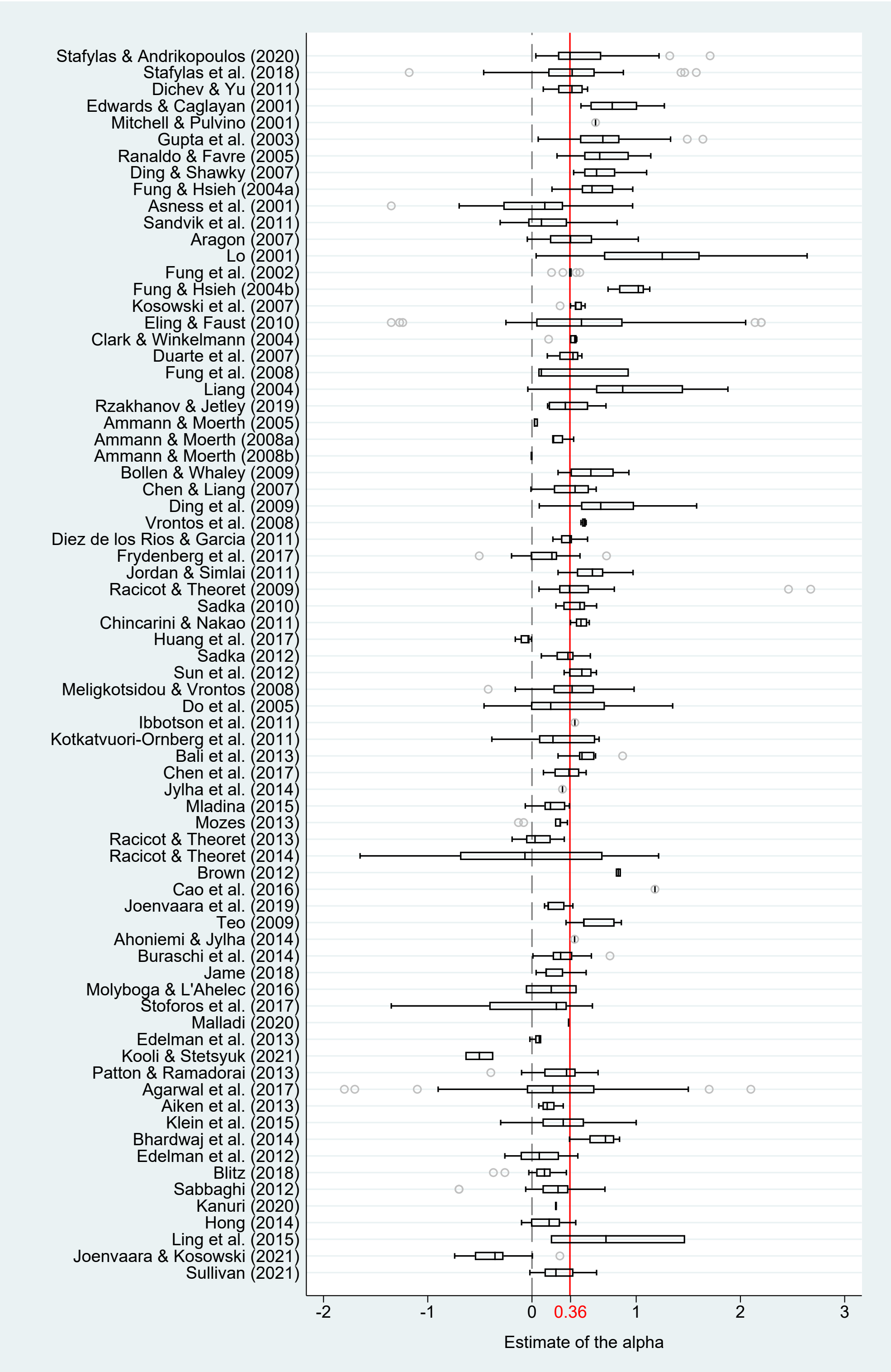
Notes: The table shows the list of 74 primary studies, from which we collect the alpha estimates that constitute our sample.

3.3.3 Sample Characteristics

In Figure 3.2, we examine how the range of reported alpha coefficients vary across the individual primary studies. For each primary study, the figure shows the median alpha estimate and their interquartile range represented by the box. The whiskers denote the minimum and the maximum values within the 1.5 times the range between the upper and lower quartiles. Figure 3.2 shows considerable variation in the reported alpha coefficients both within and across studies. While the interquartile ranges for some studies are fairly narrow, other studies exhibit interquartile ranges that exceed 1 percentage point of monthly returns, which corresponds to an annual return of 12%. Furthermore, interquartile ranges of some studies do not cross the vertical line representing the unconditional sample mean of 0.36%, which implies that the alpha coefficients reported in these studies substantially deviate from the values typical in the entire pool of research on hedge fund performance.

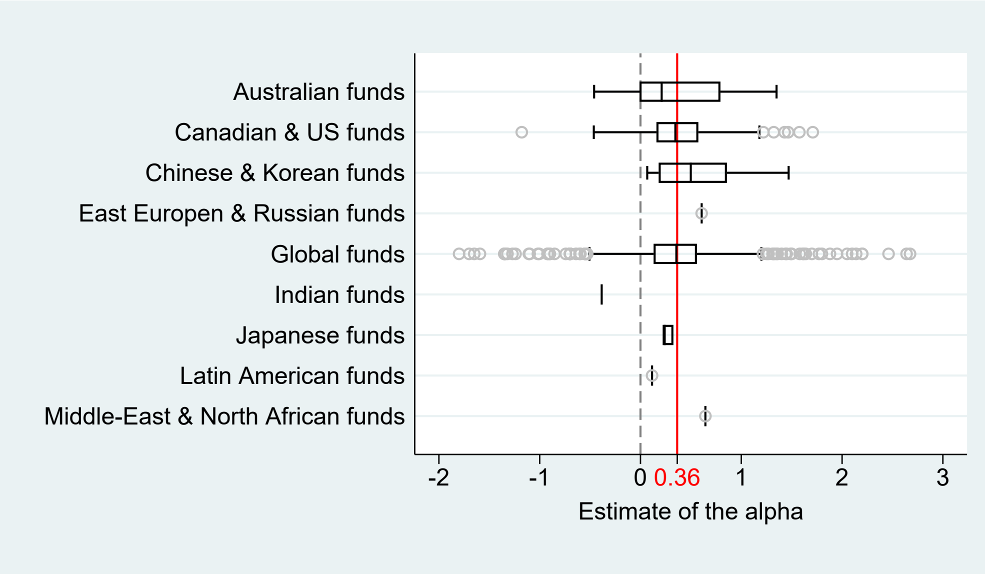
In Figure 3.2, the 74 primary studies are sorted by the median age of the underlying data, with the oldest samples at the top and the newer samples at the bottom of the figure. This figure thus provides the first preliminary evidence suggesting that hedge fund performance declined over time. The interquartile ranges of many of the studies

Figure 3.2: Distribution of Alphas in Primary Studies



Notes: The figure depicts the distribution of the alpha estimates in the individual primary studies sorted by the age of the underlying data. The length of each box represents the interquartile range (percentile 25, percentile 75). The vertical line inside the box depicts the median value. The whiskers represent the highest and lowest data points within 1.5 times the range between the upper and lower quartiles. The vertical line denotes sample mean. For ease of exposition, outliers are excluded from the figure but included in all statistical tests.

Figure 3.3: Distribution of Alphas in Geographic Scopes



Notes: The figure depicts the distribution of the alpha estimates across various geographic scopes covered by hedge funds. The length of each box represents the interquartile range (percentile 25, percentile 75). The vertical line inside the box depicts the median value. The whiskers represent the highest and lowest data points within 1.5 times the range between the upper and lower quartiles. The vertical line denotes sample mean. For ease of exposition, outliers are excluded from the figure but included in all statistical tests.

using older samples exceed the unconditional sample mean, while for the more recent studies, the interquartile ranges tend to be below it. We explore the tendency of primary studies based on newer data sets to report lower alphas further in the following analysis.

Figure 3.3 visualizes the distribution of alpha estimates generated by hedge funds covering various geographic areas. Most of the hedge funds we study are global funds. Thus, unsurprisingly, the median alpha estimate for global funds virtually coincides with the unconditional sample mean. Furthermore, alphas generated by the global funds have a relatively narrow interquartile range that is below 0.5%. This implies that most global funds in our sample generate abnormal annual returns between 0.2% and 0.6%. Similarly, funds that concentrate on the U.S. and Canada have a median return very close to the full sample mean and a fairly narrow interquartile range. Figure 3.3 also provides some indication that Australian, Indian, Japanese, and Latin American funds tend to generate somewhat lower alphas than their global counterparts. In contrast, Chinese, Korean, East European, Middle-Eastern, and North African funds, on average, generate somewhat higher alphas. However, most of these findings are based on a rather small number of observations. Furthermore, these descriptive statistics do not control for hedge funds' fundamental characteristics and differences in research design choices that may vary systematically across the primary studies. Thus, these findings should be considered preliminary. We delay drawing stronger conclusions about these characteristics till Section 3.4, where we perform a comprehensive analysis that investigates the combined effect of these characteristics.

3.3.4 Conditioning Variables

We consider several measures related to hedge fund fundamental characteristics and research design choices in the primary studies that may be relevant for explaining the cross-sectional variation in the value generated by hedge funds. Table B1 provides the definition and descriptive statistics for the explanatory variables that we use in our regression analysis. Since the determinants of hedge fund performance are not *a priori* known, we consider several "candidate" variables, and we examine how effective various combinations of these variables are in explaining the heterogeneity in the alphas reported in primary studies. For each variable Table B1 includes the definition, the unweighted mean value (Mean), the standard deviation (SD), and the mean weighted by the inverse of the number of estimates reported per study (WM), which gives all of the 74 primary studies equal weight. Many of our independent variables are indicators, and so the mean values represent the proportion of alpha estimates for which a given variable is coded as 1.

Consistent with our earlier findings, Table B1 shows that the mean value of the alpha estimates in our sample is 0.362. The mean value does not substantially change when the individual observations are weighted by the inverse of the number of estimates reported per study (WM = 0.365). The distribution of alpha estimates is fairly dispersed, with the standard deviation of 0.477. To adjust for a potential publication selection bias, we collect from the primary studies the alpha estimates' standard error (SE). Since prior literature shows that estimates based on instrumental variables (IV) tend to be less precise than coefficients estimated using different techniques Brodeur *et al.* (2020), we interact SE with an indicator variable equal to 1 for alpha coefficients estimated using IV.

Table 3.2 shows summary statistics for groups of alpha estimates determined by our conditioning variables. In the left panel, the individual alpha estimates are weighted equally. In the right panel, the alphas are weighted by the inverse of the number of estimates reported in a given study, which gives each of the 74 primary studies (rather than each of the 1,019 alphas estimates) equal weight in computing the mean value and the 95% confidence interval.

Table 3.2 shows some variation in the reported alpha estimates based on how the primary study aggregates hedge fund returns. Specifically, treating all estimates in our sample equally, reported alphas for value-weighted hedge fund indices tend to be lower than those documented for individual funds. In contrast, reported alpha estimates for equally weighted hedge fund indices are, on average, somewhat higher. Since returns on value-weighted hedge fund indices are disproportionately driven by the performance of

Table 3.2: Summary Statistics

		Unweighted			Weighted		
	No. of observations	Mean	95% conf. int.		Mean	95% conf. int.	
Aggregation of returns							
Individual funds	175	0.385	0.326	0.444	0.317	0.263	0.370
Equal-weighted funds	503	0.427	0.380	0.474	0.393	0.353	0.433
Value-weighted funds	341	0.256	0.213	0.298	0.357	0.314	0.399
Treatment of fees							
Net-of-fee returns	984	0.348	0.319	0.377	0.335	0.310	0.360
Gross returns	35	0.757	0.534	0.980	0.817	0.655	0.979
Data structure							
Cross-section data	855	0.355	0.322	0.388	0.379	0.349	0.409
Longitudal data	164	0.401	0.345	0.456	0.324	0.271	0.377
Source database							
Database: default	537	0.343	0.309	0.376	0.301	0.272	0.330
Database: CST	256	0.218	0.166	0.271	0.273	0.229	0.316
Database: CISDM	178	0.434	0.354	0.514	0.386	0.319	0.452
Database: hand-collected	22	0.411	0.263	0.559	0.576	0.356	0.796
Database: other	167	0.409	0.311	0.507	0.460	0.373	0.548
Market coverage							
Developed markets	140	0.388	0.315	0.461	0.462	0.393	0.530
World markets	879	0.358	0.326	0.390	0.350	0.322	0.378
Market conditions							
Bull market	39	0.272	0.199	0.346	0.277	0.206	0.347
Bear market	39	0.103	-0.006	0.211	0.097	-0.007	0.201
Hedge fund strategy							
Strategy: all funds	243	0.303	0.255	0.351	0.298	0.258	0.339
Strategy: equity hedge	229	0.352	0.286	0.418	0.360	0.296	0.424
Strategy: event driven	113	0.474	0.378	0.570	0.586	0.496	0.676
Strategy: relative value	94	0.294	0.215	0.373	0.401	0.329	0.473
Strategy: global	156	0.451	0.354	0.549	0.431	0.341	0.520
Strategy: fund of funds	67	0.298	0.208	0.389	0.243	0.156	0.331
Strategy: multi	40	0.347	0.198	0.496	0.345	0.209	0.481
Strategy: other	77	0.383	0.288	0.477	0.385	0.309	0.462
Risk model							
1-factor model	167	0.461	0.393	0.530	0.420	0.358	0.482
3-factor model	71	0.463	0.344	0.581	0.446	0.347	0.545
4-factor model	205	0.247	0.183	0.311	0.313	0.250	0.376
7-factor model	298	0.289	0.237	0.342	0.297	0.254	0.340
Modeling model uncertainty	142	0.313	0.240	0.385	0.392	0.325	0.459
Asset-based model	80	0.324	0.255	0.392	0.250	0.185	0.314
Other model	56	0.933	0.799	1.067	0.906	0.773	1.039
Treatment of biases							
Survivorship treated	587	0.330	0.289	0.371	0.321	0.286	0.357
Backfilling treated	307	0.264	0.205	0.322	0.315	0.265	0.365
No bias treated	414	0.412	0.369	0.455	0.450	0.411	0.489
Some bias treated	605	0.329	0.289	0.368	0.318	0.284	0.352
Estimation technique							
IV method	46	0.463	0.340	0.586	0.435	0.302	0.568
non-IV method	973	0.358	0.327	0.388	0.364	0.337	0.390
All estimates	1,019	0.362	0.333	0.392	0.365	0.339	0.391

Notes: The table reports summary statistics for the different subsets of alpha estimates reported the literature. The definition of the the individual variables is available in Table B1. In the left panel, the individual alpha estimates are weighted equally. In the right panel, the alphas are weighted by the inverse of the number of estimates reported in a given study. Each panel shows the mean value and the 95% confidence interval.

large hedge funds, this finding suggests smaller hedge funds tend to generate larger alphas than larger hedge funds. We consider this finding rather intuitive and consistent with the prediction of a rational model of active portfolio management proposed by Berk & Green (2004) that assumes a differential ability to identify profitable investment opportunities across fund managers, but decreasing returns to scale in deploying these abilities. The model suggests that for any given level of a fund manager's ability, the likelihood of generating abnormal returns decreases in the volume of resources that these managers allocate. In other words, managers of smaller hedge funds may find it easier to implement their investment strategy because suitable investment targets are easier to identify when the scope of their investment is smaller. Hence, smaller hedge funds may outperform larger funds.

Table 3.2 also suggests that the reported alpha estimates greatly vary with the treatment of hedge fund fees in the research design of a primary study. Most of the primary studies from which we collect our data sample report alpha estimates on a net-of-fee basis. The unweighted (weighted) mean value of these net-of-fee alphas is 0.348 (0.335). In comparison, the alpha estimates based on gross returns are more than twice as large. Specifically, the unweighted (weighted) mean gross alphas is 0.757 (0.817). This finding is consistent with prior literature that points out the substantial fees that hedge funds charge (Connor & Woo 2004; Malkiel & Saha 2005). These fees typically consist of a flat management fee of 1% to 2% of assets under management (AUM) and a variable performance fee usually 20% of realized returns above the risk-free rate (Fung & Hsieh 1999; Connor & Woo 2004; Stulz 2007; Kouwenberg & Ziemba 2007; Getmansky *et al.* 2015). The performance fee tends to be paid only after reaching the so-called "high water mark", i.e., the minimum level of absolute performance over the entire investment lifetime (Asness *et al.* 2001; Goetzmann *et al.* 2003; Lim *et al.* 2016; Stulz 2007), that is to say, only after recovering any previously incurred losses. However, managers of unsuccessful hedge funds may opt to close the fund down, which renders any "high water mark" provision irrelevant (Stulz 2007). Our results show that unweighted net-of-fee returns account for 46 percent of gross returns while the weighted proportion is around 41 percent. The implied performance fee is slightly higher than 50 percent of gross returns, which is close to the estimation that the effective performance fees approach 64 percent of the aggregate gross profits in Ben-David *et al.* (2020).

Table 3.2 also shows that, on average, the magnitude of reported alpha estimates is not dramatically affected by the structure of the data used for the empirical tests in the primary studies. Both the cross-sectional and longitudinal data yield similar alpha estimates (0.355 and 0.401 on an unweighted basis and 0.379 and 0.324 on a weighted basis). The alpha estimates based on longitudinal data exhibit some difference between the simple unweighted and the weighted mean, which implies that the unweighted mean is affected by several studies that report high alphas.

Given the voluntary nature of reporting information on hedge funds, we consider it likely that prior empirical results might be affected by the choice of the database used by the researchers to obtain hedge fund performance data. Table B1 indicates that many of the primary studies are based on data from a single database. The mean number of databases that our alpha coefficients are based on is 1.366. This suggests only a limited overlap between data samples in various studies, and it underscores the benefit of aggregating and integrating prior empirical results on hedge fund performance based on these diverse samples.

Table 3.2 also reveals some differences in reported alpha estimates resulting from the use of databases used as a source of hedge fund performance data in individual primary studies. Most alphas in our sample are based on four commonly used databases: (1) Thomson/Refinitiv Lipper Hedge Fund (TASS), (2) Hedge Fund Research (HFR), (3) BarclayHedge, and (4) EurekaHedge database. TASS is a popular database in hedge fund research as it provides data starting from 1990, and it is available to many academics through their institutional data sources and libraries (e.g., Princeton University Library, Wharton Research Data Service of the University of Pennsylvania). HFR was established in 1992. It provides a detailed hedge fund strategy classification that is used in many studies that analyze the performance of various subsets of hedge funds. Prior empirical research also uses the multiple industry and regional hedge fund performance indices that are provided by HFR. BarclayHedge was founded in 1985. Returns on alternative investments and information on hedge fund performance are among the key data types provided in the database. EurekaHedge was established more recently in 2001, but it offers wider coverage of live hedge funds than the competing data providers. Hence, it is frequently used in empirical studies covering international hedge funds. More than half of our hedge fund alpha estimates (specifically 537) use one of the four main databases as a data source. Due to their popularity in prior empirical research, we classify into one category all alpha estimates that are based on the data sourced from these four main databases.

Furthermore, about one quarter of the alpha estimates in our sample (specifically, 256) are based on the data from the Dow Jones Credit Suisse Hedge Fund Index (formerly known as the Credit Suisse/Tremont Hedge Fund Index) (CST) database. Less than a fifth of alphas (specifically, 178) are based on the Morningstar Center for International Securities and Derivatives Markets database (CISDM), which is affiliated with the Isenberg School of Management, and it is also accessible through Wharton Research Data Service by many academic researchers. It provides data after 1994 but it is updated only twice a year. Our Data set also comprises 22 alpha estimates based on hand collected data End additional 167 estimates based on other than aforementioned databases.

Table 3.2 shows that alpha estimates based on the four most popular databases in hedge fund research are very close the the unconditional sample mean of 0.36 discussed above. The unweighted mean of alphas based on these four databases is equal to 0.343. When we weigh the alpha coefficients in our sample by the inverse of the number of estimates reported in a given primary study, we observe a slightly lower mean value of 0.301. Relative to the alphas based on the four most popular databases, estimates based on CST are somewhat lower (0.218 on an unweighted basis, and 0.273 on a weighted basis). In contrast, alphas based on CISDM, on other databases, and also those based on hand-collected data tend to be higher. These findings suggest that, except CISDM, alpha estimates based on less frequently used databases tend to be higher.

Table B1 also shows that 86% of alpha estimates are based on hedge funds that do not restrict the geographic scope of their investment, whereas 14% are based on funds focused on the developed markets as classified by the International Monetary Fund (IMF). Table 3.2 shows slightly better performance for funds that invest in developed markets as classified by the International Monetary Fund (IMF) relative to those that do not explicitly restrict their scope to a specific geographical location. This result may be considered surprising given that more developed markets may contain fewer mispriced assets and offer fewer opportunities to earn abnormal returns. Our regression results in Section 3.4 show that this difference is not statistically significant in a multivariate setting.

We also observe in Table 3.2 that alpha estimates based on "bear" (i.e., declining) markets are lower (mean values of 0.103 and 0.097 on the unweighted and weighted basis, respectively) than studies that concentrate on "bull" (i.e., rising) (mean values of 0.272 and 0.277 on the unweighted and weighted basis respectively). Thus, despite their name, hedge funds do not seem to hold investment positions that make their returns immune to general stock market movements (i.e., to be market-neutral).

Table 3.2 also exhibits some differences in the alphas generated by various hedge fund strategies. Our data set comprises 243 alpha estimates based on the data of all funds. The mean values in the most frequent category of alphas are slightly below the unconditional mean of 0.36 (mean values of 0.303 and 0.298 on the unweighted and weighted basis, respectively). Equity hedge funds constitute the largest category of specialized hedge funds. We collect 229 alpha estimates for this type of funds. The mean alphas in this category are very close to the unconditional mean of 0.36 (mean values of 0.352 and 0.360 on the unweighted and weighted basis, respectively). This suggests that equity the performance of equity hedge funds corresponds to the overall performance of all hedge fund categories.

In comparison, we observe that event-driven hedge fund strategies, on average, generate higher alphas (mean values of 0.474 and 0.586 on the unweighted and weighted basis, respectively), followed by global strategies (mean values of 0.451 and 0.431 on the unweighted and weighted basis, respectively). In contrast, reported alpha estimates based on the funds of funds tend to be lower (mean values of 0.298 and 0.243 on the unweighted and weighted basis, respectively). This finding may be driven by the additional layer of fees that are charged by the funds of funds or by the lower effect of the backfilling and survivorship biases that may inflate some of the alpha estimates based on the individual funds.

Table 3.2 also suggests that the choice of the risk model used in primary studies to adjust for the normal rate of returns may be consequential for the documented alphas. Fung & Hsieh (2001; 2004a) and Fung *et al.* (2008) observe that hedge funds typically exhibit risk exposures that are not typical for other asset classes, such as common equities and fixed-income securities. They propose a seven-factor model that reflects risk factors that are intended to capture risk dimensions that are relevant to common hedge fund investment strategies. The authors argue that the multiplicity of these risk dimensions makes the seven-factor model suitable for measuring abnormal returns across a wide range of hedge fund strategies. Being designed specifically for measuring hedge fund performance, the seven-factor model has been extensively used in prior empirical research. Table B1 shows that 29% of alpha estimates in our sample are based on the seven-factor model (e.g., Fung & Hsieh 2004a; Buraschi *et al.* 2014; Fung *et al.* 2008; Kosowski *et al.* 2007). Table 3.2 shows that the mean values of these alpha estimates are slightly below the unconditional mean of 0.36 (mean values of 0.289 and 0.297 on the unweighted and weighted basis, respectively).

Prior hedge fund research also frequently reports alpha estimates based on several other asset pricing models that are commonly used to measure abnormal returns. The Jensen (1968) alpha based on the Capital Asset Pricing Model (CAPM) (Sharpe 1966; Lintner 1965; Mossin 1966; Black 1972) uses the equity market excess return ($R_m - R_f$) as the sole risk factor. Conceptually, the intercept term alpha represents the abnormal return to a well-diversified investor. On the one hand, this approach is simple, well-founded in financial theory, and universally applicable. On the other hand, the assumptions this approach is based on may not be suitable for measuring the performance of hedge funds that engage in complex and dynamic investment strategies that are likely to exhibit various forms of exposure to systematic risk. In this respect, the three-factor, and the four-factor models capture additional risk dimensions that may not be easy to conceptualize in a financial modeling framework, but that may still be relevant to investors due to financial market imperfections and microstructure considerations (e.g., limited liquidity of traded assets).

Table B1 shows that 20% of alphas are based on the four-factor model (Eling & Faust 2010; Stoforos *et al.* 2017; Fung & Hsieh 2004b), 7% are based on the three-factor model (Dichev & Yu 2011; Ding & Shawky 2007), and 16% are based on the 1-factor model (Ranaldo & Favre 2005; Ding & Shawky 2007; Gupta *et al.* 2003). Table 3.2 shows that, relative to the seven-factor model, the alpha coefficients estimated with the use of pricing models using fewer risk factors are typically higher. The difference is particularly pronounced for the one-factor (mean values of 0.461 and 0.420 on the unweighted and weighted basis, respectively) and three-factor models (mean values of 0.463 and 0.446 on the unweighted and weighted basis, respectively). In contrast, the alpha coefficients based on the four-factor model (mean values of 0.247 and 0.313 on the unweighted and weighted basis, respectively) are comparable to the ones based on the seven-factor model. Furthermore, we also observe rather high values for the 56 alpha coefficients reported in primary studies that use other pricing models (mean values of 0.933 and 0.906 on the unweighted and weighted basis, respectively). As the choice of the pricing model is likely related to other research design choices, we delay drawing stronger conclusions from these findings to Section 3.4, where we examine the effect of these conditioning factors in combination.

Prior research frequently mentions concerns that the measurement of hedge fund performance may be distorted by the survivorship and backfilling biases (Fung & Hsieh 2000; 2002; 2004a; Fung *et al.* 2008). Table 3.2, indeed, shows that the alpha estimates reported in primary studies tend to be higher when the authors do not explicitly adjust for the survivorship and backfilling bias (mean values of 0.412 and 0.450 on the unweighted and weighted basis, respectively) as compared to the alpha estimates, for which at least one of the biases is addressed (mean values of 0.329 and 0.318 on the unweighted and weighted basis, respectively). This finding suggests that the commonly voiced concerns about the impact of those biases are indeed warranted, and they may indeed have a substantial impact on the inferences about hedge fund performance.

Finally, we also observe some variation in the magnitude of reported alphas depending on the use of various estimation techniques. Brodeur *et al.* (2020) suggests that estimates based on instrumental variable (IV) techniques often exhibit greater publication selection bias. They argue that using IV gives researchers an additional layer of discretion because the pool of potentially relevant instruments is rather broad. Thus, researchers may choose instruments that yield results that support their *a priori* predictions or that are otherwise attractive for publication. This approach may induce a greater selectivity in coefficients that eventually get published. Consistent with this proposition, Table 3.2 shows that alphas reported in the primary studies tend to be higher when estimated based on IV (mean values of 0.463 and 0.435 on the unweighted and weighted basis, respectively) relative to those estimated using other techniques (mean values of 0.358 and 0.364 on the unweighted and weighted basis, respectively). However, the IV-based estimates are less precise so the 95% confidence intervals are rather wide, and they include the unconditional mean of 0.36 both on the unweighted basis (0.340, 0.586) and on the weighted basis (0.302, 0.568).

We also consider the impact factor of the journal where a given primary study is published and the number of times it is cited in the research literature as additional potentially relevant explanatory variables. The impact factor of a research journal and the number of citations can both be seen as proxies of publication quality. We expect studies published in more impactful journals and those that are more frequently cited to be more influential in shaping the public perception of the value generated by hedge funds. Table B1 shows that the mean number of databases that Primary studies, from which we source our alpha estimates, are published in research journals with the average discounted recursive impact factor by Research Papers in Economics (RePEc) of 4.0, and they are on average cited 5.9 times (=exp(1.773)).

Figure 3.4 shows histograms for subsets of reported alpha estimates with specific characteristics related to the estimation method, sources of data, and hedge fund strategies. Panel (a) shows the greater dispersion of alpha estimates based on IV. Panel (b) depicts the dispersion of alphas based on hedge funds that concentrate on developed markets. Panel (c) shows lower alpha estimates based on value-weighted hedge fund indices. Panel (d) depicts lower alphas reported after explicitly adjusting for the survivorship and/or backfilling biases. Panels (e) and (f) show the distribution of alpha coefficients for the various asset pricing models and the various hedge fund investment strategies.

3.3.5 Methodology

Since prior research does not provide clear guidance about the nature of hedge fund performance determinants, we treat the variables described in Table B1 as potentially relevant for explaining the heterogeneity in hedge fund performance. We examine their explanatory power using the Bayesian Model Averaging (BMA) technique. BMA considers various combinations of variables and evaluates their relevance for explaining the variation in the dependent variable. Explanatory variables that are consistently associated with the dependent variable across a multitude of regression model specifications are then identified as relevant for explaining it. BMA allows researchers to address the model uncertainty problem and to consider a fairly large number of potentially relevant variables while avoiding multi-collinearity issues that naturally arise when numerous similar variables are included in a single regression specification.

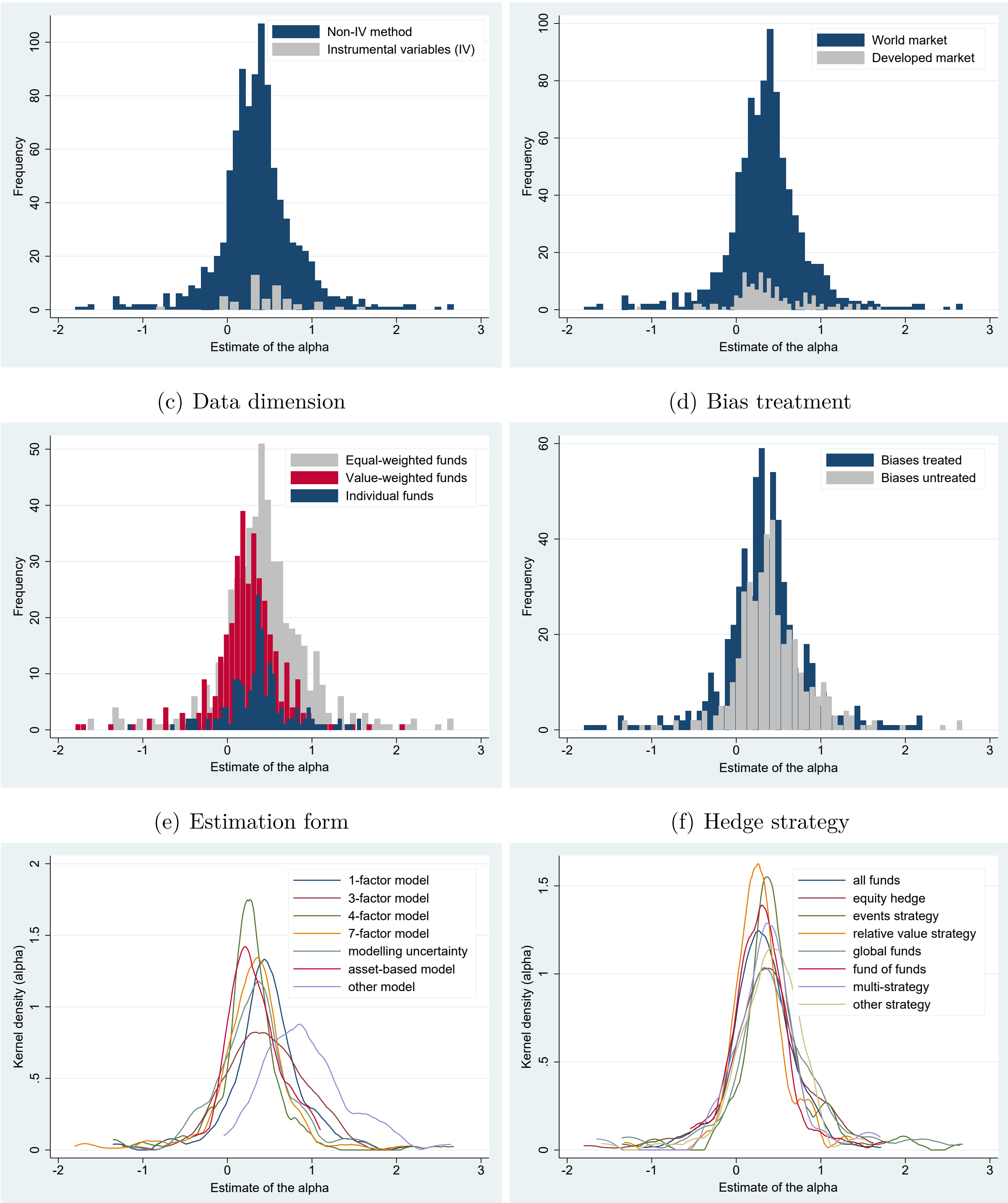
In BMA, the ability of the individual variables to explain the variation in the dependent variable is measured by their posterior inclusion probability (PIP). PIP close to 1.0 indicates that a particular variable is present in most regression models that are effective in explaining the variation in the dependent variable. In contrast, PIP close to 0.0 indicates low explanatory power of a given variable across various regression specifications. To interpret our results, we follow Jeffreys (1961) and Raftery (1995), who propose cutoff levels for PIP that can be used to evaluate how relevant a given variable is for explaining the variation in the dependent variable. They argue that PIP greater than 0.99 indicates that the variable is "decisive" for explaining the variation in the dependent variable, PIP greater than 0.95 suggests that the variable has a "strong" effect, variables with PIP

Figure 3.4: Distribution of Alphas with Specific Characteristics

(a) Method usage (b) Development level

(c) Data dimension (d) Bias treatment

(e) Estimation form (f) Hedge strategy



Notes: The figure shows histograms for subsets of reported alpha estimates with specific characteristics related to the estimation method, sources of data, and hedge fund strategies. We use the IMF definition to classify countries as developed or developing.

greater than 0.75 can be considered to have an effect on the dependent variable, and variables with PIP greater than 0.50 to have a “weak” effect. We use these cutoff levels in interpreting our empirical results.

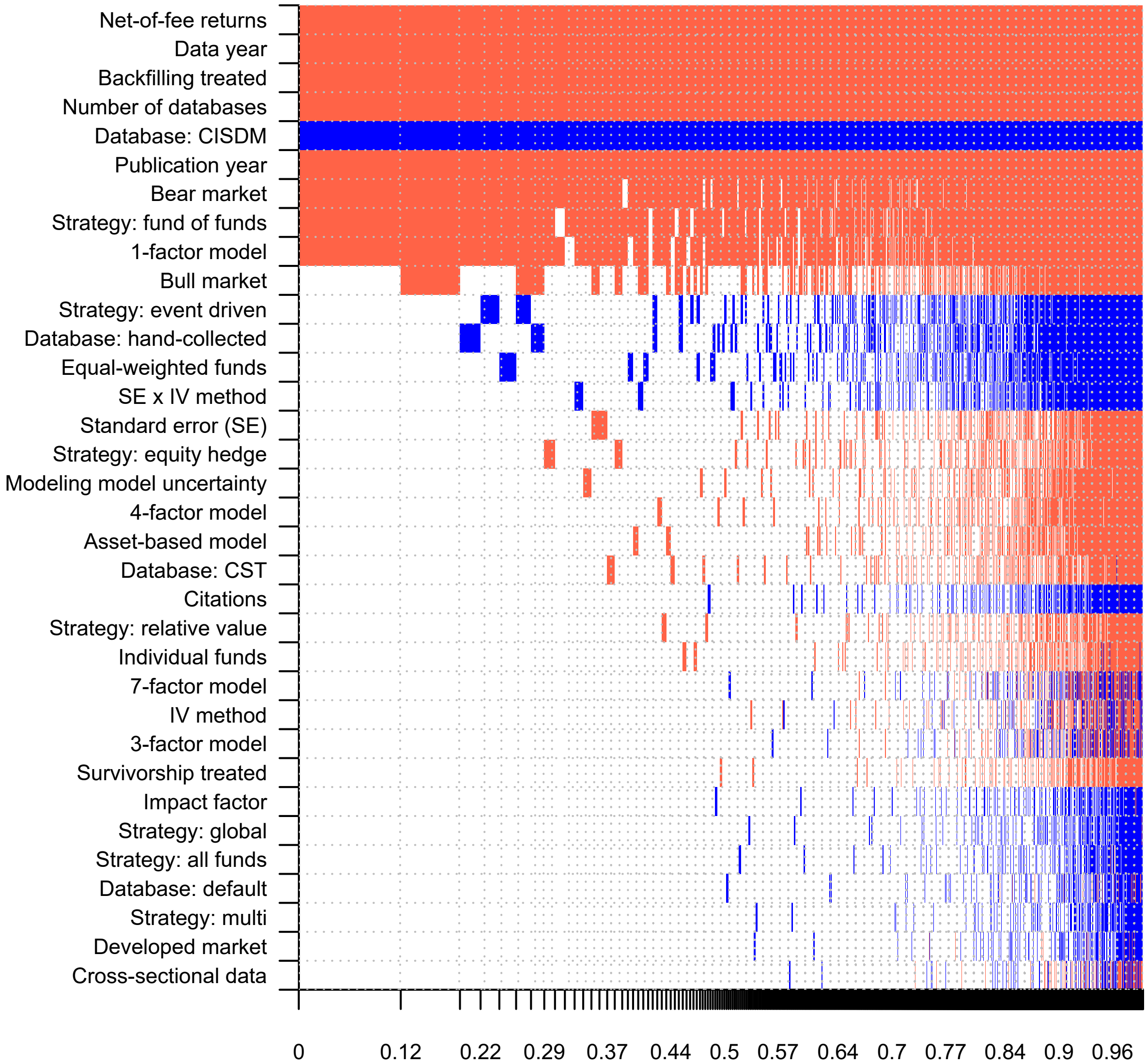
3.4 Results

3.4.1 Main Regression Results

Figure 3.5 provides a visualization of our results from the BMA. The columns in the figure denote alternative regression specifications that feature various combinations of explanatory variables. BMA weighs the individual regression models by their posterior model probabilities, which measure the models’ goodness of fit. The posterior model probability is represented by the width of individual columns. We order the models based on their posterior probability so that the models with the best fit are represented by the widest columns in the left part of Figure 3.5. Similarly, explanatory variables in individual rows are sorted based on their PIP with the most relevant variables listed at the top. The nature of the association between an explanatory variable and the dependent variable in a given regression model is depicted by the color of the corresponding cell. A blue cell (equivalent to darker shading in grayscale) implies a positive impact of a given explanatory variable on hedge fund alphas in a particular regression specification and a red cell (lighter in grayscale) denotes a negative sign of the estimated coefficient. Blank cells represent variables that are not included in a given regression model.

Figure 3.5 features 34 potential explanatory variables that reflect differences in hedge fund types, various aspects of research design, and data samples used in primary studies, as well as their publication characteristics. Figure 3.5 shows that most of the considered explanatory variables exhibit a consistently positive or consistently negative association with hedge fund alphas across all model specifications. This implies that these associations are robust to the inclusion of additional explanatory variables. Figure 3.5 also shows that the model with the best fit (i.e., the one with the highest posterior probability based on the BMA) features only 9 of the 34 considered variables. All of these 9 explanatory variables exhibit a consistent sign in all the models that comprise them. Our model with the best fit thus suggests that the reported alpha coefficients tend to be lower when hedge fund returns are computed net of the fees hedge funds charge their investors, when the backfilling bias is treated in the primary study, and when estimated based on the 1-factor model. Furthermore, we observe that primary studies report lower alphas when the estimation is based on data from a larger number of databases, and when the CISDM database is not used as a data source. The alphas also tend to be lower for the funds of funds and in bear markets. Remarkably, there seem to be strong negative associations between the reported alpha coefficients and both the data year and

Figure 3.5: Bayesian Model Averaging Visualization



Notes: This figure provides a visualization of our results from the BMA. On the vertical axis the explanatory variables are ranked according to their posterior inclusion probabilities from the highest at the top to the lowest at the bottom. The horizontal axis shows the values of cumulative posterior model probability. Blue color (darker in grayscale) denotes that the estimated parameter of a corresponding explanatory variable is positive in a given regression specification. Red color (lighter in grayscale) shows that the estimated parameter of a corresponding explanatory variable is negative. No color indicates that the corresponding explanatory variable is not included in the model. Numerical results are reported in Table 3.3. All variables are described in Table B1.

the year of publication. In particular, studies that use datasets with later-year midpoints, and those that are published more recently report lower alphas. We further discuss this finding below.

To quantify the magnitude and the variability of regression coefficients represented in Figure 3.5 by the blue or the red color, BMA expoits the characteristics of the distribution of coefficients generated by estimating various regression models. In Table 3.3, we report the posterior mean of the distribution of regression coefficients (P. mean), which represents the typical value of the regression coefficient across various regression specifications. Furthermore, the standard deviation of the posterior coefficient distribution (P. SD) shows how the estimated coefficients vary across various regression specifications. Table 3.3 also specifies PIP of individual variables, which indicates how likely each variable is to be present in the "true" explanatory model. We use the P. mean, P. SD, and PIP as the main measures to quantify the effect of the individual explanatory variables on the reported hedge fund alphas. To check the robustness of the BMA results, we also report in the right panel of Table 3.3 the ordinary least squares (OLS) estimates, which are based on a single regression model including a set of the most relevant variables identified by BMA. The frequentist OLS approach is better comparable to prior study results and thus may enhance the understanding of the effect of these relevant factors. For the OLS estimation, we report the regression coefficients for the individual explanatory variables (Coef.), their standard errors (SE), and the corresponding p-values.

The numerical results presented in Table 3.3 show that the PIP for the nine explanatory variables included in the model with the best fit range between 0.876 and 1.000, which suggests that all of these variables are important for explaining the variation in reported hedge fund alpha coefficients. Out of these nine PIPs, six are above the 0.950 cutoff that is commonly interpreted as denoting a "strong" effect of the corresponding variable. Furthermore, the PIP for the remaining three indicator variables denoting the alpha coefficients (i) estimated based on a 1-factor model, (ii) reported for funds of funds, and (iii) estimated for bear markets only are equal to 0.876, 0.905, and 0.931. This implies that all of these coefficients remain comfortably above the 0.750 cutoff proposed for the existence of a relationship between the explanatory and the dependent variables. In the OLS regression, all of the nine variables included in the BMA model with the best fit are also significant at a better than 5% level. The corresponding p-values range between 0.000 and 0.019. Thus, both the BMA and OLS estimates provide evidence in support of the relevance of these nine variables to explain the variation in hedge funds alphas. In contrast to the nine variables included in the BMA model with the best fit, the PIP of all the remaining variables is below 0.35, which indicates that they are unlikely to be relevant for explaining the variation in the alpha coefficients reported in prior empirical research. Thus, our results identify nine key characteristics that are essential for explaining the heterogeneity in reported alphas. And the R-squared in the OLS regression indicates that these nine variables explain 23% of the variability.

Six out of these nine explanatory variables are indicators that take the value of zero or one. We can thus easily compare the magnitude of the corresponding coefficients. We observe the largest coefficient for the variable denoting alpha estimates computed net of hedge fund fees. This finding is consistent with prior research that suggests that the effect of hedge fund fees on the return generated for investors may indeed be rather substantial. Hedge funds typically charge a flat management fee of 1% to 2% of AUM and

Table 3.3: Regression Results

	Bayesian model averaging			Ordinary least squares		
Variable:	P. mean	P. SD	PIP	Coef.	SE	p-value
Constant	1.851	NA	1.000	1.858	0.113	0.000
Standard error (SE)	-0.008	0.030	0.085			
SE * IV method	0.065	0.233	0.107			
Dependent variable						
Individual funds	-0.003	0.017	0.041			
Equal-weighted funds	0.010	0.026	0.144			
Net-of-fee returns	-0.439	0.075	1.000	-0.439	0.081	0.000
Data characteristics						
Cross-section data	0.000	0.010	0.013			
Data year	-0.248	0.031	1.000	-0.239	0.034	0.000
Database: default	0.000	0.006	0.016			
Database: CST	-0.004	0.020	0.060			
Database: CISDM	0.224	0.049	0.998	0.236	0.064	0.000
Database: hand-collected	0.040	0.093	0.180			
Number of databases	-0.085	0.017	0.999	-0.085	0.022	0.000
Structural variation						
Developed markets	0.001	0.008	0.014			
Bull market	-0.067	0.106	0.323			
Bear market	-0.264	0.104	0.931	-0.280	0.072	0.000
Hedge fund strategy						
Strategy: all funds	0.001	0.007	0.017			
Strategy: equity hedge	-0.006	0.021	0.085			
Strategy: event driven	0.019	0.043	0.182			
Strategy: relative value	-0.003	0.018	0.041			
Strategy: global	0.001	0.008	0.018			
Strategy: fund of funds	-0.181	0.079	0.905	-0.202	0.050	0.000
Strategy: multi	0.001	0.011	0.014			
Estimation technique						
IV method	-0.006	0.045	0.031			
1-factor model	-0.142	0.074	0.876	-0.159	0.068	0.019
3-factor model	-0.003	0.032	0.026			
4-factor model	-0.009	0.048	0.070			
7-factor model	-0.004	0.039	0.032			
Modeling model uncertainty	-0.012	0.054	0.075			
Asset-based model	-0.010	0.052	0.068			
Survivorship treated	-0.001	0.008	0.023			
Backfilling treated	-0.196	0.034	1.000	-0.198	0.062	0.001
Publication characteristics						
Publication year	-0.126	0.028	0.994	-0.138	0.056	0.013
Citations	0.001	0.007	0.046			
Impact factor	0.000	0.001	0.021			
Observations	1,019			1,019		
Studies	74			74		
R2				0.23		

Notes: The table show the main results based on BMA (left panel) and ordinary least squares (OLS) regression that includes the nine explanatory variables identified by the BMA as most relevant for explaining the variation in reported hedge fund alphas (right panel). The response variable are the hedge fund alpha estimates. P. mean represents the posterior mean of the distribution of regression coefficients. P. SD represents the posterior standard deviation of the distribution of regression coefficients. PIP denotes the posterior inclusion probability of a given variable in the "true" explanatory model. Coef. denotes the slope coefficient based on OLS. SE shows the standard error of the slope coefficient in the OLS regression model. The p-value show the probability of obtaining the result for a given explanatory variable under the assumption that the variable has no explanatory power (i.e. the null hypothesis is correct). BMA employs uniform model prior (Eicher *et al.* 2011) and dilution prior suggested by George (2010), which accounts for collinearity. The frequentist check (OLS) includes the variables recognized by BMA to comprise the best model and is estimated using standard errors clustered at the study level. All variables are described in Table B1.

a variable performance fee, usually 20% of realized returns above the risk-free rate (Fung & Hsieh 1999; Connor & Woo 2004; Stulz 2007; Kouwenberg & Ziemba 2007; Getmansky *et al.* 2015). The performance fee is usually paid only after reaching the so-called "high water mark," i.e., the minimum level of absolute performance over the entire investment lifetime (Asness *et al.* 2001; Goetzmann *et al.* 2003; Lim *et al.* 2016; Stulz 2007). Due to the conditional nature of some of these fees, their effective impact on the value hedge funds generate for their investors is not trivial to quantify. Less successful hedge funds are more likely terminated, and investors cannot offset gains and losses across various hedge funds. Ben-David *et al.* (2020) estimates that, on average, hedge funds appropriate in fees almost two-thirds of the excess return they generate.

Our approach offers an alternative way of estimating the effective fees paid by hedge fund investors. We exploit the composition of our sample that includes both alphas estimated using gross returns and alphas estimated net of fees. Thus, we are able to quantify the effect of hedge fund fees after controlling for all other relevant characteristics that affect the magnitude of reported alpha estimates. Our results suggest that after adjusting for all other relevant hedge fund alpha determinants, we observe that monthly alphas estimated on the net-of-fee basis are by -0.439 lower than those estimated gross of fees. We also observe that the magnitude of this coefficient is essentially identical to the one obtained based on the OLS that we report in the right panel of Table 3.3. These findings suggest that the combined effect of the management and performance fees is indeed large. Hedge funds seem to charge their investors more than 5.0% annually.

Considering the slope coefficients at other indicator variables we observe that the hedge fund alphas for bear markets are by -0.246 lower relative to our benchmark case. The magnitude of this coefficient is very similar to the one based on the OLS estimation of -0.280. There is a controversy in prior research literature about the relative performance of hedge funds in bull and bear markets. Our results suggest that hedge funds generate substantially lower alphas when stock market prices decline when they seem to underperform their typical performance by about -3.2% per annum. Thus, despite their name, hedge funds do not seem to hold investment positions that make their returns immune to general stock market movements (i.e., to be "market-neutral").

We also observe that the alpha coefficients reported in primary studies are related to the databases, from which the primary data are sourced. Specifically, we document that the reported alpha estimates tend to be lower when based on more source databases. We observe virtually identical coefficients corresponding to the inclusion of one additional database for obtaining the primary study sample based on the BMA and OLS (in both cases, -0.085 after rounding). Both of these coefficients are highly statistically significant. These findings suggest that using more comprehensive datasets tends to be associated with lower reported hedge fund performance estimates. Furthermore, alphas based on samples that include the CISDM as one of the source databases tend to be higher. The posterior mean coefficient based on the BMA is equal to 0.224, and it is very close to the OLS estimate of 0.236. Thus, researchers should be aware that alpha estimates based on the CISDM database tend to overstate hedge fund performance relative to studies based on the other databases.

Furthermore, we find that the alpha estimates reported in prior research tend to be lower when estimated for the funds of funds rather than for the individual hedge funds and when explicitly adjusted for the backfilling bias. The posterior mean of the coefficient at the indicator variable denoting alphas estimated for the funds of funds -0.181, which is fairly comparable to the corresponding slope coefficient based on the OLS estimation of -0.202. Similarly, alpha estimates explicitly adjusted for the backfilling bias are lower by -0.196, which is again comparable to the slope coefficient of -0.198 based on the OLS. These findings suggest that the backfilling bias and the selection biases addressed by estimating performance for funds of funds are indeed rather consequential for the reported results. Thus, the frequently voiced concerns that data biases in some prior empirical studies may affect inferences about hedge fund performance are indeed warranted.

Finally, we document strong negative associations between the magnitude of reported alpha coefficients on the one hand and the mid-year of the data sample and the publication year on the other. Increasing the data midpoint year by one, on average, reduces reported alphas by -0.248 based on BMA or by -0.239 based on OLS. In both cases, this result is strongly statistically significant, with the PIP approaching 1.0 and the p-value below 1%. In addition, studies published more recently also report lower alphas. Specifically, increasing the year of publication by one is associated with a reduction in reported alpha estimates by -0.126 based on the BMA or by -0.138 based on the OLS. The effect of the publication year is incremental to the effect of the data sample mid-year discussed above. These findings imply that studies based on newer datasets and studies published more recently tend to report substantially lower hedge fund alpha estimates. This suggests that hedge fund performance has substantially decreased over time. In Subsection 3.4.3, we further elaborate on these findings, and we show that due to the declining time trend, the current estimate of hedge fund performance is not reliably different from zero.

We further observe that the absolute value of the posterior mean of all the other indicator variables that are not included in the model with the best fit as identified by BMA is below 0.070. This implies that their effect on hedge fund performance is less than 1% per annum. In other words, it seems that the BMA approach identified nine variables relevant for explaining variation in hedge fund alphas. The effect of other variables is likely to be fairly marginal.

3.4.2 Sensitivity Analysis

Notwithstanding the BMA's advantages for analyzing research questions where the set of potential explanatory variables is not *a priori* given, the BMA method may be affected by the priors used as a point of departure for Bayesian estimation. To investigate how robust our results are to the modification of these priors, we recompute them using several different priors proposed in prior literature. We examine the extent to which the use of different priors alters our inferences about the power of the individual variables to explain the variation in the alpha coefficients reported in the primary studies on hedge fund performance.

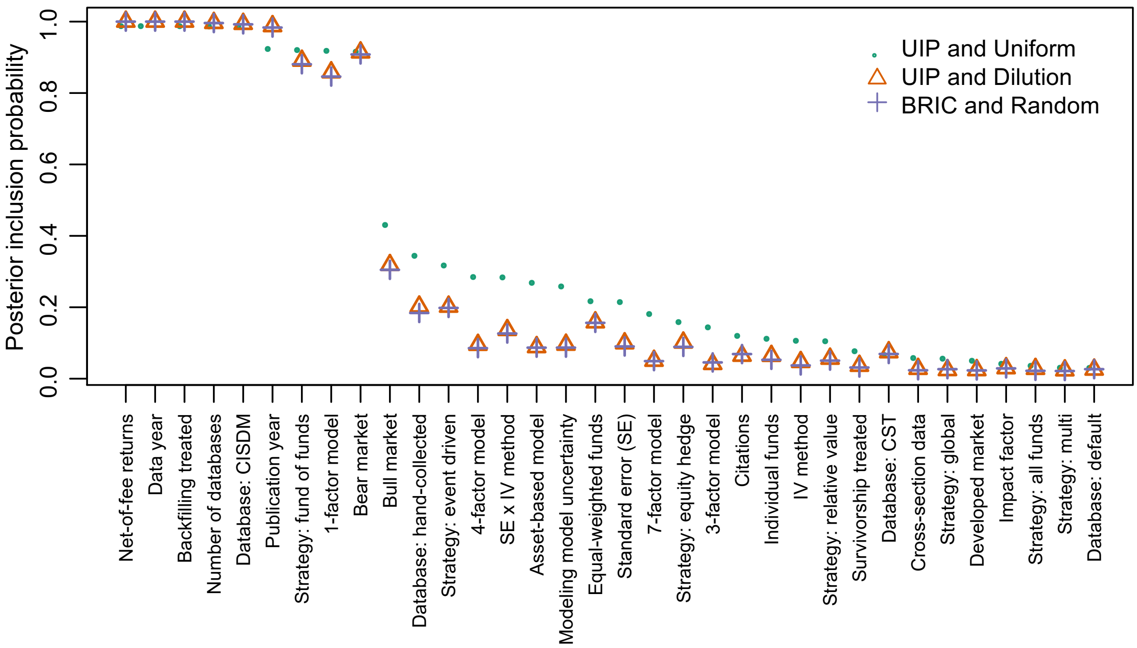
Figure 3.6 depicts the results of our sensitivity analysis. Again, we order the individual explanatory variables based on their estimated relevance in our main test. Figure 3.6 indicates that the choice of priors in the BMA is indeed somewhat relevant for the numerical values of our results. Nevertheless, the use of different priors does not dramatically alter our main conclusions that we discuss above. For most of the explanatory variables, the estimates based on different priors are placed rather close to one another, which implies that a different choice of priors would not dramatically alter the inferences about the prominence of the nine key explanatory variables that we identify as fundamental for explaining the variation in the reported alpha coefficients.

We observe that the unit information priors (UIP) and dilution priors that are recommended by George (2010) produce virtually identical estimates as the BRIC and random that represent a g-prior proposed by Fernandez *et al.* (2001). In comparison, the UIP and uniform priors recommended by Eicher *et al.* (2011) yield slightly higher estimates for most of the variables that are not included in our BMA model with the best fit. Nevertheless, the sizable gap in relevance between the nine explanatory variables included in our model with the best fit and the remaining variables clearly stands out regardless of the set of priors we use. Thus, we conclude that our findings are fairly robust to the choice of priors in our BMA estimation.

3.4.3 Best Practice Estimate

In Subsection 3.4.1 we analyze the impact of variables that can potentially explain the variation in the hedge fund alpha estimates reported in the primary studies. In this section, we provide an estimate of current hedge fund performance based on the best practices of estimating alphas. Below, we motivate our choice of methodological approaches that we believe constitute the best practices in this field of research. Even though the choice of these parameters inevitably involves a subjective judgment, we closely follow arguments raised in the research discourse on the appropriate methodology and its limitations in research on hedge fund performance. Based on these arguments, we set the corresponding variables in our empirical model to values that we argue constitute the best

Figure 3.6: Sensitivity of BMA to Different Priors



Notes: This figure shows the sensitivity of our results on the relevance of the individual variables for explaining the variation in the alpha coefficients reported in the primary studies to the various priors used in BMA. UIP stands for the unit information priors. UIP and Uniform represent the priors recommended by Eicher *et al.* (2011). UIP and Dilution represent the priors recommended by George (2010). BRIC and Random represent a g-prior proposed by Fernandez *et al.* (2001) for parameters with the beta-binomial model prior (Ley & Steel 2009) for model space; this ensures that each model size has equal prior probability.

practices for this estimation. We believe that the best practice approach likely generates the most reliable alpha estimates that are relevant for current investment decisions. Below, we discuss and motivate our choices concerning the individual variables.

First, we argue that an ideal study on hedge fund performance should be free of data and publication biases. In the results discussed above, we document a substantial impact of the backfilling bias in estimating hedge fund alphas. In our best practices model, we thus plug in one for the indicator variable that captures that the survivorship and backfilling biases are treated in the design of the primary study. Furthermore, Yang *et al.* (2023) investigate the impact of the publication selection bias in hedge fund research. Following this study, we plug zero for the measure of a hedge fund alpha's standard error and also for the respective interaction term in our best practices model. This treatment ensures that our best practices estimate is free of any publication selection bias in the primary studies from which we source our dataset.

Second, we argue that from an investor's perspective, it is relevant to measure hedge fund alphas net of any management and performance fees. Fees retained by the hedge funds do not constitute realized returns that accrue to investors. Therefore, any portion of return generated by hedge funds that is retained in the form of fees should be irrelevant for computing hedge funds' effective performance from investors' perspective. Prior research argues that these fees can indeed be rather substantial (Ben-David *et al.* 2020). Consistent with these propositions, our results also suggest a sizeable difference between the alpha coefficients estimated on the gross basis and those that are net of all fees. Hence, in our best practices model, we plug in one for the variable, indicating that the corresponding alpha is estimated on the net-of-fees basis.

Third, we expect investors to be particularly interested in the most recent estimates of hedge fund performance that likely closely reflect the investment opportunities that are currently available. The hedge fund industry has undergone substantial development over time. The number of funds and the value of resources they manage has surged over the past decades (Stulz 2007). These days, more hedge funds compete to identify profitable investment opportunities and attract investors. The more intensive competition likely impacts the returns that hedge funds are able to generate today relative to the returns they generated in the past. Furthermore, the greater regulation of the hedge fund industry may have also limited their ability to generate superior returns to investors (Shi 2017; Cumming *et al.* 2020; Aragon *et al.* 2013). We thus set both the data year and the year of publication to the maximum values these measures have in our sample.

Finally, we expect investors to value studies that are well-published and well-cited. Therefore, we plug in our sample maxima for the impact factor and the number of citations. We set the remaining variables to their sample means.

Table 3.4 shows our best practice estimates of alpha coefficients jointly for all hedge fund types, as well as the separate expected alphas for the individual hedge fund types. Table 3.4 shows that, relative to the unconditional sample mean of 0.36 discussed above, the overall best practice estimate based on all hedge fund returns is small and negative, i.e., -0.079. The corresponding 95% confidence interval is fairly wide (-0.393, 0.235), and it includes zero. Thus, judging based on the most up-to-date best practice alpha estimate, we cannot reject the null hypothesis that hedge funds currently generate no abnormal after-fee return for their investors.

The remaining rows in Table 3.4 report the best-practice estimates for eight main hedge fund investment strategies. Similar to the overall best-practice alpha estimate computed for the pooled sample of all hedge funds, all eight alpha estimates for the individual hedge fund types are negative, and the corresponding 95% confidence intervals all include zero. Hence, based on our evidence we are unable to document reliably positive alphas for any of the common hedge fund investment strategies. These findings suggest that after controlling for methodological imperfections and after considering the trend over time in the reported alpha estimates, no type of hedge funds generates reliably positive after-fee abnormal returns for investors.

We observe the most negative alpha estimate of -0.249 for the funds of funds. The 95% confidence interval for this approach of measuring hedge fund performance is also fairly wide (-0.590, 0.092), which prevents us from drawing stronger inferences. However, we

Table 3.4: Best Practice Estimates

	Mean return	95% conf. int.	
All strategies	-0.079	-0.393	0.235
Strategy: all funds	-0.067	-0.372	0.237
Strategy: equity hedge	-0.073	-0.418	0.271
Strategy: event driven	-0.049	-0.376	0.277
Strategy: relative value	-0.071	-0.380	0.238
Strategy: global	-0.067	-0.391	0.257
Strategy: fund of funds	-0.249	-0.590	0.092
Strategy: multi	-0.067	-0.368	0.235
Strategy: other	-0.068	-0.373	0.238

Notes: The table shows the best practice alpha estimates from our BMA model for the hedge funds in general and for the individual hedge fund strategies. The mean return represents the expected alpha coefficient conditional on the inputted values of explnatory variables that we consider to represent the best practice in hedge fund performance research. We provide the motivation for our choices in the main body text. The 95% confidence intervals in parentheses are constructed using the standard errors estimated by OLS with standard errors clustered at the study level.

observe that the confidence interval approaches being entirely below zero, which would indicate a reliably negative after-fee abnormal return. Estimating the alphas for the funds of funds may be viewed as one of the ways of correcting for the survivorship bias in hedge fund data. Hence, fund of funds' returns may constitute a realistic estimate of hedge fund performance plausibly achievable for investors. Furthermore, investing in the funds of funds might seem attractive for investors who want to diversify away some of the risks they take by investing across several hedge funds. Nevertheless, investing in funds of funds also entails another layer of fees. Our evidence suggests that the best-practice estimate for the fund of funds' abnormal return is indistinguishable from zero, and it approaches being significantly negative.

Table 3.5 shows the economic significance of key variables included in our best-practice model. The table provides insights about the relative importance of these variables for our quantification of the best best-practice estimates of hedge fund alphas. The left panel of Table 3.5 shows how one-standard-deviation change in a given explanatory variable effects the best-practice alpha estimate both in absolute terms and as a percentage of the best-practice estimate. In the right panel, we show the corresponding change in the best-practice alpha estimate that would result from a change in a given explanatory variable from its minimum to the maximum value in our sample.

Consistent with our previous analysis, Table 3.5 shows that several explanatory variables have a substantial impact on the magnitude of the best-practice alpha estimate. We observe the largest effect for the midpoint year in the dataset used in a given primary study. Increasing the data sample midpoint year by one standard deviation reduces the monthly alpha estimate by -0.149 percentage points. Alternatively, after having controlled

Table 3.5: Economic Significance of Key Variables

	One-std.-dev. change		Maximum change	
	Effect on σ	% of best practice	Effect on σ	% of best practice
Net-of-fee returns	-0.080	101%	-0.439	557%
Data year	-0.149	189%	-0.802	1,017%
Database: CISDM	0.085	-108%	0.224	-284%
Number of databases	-0.089	112%	-0.592	750%
Bull market	-0.013	16%	-0.067	85%
Bear market	-0.051	64%	-0.264	335%
Strategy: all funds	0.000	0%	0.001	-1%
Strategy: equity hedge	-0.002	3%	-0.006	7%
Strategy: event driven	0.006	-7%	0.019	-23%
Strategy: relative value	-0.001	1%	-0.003	4%
Strategy: global	0.000	0%	0.001	-1%
Strategy: fund of funds	-0.045	57%	-0.181	229%
Strategy: multi	0.000	0%	0.001	-1%
1-factor model	-0.053	67%	-0.142	180%
Survivorship treated	0.000	1%	-0.001	1%
Backfilling treated	-0.090	114%	-0.196	248%
Publication year	-0.090	114%	-0.384	487%

Notes: The table shows the results of our analysis of the economic significance of key variables included in our best-practice model. The left panel quantifies how much one standard deviation change in a given explanatory variable affects the best-practice alpha estimate both in absolute terms and as a percentage of the best-practice estimate. The right panel shows the corresponding change in the best-practice alpha estimate resulting from changing the value of the explanatory variable from its sample minimum to its sample maximum. A detailed description of the variables is available in Table B1.

for all study and hedge fund characteristics, the best practice-alpha estimates based on the oldest and the most recent dataset differ by -0.802 percentage points. Furthermore, we also document a substantial effect of the year of publication. A one-standard deviation increase in the publication year is associated with a reduction in the practice-alpha estimates by -0.090 percentage points. The most recent studies in our sample report alpha estimates that are *ceteris paribus* lower by -0.384 percentage points relative to the oldest studies in our sample. These findings provide strong evidence suggesting that the abnormal returns generated by hedge funds decreased over time.

Table 3.5 also underscores the importance of the data sources and method choices for the magnitude of the best-practice alpha estimates. *Ceteris paribus*, increasing the number of databases used in a primary study by one tends to be associated with a reduction in the best-practice alpha estimate by -0.089 percentage points. The more comprehensive studies that pool their data from several source databases may be more effective in covering the complete universe of all existing hedge funds. Hence, their conclusions may be more representative of the entire hedge fund population. Hence, the number of source databases may be viewed as one aspect of a study's quality. We document that more comprehensive studies tend to report lower alphas.

Furthermore, adjusting for the backfilling bias, on average, reduces the alpha estimates by -0.196 percentage points (for indicator variables, we interpret the change from the minimum value of zero to the maximum value of one). In a similar vein, computing the alphas for the funds of funds implies a reduction in the estimate by -0.181 percentage points. Finally, using a 1-factor risk model is *ceteris paribus* associated with best-practice alpha estimates that are lower by -0.142 percentage points. Since the 1-factor risk model may not be able to effectively adjust for the systematic risk that the hedge fund strategies entail, using more complex models may also be viewed as an indication of a study's quality.

Finally, Table 3.5 also documents a substantial effect of adjusting for hedge fund fees and of limiting the estimation on bear markets, which we discuss above. Overall, the quantification of the effect indicates that the above-discussed variables indeed have an economically substantial effect on the best-practice estimates of alpha coefficients.

3.5 Conclusion

We analyze prior empirical evidence on hedge fund performance published in academic journals between 2001 and 2021. In recent years, the amount of capital in the economy allocated by hedge funds has surged. Their growing economic prominence, as well as the macroeconomic impact of some of their notorious failures, prompted calls for greater insight into the determinants of their performance. Measuring the value hedge funds generate for their investors is complicated by data fragmentation resulting from the voluntary nature of many hedge fund disclosures and the plurality of estimation approaches used in prior empirical research. To aggregate and synthesize this pool of diverse empirical results, we conduct a meta-analysis of 1,019 alpha coefficients from regressions of hedge fund returns on risk factors collected from 74 studies. We examine how the reported alpha estimates vary over time and across hedge fund characteristics, and we study how they are affected by research design choices in the primary studies.

We show that the value generated by hedge funds is substantially diminished by the fees they charge. Furthermore, we document a strong declining trend in the reported hedge fund alphas over time. Our best practices alpha estimates of current hedge fund performance are not reliably different from zero. Furthermore, when we classify hedge funds into common categories based on the nature of their investment strategies, we observe that the best-practice estimate of their current performance is not significantly positive for any of these categories. All of these estimates are negative. For one of the categories – the fund of funds – the 95% confidence interval approaches being fully below zero.

In addition, we identify several research design characteristics that affect the reported alphas. The published alpha estimates tend to be lower (i) when adjusted for the backfilling bias, (ii) when estimated for the fund-of-funds, (iii) when estimated based on the 1-factor model, (iv) when estimated for the declining "bear" markets, (v) when more source databases are used, and (vi) when the CISDM database is not used as a data source.

Our findings have important implications for investors who consider alternative investment strategies, for regulators who seek the optimal design of the regulatory framework, as well as for researchers who analyze hedge fund performance. Investors may benefit from a better understanding of the level of return they may expect in various hedge fund types. Our findings suggest that even though hedge funds used to generate positive value for investors in the past, on average, they do not do so anymore. This finding is also relevant for regulators as it is likely related to the intensity of competition in the hedge fund market and the impact of the regulatory requirements. The number of hedge funds has steeply increased over time, which may have intensified the competition among them and diminished abnormal returns that the early hedge funds were able to achieve. The decline in the value hedge funds generate for investors may have also been driven by the influx of resources hedge funds manage and by the decreasing returns to scale of managerial ability to identify profitable investment opportunities. Hedge fund performance may have also changed over time due to progressively tighter regulation requiring greater hedge fund transparency, which may complicate their ability to fully exploit their proprietary investment strategies. We call for more research to distinguish between these potential underlying causes, and we provide systematic evidence to researchers about how their research design choices affect the reported alpha estimates.

Chapter 4

The Impact of Regulatory Change on Hedge Fund Performance

Fan Yang

Abstract We examine the impact of a regulatory change that mandated greater transparency of hedge funds in Europe on their performance. In 2011, the European Union (EU) adopted the Alternative Investment Fund Managers Directive (AIFMD, 2011/61/EU) that imposed stricter regulatory and disclosure requirements on *inter alia* hedge funds domiciled or marketed in the EU. We exploit the fact that the requirements for more comprehensive disclosure stipulated by the AIFMD only apply to hedge funds above a specified size threshold, and we use the difference-in-difference-in-differences (DDD) method to document a modest drop in monthly abnormal returns generated by the treated hedge funds ranging from 0.21 to 0.24 percentage points, which corresponds to a loss in annual "alpha" ranging from 2.6 to 2.9 percentage points. In support of the causal effect of the new regulation, we further observe that the drop in performance was less pronounced for hedge funds subjected to a stricter regulatory environment before the implementation of the AIFMD. Our results have important implications for investors considering alternative investments, for regulators seeking optimal ways of regulating hedge funds, as well as for researchers who study the impact of tighter regulation and greater transparency on economic outcomes.

The study was supported by the Charles University Grant Foundation (GA UK), project No. 344021. An earlier version of this paper was published in IES WP series 7/2023.

4.1 Introduction

The economic significance of hedge funds as an investment device and capital allocation mechanism has recently dramatically increased. Barth *et al.* (2020) and Stulz (2007) estimate that over the past three decades, the assets under management (AUM) invested in hedge funds increased more than 100 times, and the gross value assets on hedge funds' balance sheets now exceed $8.3 trillion. Traditionally, hedge funds were lightly regulated because they explicitly target sophisticated investors (high net-worth individuals and institutions) who are well-positioned to evaluate the rewards and risk of their investments and thus less in need of protective regulation than smaller retail investors (Engert 2010). Notwithstanding their name, hedge funds often hold complex positions that may exhibit substantial exposures to economic risks. Furthermore, these sensitivities may be hard to *a priori* quantify. These non-trivial risk exposures may be particularly relevant because, over time, the hedge funds' investor base has grown much broader; it now frequently involves even investors with a much lower appetite to take risks, e.g., pension funds (Lukaj & Healy 2007). The increasingly prominent role hedge funds play in the economy, the broadening investor base, and the complexity of their risk exposures, led to concerns about their potentially destabilizing macroeconomic impact, which only intensified after the 2008 financial crisis. Such considerations opened questions about the optimal way of regulating hedge funds (Bianchi & Drew 2010; Buller & Lindstrom 2013; Kaal *et al.* 2014; Johnston 2015).

The European Union (EU) responded to the calls for tighter regulation of hedge funds and other alternative investment vehicles by adopting the Alternative Investment Fund Managers Directive (AIFMD, 2011/61/EU) (Kamal 2012). The AIFMD requires hedge fund and other alternative investment fund managers whose AUM are above specified size thresholds to obtain authorization from their home country regulator and to follow a code of ethical conduct. In addition, the AIFMD requires hedge fund managers to implement the following procedures. First, for each fund managers are obliged to choose a depository that monitors the hedge fund assets and the flows in these assets. Second, hedge funds are obliged to request a periodic independent valuation of their assets and a calculation of the net asset value per unit or share. Third, hedge funds are obliged to provide periodic reports to the regulatory authorities that include information on the fund's financial statements, investment activities, and aggregate compensation paid to senior management and members of the staff. Fourth, the AIFMD also obliges subject hedge fund managers to disclose to the investors their investment policy, strategy, and objectives, as well as the expected level of leverage they consider sustainable for their investment strategy.

In this paper, we empirically investigate the impact of the AIFMD on hedge fund performance. We argue that the expected impact is not obvious. On the one hand, the enhanced transparency may facilitate oversight of hedge funds by investors and regulatory authorities, which may strengthen managerial motivation to perform and curb self-interested behavior that may arise from the agency problem (Frumkin & Vandegrift 2009). On the other hand, the more comprehensive disclosure may also reveal vital proprietary information to competitors who may imitate successful strategies and outsider investors who predict the trades and then trade against the fund, thus eroding hedge funds' competitive advantage (Cumming & Dai 2010; Shi 2017; Agarwal *et al.* 2013b). This may discourage managers from implementing some of the profitable investment strategies (Bianchi & Drew 2010) and disincentivize them from innovating and seeking new investment approaches (Bianchi & Drew 2010). Furthermore, activities required to ensure compliance with the new directive and the new reporting procedures may be costly in terms of managerial time and attention. They may distract managers from investment decisions, which may lead to less optimal investment and inferior performance (Restrepo 2020). Since *a priori* it is not clear which of the two effects shall prevail, we argue that it is worthwhile to empirically analyze the impact the AIFMD had on the value European hedge funds generate for their investors, which is important for evaluating the benefits and the costs of this piece of financial regulation.

Further research on this question is also important because prior studies that examine similar regulation in different contexts, e.g. the Dodd-Frank Act that is applicable to hedge funds active in the U.S., often produce conflicting findings. For instance, Cumming *et al.* (2020) compare the pre-Dodd-Frank period and after-Dodd-Frank period and conclude that U.S.-based funds have lower alpha after the implementation of the Dodd-Frank Act. Furthermore, Cumming *et al.* (2020) document a differential impact of the Dodd-Frank Act on hedge funds with different strategies. In contrast, Kaal *et al.* (2014) document a positive relationship between the Dodd-Frank Act and hedge fund performance even though they find the effect not to be persistent beyond several months following the implementation of the regulation. Thus, even though the preponderance of *a priori* arguments suggest that the new regulation that requires more extensive disclosures may impair the managerial ability to generate value for investors, whether this prediction actually materializes is far from obvious and, therefore, merits an empirical investigation.

Compared to the U.S. regulation, the AIFMD has similar requirements for registration and transparency but a wider scope of application, stricter risk control, and easier availability of information disclosure. These provisions may involve greater compliance costs and give the competitors easier access to information on hedge fund strategies, which may make the European regulation more impactful and costly than the U.S. regulation. Clarkson *et al.* (2014) argues that studying the impact of regulation both in the U.S. and the EU is important because they constitute the two largest hedge fund markets that can "set the tone" for appropriate hedge fund regulation by becoming "notable examples to follow" even for regulators in the other parts of the world. Johnston (2015) suggest that the AIFMD creates some pressure for regulatory convergence since non-EU hedge funds can only obtain the EU-wide passport when the EU authority believes that there is no significant difference in investor protection, market disruption, and the monitoring of systematic risk between two jurisdictions. Given the importance of the European hedge fund regulation, the conflicting prior empirical findings on the impact of similar types of regulation, and the notable differences between this regulation and the AIFMD, we consider it worthwhile to study how the AIFMD affected the value hedge funds generate for their investors.

Our main empirical analysis relies on the difference-in-difference-in-differences (DDD) method. We exploit the fact that the AIFMD does not apply to all hedge funds active in Europe. Funds whose size falls below thresholds specified in the directive, as well as funds *ex ante* compliant with other compatible regulations, are not subject to the AIFMD. Our treatment group thus comprises hedge funds that are domiciled and/or marketed in the EU and, at the same time, are not exempted from compliance with certain provisions of the AIFMD. Our control groups consist of (i) hedge funds that are not active in Europe but would meet the non-exemption requirements if they were, and (ii) hedge funds that are active in Europe but are exempted from the AIFMD requirements. The DDD approach identifies a change in the performance differential between the non-exempted and exempted hedge funds in Europe and elsewhere around the implementation of the AIFMD. The combination of the "geographic" dimension and the "size" dimension makes our results more robust to potential alternative explanations.

Observing a performance decline in large hedge funds relative to smaller hedge funds may be driven by different sensitivities that these hedge funds have to macroeconomic development rather than by the new regulation. Similarly, even though hedge funds tend to be unconstrained in their investment strategies, the actual investments of European hedge funds may be geographically biased towards European assets. Hence, any difference in performance between the European and non-European hedge funds may potentially be driven by the variation in geographic coverage of their holdings. Using the DDD approach allows us to identify the effect specifically in hedge funds that are active in Europe and are not exempted from the AIFMD. Relative performance changes in this narrowly defined treatment group are unlikely to be driven by any other confounding events, which allows us to draw stronger causal inferences about the impact of the European directive on hedge fund performance.

To further increase the confidence in our empirical results we use the propensity score matching (PSM), which ensures greater similarity of our control group relative to the treatment group. In the PSM, the control group is constructed so that its observable characteristics are similar to the corresponding characteristics of the treatment group. This limits the risk that the treatment group and the control group are fundamentally different and the observed changes in the performance are driven by these fundamental differences. Thus, providing a complementary result based on PSM increases our confidence that our main results based on DDD are not driven by fundamental differences between the treatment and control groups.

Based on the *a priori* arguments on the likely impact of mandating more extensive disclosure on hedge fund performance that we discuss above, we expect the new European regulation to be negatively associated with the value hedge funds generate for their investors. Consistent with our expectations, we observe that, relative to the funds that were not subject to the AIFMD, the treated hedge funds experienced a performance decline after the implementation of this regulation. Our results based on the triple difference approach (i.e., DDD) indicate that, relative to the benchmark, the drop in monthly abnormal returns in treated hedge funds ranges from 0.21 to 0.24 percentage points depending on the risk model used to estimate the "alpha". This corresponds to a loss in annual value generated for investors ranging from 2.6 to 2.9 percentage points. This relative decline in performance is not only statistically significant but also economically meaningful. In support of the causal effect of the new regulation, we further observe that the drop in performance was less prominent for hedge funds subjected to a stricter regulatory environment before the implementation of the AIFMD. Furthermore, we focus on the fund name list collected from the European Securities and Market Authority (ESMA) and their authorization dates to compare the performance changes of registered funds with unregistered funds matched through PSM. Our results from the matched groups confirm our main results. After the implementation of the AIFMD, the registered funds experienced a 0.21 percentage point drop in hedge funds' performance compared to unregistered funds, which is comparable to our main results based on DDD.

Our research makes several important contributions to prior research literature that are also relevant for investors and policymakers. First, this study provides novel evidence from a hitherto underexplored context on the impact of regulatory changes on hedge fund performance. There has been an extensive debate about the optimal way of regulating hedge funds and other alternative investment funds. Since their very origin, hedge funds were explicitly structured to benefit from limited regulatory requirements, which allow them to implement complex and dynamically evolving investment strategies and remain relatively secretive about them to protect and exploit their potentially unique proprietary expertise (Connor & Woo 2004; Stulz 2007; Brown *et al.* 2018). With the increasing prominence of hedge funds in the economy and the broadening base of hedge fund investors (Barth *et al.* 2020), this lenient regulatory environment seems to be no longer tenable. Greater transparency may be required for better oversight of hedge funds. Nevertheless, it is reasonable to expect that the requirement for greater transparency will also communicate important information about the characteristics of hedge fund trading strategies and invite imitation by competitors, which may disincentivize hedge fund creation and innovation and harm the industry. Thus, striking the right balance in the disclosure requirements requires a delicate balancing act of the corresponding benefits and costs. Our research informs the regulator about one aspect that is relevant to reaching this trade-off. The use of the DDD methodology also allows us to reconcile some of the conflicting findings in prior literature on this topic. Kaal *et al.* (2014) and Frumkin & Vandegrift (2009) use the monthly return to compare the before-rule and after-rule hedge fund performance and find a positive relationship between the regulation and hedge fund performance. Our results let researchers evaluate the level of potential bias in prior studies that use the DiD and DDD methods to measure hedge fund performance through raw returns that are not adjusted for risk exposures of hedge funds. Our use of several risk-adjusted performance measures allows researchers to evaluate the robustness of the results to the modification of performance measurement methodology.

Second, our study is particularly relevant to hedge fund regulation in the European context and for investors who consider investing in European hedge funds. It differs from related prior research by its emphasis on analyzing the impact of European regulation and the AIFMD in particular. Naturally, the European regulation exhibits some similarities with the regulatory changes in other geographic regions, e.g., the Dodd-Frank Act, that have been more extensively studied before (Aragon *et al.* 2013; Cumming *et al.* 2020; Shi 2017). For example, both regulations require hedge fund manager registration with the regulator and greater information transparency. Nevertheless, other provisions are unique to the European regulation, e.g., the requirement of independent valuation and depositories. Hence, it is unclear to what extent prior empirical findings about the impact of regulatory changes in different settings would generalize to the European context. The paucity of systematic research on the impact of the AIFMD is surprising, given the important role European regulation has in shaping general standards for hedge fund operations (Clarkson *et al.* 2014). These insights are important because the AIFMD is still subject to potential amendments[1]. For the policy decisions concerning the optimal level of regulatory "tightness", it is important to quantify the impact the original regulation had on hedge fund performance. The aim of the new EU legislation is to foster or develop the alternative industry in Europe. That is the reason why the EU commissions invited various groups of industry experts to suggest ways of improving hedge fund activities in the markets[2]. But whether the result is preferable raises doubt among the groups of opponents. Other jurisdictions may need such information to understand and develop hedge fund regulations.

Third, our study also contributes to the more general discussion about the advantages and drawbacks of corporate transparency. It is commonly argued that more transparent companies reduce the information asymmetry between the investors and the managers

[1]Source: https://www.europarl.europa.eu/thinktank/en/document/EPRS$_{B}RI$(2022)729321.
[2]Source: https://www.hedgeweek.com/european-commission-appoints-experts-hedge-funds-group/.

and thereby limit the consequences of the agency problem (Breuer 2021). Thus, companies that provide better disclosures are expected to raise capital from investors at better terms. Nevertheless, notwithstanding these capital market advantages, managers sometimes argue that mandating extensive disclosures about a company's performance and financial position may compromise its position vis-a-vis competitors and compromise its competitive position in product markets. Such an effect may have important implications for economic growth and welfare in general. Our study contributes to this debate by providing evidence from a very specific capital market where the proprietary cost of information might be particularly salient and where the nature of the industry allows us to study the effect with considerable precision.

The remainder of this chapter is organized as follows. Section 4.2 discusses the institutional background related to hedge fund regulation in general and the AIFMD in particular. Section 4.3 provides an overview of prior research literature on hedge fund performance. Section 4.4 explains our methodology and the data sample. Section 4.5 discusses our empirical results and Section 4.6 concludes.

4.2 Institutional Background

4.2.1 Origin of hedge funds

It is commonly believed that the first hedge fund was created in 1949 by a former Fortune magazine writer Alfred Winslow Jones (Connor & Woo 2004; Stulz 2007). Even though the financial industry has undergone dramatic development over the seven decades that have passed since then, many of the features of this first hedge fund resemble common hedge fund characteristics today. First, A. Jones structured the fund to be exempt from the Investment Company Act of 1940, which was the main pillar of the Security and Exchange Commission (SEC) regulations of investment entities at the time (Connor & Woo 2004). This exemption gave the fund greater flexibility in the use of investment techniques. Second, the fund made a relatively concentrated (rather than well-diversified) investment in a limited number of stocks that it considered undervalued and it hedged some of its risks by short selling other stocks. The long-short equity strategy still remains one of the most popular hedge fund strategies. It is also a strategy that gave "hedge" funds their name. Third, to build investors' confidence A. Jones co-financed a substantial portion of the fund's assets (40%) with his own money (Stulz 2007). Fourth, A. Jones used financial leverage to increase risk and simultaneously enhance the fund's ability to earn a higher return on the base capital. Fifth, A. Jones charged the investors a performance fee of 20% of returns earned (Connor & Woo 2004). All of these features are quite common in hedge funds even nowadays.

As of today, hedge funds typically aim at pursuing high returns on capital contributions for investors using sophisticated trading strategies in securities, currencies, and derivatives. They usually require investors to commit a large amount of initial capital for fixed periods when they are not allowed to withdraw investments. Some funds also require a notice several months prior to a redemption period (Aragon 2007). The "lockup period" and "redemption notice period" give managers more freedom for their specific strategies to aggressively exploit a wide range of investment opportunities in the market. Hedge fund managers receive generous rewards tied to fund performance (Malkiel & Saha 2005). In general, they have 20% of profits if the fund exceeds its previous high-water mark except for regular fees of 2% of assets (Guasoni & Obłój 2016). The incentive fee is important because it aligns the managers' interest with investors, while the high-water mark ensures that managers would not be satisfied with merely recovering previous losses.

To achieve superior returns, many hedge funds have been explicitly organized to be exempt from several related regulations in the US. Typical hedge funds target wealthy individuals or institutional investors to meet the exemption for "private offerings" under the Securities Act of 1933 and the "safe harbor" provision in Rule 506 of Regulation D. They avoid being classified as a financial player under the regulation by federal legislation (Kaal & Oesterle 2016). Also, hedge funds avoid being defined as an investment company under the Investment Company Act of 1940 based on the two exclusions: they either have fewer than 100 investors or only "qualified purchasers" who are high net-worth individuals or institutional investors (Kaal & Oesterle 2016). In addition, since hedge funds are not publicly traded companies and hedge fund managers are not public investment advisers, they are not subject to periodic reporting requirements according to the Securities and Exchange Act of 1934 or registration requirements of the Investment Advisers Act of 1940 (Kaal & Oesterle 2016).

4.2.2 Regulation of hedge funds in the US

The flexibility hedge funds enjoy puts them in a strong position to aggressively exploit any mispricing and contribute towards enhancing financial markets' efficiency. Nevertheless, this freedom to pursue diverse investment strategies implies that hedge funds may explore uncharted investment territories and be more likely to fail than conventional investment funds. Since the hedge fund industry has dramatically increased in size, these failures may have severe macroeconomic consequences, which raises questions on the optimal ways of regulating hedge funds.

The arguably most famous hedge fund, Long-Term Capital Management (LTCM), pursued a fixed-income arbitrage strategy that exploited small interest rate spreads between various debt securities. The strategy was once lauded for the sophisticated use of techniques to moderate risk exposure and for promoting the efficiency of bond markets. The highly leveraged positions became unsustainable during the "flight to safety" precipitated by the Russian debt crisis of 1998, which necessitated a massive bailout organized by the Federal Reserve Bank (FED) to prevent a more general financial meltdown (Stulz 2007). After the fall of LTCM in 1998, the authority realized the risk of hedge funds for the global financial system. In 2004, the SEC issued Rule 203(b)(3)-2 requiring more hedge fund advisers to register under the Investment Advisers Act of 1940 by tightening exemption standards(Kaal & Oesterle 2016). The amendments required hedge fund advisors to register with the SEC except those that have less than \$25 million assets under management (AUM) or with a lockup longer than 2 years. And the net worth requirement for accredited investors was raised to \$1.5 million. The new regulations also require the investment advisors to file Form ADV to disclose publicly information about fund characteristics such as potential conflicts of interest and past legal problems (Brown *et al.* 2008). Before the new rule, only limited managers from large hedge funds had registered in the SEC and so the hedge fund industry expressed strong opposition to the requirements. In 2006, the D.C. Circuit in *Goldstein v. SEC* vacated the hedge fund rule as it is contrary to the Congressional intent not to regulate hedge fund investment advisers (Frumkin & Vandegrift 2009; Kaal & Oesterle 2016).

The Dodd-Frank Act was proposed in June 2009 and became effective in March 2012 in the aftermath of the financial crisis. The Act authorized the SEC to promulgate registration rules and disclosure requirements (Kaal & Oesterle 2016). It requires smaller advisors (with \$25-\$100 million AUM) to register with states instead of the SEC, which prevents small advisors from registration exemption. It also imposes a number of record-keeping and reporting obligations for sensitive and proprietary information, including the types of investments managed, net asset values, investment strategies, performance and changes in performance, and positions held through filing the SEC Form PF. Furthermore, the information of clients and employees and related practices to conflict of interest should be disclosed (Cumming *et al.* 2020). The Dodd-Frank Act substantially improves the regulatory oversight of hedge funds.

4.2.3 Comparison of EU regulations before AIFMD among countries

Hedge funds are mostly managed in the US, which accounted for 70% of the hedge fund industry in late 2012, and the EU is in second place and accounts for 21% (Clarkson *et al.* 2014). Within the EU, the UK plays a dominant role in hedge fund management, with other individual members having less than 10% distributions. However, the location of management could differ from the authority of registration (Woll 2012). Some regions, such as the Cayman Islands, British Virgin Islands, Delaware (US), Bermuda, and Guernsey, have relatively lax regulations with the lowest tax rates, and so attract hedge funds as offshore domiciles. These offshore locations usually have no corporate taxes, lax regulations and further help avoid personal tax in corresponding onshore locations. One example is that non-US investors are exempted from tax obligation if the hedge fund is not domiciled in the US (Kamal 2012). Hedge funds set offshore sites to lower their tax obligation, simplify their operation and management, and thus allow for higher returns for investors.

The centers of hedge fund management, UK and US, used to allow the variation of domicile and management. Even two regions have distinctions in their regulation models and executions before 2010, for instance, US hedge fund managers could use the legal exemption and avoid registration as managers with the SEC while the UK managers should be accredited by the Financial Services Authority, hedge funds themselves did not see the coerciveness to be registered in nether regions (Fioretos 2010). In 2009, around 78% hedge funds were incorporated in offshore tax havens. Other than those offshore countries or regions, the main domiciles in the EU are Ireland and Luxembourg (Woll 2012).

It is expected that domiciles in some traditional choices, like Ireland and Luxembourg would increase after the implementation of the AIFMD when some non-EU hedge funds do not want to give up the EU market and consider to redomicile. In fact, the proportion of hedge fund domicile in the Caymans has already dropped by 7.5% (from 40% to 37%) while the proportion of Ireland and Luxembourg has increased by 60% (from 4.5% to 7.3%) from 2008 to early 2010s. It seems that after the effectiveness of the directive, it is hard for hedge funds seeking global investors including the EU investors to register in the offshore sites and escape from hedge fund regulations (Kamal 2012).

Hedge funds were not regulated at the EU level before the adoption of the EU Directive. But hedge fund managers might be regulated in member states by a combination of financial and company law regulations, accompanied by industry-developed standards in some sectors (Moschella 2011). In continental Europe, there were regulations about registration, disclosure, and reporting requirements (Woll 2013). Some countries, such as France, Italy, Spain, and Germany, regulate the fund as an onshore vehicle (Quaglia 2011). However, most countries accounted for only a small part of the hedge fund business in the EU. We focus on the analysis of regulation changes in three countries as examples, including the United Kingdom, Ireland, and Luxembourg. These countries, together with France, account for over 50% managers authorized in the EU (European Securities and Markets Authority 2019).

The UK, as the most popular financial market in the EU, hosted four-fifths of hedge funds in the EU (Quaglia 2011). It mostly relied on indirect regulations that were imposed on the counterparts instead of hedge funds (Woll 2013). The Financial Services Authority (FSA) has a principle-based approach to regulation. But the principles are very broad and allow the FSA to change the approach without enacting any legislation. The UK requires authorization of persons if relevant hedge funds products and services are to be traded in the public, it also needs information about investment strategies. However, most hedge funds were not authorized because they did not want to release their strategies (Sami 2009). Other relevant regulations in the UK were too general and not specific to hedge funds. And so it may cause problems in implementing the rules.

In Ireland, several forms- the Unit Trusts, Investment Limited Partnerships, Common Contractual Funds, and Variable or Fixed Asset Company can be used as the legal structures of AIFs and different rules were evolved for the different legal structures. For example, the Investment Limited Partnerships Act 1994 regulates investment limited partnerships while the Investment Funds, Companies, and Miscellaneous Provisions Act 2005 regulates Common Contractual Funds. Commonly it requires an authorization process for the AIF itself and the authorization process for its promoter and service providers. In addition, adequate information on the expertise and the reputation of the proposed directors of the management company must be provided, and the minimum capital of the management company is specified. There is no limit on the leverage, but prospectus disclosure is necessary. Annual audited financial statements are required. The monthly report should provide information about net asset value and net subscription and redemptions in the fund units/shares during the month, and the method of valuation of assets should be disclosed in the fund's constitutive document and prospectus (Fagetan 2020).

Luxembourg has various regulatory frameworks that can be applied by hedge funds that are distinguished by the strictness of regulations. The vast majority of hedge funds before the AIFMD are set up as the Specialized Investment Funds that are subject to a lighter regulation. The lighter set of rules needs the demonstration of managers' qualifications and reputation for approval of the launch of entities. There is no minimum capital requirement, but the minimum must be reached within a period of twelve months following its authorization. Depositories and annual audited financial statements are required, but the publication of net asset value is not necessary (Association of the Luxembourg Fund Industry 2014).

Overall, these countries had their own regulations before the AIFMD, while they either did not have the strong force for hedge funds to register, such as the UK, or had different sets of regulations varying in the strictness for a certain scope of hedge funds, such as Luxembourg or Ireland. Differences exist in the individual requirements. For instance, Luxembourg has rules about depositories while the UK does not have relevant rules. Different countries are various in the force of implementation, the specification or the strictness of authorization, operation, valuation, and disclosure requirements, but they have some similar rules in these aspects.

4.2.4 How does AIFMD differ from the Dodd-Frank Act?

The European Union adopted the Directive 2011/61/EU on Alternative Investment Fund Managers (AIFMD) that authorizes, supervises and regulates the managers of a range of alternative investment funds, such as hedge funds and private funds. The Directive applies to all EU alternative fund managers managing EU or non-EU funds and to non-EU alternative investment managers who manage EU funds or markets in the EU. Even though the Directive aims to regulate the managers, it is disproportionate to regulate the structure or composition of the portfolios of AIFs managed by AIFMs at Union level (The European Parliament and the Council of the European Union 2011; Ferran 2011). The following main aspects are summarized and compared to analyze the potential effects caused by the AIFMD.

Authorization requirements. The AIFMD has some conditions for the qualification of hedge fund managers by requiring that managers that apply for the authorization need to meet the requirements such as sufficient initial capital, good reputaion and experience(The European Parliament and the Council of the European Union 2011). Most hedge funds in the regions without capital requirements in the past only hold the capital needed for managing the investment (Kamal 2012). The managers that do not fulfill the capital requirements may attempt to solicit more investments or be eliminated. Thus, the rule may screen out the undesirable funds that are not expected to perform well without sufficient assets or competent managers while reducing the number of small hedge funds without the ability to raise adequate capital for the operation.

Unlike the US laws allowing for the existence of offshore hedge funds marketing in the US but not subjected to the US requirements, the AIFMD has constraints on all hedge funds that market in the EU. Then non-EU funds domiciled in the third country which considers EU investors would be faced with two choices. First, non-EU funds overcome hurdles that are required for outside fund marketing in the EU. Such hurdles involve cooperation arrangements between the third country and EU member states and agreements about the effective exchange of tax information. Therefore, the hurdles put the non-EU hedge fund in an inferior position. Reaching such requirements highly depends on relevant local government policy and systems instead of hedge funds themselves. Some hedge funds domiciled in popular hedge fund offshore locations such as Cayman Islands, British Virgin Islands, Guernsey would be excluded from the EU market due to not being able to comply with the tax model (Kamal 2012).

Second, hedge funds may re-domicile in one of the member states in the EU. The advantage is that the relocation makes it possible to have the largest pool of investors including EU investors and global investors. Moving funds from offshore countries to onshore Europe contributes to the dynamic market in the EU. However, the funds that choose to relocate have to bear double costs involving the cost of relocation and the cost of meeting the requirements of AIFMD. The cost of restructuring hedge funds to meet the current needs of investors is estimated to be significant (around €1.4 billion) (Kamal 2012).

Independent valuation and depositaries. Unlike the US which depends mainly on the authority to monitor the risk, the EU also utilizes third parties such as credit institutions or external valuers. The fund manager should ensure the independent valuation of assets at least once a year and need a single depositary for the safekeeping of the assets and monitoring cash flows (The European Parliament and the Council of the European Union 2011). The depository is responsible for assuring cash flows and verifying the ownership of all other assets by the AIF. And it is liable for the loss if the financial assets are held in custody unless due to some external events beyond the control.

Transparency and disclosure. Similar to the Dodd-Frank Act, the fund managers need to disclose relevant information about the investment strategies, performance, and assets. According to the AIFMD, the fund needs to prepare an annual report for investors on request and the competent authorities of the home Member State and the accounting information given in the annual report shall be audited. Investors have the right to know the investment strategy and objectives of the AIF, a description of the types of assets in which the AIF may invest, the techniques it may employ and all associated risks, any applicable investment restrictions, the extent of leverage, the maximum level of leverage, the circumstances in which the AIF may use leverage, etc (The European Parliament and the Council of the European Union 2011). The difference is that the US intends to disclose this information to the authorities instead of the investors to analyze and monitor the risk while the EU requires not only initial disclosure of relevant information but also the annual reports which are available to both the investors and the authority. And the EU regulators pay special attention to imposing pressure to use high leverage.

AIFMD passport. Since the aim of the AIFMD is to harmonize financial services and to facilitate marketing across the EU, the authorized AIFMs have the passport to provide financial services to AIFs domiciled in any member state and to offer their products to "professional investors" within the EU. However, this passport, in the short-term, is only accessible for AIFMs marketing EU-domiciled AIFs due to the concerns for competition and possible regulatory differences (Johnston 2015). It is an attractive benefit for EU managers managing their EU hedge funds without the constraints of different legislative requirements. Conversely, it provides unequal benefits compared to costs for non-EU funds or fund managers to be authorized in one of the member states in the EU-wide context.

4.3 Prior research on hedge fund performance

In theory, the relationship between hedge fund regulation and hedge fund performance is ambiguous (Cumming & Dai 2010). The enhanced regulation could improve the quality of hedge funds, reduce information asymmetry between the managers and investors or authorities, restrict managers from unethical and compensation-oriented actions and possibly improve the performance of the hedge funds (Cumming & Dai 2010; Frumkin & Vandegrift 2009). The lack of regulatory oversight makes it possible for managers to merely chase high compensation and disguise investment schemes, which is a part of the agency problem. One example in Cumming & Dai (2010) is that two funds under the control of the same managers could have strategies of shorting the S&P index and going long on S&P separately. The result would be one wins and one loses but managers still have high compensation from fixed management fees and carried interest performance fees. Neither investors nor regulatory authorities would know the true nature of these hedge funds. Given the improved regulation and oversight, hedge fund structure and performance may be enhanced by preventing managers from such behavior. However, the regulations and rules may hamper hedge fund performance because managers lose the freedom to contract and organize resources in the most efficient way. Under the pressure of regulations, managers may behave in some conservative ways to form strategies and would not get some “bold” rewards. The common regulations such as registration, restrictions on minimum hedge fund size, restrictions on the location of key service providers, and market channels for hedge fund distributions not only set barriers to entry or participation and to choose efficient human resources but also impose high compliance costs on hedge funds and influence their instant and continuous performance. Thus, the regulations may lead to worse performance and less efficient hedge fund structures (Cumming & Dai 2010).

To have a clear understanding of the relationship between regulations and hedge fund performance. Cumming & Dai (2010) collectively investigate 29 countries’ hedge fund data. Some countries, such as the United Kingdom, have minimum capital requirements to operate as hedge fund managers, and restrictions on marketing channels (banks, fund distribution companies, other financial service institutions, etc.). Besides the above related regulations, countries like Canada and Germany have restrictions on the location of key service providers. The results show that the requirements like locational restrictions of key service providers give rise to lower performance. But minimum capital requirements and locational restrictions of key service providers are associated with lower standard deviations of returns. Therefore, the requirements could lower risks in the market.

Besides international differences in hedge fund regulations, several papers focus on individual aspects of hedge fund regulation, such as transparency, in the United States. Section 13(f) of the Exchange Act (adopted by the SEC in 1978) requires hedge fund managers who exercise investment discretion over accounts holding at least $100 million in publicly traded companies, convertible bonds or options to make quarterly disclosures of portfolio holding in these securities to SEC on Form 13F within 45 days of the quarter end. However, managers could request confidential treatment to delay public disclosure of some or all of the holdings reported on Form 13F. Then Form 13F "add new holdings" Amendment should be filed within six days of the end of the confidential treatment period (Aragon *et al.* 2013). More transparency is beneficial for investors' decision making, but it may reduce the incentives of hedge fund managers since revealed information makes competitors identify their strategies or free-ride on their efforts. Shi (2017) uses TASS data from 1994 to 2010 and finds that the drop in alpha is concentrated among funds that disclose a greater fraction of their assets. The return correlations between the disclosing funds and other hedge funds that have the same investment style increase after the disclosure. That implies that after a fund discloses, other funds take similar positions. Aragon *et al.* (2013) find that managers are more likely to seek confidential treatment for positions if they perform well in the past. And securities that are kept confidential earn significantly positive abnormal returns over the post-filing confidential period while securities disclosed originally do not have abnormal stock price performance over the same period. Agarwal *et al.* (2013b) also compare confidential holdings and original holdings and find the confidential holdings have higher benchmark-adjusted returns than the original holdings up to 12 months. All these studies suggest that Form 13F and complete public disclosure may encourage free-riding activities and negatively influence fund performance.

Some studies investigate the US regulation's overall effect, including the registration and disclosure requirements. For example, Frumkin & Vandegrift (2009) examine the effect of the amendment of the Investment Advisers Act of 1940 by the SEC in 2004 when the range of hedge fund registration was extended and additional general information about hedge funds was required to be disclosed. They expect the rule to reduce advisors' fraudulent or unethical behavior and improve investors' average quality. Therefore, the registration rule would improve the performance of hedge funds. Using a regression of excess monthly returns over S&P 500 returns on size, age, volatility, and registration rule, they find registration increases hedge fund returns by 11.6 percent compared to before the registration period and registration in-effect period.

Nonetheless, prior studies discussing the influence of the Dodd-Frank Act show inconsistent results. The Dodd-Frank Act (effective in 2012) requires smaller advisors (with $25-$100 million AUM) to register with states instead of the SEC, which prevents small advisors from registration exemption. Compared with the 2006 rule of hedge funds, it imposes several more extensive disclosure requirements including the types of investments managed, net asset values, investment strategies, financing information, products used by advisors, risk metrics, performance and changes in performance, positions held through filing the Form PF hedge fund returns (Cumming *et al.* 2020; Kaal *et al.* 2014). Because the Dodd-Frank Act improves oversight, Cumming *et al.* (2020) expect some risk-averse managers to change their activities to meet requirements. And the cost of compliance expenditure may lower the returns. They compare the pre-Dodd-Frank period and the after-Dodd-Frank period and find evidence that US funds have lower alpha after the effectiveness of the Dodd-Frank Act. Kaal *et al.* (2014) discuss a similar topic but they see no evidence about the negative relationship between hedge fund performance and the Dodd-Frank Act. Moreover, they find a temporary but positive relationship between monthly return and the coefficient of the treated interaction term. These studies have conflicted results regarding the influence of hedge fund registration and transparency on performance.

As prior studies have similar conclusions that greater transparency is related to lower hedge fund performance by analyzing Form 13F under Section 13(f) of the Exchange Act, we expect that the negative effect of leaking managers' strategies and weakening their incentives outweighs the positive effect of reduction of managers' misconducts. Comparing the AIFMD with the US regulation, the EU Directive requires preparing audited annual reports for investors and authorities on demand instead of merely disclosing information to the authorities for monitoring. Therefore, competitors can more easily access investment techniques and enhance the negative influence caused by transparency. As the EU rules pay special attention to the use of leverage, specifying the disclosure of maximum leverage, the extent of leverage, and its application in certain circumstances. It seems the EU hedge fund manager would face greater pressure to operate their business with high leverage to chase superior revenues. From the perspective of information disclosure, the EU regulation should have a greater negative effect on performance than the US regulation.

However, prior results show inconsistent overall effects of the combination of registration, transparency, and related compliance costs. Since transparency probably leads to a negative impact, as we discussed above, and the compliance costs lower the earnings, the overall positive relationship between the US regulation, such as the Dodd-Frank Act, and the hedge fund performance may suggest that the undernourished hedge funds decrease due to the strict rule to register. Then the overall industry performance is expected to grow (Frumkin & Vandegrift 2009) after removing costs resulting from transparency and other compliance fees. The EU Directive requires minimum capital amounts and qualified managers to assure the industry's quality, which may positively influence hedge fund performance but raises the capital costs for some small funds. The quality improves hedge fund performance when the diminution of underperforming hedge funds exceeds that of wellperforming small funds. The improvements might exist based on the US regulation, while the EU regulation involves more cost as the hedge funds that are located in third countries are expected to have an additional one-time large cost for relocating to a member state of the EU; otherwise, they are not able to market in the EU. The cost of authorization and restriction is probably greater than the benefits from improving the overall hedge fund quality.

Besides, we observe that two studies, Frumkin & Vandegrift (2009) and Kaal *et al.* (2014) that reveal the positive influence of US regulations on hedge fund performance, use returns for investigation, which may not properly measure hedge fund performance since various risk factors may affect the results. Thus, we rely more on studies using risk-adjusted return and predict that the cost of constraints on hedge fund operation and disclosure requirements outweighs the benefit of improving the hedge fund quality by the oversight based on the strictness of US regulations. With additional relocation costs and stricter disclosure requirements, the EU regulation is more likely to negatively influence hedge fund performance.

Except for the similar rules to the US regulation, the AIFMD requires hedge funds to have periodic independent valuations and depositories. The depository could increase investors' confidence by reducing asymmetric information between the manager and investors (Kamal 2012). However, failing to correctly value the assets of hedge funds may lead to sanctions by the authorities for valuation shortcomings. It is effective in continuously controlling risk, while hedge funds would unavoidably distract attention and require more investment in their operation and management in case of any violation of the AIFMD. The compliance cost from the valuation and depository is continuous and may constitute another important part of the total cost.

Combining the analysis of individual rules in EU regulation based on prior literature, we expect an overall negative relationship between the AIFMD and hedge fund performance. Furthermore, Kehoe (2013) points out that the rigid approach of AIFMD seems to make hedge funds behave like mutual funds. Prior studies show some evidence that hedge funds provide more risk-adjusted return than mutual funds (Eling & Faust 2010). If the authorities make efforts to regulate hedge funds towards mutual funds and abandon characteristics such as the freedom to use high leverage and to be less transparent to hide special strategies, it is reasonable to rethink the earning ability of hedge funds. Based upon all above reasons, we have our first hypothesis as follows:

H1: EU hedge funds have lower risk-adjusted performance after the AIFMD was transposed into EU countries' national laws.

To some extent, the AIFMD reflects many similar aspects, including initial capital, depository, transparency, and investor protection, required in different degrees by the pre-existing regulatory systems in various member states(Fagetan 2020). After the AIFMD, the hedge funds would be subjected to the standard rules as the minimum requirements, and it is convenient for EU hedge fund managers to market in the EU region. The AIFMD is more comprehensive in the regulated activities and execution (Kim 2014) and the scope of the new regulation covers more entities and avoids most off-shore funds operating in the EU markets.

We expect the AIFMD to have different influences on different member states based on their pre-existing regulatory environment. If a country had stricter regulations, the new rule would not be that influential when the country has its special considerations and chooses not to downgrade its requirements. For member states with similar levels or slightly lower levels of strictness, like France, Fagetan (2020) suggest that the AIFMD should not have a significant influence because when the prior regulations are strict enough, the number of hedge funds would be restricted and the quality of the managers and funds were ensured to a large extent. As a result, we expect the AIFMD to have a greater influence on the hedge funds of two sources. The first source is the hedge funds existing in the EU supervised by local authorities or industry standards with unstrict regulations. Another source is hedge funds coming from offshore sites with lax regulations. The influence of hedge funds in third countries would be greater since they would be subject to both relocation costs and other compliance costs. Within the EU context, even jurisdictions in different areas had different regulations in terms of authorization and transparency before the EU directive, in general, the enforcement of AIFMD, as minimum requirements, would have greater influence in the funds that were domiciled in regions that had lax regulations. And thus our second hypothesis is:

H2: EU Hedge funds domiciled in countries with strict regulations before the AIFMD would experience a less drop than the hedge funds domiciled in countries with lax regulations.

4.4 Research Design

4.4.1 Data Collection

We obtain data from the Eurekahedge database and European Securities and Markets Authority database. Eurekahedge is a commercial hedge fund database that provides a comprehensive and global data of hedge funds including monthly returns, AUM and main characteristics of funds such as country focus, domicile, leverage, management fee, performance fee, hurdle rate, main strategies, and management company. By the end of November 2021, the database has 26,862 global hedge funds. Among them, 19,620 hedge funds provide monthly returns or AUM as time-series information. We exclude 8,864 US hedge funds because relevant hedge fund regulations (Dodd-Frank Act) became effective in 2012. And Cumming *et al.* (2020) has provided evidence that the rule has an effect on the US hedge fund regulations..

We did the following procedures to mitigate the survivorship bias and backfilling bias. We checked all data involved and found our data are less likely to suffer from survivorship bias because they are all post 1994, when the database includes both live and dead hedge funds. Backfilling bias occurs when a new hedge fund enters a database, it has an incubation period to record past performance. If the past performance is good, the record is backfilled. If the past performance is bad, the hedge fund probably cease the operation. This situation makes average returns in databases biased upward (Fung & Hsieh 2004a). We only include the observations that post the initial date of adding the information of hedge funds to mitigate the backfilling bias.

To further test the authorized entities' performance, we need the entities listed on the official website of the European Securities and Markets Authority. Using the list has two advantages: 1) We can identify the exact time when each entity started to comply with the AIFMD and have the comparison within each entity instead of utilizing the uniform date of effectiveness of the rule; 2) the information of the database is based on the provided information from hedge funds, we could not rule out the possibility that some hedge funds that used to specialize in the EU market shifted their focuses and leave the EU market after the enactment of EU hedge fund regulations but did not update the relevant regional information. The survey conducted by Open Europe indicated that 27% had a preference to give up the EU market if the Directive became effective (Kamal 2012). Nonetheless, the information could be inaccurately updated or reflected on the website since some key words of "country focus" in the database involve multiple regions and the EU is in their list of regions. The fact is we did find in the database some hedge fund domiciling in the offshore sites (e.g., Cayman Island) show their European "country focus". As we discussed the potential choice of hedge funds in the third countries, theoretically, it is possible but hard for hedge funds to operate in offshore locations and so it is reasonable to question that some geographical information of hedge funds was not correctly updated.

Since the entities include funds of funds, real estate funds, hedge funds, and private equity funds, it is necessary to identify the hedge funds by searching for the relevant information from a commercial hedge fund database. We downloaded the AIFM list, which involves 3,696 entities by the end of 2021. We identify the authorized hedge funds by searching the entity name on the platform of Eurekahedge corresponding to the management company in the Eurekahedge. If the entity name is close to the management company but not the same, we confirm whether the address is the same. For the management companies that have multiple funds located globally, we only include the funds that are either domiciled or marketed in the EU in case our list involves other funds that are less likely to apply for authorization. Moreover, we do not include the funds managed by the EU managers but not marketed or domiciled in the EU since they are exempt from strict rules such as the requirements in respect of depositary and fund annual reports.

Table 4.1: Summary statistics on hedge fund characteristics

Panel A Hedge Fund Returns, Flows, and AUM

	EU			NonEU			Total		
	Before AIFMD	Post AIFMD	Totals	Before AIFMD	Post AIFMD	Totals	Before AIFMD	Post AIFMD	Totals
AUM (millions)									
Obs	192,732	386,018	578,750	231,785	115,965	347,750	424,517	501,983	926,500
Percent	20.80%	41.66%	62.47%	25.02%	12.52%	37.53%	45.82%	54.18%	100.00%
Mean	365.02	620.55	535.45	155.9	263.91	191.92	250.84	538.16	406.51
Median	89	147	123	38	53	43	56	117	84
St.d.	906.93	1444.81	1296.43	529.79	1137.76	788.26	733.16	1388.12	1144.89
Returns (%)									
Obs	274,799	426,521	701,320	303,055	134,236	437,291	577,854	560,757	1,138,611
Percent	24.13%	37.46%	61.59%	26.62%	11.79%	38.41%	50.75%	49.25%	100.00%
Mean	0.48	0.31	0.37	0.73	0.57	0.68	0.61	0.37	0.49
Median	0.44	0.26	0.32	0.68	0.51	0.63	0.56	0.31	0.42
St.d.	4.45	3.64	3.98	5.3	5.87	5.48	4.92	4.29	4.62
Flows (millions)									
Obs	188,086	381,303	569,389	227,126	115,013	342,139	415,212	496,316	911,528
Percent	20.63%	41.83%	62.47%	24.92%	12.62%	37.53%	45.55%	54.45%	100.00%
Mean	3.09	-0.02	1.01	0.67	-0.69	0.22	1.77	-0.17	0.71
Median	0.01	-0.08	-0.02	0.00	-0.04	0.00	0.00	-0.07	0.00
St.d.	61.39	86.44	79.07	48.3	36.6	44.71	54.63	77.79	68.23

Panel B Fund Characteristics

	Obs	Mean	St.D.	First quartile	Second quartile	Third quartile
Hurdlerate	16,028	0.28	0.45	0	0	1
Watermark	16,497	0.77	0.42	1	1	1
Leverage	8,423	0.82	2.01	0	0	0.5
Lockup	16,201	0.11	0.32	0	0	0
PerformanceFee (percent)	16,923	15.85	7.69	15	20	20
ManagementFee (percent)	16,955	1.43	0.63	1	1.5	2
UCITS	17,260	0.38	0.48	0	0	1
SharpeRatio	13,945	0.84	26.02	-0.02	0.39	0.82
SortinoRatio	13,740	1.73	17.86	-0.04	0.58	1.33
AnnualisedReturn	13,963	4.42	24.34	-0.14	3.4	7.63

(continued on next page)

Table 4.1: Summary statistics on hedge fund characteristics (continued)

Panel C Distribution of the Number of Authorization of the Entity

	Authorization	Percent
2013	9	2.53
2014	177	49.72
2015	65	18.26
2016	34	9.55
2017	31	8.71
2018	14	3.93
2019	16	4.49
2020	8	2.25
2021	2	0.56
Total	356	100

Panel D Distribution of the Number of Withdrawal of the Entity

	Withdrawal	Percent
2017	6	3.35
2018	3	1.68
2019	4	2.23
2020	162	90.5
2021	4	2.23
Total	179	100

Notes: Panel A reports the hedge fund returns, flows, and AUM divided by the pre- AIFMD / post- AIFMD period and EU / NonEU region. NonEU region excludes the US because the Dodd-Frank Act became effective in 2012. Panel B shows other fund characteristics provided by the Eurekahedge. Hurdlerate, watermark, leverage, lockup are dummy variables that indicate whether the hedge fund uses hurdlerate, watermark, leverage, lockup in the operation. Panel C & D list the distribution of authorization and withdrawal in each year according to the ESMA (European Securities and Markets Authority).

Table 4.1 shows the descriptive characteristics of hedge funds. Panel A reports the hedge fund returns, flows, and assets under management (AUM) from 1994 to 2021 categorized by the pre- AIFMD / post- AIFMD period and EU / NonEU region. The returns are net of fees. The inflows are measured by $AUM_{t+1} - AUM_t * (1 + Return)\%$. The dataset excludes the US because the Dodd-Frank Act became effective in 2012. Panel A shows that the average AUM for EU funds increases from 365 million dollars before AIFMD to 620 million dollars after the AIFMD and NonEU funds have a similar 70% increase. In terms of the average returns, Both EU and NonEU funds experience a drop after the enforcement of the AIFMD. It is consistent with the finding from Bollen *et al.* (2021) that shows the decline in hedge fund performance over the past decades which is possibly associated with the market characteristics and overall economic environment. The EU has a 35% decrease in returns (from 0.48% to 0.31%) while the NonEU drops less by 22%. However, the EU returns have lower standard deviation in the period after the rule compared with the period before the rule, which is opposite to the NonEU funds. After 2013, the inflows are negative for both EU and NonEU regions, However, the NonEU hedge funds have more withdrawals.

Panel B shows other fund characteristics provided by the Eurekahedge. Hurdlerate, watermark, leverage, lockup are dummy variables that indicate whether the hedge fund uses hurdlerate, watermark, leverage, lockup in the operation. In addition, the fee structure, including the management fee and performance fee for hedge funds, the annualized returns, the risk-adjusted returns based on the Sharpe ratio and Sortino ratio are summarized. We also collect the variable whether the hedge funds are subject to the EU Undertakings for Collective Investment in Transferable Securities (UCITS) Directives, a rule with its original purpose to protect retail investors primarily in mutual funds, that covers some hedge funds before the AIFMD is enacted for identifying the hedge funds' obligation to AIFMD since it is not compulsory for UCITS funds to re-register under the AIFMD.

Panel C & D list the distribution of authorization and withdrawal under the AIFMD in each year according to the ESMA (European Securities and Markets Authority). The entity name corresponds to the management company in Eurekahedge and an entity may operate and manage several hedge funds. Panel C shows that the majority of EU hedge fund entities have been registered within one year after the AIFMD transposed to the EU. Panel D shows that until 2016, no entities withdrew from the regulatory regime while about 45% entities withdrew in 2020 as a peak.

4.4.2 Methodology

Measurement of hedge fund performance

Prior studies utilized different methods to measure hedge fund performance in the relevant literature. Frumkin & Vandegrift (2009) use the monthly returns in excess of S&P 500 returns to represent the hedge fund performance possibly due to the short effectiveness of the period of the 2006 amendment of the Investment Advisers Act by the SEC. And Kaal *et al.* (2014) simply use the monthly return to compare the instant or short-term influence of performance. Both studies lead to a positive relationship of hedge fund performance and enforcement of regulations. However, most studies use the alphas through the linear aggression of hedge fund returns on risk factors to evaluate the performance and they provide evidence that the regulation negatively influences the hedge fund performance.

The conflicted results from prior research may come from the different measurements of the hedge fund performance. Using monthly returns to reflect the hedge fund performance could be inaccurate because the returns include some risk-factor returns that are not representative of hedge fund earning ability. The increase in the returns may come from the increase of the general market returns influenced by the market characteristics. Many prior studies regress returns on several risk factors to estimate the extra profits from hedge funds but they have inconsistent choices and combinations of risk factors in evaluating the hedge fund performance. To have a valid comparison, we use several common methods including the excess return, one-factor model that follows the CAPM, the three-factor model (Fama & French 1998), four-factor model (Carhart 1997), and the eight-factor model that consists of the Fung & Hsieh (2004a) seven factor model and the emerging market factor (MSCI Emerging Market index monthly total return) (Fung *et al.* 2006). The Fung & Hsieh model involves bond, currency, and commodity trend-following risk factors, equity market factor (Standard & Poors 500 index monthly total return), the size spread factor (the difference between Russell 2000 index monthly total return and Standard & Poors 500 monthly total return), the bond market factor (monthly change in the 10-year treasury constant maturity yield), the credit spread factor (difference between monthly change in the Moody's Baa yield and 10-year treasury constant maturity yield). These factors can explain above 50% of hedge funds (Fung & Hsieh 2004a). Above methods have been extensively used to measure the performance in prior studies (Bali *et al.* 2013; Kosowski *et al.* 2007; Teo 2009). Following Shi (2017), we use the non-overlap two-year window of individual hedge funds as the periodical performance and regress the monthly return over the two-year window on the risk factors. The unexplained return observed by the intercept (or alpha) represents the abnormal return obtained by managers' particular skills.

Difference-in-differences method (DiD)

Table 1 indicates that the decrease in the return may exist due to the overall economic forces indistinguishable from regulated and unregulated regions. To evaluate the impact of AIFMD on hedge funds, we use the difference-in-differences method commonly utilized by prior studies that want to evaluate the influence by a certain hedge fund regulation (e.g., Shi 2017; Cumming *et al.* 2020). According to the Directive, the rule should apply to AIFMs managing funds not covered by UCITS. In addition, AIFMs manage portfolios of AIFs whose AUM do not exceed 100 million euros or between 100 million euros and 500 million euros and at the same time, are unlevered and have lockup requirements of at least 5 years are exempted from strict regulations. They are only required to register and provide information on strategies and instruments. We identify the UCITS funds or the smaller funds that can be exempted from strict rules as the funds with exemption and the remaining funds as the funds with no exemption. First, we compare the change in performance after the AIFMD was imposed into national laws within the European Union between funds with no exemptions and funds with exemptions:

$$\begin{aligned} Performance_{i,t} = \gamma_i + \gamma_t + \beta_0 + \beta_1 \text{nonexemption} + \beta_2 \text{post} + \\ \beta_3 \text{nonexemption\_post}_{i,t} + controls_{i,t} + \epsilon_{i,t} \end{aligned} \tag{4.1}$$

where the dependent variable is the hedge fund performance in period t of fund i. The nonexemption_post is a dummy variable that equals to 1 if the manager of fund i is likely to be subject to AIFMD in period t post July 2013 when the AIFMD is transposed into national laws and 0 otherwise. If a period of 24 months covers July 2013, we treat it as before the AIFMD. Similar to Shi (2017), control variables are lagged log fund size or the linear and quadratic terms of lagged fund size. We assume each fund has its own fixed characteristics such as management fee, lockup requirements, incentive fee, and regional focus that affect its influence and the time trend would cause overall influence in worldwide markets. The γ_i represents fund fixed effects and γ_t reflects time fixed effects, and $\epsilon_{i,t}$ is an idiosyncratic error.

Second, we compare the change in performance for funds with no exemption between the European Union and other regions (we exclude the US funds because the Dodd-Frank Act became effective for US funds in March 2012). The control group is NonUCITS funds or large funds that do not meet relevant exemption rules in other regions:

$$Performance_{i,t} = \gamma_i + \gamma_t + \beta_0 + \beta_1 \text{eu} + \beta_2 \text{post} + \beta_3 \text{eu\_post}_{i,t} + controls_{i,t} + \epsilon_{i,t} \tag{4.2}$$

Where eu_post equals 1 if the fund i is an European fund or a fund marketing in the EU in period t post AIFMD and 0 otherwise. As in equation (4.1), β_3 is the key indicator of the interaction term that reflects the influence.

However, limitations remain within the DiD method in exploring the effect of the regulations. Cumming *et al.* (2020) use the April 2012 to divide the pre-regulation and post-regulation periods and regress alpha on the interaction of post period and US hedge fund. The negative coefficient can show the US hedge funds experience a negative change or less development relative to other areas. But it is still possible to explain the result by other reasons than the regulations. For instance, the US market is saturated and has fewer marginal returns (suppose the asset inflows are similar for the treatment and control groups). Or hedge funds in other regions are more sensitive to the development of the financial market after the global crisis. These factors make the assumption of similar trends for control and treated groups invalid and may influence the DiD result and cause a similar effect as the regulations. It would be better to introduce a third confounding factor that allows us to compare not only the EU hedge funds with other global hedge funds but also a subgroup of EU hedge funds subjected to the AIFMD with another group of the EU hedge funds not subjected to the AIFMD. And so, we use the DDD method.

Difference-in-difference-in-differences method (DDD)

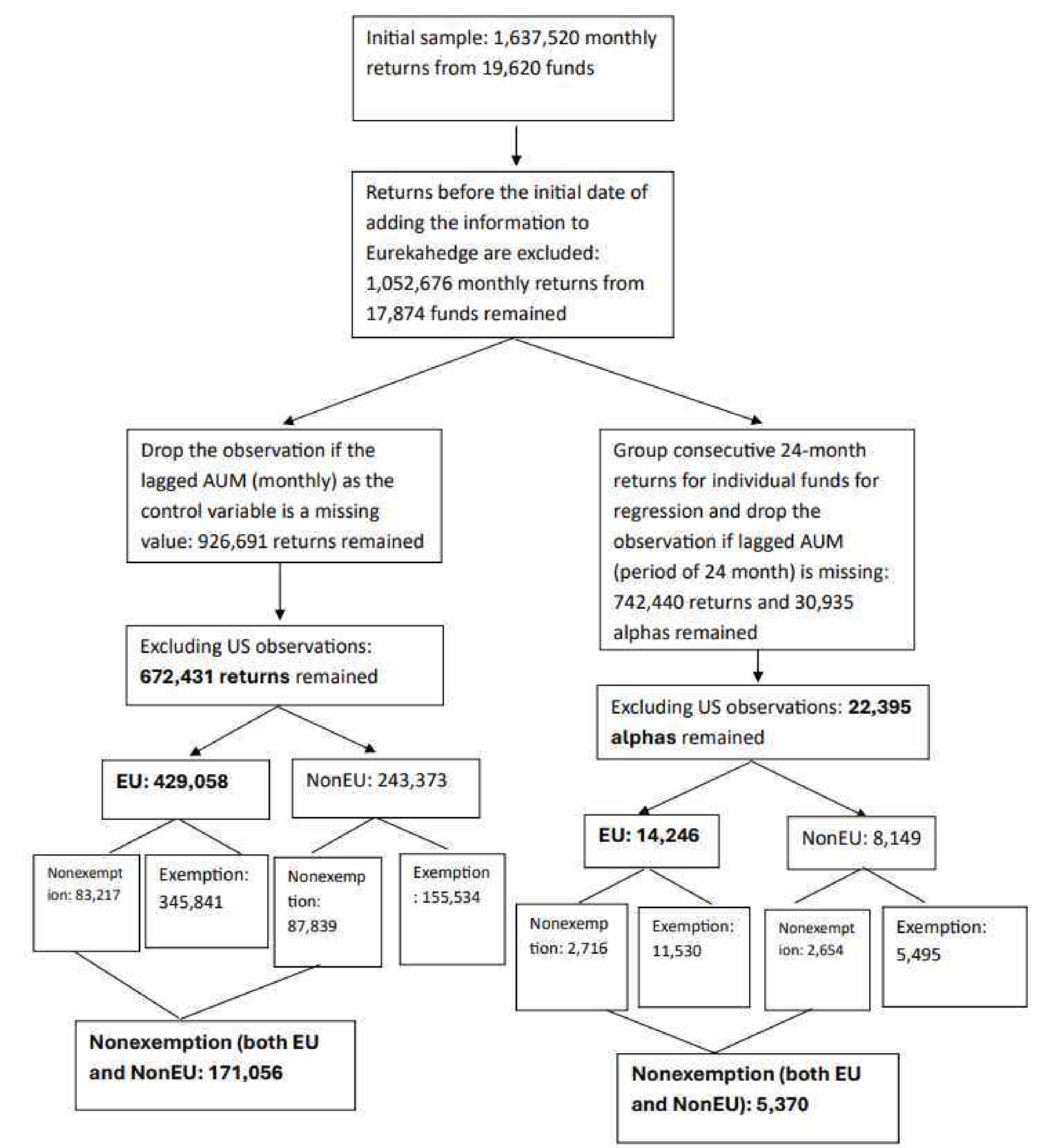
We combine the two DiDs and form DDD to investigate the impact of AIFMD on hedge funds. The triple difference method does not require two parallel trend assumptions to make a causal interpretation. Instead, it assumes specified bias existing in two comparisons of treatment group and control group in DiD. And subtracting the second DiD makes the bias differenced out (Olden & MĂ¸en 2022). And we would assume, reasonably, that the general economic difference or market difference does not affect the relative outcome of the funds with and without exemption. It allows evaluation of differences before and after AIFMD comparing relevant funds from EU and other regions, at the same time, it accounts for differences between funds with exemption and no exemption:

$$
\begin{aligned}
Performance_{i,t} = \gamma_i + \gamma_t + \beta_0 + \beta_1 \text{eu} + \beta_2 \text{post} + \beta_3 \text{nonexemption} + \\
\beta_4 \text{eu\_post}_{i,t} + \beta_5 \text{nonexemption\_post}_{i,t} + \beta_6 \text{eu\_nonexemption} \\
+ \beta_7 \text{eu\_nonexemption\_post}_{i,t} + controls_{i,t} + \epsilon_{i,t}
\end{aligned}
\tag{4.3}
$$

In equation (4.3), β_7 measures the double-interaction influence among the EU, post, and funds with no exemption. The results show the value of the difference of performance changes after the regulation between funds with exemption and non-exemption in the EU minus the difference of performance changes between funds with exemption and non-exemption in other regions.

Figure 5.1 shows our sample selection process for both excess returns and alphas as measurements of hedge fund performance. We have the Eurekahedge initial sample with 1,637,520 monthly returns from 19,620 funds. First, we exclude the monthly returns

Figure 4.1: Sample Selection Process



Notes: The figure shows our sample selection process.The left side is the process of cleaning data to obtain excess returns as a measurement of hedge fund performance. The right side is the process of cleaning data to obtain alphas as measurements of hedge fund performance.

if they are prior to the beginning date of hedge funds to add the information to the database to mitigate the backfilling bias. 1,052,676 monthly returns are left after the process. Second, we exclude the observations if they have missing values of lagged AUM that is considered as the control variable in the regression when we use excess monthly return over a risk-free rate to measure hedge fund performance. We exclude observations if they do not have lagged AUM of a 24-month period when we use the alpha as the measurement of hedge fund performance. Third, We remove the US observation in the comparison. As a result, we have 672,431 returns for using excess monthly returns over a risk-free rate to measure hedge fund performance and 30,935 alphas as abnormal returns for hedge funds.

Since we expect that the affected group is the nonexempt EU group, and the other three groups, including the exempt EU group, nonexempt NonEU group, and exempt NonEU group, are less likely to be influenced by the AIFMD, we do not use the constant sample for investigation. Instead, we use the subsamples to compare them with the treated group. Using the excess return over a risk-free rate as an example in the left side of Figure 5.1, we compare the nonexempt group (83,217) and exempt group (345,841) under the EU classification in Equation 4.1. And we compare the EU group (83,217) and the NonEU group (87,839) under the classification of Nonexemption in Equation 4.2. We use the whole sample-672,431 returns in Equation 4.3.

Similarly, we use another triple difference to measure the influence from the strictness of pre-existing rules for protecting the investors and monitoring the market risk. The following equation is used to test our second hypotheses:

$$\begin{aligned} Performance_{i,t} = {} & \gamma_i + \gamma_t + \beta_0 + \beta_1 \text{disclosure} + \beta_2 \text{post} + \beta_3 \text{nonexemption} \\ & + \beta_4 \text{disclosure\_post}_{i,t} + \beta_5 \text{nonexemption\_post}_{i,t} \\ & + \beta_6 \text{disclosure\_nonexemption} \\ & + \beta_7 \text{disclosure\_nonexemption\_post}_{i,t} + controls_{i,t} + \epsilon_{i,t} \end{aligned} \tag{4.4}$$

As we discussed in the background of pre-existing hedge fund rules in individual member states, it is challenging to quantify their overall strictness by analyzing specific requirements since various countries may have different levels of enforcement and specifications in a certain aspect of regulation and some countries have multiple regulatory frameworks that differ in demands. In equation (4.4), we use the business extent of the disclosure index from World Bank Open Data to reflect the strictness of pre-existing hedge fund regulation. The data is collected from more than 9,000 local experts, such as lawyers, accountants, business consultants, government officials who routinely advise or administer on legal requirements, through a standardized survey provided by the World Bank. And two aspects of the information, data from readings of laws and regulations and data on time and motion indicators that measure efficiency in achieving a regulatory goal are used to identify the disclosure index which ranges from 0 to 10 to measure the extent to which investors are protected through disclosure of ownership and financial information from individual countries. It cannot quantify the whole strictness specifically for hedge fund rules, however, can partly represent the pre-existing degrees of requiring necessary procedures and transparency for investor protection. The dummy variable disclosure equals 1 if the index is above 6 as the median index of our EU hedge fund sample and equals 0 otherwise. And β_7 represents the double-interaction influence among the pre-existing disclosure, post, and funds with no exemption.

Propensity score matching

Above DiD and DDD methods are based on the single point- July 2013 as the required deadline for transposing into national laws for dividing the timeline into pre-AIFMD period and post-AIFMD period. However, in practice different member states may not take actions coinciding with the deadline and they usually allow hedge funds to register within one year after the law became effective. And so the real situation involves date differences to make the comparison based on the deadline of enforcement. In addition, it is unavailable to observe the relocation information of hedge funds since such information is not frequently changing and the hedge funds do not always update in a timely manner. We presume a part of the hedge funds would relocate from off-shore sites to the EU if they do not want to give up the EU market or from EU to other places if they plan to avoid the EU regulations. And we do find some hedge funds domiciled in off-shore sites claim their marketing scope includes the EU, which is theoretically difficult to achieve. Our sample in the above regression does not take the changes of domiciles into consideration.

To address the two concerns mentioned above, we use the data from ESMA which directly shows the authorization date and withdrawal date for entities under AIFMD, which makes our sample of EU hedge funds and the influential date valid to test. Specifically, we use the PSM to match the periodical FH alphas from registered hedge funds with the periodical FH alphas from unregistered hedge funds based on the similarity of the fund characteristics. Following Shi (2017), we exactly match the investment strategies and the time variable. The hedge funds are commonly analyzed by investment strategies in prior studies and different strategies have been shown to differ in their performance (Aggarwal & Jorion 2010; Bali *et al.* 2013). And the time variables affect the general hedge fund returns especially when the market has fluctuations (Agarwal *et al.* 2017).

We include other variables that are both likely to affect the performance and the authorization. Prior studies provide evidence that size is related with the hedge fund performance even though their results are inconsistent. In our context, the size of a hedge fund directly also relates to the obligation of authorization. The flows with the prior assets under management are correlated with size, and hence the flows are included in our variables for selecting appropriate control observation.

Considering the scope of the AIFMD, we add the EU/NonEU and UCITS/ NonUCITS for the nearest match. The UCITS hedge funds are not required to apply to be AIFMD hedge funds, but some UCITS hedge funds still register as AIFMD hedge funds. These funds are convenient in the application since the two sets of rules share some similarities and authorized managers of a UCITS are entitled to request to be authorized as AIFM subject only to complying with any relevant additional requirements for the new authorization. It is necessary to match based on close groups for consistent comparison. We expect the EU hedge funds and NonUCITS funds are more likely to apply for the registration and the different groups may differ in their performance. Furthermore, we have other fund characteristics that are possible to influence the performance, such as leverage, lockup, high watermark, hurdle rate, management fee and incentive fee, and further influence their size and possibility to be under the control of AIFMD.

After the matching, we use the DiD method to test the influence of authorization based on the propensity matching results:

$$Performance_{i,t} = \gamma_i + \gamma_t + \beta \text{treated\_post}_{i,t} + controls_{i,t} + \epsilon_{i,t} \tag{4.5}$$

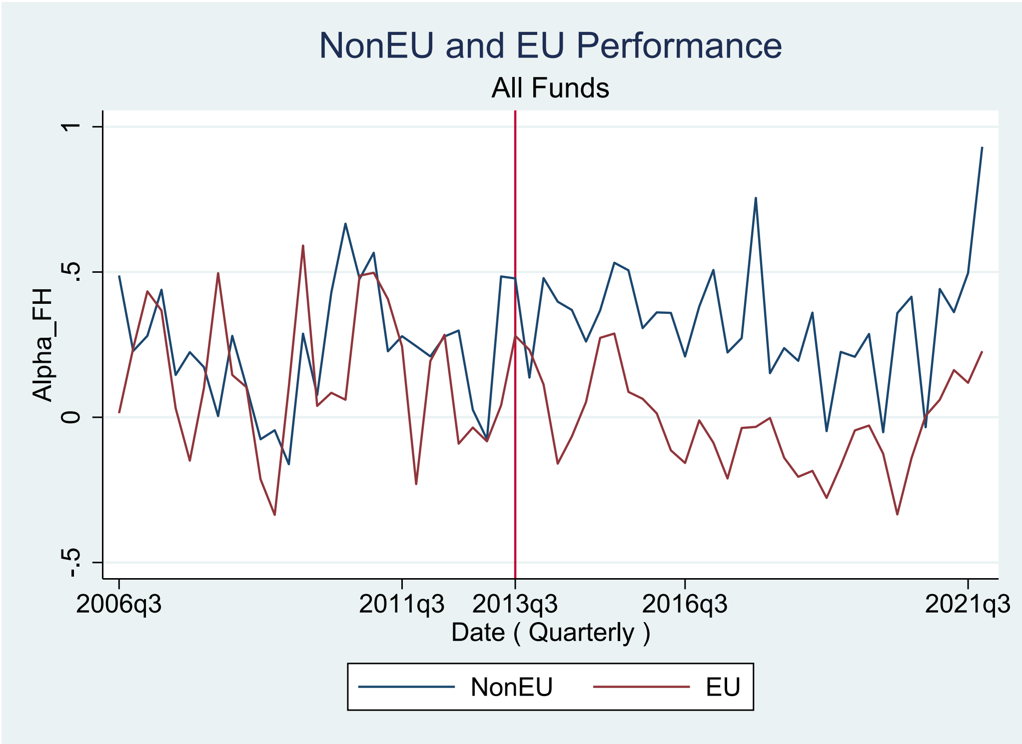
4.5 Results

Figure 4.2 describes overall EU and NonEU hedge fund performance measured by FH alphas from 2006 to 2021. The x-axis represents the timeline that ends the two-year window of individual hedge funds. The figure shows that before the AIFMD was imposed, there was not much difference between EU and NonEU alphas. Instead, they overlapped to a large extent in their performance. However, the EU funds performed worse than the after-2013 period while the NonEU funds seem better than the pre-2013 period. And so EU hedge funds were about 0.5% lower in alpha than NonEU hedge funds on average after 2013.

Figure 4.3 also presents the performance of EU and NonEU groups, but the hedge funds are those constrained by the AIFMD according to the size and NonUCITS identity. The results show that before 2013, they were still similar but differed in fluctuations. Neither EU nor NonEU hedge funds outperformed the other group all the time before the AIFMD became effective. After 2013, EU funds' alpha became lower than NonEU funds' alpha. However, the difference is not huge.

Table 4.2 shows the DiD results on the interaction between the funds with no exemption and Post-AIFMD. The sample is restricted to all hedge funds domiciled in the EU or marketing in the EU. From column 1 to column 5, the table lists different estimates obtained from different hedge fund measurements. Our results, obtained from excess returns, three-factor model, and four-factor model as the measurements of hedge fund performance show that within the EU, the managers of larger funds and nonUCITS funds that are subject to AIFMD experienced a negatively significant influence from 0.03 to 0.15 compared with managers of funds with exemption while the one-factor model and

Figure 4.2: NonEU and EU Performance

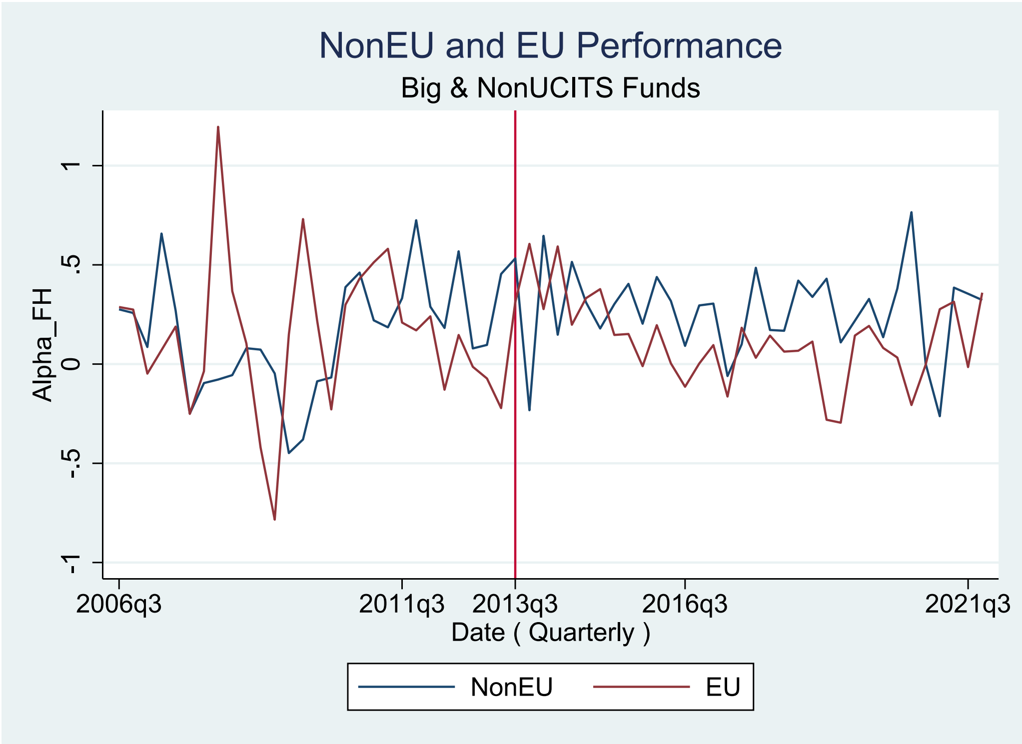


Notes: The table shows all EU and NonEU hedge fund performance measured by FH alphas from 2006 to 2021. The x-axis represents the timeline that ends the two-year window of individual hedge funds.

FH model do not convey such significant result to support H1. Relying solely on the coefficient of interaction in equation (4.2), the results are insufficient to prove the relationship between the EU regulation and hedge fund performance. However, the insignificance may come from the spillover effect from funds without exemptions to funds with exemptions.

Table 4.3 shows the DiD results on the dummy variable EU and Post. The sample is restricted to all hedge funds with no exemption. Due to the fixed effects included in the regression, the constant variable EU is omitted for collinearity. The coefficient of EU_Post captures the effect of AIFMD on hedge fund performance comparing the EU and other regions. It shows, based on the alphas from risk factor models, the EU funds without exemption experienced a drop varying from 0.06 to 0.29 in performance after the rule compared with the corresponding NonEU funds of other areas. The result is significant at 1%. If we simply measure the excess returns, we would get a negative but insignificant result. The inconsistent results also support our previous opinion that the influence differs possibly because of the different choices of measurements of hedge funds. In general, the majority estimates support H1. However, we cannot rule out the possibility that EU funds and nonEU funds have different sensitivities to the market situation or time trend. If the macro factors cause different influences between the treated and control groups, we could also obtain similar significant results.

Figure 4.3: Performance of EU and NonEU Funds without Exemption



Notes: The table shows nonexempted EU and NonEU hedge fund performance measured by FH alphas from 2006 to 2021. The x-axis represents the timeline that ends the two-year window of individual hedge funds.

Table 4.4 shows the results of triple differences after we include all factors. All estimates of EU_Nonexemption_Post provide strong evidence of a significant decline of around 0.2% in relevant hedge funds' performance between treated group and control group within the EU comparing the corresponding difference outside the EU. The coefficient of EU_Post denotes the reverse effect of DiD interpreted as the difference in performance change of funds with exemption between the EU and NonEU regions. Our results show 2 out of 5 estimates of EU_Post indicate a positively significant relationship suggesting the EU funds with exemption have better performance than NonEU funds with exemption while the other 3 estimates show an insignificant relationship. The coefficient of EU_Nonexemption represents the performance of two groups within each EU or NonEU region prior to the enforcement of the AIFMD. And all alpha results for EU_Nonexemption suggest insignificant performance differences between the two groups between EU and NonEU, which provides evidence to confirm our assumption about the parallel trend in the performance differences of two subgroups in two areas. The majority of coefficients of Nonexemption_Post reflect a significant increase in performance between funds without exemption and funds with exemption in nonEU regions, which is opposite to the effect in the EU. All these results show that starting from a similar pattern for non-exempt funds relative to exempt funds in trends before AIFMD, in NonEU areas,

Table 4.2: Difference-in-differences on Nonexemption and Post

	(1)	(2)	(3)	(4)	(5)
			With fixed effects		
	Excess Return	Alpha_1 factor	Alpha_3 factors	Alpha_4 factors	Alpha_FH model
Post	5.988	0.843	1.341***	1.459***	0.407
	(1.40)	(1.63)	(2.82)	(3.31)	(0.64)
Nonexemption	-0.219***	-0.215***	-0.175***	-0.173***	-0.120*
	(-4.30)	(-4.19)	(-3.32)	(-3.30)	(-1.77)
Nonexemption_Post	-0.0803**	-0.0287	-0.151***	-0.150***	-0.0995
	(-2.14)	(-0.60)	(-3.21)	(-3.22)	(-1.61)
aum_euros	-1.72×10^{-4}***	-1.12×10^{-4}***	-8.61×10^{-5}***	-1.08×10^{-4}***	6.91×10^{-5}**
	(-5.56)	(-3.84)	(-3.22)	(-3.89)	(-2.28)
lnaum	-0.117***	-0.0700***	-0.0803***	-0.0739***	-0.0624***
	(-8.10)	(-4.12)	(-4.53)	(-4.27)	(-2.71)
aumsq	6.52e-09***	4.59e-09**	4.25e-09***	5.20e-09***	3.58e-09**
	(3.46)	(2.54)	(3.00)	(3.29)	(2.35)
_cons	-2.786	0.291	0.111	-0.0822	0.716
	(-0.65)	(0.70)	(0.31)	(-0.26)	(1.48)
N	429058	14246	14246	14246	14246
R2	0.157	0.162	0.171	0.166	0.131

Notes: The table shows our DiD results from Equation 4.1. From column (1) to Column (5), the hedge fund performance is measured by excess return, CAPM model, 3-factor model(Fama & French 1998), 4-factor model(Carhart 1997), and Fung *et al.* (2008) 8-factor model respectively. T-statistics are reported in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels.

non-exempt funds outperform in trend than exempt funds after 2013 while in the EU they underperform. The evidence supports H1 and we find that even in DiD the choice of measurement matters for the effect of regulation, it does not have significant difference when we use the DDD and only assume the relative parallel for comparing subgroups in EU and NonEU regions.

In Table 4.5, we present the influence from the pre-existing disclosure index in member states on the regulation effect. The results, on average, show that pre-existing disclosure index is significantly and positively relevant to hedge fund performance. Combining the negative relationship between the EU regulation and hedge fund performance, the results indicate that the more transparency a member state required before the AIFMD, the less negative influence the member state would experience after the rule was transposed into national laws. And the difference of the regulation effect between countries with low pre-existing disclosure index and high index is quantified as 0.3 on average. Our results provide strong evidence for H2.

All above results are obtained through the comparison of hedge fund performance before and after July 2013, which is considered as the representative point for enforcement of the EU regulation. However, the representative point does not apply to all entities. Besides the Eurekahedge, we use the ESMA data with specific dates of authorization

Table 4.3: Difference-in-differences on EU and Post

	(1)	(2)	(3)	(4)	(5)
			With fixed effects		
	Excess Return	Alpha_1 factor	Alpha_3 factors	Alpha_4 factors	Alpha_FH model
EU	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
Post	-0.159	-0.413**	-0.236	-0.179	0.0000287
	(-0.33)	(-2.35)	(-1.48)	(-1.15)	(0.00)
EU_Post	-0.0574	-0.170***	-0.257***	-0.293***	-0.211***
	(-1.00)	(-2.71)	(-4.09)	(-4.53)	(-2.99)
aum_euros	1.41×10^{-4}***	1.15×10^{-4}**	9.44×10^{-5}	9.70×10^{-5}*	6.96×10^{-5}
	(3.99)	(2.28)	(1.64)	(1.77)	(1.47)
lnaum	-0.594***	-0.474***	-0.443***	-0.445***	-0.341***
	(-17.56)	(-9.49)	(-8.44)	(-8.53)	(-8.34)
aumsq	-3.21e-09***	-2.82e-09**	-1.70e-10	-5.19e-10	1.47e-09
	(-4.96)	(-2.47)	(-0.13)	(-0.40)	(1.36)
_cons	12.29***	9.540***	9.007***	9.035***	6.676***
	(15.85)	(9.90)	(8.92)	(9.01)	(8.43)
N	171056	5370	5370	5370	5370
R2	0.173	0.271	0.258	0.242	0.155

Notes: The table shows our DiD results from Equation 4.2. From column (1) to Column (5), the hedge fund performance is measured by excess return, CAPM model, 3-factor model(Fama & French 1998), 4-factor model(Carhart 1997), and Fung *et al.* (2008) 8-factor model respectively. T-statistics are reported in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels.

and withdrawal, on one hand, to address the concern that the dates of authorization of different entities differ because of the time difference of effectiveness of AIFMD in national laws and the application date, on the other hand, to ensure that our treated group includes only hedge funds that have been registered by the EU authority considering that some information like country focus or location in the commercial database may not be updated in time.

We use PSM to match the sample performance of hedge funds based on the funds' characteristics, the sample date, and the investment strategy. First, we exactly match the date and investment strategy to make sure the significant factors, including the time period and strategy, do not influence the funds' performance. Then, we use the funds' characteristics to obtain their propensity score to match the treated observations with control variables using the nearest-neighbor method without replacement. To ensure the quality of matches, we set a caliper of 1%. Figure 5.2 shows the comparison results of propensity scores of treated and control groups before and after matching. Without the effect of the regulation, we expect the trend of performance is similar and so we analyze the performance trend of registered funds with their peers using their specific dates of authorization and withdrawal to define the valid regulatory period.

Table 4.4: Difference-in-difference-in-differences on EU, Post, and Nonexemption

	(1)	(2)	(3)	(4)	(5)
			With fixed effects		
	Excess Return	Alpha_1 factor	Alpha_3 factors	Alpha_4 factors	Alpha_FH model
EU	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
Post	-1.221	-2.053**	-1.287	-0.946	-1.586*
	(-0.73)	(-2.07)	(-1.44)	(-1.15)	(-1.76)
Nonexemption	-0.394***	-0.308***	-0.268***	-0.268***	-0.228***
	(-8.28)	(-5.52)	(-4.45)	(-4.38)	(-3.89)
EU_Post	0.210***	0.111**	0.0185	-0.0241	0.0278
	(4.80)	(2.33)	(0.35)	(-0.44)	(0.44)
Nonexemption_Post	0.171***	0.195***	0.124**	0.0913	0.137**
	(3.34)	(3.36)	(1.99)	(1.44)	(2.00)
EU_Nonexemption	0.135**	0.0872	0.0519	0.0632	0.110
	(2.08)	(1.28)	(0.71)	(0.85)	(1.48)
EU_Nonexemption _post	-0.222***	-0.214***	-0.240***	-0.215***	-0.236***
	(-3.51)	(-3.01)	(-3.24)	(-2.87)	(-2.72)
aum_euros	-1.77×10^{-4}***	-1.42×10^{-4}***	-1.46×10^{-4}***	-1.55×10^{-4}***	-9.54×10^{-5}***
	(-5.63)	(-5.30)	(-5.79)	(-6.11)	(-4.56)
lnaum	-0.104***	-0.0668***	-0.0631***	-0.0598***	-0.0622***
	(-9.94)	(-5.57)	(-4.95)	(-4.55)	(-4.04)
aumsq	5.20e-09**	4.57e-09**	6.03e-09***	6.31e-09***	4.78e-09***
	(2.56)	(2.49)	(4.14)	(4.10)	(4.91)
_cons	3.821**	2.974***	2.422***	2.106**	2.676***
	(2.27)	(3.00)	(2.70)	(2.54)	(2.96)
N	672431	22395	22395	22395	22395
R2	0.142	0.152	0.133	0.119	0.080

Notes: The table shows our DDD results from Equation 4.3. From column (1) to Column (5), the hedge fund performance is measured by excess return, CAPM model, 3-factor model(Fama & French 1998), 4-factor model(Carhart 1997), and Fung *et al.* (2008) 8-factor model respectively. T-statistics are reported in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels.

Table 4.6 shows the logit regression for estimating the relationship of the authorization and the characteristics we use in matching. As we expected, the size and the flows are significantly positive with the hedge fund possibility to apply for registration. And EU funds have a high possibility to be registered. The lockup is negatively related with the registration probably due to the rule that smaller funds, which have AUM between 100 million euros and 500 million euros with lockup greater than 5 years, are exempted from strict regulations. The rule allows smaller funds with longer lockups not constrained strictly by the AIFMD. Unlike our expectation, the UCITS group which is exempted from the authorization has a higher possibility to apply for AIFMD hedge funds. The potential reason could be the convenience of application for UCITS hedge funds due to

Table 4.5: Difference-in-difference-in-differences on Pre-disclosure index, Nonexemption, and Post

	(1) Excess Return	(2) Alpha_1 factor	(3) Alpha_3 factors	(4) Alpha_4 factors	(5) Alpha_FH model
Disclosure	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
Post	9.800**	0.613	1.132*	1.392***	-0.687
	(2.12)	(0.82)	(1.90)	(2.71)	(-0.98)
Nonexemption	0.00804	-0.212**	-0.102	0	-0.145
	(0.06)	(-2.33)	(-1.12)	(.)	(-1.54)
Disclosure_Post	-0.0492	0.0501	0.0276	-0.0320	-0.0755
	(-1.33)	(1.06)	(0.58)	(-0.70)	(-1.29)
Nonexemption_post	-0.339***	-0.299***	-0.395***	-0.578***	-0.193*
	(-2.97)	(-2.63)	(-3.63)	(-6.27)	(-1.77)
Disclosure_Nonexemption	-0.283*	0.0333	0.0615	0	-0.0449
	(-1.69)	(0.24)	(0.42)	(.)	(-0.30)
Disclosure_Nonexemption _post	0.300**	0.393***	0.295**	0.269**	0.300**
	(2.19)	(2.73)	(2.16)	(2.35)	(1.99)
aum_euros	-1.35×10^{-4}***	-9.17×10^{-5}***	-7.95×10^{-5}***	-1.02×10^{-4}***	-4.96×10^{-5}*
	(-5.24)	(-3.16)	(-3.04)	(-3.79)	(-1.93)
lnaum	-0.129***	-0.0654***	-0.0720***	-0.0677***	-0.0556***
	(-8.54)	(-3.29)	(-3.62)	(-3.55)	(-2.82)
aumsq	4.56e-09***	3.13e-09**	3.36e-09***	4.34e-09***	2.10e-09*
	(3.51)	(2.07)	(2.85)	(3.24)	(1.91)
_cons	-6.505	0.334	0.0771	-0.152	1.669***
	(-1.41)	(0.51)	(0.16)	(-0.39)	(2.74)
N	353101	11514	11514	11514	11514
R2	0.193	0.197	0.203	0.207	0.174

Notes: The table shows our DDD results from Equation 4.4. From column (1) to Column (5), the hedge fund performance is measured by excess return, CAPM model, 3-factor model(Fama & French 1998), 4-factor model(Carhart 1997), and Fung *et al.* (2008) 8-factor model respectively. T-statistics are reported in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels.

the similarity of the sets of regulations.

Table 4.7 compares the treated group and the matched control group after the PSM. The results show that the two groups have similar characteristics. Even though the hurdle rate and lockup show a difference at 5% significance, the hurdle rate and lockup are measured as dummy variables, the hedge funds with hurdle rate and lockups represents a rather small portion both in the treated group and control group. Thus, the matching has relatively high quality.

In Table 4.8, the DiD results after PSM also provide strong evidence that funds experienced a 0.2% drop in alpha after the authorization process and being imposed by the relevant rules compared with the corresponding control group. The drop is slightly lower compared to our DDD results but these two results are very close. The regression

Table 4.6: Logit regression in the propensity score matching

Variables	authorization	t-statistic
lnaum	0.0807***	(6.57)
Flow_period	0.000116**	(2.51)
Hurdlerate_Dummy	-0.265***	(-4.94)
Watermark_Dummy	0.0635	(0.96)
Leverage	0.0738***	(5.46)
Management Fee	0.00602	(0.14)
Performance Fee	-0.0241***	(-6.42)
Lockup_Dummy	-0.581***	(-5.57)
EU	1.763***	(22.09)
NonUCITS	-0.506***	(-8.62)
_cons	-3.314***	(-13.01)
N	12107	

Notes: The table shows the logit regression results for estimating the relationship of the authorization and the characteristics we use in matching. T-statistics are reported in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels.

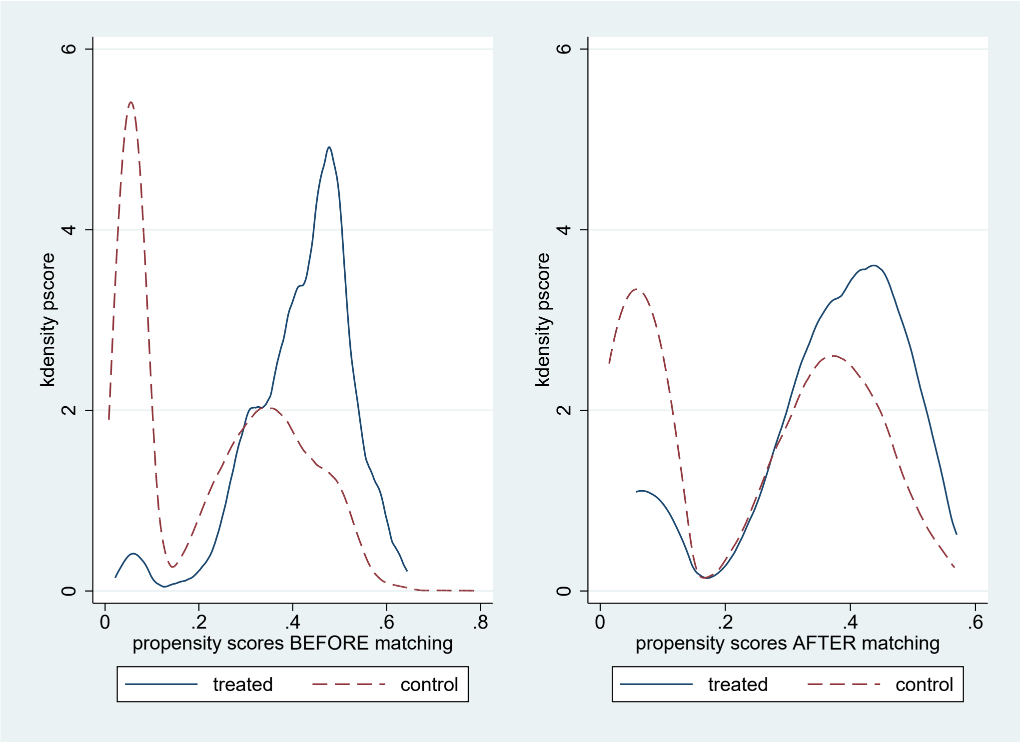
provides further evidence for the negative relationship between the AIFMD and hedge fund performance. We conclude that hedge funds are negatively affected by the EU regulation by around 0.2% in their performance.

4.6 Conclusion

In this article, we investigate the effect of the EU AIFMD on the hedge fund performance. Through analyzing the main aspects of change in the regulations compared with the US regulations, we find the EU regulation has more extensive constraints, compliance costs and wider scope of disclosure to the public. Combining the analysis of prior methods and findings and the fact that the EU regulation is stricter than the US rules, we expect the hedge fund performance is negatively related with the EU regulation.

We develop based on the common DiD method and introduce the third factor to

Figure 4.4: Comparison of samples before and after PSM



Notes: The figure shows the comparison results of propensity scores of treated and control groups before and after matching.

Table 4.7: Characteristics of control group and treatment group

	(1) Control Group		(2) Treated Group		(3) Difference	
	mean	sd	mean	sd	mean	t-statistic
lnaum	18.114	2.294	18.105	2.542	0.009	(0.059)
Flow_period	-26.883	340.723	-18.99	375.857	-7.893	(-0.331)
Hurdle Rate	0.254	0.436	0.199	0.399	0.055**	(1.987)
Watermark	0.72	0.45	0.775	0.418	-0.055*	(-1.913)
Leverage	0.71	1.633	0.573	1.244	0.137	(1.415)
Management Fee	1.416	0.59	1.378	0.806	0.038	(0.810)
Performance Fee	14.475	7.801	14.15	7.493	0.325	(0.640)
Lockup	0.026	0.161	0.057	0.233	-0.207**	(-2.325)
EU	0.64	0.48	0.642	0.48	-0.002	(-0.069)
NonUCITS	0.545	0.498	0.541	0.499	0.004	(0.133)
N	453		453		906	

Notes: The table compares the treated group and the matched control group after the PSM. Column (3) is the difference between the control group and the treated group. T-statistics are reported in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels.

formulate the DDD interaction term and avoid the limitation involved in the DiD method for a more accurate result. Besides, we use the PSM approach to test further through comparing the change of performance using the effective date of the AIFMD of individual

Table 4.8: DiD regression using the matched group

	alpha_FH	t-statistic
treated_post	-0.208***	(-2.75)
aum_euros	-2.15×10^{-4}**	(-2.04)
lnaum	-0.00558	(-0.31)
aumsq	5.74e-08**	(2.45)
_cons	0.185	(0.61)
N	906	
R2	0.261	

Notes: The table shows the regression results from Equation 4.5. T-statistics are reported in parentheses. ***, **, and * denote statistical significance at 1%, 5%, and 10% levels.

hedge funds. Our hypothesis is supported by empirical evidence. Our paper does show that the DiD method may convey different results based on various measurements of hedge fund performance while the coefficient of triple interaction consistently shows that hedge fund performance is negatively affected by the AIFMD. Our PSM analysis has similar results to reflect a negative change in performance by 0.2% in alpha. Our research also supports that the more transparent a member state was, the less negative influence it had after the AIFMD. We provide evidence from another dataset and contribute to the understanding of current conflicted findings in the area, the improvement the judgment of investors of hedge funds. It is also valuable for authorities in making or amending relevant hedge fund rules.

Nonetheless, our study has some limitations. First, our data may be subject to self-selection bias since we use a single commercial database that may miss some hedge fund data due to voluntary reporting. Second, using the extent of business disclosure before the rule to represent the pre-rule strictness of the hedge fund industry in member states may be imprecise. But due to the lack of other information, we utilize the most proper data on hand. Considering the impact of the EU regulation, one of the interesting aspects is to discuss the extent of regulatory arbitrage, however, the aspect has been seldom detected because of the historically incomplete registration information and disclosure of hedge funds to the corresponding authority. Our research still follows the stream of literature discussing the relationship of hedge fund performance and regulations but further studies may analyze the regulatory arbitrage if the hedge funds provide additional information on their relocations.

Chapter 5

Response Letter

Dear Reviewers and Examination Committee Members,

Thank you for reviewing my dissertation thesis entitled *"Hedge Fund Regulation, Characteristics, and Performance"* and for providing valuable feedback for me.

I was glad to learn that all three Reviewers viewed my dissertation thesis rather favorably, and they recommended the thesis for defense without substantial changes. At the same time, I greatly appreciated your comments and suggestions. Your feedback allowed me to take a fresh look at my thesis and make changes to further enhance its quality. Your feedback also let me understand which parts benefited most from improvements and concentrate my efforts on most carefully updating specifically these parts.

In this response letter, I detail out how I responded to the individual points you made. Wherever applicable, I explain and motivate the corresponding changes. For ease of exposition, I reproduced the text of the reviews in **black** text color, and I described and motivated my responses in the revised version of my dissertation thesis in **blue** text color. To make it easier for you to find the changed text, I also use the **blue** color for the new parts of my dissertation thesis. I plan to remove the text coloring in the very final version of the thesis.

I hope you will appreciate the changes I made and you will consider them beneficial for the quality of my dissertation thesis.

Kindest regards,

Fan Yang
Doctoral Candidate
Charles University

Examination Committee

Feedback was provided orally after the pre-defense.

I thank you for the feedback you kindly provided after my pre-defense. As you suggested, I carefully studied the comments provided by the three Reviewers. In updating my dissertation thesis, I closely followed the individual comments the Reviewers made. At my pre-defense, you emphasized that especially my paper on the impact of regulatory changes on hedge fund productivity (Chapter 4) would benefit from further improvements. Following your guidance, I concentrated my efforts on updating and improving particularly that part of my thesis.

Supervisor – Professor Jiri Novak

I do not have any additional major comments on what should be improved in the dissertation thesis for the purposes of its defense. I expect Fan to improve the individual papers based on the comments she receives from her referees and from the dissertation committee.

Thank you for your positive evaluation of my dissertation thesis. I recognize the need for further improvements of the individual papers based on the comments received from the Referees and the Examination Committee. Below, I detail out how I responded to these comments.

Referee: 1 – Professor Mattias Hamberg

Introduction

Hedge funds have been around for several decades and represent an interesting phenomenon well worthy of further exploration. I feel privileged to have had the opportunity to read and learn from the proposed doctoral thesis *"Hedge Fund Regulation, Characteristics, and Performance"*. The work consists of three separate, yet closely connected, research studies.

In brief, the thesis consists of (i) a co-authored literature review of potential methodological issues connected to the assessment of hedge fund performance, (ii) a co-authored meta-analysis of past research, and (iii) a single-authored empirical study conducted by the defendant. Personally, I think that is a very appropriate mix of studies for a doctoral thesis, and I very much approve the overall set-up.

A written assessment of any advanced research topic, that is not already published and final, is tricky in the sense that the author is unable to share her knowledge and insight to me directly. I thus apologize if I make any misunderstandings. I have been suggested to be rather brief; and hence, I keep my comments as short and concise as possible. I also make no comments on language issues and minor details. They exist, and they are concentrated to the third paper.

Given that the first two research studies are conducted together with several senior researchers with a successful track record, it is no wonder that the quality of the work is high. For this reason, I believe that the most important comments relate to the defendant's third paper which I deem as the most important one to assess, and it is also the one that has the greatest opportunities for improvement in the time until the final defense.

Thank you for your positive evaluation of the topic and the composition of my dissertation thesis. I am pleased that you consider hedge fund performance an interesting phenomenon that is well worthy of further exploration. I am also glad that you find the mix of co-authored and single-authored empirical studies appropriate for a doctoral thesis. I understand that you consider the third paper less developed than the other two. Following your guidance, I concentrated my efforts on updating and improving especially that part of my thesis. Below, I detail out how I responded to your comments.

Chapter 2: Is research on hedge fund performance published selectively? A quantitative survey

In this co-authored paper, the defendant analyses whether estimates of hedge fund performance are affected by publication bias. My role is primarily to assess whether the defendant's work is sufficient for defense. I find the topic of the paper to be relevant, neatly presented, and with solid empirical analyses.

Before a final defense, this paper does not have to be changed in any material way.

However, let me just make a few extra remarks. The research could be one of the first quantifications of the impact of potential selective publication and data biases on the hedge fund performance estimates. I believe that the authors should state more upfront how their findings contrast existing research, and adds incremental knowledge. Having a good clear edge, benefits the selling of the work.

The authors may also want to reconsider the discussion of performance measures in theory. My suggestion is primarily cosmetic. Maybe retitle Section 2.2 as "Previous research on the assessment of hedge fund performance".

Thank you for these comments. I am pleased that you find the topic of the paper to be relevant, neatly presented, and with solid empirical analyses. I am also glad that you believe that this paper does not have to be changed in any material way before my final defense.

Indeed, to the best of my knowledge, our paper is the first systematic quantitative review of prior empirical research on hedge fund performance that aims at quantifying the impact of potential publication selection bias on the estimates reported in that stream of literature. I agree that we should have better motivated our original contribution relative to prior research literature. I also agree that the heading you suggested for Section 2.2 is very fitting. In response to your comments, I updated my thesis in the following way.

First, I have reformulated some paragraphs and added additional texts in the introduction of Chapter 2 to motivate the incremental contribution of this study to the hedge fund research literature and contrasting findings in publication bias literature. I copy the new pieces of text below.

"Prior empirical literature includes numerous conflicting results, which make it difficult to draw clear conclusions. Some studies analyze the potential reasons that include various data sources, data biases and measurement approaches. For instance, Liang (2000) and Fung et al. *(2006) show that survivorship bias vary from 0.39% to 2.4% across different databases. The selection bias is round 1.17% in Edwards & Caglayan (2001) and 1.5% in Fung* et al. *(2006). Several papers point out that the risks in hedge funds are complicated, and they develop multiple measurement methods to capture the hedge fund performance (e.g., Sadka 2010; Bollen & Pool 2009). However, these examples show their concentration on individual aspects or constraints to evaluate the hedge fund performance but lack a study synthesizing this pool of diverse empirical results. Besides, we find that the fragmentation of hedge fund performance data and the wide range of alternative approaches to controlling for risk give researchers considerable discretion in research design."*

"The result is inconsistent with the suggestion from Brodeur et al. *(2020) that publication selection bias is related to the researchers' discretion possibly because less tendency exists regarding the favorable hedge fund results."*

Second, following your suggestion, I changed the heading of Section 2.2 to the following to clearly indicate the contents of that section.

"Previous research on the assessment of hedge fund performance"

Chapter 3: Where Have All the Alphas Gone? A Meta-Analysis of Hedge Fund Performance

In this co-authored paper, the defendant analyses results presented in a large number of previous studies. My role is primarily to assess whether the defendant's work is sufficient for defense. I find the topic of the paper to be highly relevant, very neatly presented, and the empirical analyses are trustworthy. I have little doubt that this paper can be accepted in a good journal in the area of financial economics.

The paper does not have to be changed in any way.

Thank you for your positive evaluation of this paper. It is great to hear that you consider the topic relevant, you like the presentation of the paper, and you find the empirical analyses trustworthy.

Chapter 4: The Impact of Regulatory Change on Hedge Fund Performance

0. Overall feedback.

This paper addresses whether European regulatory changes impact the performances of hedge funds. I appreciate the work, and I have no exceptional objections to the paper as such. However, I do not believe that the work, in its current design, should be submitted to a prestigious academic journal. Improving the readability of the paper would substantially improve the chances of a worthy publication. I explain my reasoning underneath as I raise four issues. None of them is sufficiently material to hinder the defense, but they should be thoughtfully considered.

Thank you for this comment. I recognize that you consider this paper to merit the most improvement. I have concentrated my efforts between the pre-defense and the final defense on improving the readability of this chapter and other aspects that you mention in your comments below.

1. Presentation of research problem and expectations.

I am positive to the work that has been conducted, including the research question and the empirical analysis in large, but I am concerned over the structure of the presentation. I would recommend the candidate to make improvements. To explain myself, let me briefly summarize the content.

Introduction: Very long (6,5 pages), and not straight to the point. It is not until p107 that the author communicates the purpose to me: "*It is therefore important to study how hedge funds perform before and after the implementation of the EU Directive and how certain aspects of major changes in special rules or their combinations influence the overall industry.*"

Literature review: The literature review is much shorter than the introduction (2,5 pages), and despite its title, it contains fewer references than the introduction. It also ends without making me understand what it is used for.

Institutional background: This is a section that I enjoy reading as it rather clearly explains the study's setting.

Hypotheses: Then comes the author's expectations, which make sense to me, but they are not at all embedded in a review of prior literature.

My opinion is that the defendant should rethink how to design the first half of the paper. If nothing else, this is crucial for publication-success. The paper would benefit from a more concise introduction, followed by a description of the institutional background.

Subsequently, I would like to read the theory and see the defendant's expectations properly embedded in the literature. Please note that I am not suggesting that the defendant changes the research objective or the hypotheses, but works on the presentation of the work.

I recommend a shorter introduction and more references to back up arguments. An example: p. 105. "*In fact, the failure rate of hedge funds is around 10 percent but the high level of failure merely eliminated worse hedge funds rather than increasing the systematic risk.*" – Reference?

I recognize the need to refocus and clarify the sections you allude to. In response to your suggestion, I have completely restructured the first four sections of this chapter. I have restructured and refocused the introduction in Section 4.1. As you suggest, I proceed with the discussion of the institutional background in Section 4.2 before I review of prior research literature on hedge fund performance in Section 4.3. Then, I discuss my methodology and data sample in Section 4.4. See below for a more detailed discussion of the individual sections in this chapter.

Introduction. I have completely rewritten the introduction in Section 4.1. Following your suggestion, I have built up the introduction around the purpose of this chapter. I provide the context for the purpose in the first two paragraphs. In the third paragraph, I present the purpose of this chapter and discuss the controversy concerning the *a priori* predictions (see the quote below). I also motivate why it is important to investigate this research question and how it connects to and contributes to prior research in the area.

> *"In this paper, we empirically investigate the impact of the AIFMD on hedge fund performance. We argue that the expected impact is not obvious. On the one hand, the enhanced transparency may facilitate oversight of hedge funds by investors and regulatory authorities, which may strengthen managerial motivation to perform and curb self-interested behavior that may arise from the agency problem (Frumkin & Vandegrift 2009). On the other hand, the more comprehensive disclosure may also reveal vital proprietary information to competitors who may imitate successful strategies and outsider investors who predict the trades and then trade against the fund, thus eroding hedge funds' competitive advantage (Cumming & Dai 2010; Shi 2017; Agarwal* et al. *2013b)."*

Institutional background. I am glad you enjoyed reading the section on the institutional background. Following your suggestion above, I have now placed the discussion of the institutional background in Section 4.2, just after the introduction. This section informs the reader about hedge fund regulation in general and also, more specifically, about the European directive before discussing the research question and prior research on the impact of this regulation on hedge fund performance.

Literature review. I have extended the literature review in Section 4.3. It now covers about 5 pages and provides a more comprehensive overview of relevant research. the literature review section is structured so that I can use the discussion of prior research literature to motivate my hypotheses.

Hypotheses. Following your suggestion, I have now integrated hypotheses into the literature review section (4.3). The individual hypotheses that reflect my *a priori* expectations about the nature of the studied relationship are also motivated by references to prior literature.

References. Thank you for pointing out the need to better substantiate my claims with references. When revising my dissertation thesis, I have erased the specific sentence you refer to in your comment above. Nevertheless, concerning other factual statements that I make, I have now included either academic references or hyperlinks to relevant websites to substantiate the factual claims.

2. Mechanisms that connect regulatory reforms and performance.

This comment could have been presented as part of my previous comment; however, I want to emphasize that the mechanism(s) behind the studied relationship is not clearly explained to the reader. The defendant provides some explanation in association with the development of hypotheses (and at a few other places in the introduction and the Institutional background), but it would be much better to upfront explain what the phenomenon is, and why it exists.

Thank you for pointing this out. I have realized that, in the earlier version of my dissertation thesis, the mechanism was not clearly explained and the hypotheses were not well connected to the findings from the prior studies. I improved the explanation of the mechanism behind the studied relationship, as you suggested. First, in the third paragraph of the revised version of the introduction (Section 4.1), I present the purpose of this chapter, and I discuss the controversy concerning the *a priori* predictions. This discussion explains to the reader why the new European regulation is likely to have the expected effect on the value hedge funds generate for their investors. In the introduction, I also discuss why it is important to empirically investigate this research question and what we already know about it from prior research. Second, in the literature review section (Section 4.3), I provide a detailed overview of prior studies on hedge fund regulation, and I discuss what implications this prior research has on the expected outcome of our empirical analysis concerning the AIFMD. I compare the specific requirements of the AIFMD with other hedge fund regulations discussed in these prior studies to make predictions for the effects of individual requirements and then make the overall prediction. In addition, I added the reasoning part from the perspective of potential bias of prior findings resulting from the choice of method to enhance the prediction. Here, I cite the process of the prediction of individual requirements and the overall prediction.

"As prior studies have similar conclusions that greater transparency is related to lower hedge fund performance by analyzing Form 13F under Section 13(f) of the Exchange Act, we expect that the negative effect of leaking managers' strategies and weakening their incentives outweighs the positive effect of reduction of managers' misconducts. Comparing the AIFMD with the US regulation, the EU Directive requires preparing audited annual reports for investors and authorities on demand instead of merely disclosing information to the authorities for monitoring. Therefore, competitors can more easily access investment techniques and enhance the negative influence caused by transparency. As the EU rules pay special attention to the use of leverage, specifying the disclosure of maximum leverage, the extent of leverage, and its application in certain circumstances. It seems the EU hedge fund manager would face greater pressure to operate their business with high leverage to chase superior revenues. From the perspective of information disclosure, the EU regulation should have a greater negative effect on performance than the US regulation."

"However, prior results show inconsistent overall effects of the combination of authorization, transparency, and related compliance costs. Since transparency probably leads to a negative impact, as we discussed above, and the compliance costs lower the earnings, the overall positive relationship between the US regulation, such as the Dodd-Frank Act, and the hedge fund performance may suggest that the undernourished hedge funds decrease due to the strict rule to register. Then the overall industry performance is expected to grow (Frumkin & Vandegrift 2009) after removing costs resulting from transparency and other compliance fees. The EU Directive requires minimum capital amounts and qualified managers to assure the industry's quality, which may positively influence hedge fund performance but raises the capital costs for some small funds. The quality improves hedge fund performance when the diminution of underperforming hedge funds exceeds that of wellperforming small funds. The improvements might exist based on the US regulation, while the EU regulation involves more cost as the hedge funds that are located in third countries are expected to have an additional one-time large cost for relocating to a member state of the EU; otherwise, they are not able to market in the EU. The cost of authorization and restriction is probably greater than the benefits from improving the overall hedge fund quality."

"Except for the similar rules to the US regulation, the AIFMD requires hedge funds to have periodic independent valuations and depositories. The depository could increase investors' confidence by reducing asymmetric information between the manager and investors (Kamal 2012). However, failing to correctly value the assets of hedge funds may lead to sanctions by the authorities for valuation shortcomings. It is effective in continuously controlling risk, while hedge funds would unavoidably distract attention and require more investment in their operation and management in case of any violation of the AIFMD. The compliance cost from the valuation and depository is continuous and may constitute another important part of the total cost."

"Combining the analysis of individual rules in EU regulation based on prior literature, we expect an overall negative relationship between the AIFMD and hedge fund performance......"

3. Sample and empirical analyses.

Section 4.5.1 has the title "Data Collection" and it begins with the total number of hedge funds available (26,862 global hedge funds). But from there it is currently impossible to see the sample selection procedures, and how the author conducts her empirical analyses. At the moment, I really cannot make sense of the data. I suggest additional explanations.

In a similar vein: The author conducts a series of empirical tests; presented in Table 4.2 to 4.8. None of these tests are based on the same samples. I am wondering if it is possible to either homogenize the sample selection, or to better explain relations between the samples? One way to do so would be to introduce a proper table on the sample selection procedure. Currently, Table 4.1 only contains descriptive statistics.

Finally, because I do not understand the sample size...how come there are hundreds of thousands of observations in the excess return regressions? For example, in Table 4.4, the author reports 672,431 observations. – Are these daily returns? Explain.

Thank you for this comment. I agree that the sample selection procedure and the additional explanation contribute to the readers' understanding since the paper uses different samples, which may cause some confusion. The 672,431 observations are monthly returns. The monthly returns are regressed based on the consecutive base of 2 years for obtaining the "alpha" as the risk-adjusted returns. As I discussed in Chapter 4, both excess monthly returns and risk-adjusted returns are considered for evaluating the hedge fund performance. To explain in detail my sample selection process and illustrate why the different samples are used in the regressions, I added two paragraphs and a figure to show the details. I cite them below:

"Figure 5.1 shows our sample selection process for both excess returns and alphas as measurements of hedge fund performance. We have the Eurekahedge initial sample with 1,637,520 monthly returns from 19,620 funds. First, we exclude the monthly returns if they are prior to the beginning date of hedge funds to add the information to the database to mitigate the backfilling bias. 1,052,676 monthly returns are left after the process. Second, we exclude the observations if they have missing values of lagged AUM that is considered as the control variable in the regression when we use excess monthly return over a risk-free rate to measure hedge fund performance. We exclude observations if they do not have lagged AUM of a 24-month period when we use the alpha as the measurement of hedge fund performance. Third, We remove the US observation in the comparison. As a result, we have 672,431 returns for using excess monthly returns over a risk-free rate to measure hedge fund performance and 30,935 alphas as abnormal returns for hedge funds. "

"Since we expect that the affected group is the nonexempt EU group, and the other three groups, including the exempt EU group, nonexempt NonEU group, and exempt NonEU group, are less likely to be influenced by the AIFMD, we do not use the constant sample for investigation. Instead, we use the subsamples to compare them with the treated group. Using the excess return over a risk-free rate as an example in the left side of Figure 5.1, we compare the nonexempt group (83,217) and exempt group (345,841) under the EU classification in Equation 4.1. And we compare the EU group (83,217) and the NonEU group (87,839) under the classification of Nonexemption in Equation 4.2. We use the whole sample- 672,431 returns in Equation 4.3. "

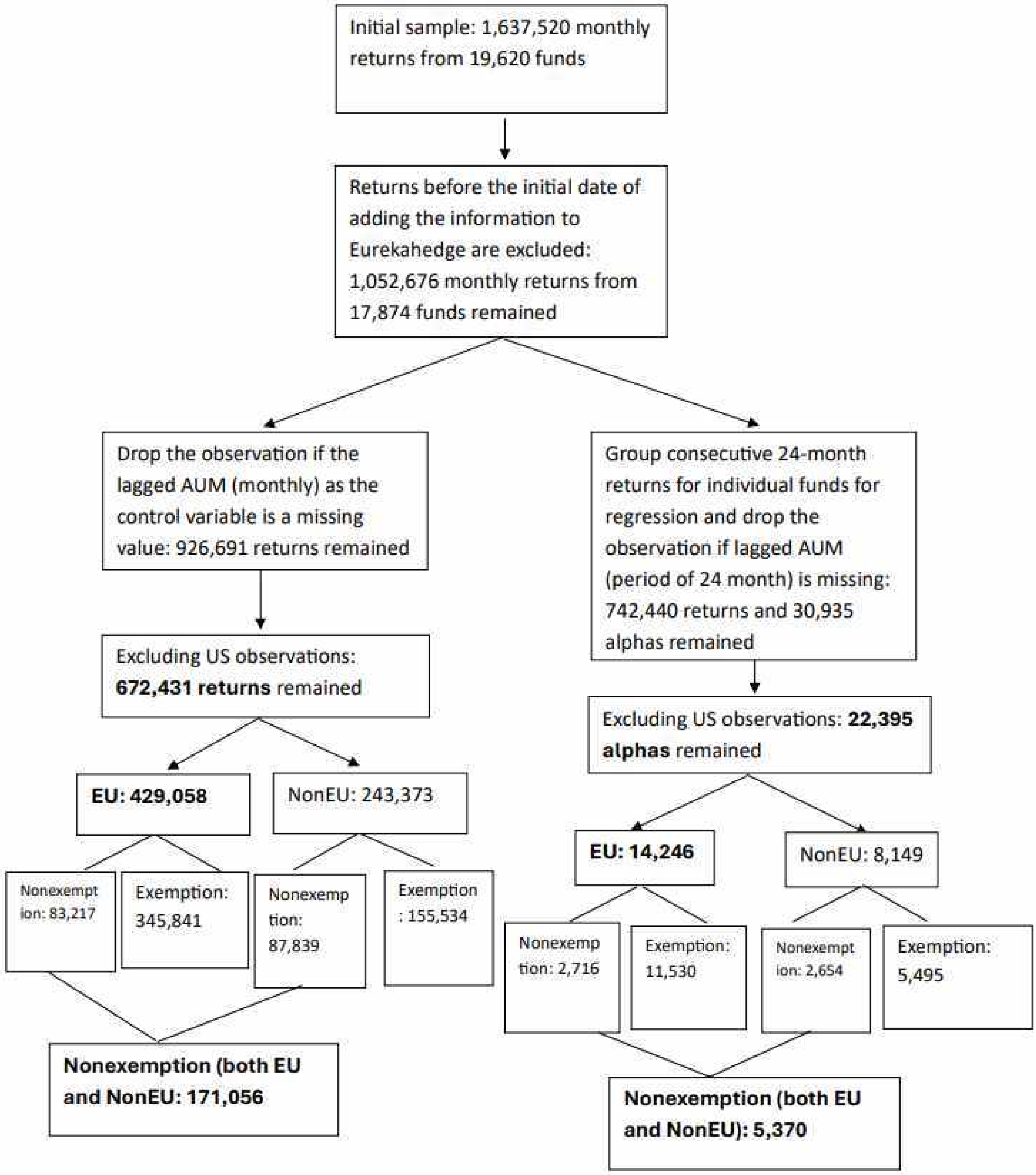
4. Measure of dependent variable.

I understand that there is a huge debate in empirical finance in terms of how to measure the risk-adjusted return. For this reason, empirical financial economics uses a large range of models (1-factor, 2-factor, 7-factor models, etc). I am not sure that presenting almost identical results from 4-5 similar models in each table enhances our understanding. Moreover, how should we interpret results when one or two models provide different results from the others? – Existing theory cannot guide us into knowing which model is the best.

My suggestion to the defendant is to decide what model(s) to use and then report that alternative models provide "qualitatively similar results." It is a rather easy fix; and as I see it, it improves readability.

Finally: It might be worth making tables self-explanatory. Or alternatively, to include a table that summarizes the variables' definitions.

Figure 5.1: Sample Selection Process



I show several models for two reasons. First, since there is no universally accepted model for measuring hedge funds (as we discussed in Chapters 2 and 3), I follow prior studies to use several risk models to provide more evidence on the potential impact model choice may have for the documented alphas. This evidence may convey additional information to readers. Second, I show these results in the main tables instead of in the robustness check due to their inconsistent results presented in the DiD method. Prior studies have conflicting results about the effect of strict regulations on hedge funds. The potential reasons could be the choice of measurement methods of hedge fund performance in DiD method. I choose the DiD method with various measurement methods to evaluate whether the measurement methods influence the effect. I do find inconsistent results in different evaluation methods. That, in turn, may make my choice of DDD valid because when I use the DDD further, I see consistent results. It indicates that the results from DDD are less likely to be impacted by the choice of evaluation methods.

Thank you for your suggestion to include a table that summarizes the variables' definitions. I added notes to the tables and figures in chapter 4 to clarify them. In addition, I added the variable description table as an Appendix to Chapter 4 (see below).

Table 5.1: Variable Description

Variable	Description
Dependent variable	
Alpha_FH model	Alpha obtained by Fung& Hsieh (2008) eight-factor model.
Alpha_1 factor	Alpha obtained by CAPM model.
Alpha_3 factors	Alpha obtained by Fama & French (1998) three-factor model.
Alpha_4 factors	Alpha obtained by Carhart (1997) four-factor model.
Excess Return	The performance is measured by the difference between excess return and risk-free return.
Control variable	
aum_euros	The lagged AUM in euros.
lnaum	The log terms of fund i's assets in euros at the end of period t-1.
aumsq	The quadratic terms of fund i's assets in euros at the end of period t-1.
Independent variable (DID and DDD)	
EU	Indicator variable equal to 1 if the fund is located or marketing in European Union and 0 otherwise.
Size	Indicator variable equal to 0 if aum_euros$\leq$100 or 100$<$aum_euros$\leq$500 & Leverage="No" & lockup is at least five years (meet exemption requirements) and 1 otherwise.
Post	Indicator variable equal to 1 if the period is after the AIFMD has be imposed and 0 otherwise.
EU_Post	Variable equal to EU*Post.

Continued on next page

Table 5.1: Variable Description (continued)

Variable	Description
NonUCITS	Indicator variable equal to 0 if the fund is subject to UCITS and 1 otherwise. Status in 2021 used for the entire period.
Nonexemption	Variable equal to Size*Nonucits.
Nonexemption_Post	Variable equal to Nonexemption*Post.
EU_Nonexemption	Variable equal to EU*Nonexemption.
EU_Nonexemption_Post	Variable equal to EU*Nonexemption*Post.
Disclosure	Variable equal to the business extent of the disclosure index from World Bank Open Data.
Disclosure_Post	Variable equal to Disclosure*Post.
Disclosure_Nonexemption	Variable equal to Disclosure*Nonexemption.
Disclosure_Nonexemption_Post	Variable equal to Disclosure*Nonexemption*Post.
PSM variable	
treated_post	Indicator variable equal to 1 if the fund is authorized and is within the authorized periods and 0 otherwise.
Hurdle Rate	Indicator variable equal to 1 if the fund has hurdle rate and 0 otherwise.
Flow_period	The flows during the 2 year window.
Watermark	Indicator variable equal to 1 if the fund uses watermark to evaluate managers' performance and 0 otherwise.
Leverage	Indicator variable equal to 1 if the fund use the leverage and 0 otherwise.
Management Fee	The management fee.
Performance Fee	The performance fee.
Lockup	Indicator variable equal to 1 if the fund has the lockup requirement and 0 otherwise.

Notes: The table provides the variable list and their descriptions.

Referee: 2 – Professor Karel Hrazdil

Introduction

Edit carefully and have a professional editor proofread your work. For example, "The chapter 4 analysis of the impact of the regulation based in the EU area is not only a plus to the existing literature." This sentence is full of grammar errors (i.e., chapter 4 analyzes) and incomplete (i.e., not only, but also).

Thank you for pointing out the need for language editing of my dissertation thesis. In response to your suggestion, I have thoroughly rewritten Chapter 1 and the first three sections of Chapter 4. In addition, I asked a professional language editor to further improve the use of the English language in my dissertation thesis. I hope you will find the improvements satisfactory.

Chapter 2: Is research on hedge fund performance published selectively? A quantitative survey

I am not sure how much to add since the paper was published in the *Journal of Economic Surveys*. One caveat about the paper is that I would caution making inferences about publication selection bias (whether present or not). Even though you claim that the paper uses several approaches to test for publication selection bias, you do not explicitly test for the bias and analyze only a small subset of empirical evidence on the topic. (, as evidenced by using a sample of only 1,019 intercept terms from regressions of hedge fund returns on risk factors (the alphas) collected from 74 studies published between 2001 and 2021).

Related to the point above, it is unclear how you perform the screening and reach the final sample. Specifically, please elaborate on why you drop 934 studies from the 1,095 originally screened studies. Do you drop over 90% of observations dues to the fact that the screened studies do not report alpha estimates, accompanied by a measure of statistical significance

Further, the choice of the tertiary search journals: *Journal of Portfolio Management*, *Journal of Financial Data Science*, *Journal of Impact and ESG Investing*, and *Journal of Fixed Income*, arguably aimed at professionals seems ad hoc. Based on my own search, journals such as: *Journal of Empirical Finance*, *Financial Analyst Journal*, *Journal of Futures Markets*, *Journal of Banking & Finance*, *European Journal of Finance*, and *European Financial Management* (to name a few) also contain papers related to hedge fund performance and their omission may bias your results, affecting the validity of the publication and data biases you analyze in this paper.

Thank you for these comments. You are right that our tests for the publication selection bias are based on the sample of alpha estimates collected from the primary studies. Essentially, all empirical tests in research are performed with the use of samples that do not comprise the entire population. That also applies to our research presented in Chapter 2 of this dissertation thesis. The main challenge researchers face is to use samples that can be reasonably considered representative of the entire population. Our sample consists of 1,019 observations that we have collected from 74 primary studies. While no data set in research is fully comprehensive, the sample size makes this study one of the larger meta-analyses in finance. Therefore, we consider our data set sufficiently large to be likely representative of the entire population and to draw meaningful inferences about the patterns in this stream of literature in general.

You are right that out of 1,095 that we initially screened as candidate studies that may be potentially relevant for our analysis, we eventually collected data from 74 prior empirical studies. However, that is not uncommon. As Figure 2.2 shows, our initial Google Scholar keyword search returns a number of working papers and earlier versions of published papers. Some of the papers identified by our keyword search overlap with results from similar searches through the top 5 finance journals and the Portfolio Management Research website. In addition, we do not collect estimates from working papers because these have not yet undergone the quality assurance resulting from the peer review process. We only collect data from the published version of a given study. Prior research argues that the magnitude of the publication selection bias in the published versions of research papers is comparable to the bias in working papers. Therefore, we believe that restricting our analysis to published papers does not compromise the generalizability of our findings. We exclude 934 papers in the screening mainly due to duplicates of paper versions, duplicates in multiple searches, and the criterion of publication. Some of the eligible 161 studies identified by our keyword search do not contain the alpha coefficients that we analyze in our meta-analysis. Furthermore, we impose several additional requirements concerning the information provided in the primary study about these estimates that allow us to analyze the estimates. For example, the estimates must be accompanied by measures of statistical significance. Thus, we do not collect the alpha estimates from studies that do not report these measures. After the "Eligibility", 74 studies are identified to meet our requirements.

I recognize that the choice of journals targeted towards practitioners is subjective and may be seen as *ad hoc*. We used these journals for our pre-screening mainly because some of the early research on hedge funds was published in these practitioner-oriented journals that may not be adequately represented in Google Scholar searches because they may be less frequently cited in conventional academic journals. For example, some of the journals listed on the Portfolio Management Research website are not ranked high based on the general academic criteria guided by citations in top academic journals and so they may be under-represented in the Google Scholar searches. We recognize that hedge fund research might be published in other academic journals you mentioned in your comment. Nevertheless, these tend to be standard academic journals that are likely to be adequately represented in our Google Scholar keyword searches. Our aim is to identify a pool of studies that are likely most influential in shaping the general perception of hedge fund performance. We hope that the combination of the practitioners' journals and a systematic Google Scholar keyword search generates such a list.

Chapter 3: Where Have All the Alphas Gone? A Meta-Analysis of Hedge Fund Performance

There are many repetitive sections in this chapter (relative to the introduction and chapter 2), especially about hedge fund performance literature review and data selection. Such repetition is reasonable for individual papers aimed for journal submission but not necessary for the thesis.

Thank you for the comment. Indeed, Chapters 2 and 3 are based on companion papers that are based on the same data set and investigate two related research questions. Therefore, some of the sections in the two chapters overlap. Following your suggestion, I have removed some repetitive sections in Chapter 3, including the issues caused by using FOF to alleviate the bias in the introduction, illustration of several biases in the background, the distribution of hedge fund studies published in top journals across years, and the distribution of the collected sample in the section of sample characteristics. I keep the repetitive section, distribution of alpha estimates in individual studies, in Chapter 3, since it illustrates the decreasing trend of alphas, relating to the finding in the corresponding Chapter. Instead, I deleted the distribution of alpha estimates in individual studies in Chapter 2 to avoid repetition. Wherever applicable, I refer to the relevant text in the other chapter of this dissertation thesis instead of preparing the text itself. The contents of Chapter 3, from the introduction to the research design, have been cut by about 5 pages.

I would highlight one caveat about the publication opportunity of this paper. It will be difficult to convince reader of the incremental contribution, since the same dataset was already utilized in chapter 2 and the resulting paper was already published.

Thank you for pointing this out. I recognize that the uniqueness of the dataset might be one of the reasons why referees and editors may wish to consider our paper for publication. Therefore, having two companion studies based on a single data set may be more challenging to publish. Nevertheless, we also believe that in our setting to studies based on the same data set are warranted because they address two different research questions.

In Chapter 2, we mainly analyze the presence of the publication selection bias in research literature analyzing hedge fund performance. Quantifying the effect of selective publication in this stream of literature may be interesting in its own right. Furthermore, it allows us to adjust the reported performance estimates for this bias and offer a corrected estimate of hedge funds' overall performance. We argue that given the considerable discretion researchers have in choosing their data samples and research methodologies, it is plausible to expect selective publication in this research literature. In contrast to this expectation, we observe very limited publication selection bias. We also demonstrate that this finding is observable in main data set partitions. We believe that this finding informs researchers, investors, and regulators about the hedge fund overall performance. Furthermore, it informs the academics about researchers' propensity to publish performance estimates selectively.

In contrast, in Chapter 3, we examine various determinants of the published hedge fund performance estimates. We identify several relevant measures that affect the reported alphas. Perhaps most notably, we document a strong declining trend in the magnitude of the performance estimates over time. We also show that this trend applies to all main hedge fund strategies. This finding lets us conclude that even though hedge funds used to generate value for investors in the past, on average, they do not do so anymore. Furthermore, the analysis in Chapter 3 also allows us to quantify the impact of performance and management fees charged by hedge funds. This study also informs researchers about the consequences of research design choices, such as adjusting for backfilling biases, using particular risk models to estimate abnormal returns, etc. These findings are important for future research, investment choices, and regulatory decisions.

I hope that our referees will recognize that even though the two studies are based on an identical data set, the research questions we investigate are different. Thus, the two studies propose different conclusions that make different kinds of contribution to prior research literature.

What are the fees charged by hedge funds over time? If the fees are increasing, and if this is the real reason behind the decline in alpha, then what does the meta analysis really suggests (that fees are going up)? Does the increased regulation help justify the increased fees? Can you provide some brief meta evidence of fund fees over time?

Thank you for these ideas. We are aware that the declining trend in hedge fund performance is a remarkable finding that merits investigating potential explanations. We agree that both the changes in the fees charged by hedge funds and that changes regulatory requirements, that you mention in your report, may contribute to this trend. Unfortunately, we believe that the limitations of our data set prevent us from providing additional insights and to discriminate between these explanations.

We agree that an increase in hedge fund fees should *ceteris paribus* be associated with lower after-fee returns, i.e., lower alphas. Due to the conditional nature of hedge fund fees, their effective impact on the net-of-fees alphas is difficult to measure. Nevertheless, the hedge fund industry has grown more competitive over time. One may expect such development to be more likely associated with lower rather than higher fees as the new hedge funds must compete more forcefully to attract investors. We recognize that documenting changes in hedge fund fees and examining whether these changes affect the net-of-fees alphas could be potentially interesting. Nevertheless, our data set contains only 35 observations of alphas estimated on a gross basis. Thus, unfortunately, our data set is insufficient to document trends in the magnitude of hedge fund fees.

We also recognize that the changes in the regulatory environment may have impacted hedge fund performance. At the same time, these regulatory changes often coincide with other economic developments. Furthermore, the estimates that we collected from the primary studies are typically estimated based on sample periods that span the time before and after regulatory changes. Thus in Chapters 2 and 3, we are unable to use this specific data set to draw conclusions about the impact of these regulatory changes. We also believe that to isolate the specific effect of regulatory changes it is advisable to use "sharper" identification methods. To that end, in Chapter 4, we use the difference-in-difference-in-differences (DDD) and the propensity score matching (PSM) methods. We believe that these approaches are better suited for identifying the impact of specific regulatory changes.

Are studies post-2015 different from pre-2015, in terms of research design and controls?

You are right that the research design and the use of control variables varies across the 74 primary studies, from which we collect our sample of alpha estimates. We have not systematically examined the differences in research design choices between studies published before and after 2015. Nevertheless, given the evolution of the research field, we cannot rule out the possibility that the newer studies have a tendency to use somewhat different methodological approaches than older studies. At the same time, the methodological framework that we use for our meta-analysis allows us to codify the research design and the use of control variables used in the primary studies. We use these coded variables in our Bayesian Model Averaging (BMA) estimation. For example, we use an indicator variable for each of the risk models that a given alpha estimate is based on. In our empirical analysis, we also include additional variables that capture the age of the data set and the year of publication of a given primary study. We believe that any systematic trend in the use of research design over time would be captured by these control variables that we use in our BMA. Thus, we do not expect these differences to affect the inferences we draw from our empirical results.

Chapter 4: The Impact of Regulatory Change on Hedge Fund Performance

To be consistent with other chapters, I would write this chapter using the present tense (i.e., what you contribute to literature, as opposed to what your paper will contribute to). This chapter will require a significant rewrite and editing before it can be submitted as a stand-alone paper ready for a journal submission. For example, the introduction is too descriptive. Second, rather than describing the potential effects of regulation on hedge fund performance from the start, highlight the differences between EU regulatory framework and US regulation investigated in prior literature to motivate the analysis of why EU regulation would have a negative impact on hedge funds relative to US regulation.

Thank you for the comment. I recognize that this chapter required a significant rewrite and editing. I have completely rewritten and restructured the first four sections of this chapter. I agree that in the previous version, the introduction was too descriptive and it needed better highlighting of the differences between the EU regulatory framework and the US regulation investigated in prior literature.

Following your suggestions, I have concentrated my efforts between the pre-defense and the final defense on improving the readability of this chapter. I have condensed the contents in the introduction, discussing the potential effects of regulation (Section 4.1). I have included a discussion of the institutional background in Section 4.2. This section informs the reader about hedge fund regulation in general and also, more specifically, about the European directive. I now present the discussion of the EU Directive; I discuss the inconsistent prior results on similar regulations from different contexts and compare the US and EU regulations. I motivate the analysis of why EU regulation would have a negative impact on hedge funds relative to US regulation. I have extended the literature review in Section 4.3. The literature review section is structured so that I can use the discussion of prior research literature to motivate my hypotheses reflecting the *a priori* expectations about the impact of the EU regulation on hedge fund performance. Below, I cite the parts of prior literature results, the comparison of US and EU regulations, and the prediction on the impact of the EU regulation in the introduction.

Thank you for pointing out the issue concerning the consistent use of tenses. Since Chapter 4 concerns a new piece of regulation that was enacted in 2011, it is sometimes natural to use the past tense. Nevertheless, in the new version of this chapter, I try to use the present tense wherever possible to make it consistent with the remainder of the thesis.

In addition, I asked a professional language editor to polish the text further.

"Further research on this question is also important because prior studies that examine similar regulations in different contexts, e.g., the Dodd-Frank Act that is applicable to hedge funds active in the U.S., often produce conflicting findings. For instance, Cumming et al. *(2020) compare the pre-Dodd-Frank period and after-Dodd-Frank period and conclude that U.S.-based funds have lower alpha after the implementation of the Dodd-Frank Act. Furthermore, Cumming* et al. *(2020) document a differential impact of the Dodd-Frank Act on hedge funds with different strategies. In contrast, Kaal* et al. *(2014) document a positive relationship between the Dodd-Frank Act and hedge fund performance even though they find the effect not to be persistent beyond several months following the implementation of the regulation. Thus, even though the preponderance of* a priori *arguments suggest that the new regulation that requires more extensive disclosures may impair the managerial ability to generate value for investors, whether this prediction actually materializes is far from obvious and, therefore, merits an empirical investigation."*
(...)

"Compared to the U.S. regulation, the AIFMD has similar requirements for registration and transparency but a wider scope of application, stricter risk control, and easier availability of information disclosure. These provisions may involve greater compliance costs and give the competitors easier access to information on hedge fund strategies, which may make the European regulation more impactful and costly than the U.S. regulation."
(...)

"Based on the a priori *arguments on the likely impact of mandating more extensive disclosure on hedge fund performance that we discuss above, we expect the new European regulation to be negatively associated with the value hedge funds generate for their investors."*

I agree that the relationship between hedge fund regulation and hedge fund performance is ambiguous; however, what matters is the type of regulation and its effects on performance. Regulation promoting transparency and accountability is different than regulation restricting leverage levels and levels of asset under management. Rather than describing the effects of regulation on performance, I would describe and focus on specific components of EU regulation and its predictive impact on performance.

Thank you for pointing out the importance of considering the specific provisions of the regulatory rules in evaluating the expected income the new regulation is likely to have on hedge fund performance. Initially, I worked towards investigating the specific provisions in the AIFMD and analyzing their likely effect. Recognizing that the EU member states have different hedge fund rules in the past, through comparison of the current rule and previous rules, I may discuss the effect of a certain requirement, such as transparency and restriction on the leverage. However, during the process of learning their previous individual requirements, I found that it is difficult to categorize or quantify their individual requirements because the details and extent of requirements vary. For some countries with multiple systems for registration, it is unable to connect a certain hedge fund with a certain system. As a result, the current work utilizes the available information to discuss the overall influence.

However, it is still helpful to include the analysis of specific components of EU regulation for predicting individual impact on performance, thus developing the hypotheses for the overall performance. Based on your suggestion, I added the analysis of specific components in section 4.3 to develop the overall prediction. I cite them below:

“As prior studies have similar conclusions that greater transparency is related to lower hedge fund performance by analyzing Form 13F under Section 13(f) of the Exchange Act, we expect that the negative effect of leaking managers’ strategies and weakening their incentives outweighs the positive effect of reduction of managers’ misconducts. Comparing the AIFMD with the US regulation, the EU Directive requires preparing audited annual reports for investors and authorities on demand instead of merely disclosing information to the authorities for monitoring. Therefore, competitors can more easily access investment techniques and enhance the negative influence caused by transparency. As the EU rules pay special attention to the use of leverage, specifying the disclosure of maximum leverage, the extent of leverage, and its application in certain circumstances. It seems the EU hedge fund manager would face greater pressure to operate their business with high leverage to chase superior revenues. From the perspective of information disclosure, the EU regulation should have a greater negative effect on performance than the US regulation.”

“However, prior results show inconsistent overall effects of the combination of registration, transparency, and related compliance costs. Since transparency probably leads to a negative impact, as we discussed above, and the compliance costs lower the earnings, the overall positive relationship between the US regulation, such as the Dodd-Frank Act, and the hedge fund performance may suggest that the undernourished hedge funds decrease due to the strict rule to register. Then the overall industry performance is expected to grow (Frumkin & Vandegrift 2009) after removing costs resulting from transparency and other compliance fees. The EU Directive requires minimum capital amounts and qualified managers to assure the industry’s quality, which may positively influence hedge fund performance but raises the capital costs for some small funds. The quality improves hedge fund performance when the diminution of underperforming hedge funds exceeds that of wellperforming small funds. The improvements might exist based on the US regulation, while the EU regulation involves more cost as the hedge funds that are located in third countries are expected to have an additional one-time large cost for relocating to a member state of the EU; otherwise, they are not able to market in the EU. The cost of authorization and restriction is probably greater than the benefits from improving the overall hedge fund quality.”

(...)

> *"Except for the similar rules to the US regulation, the AIFMD requires hedge funds to have periodic independent valuations and depositories. The depository could increase investors' confidence by reducing asymmetric information between the manager and investors (Kamal 2012). However, failing to correctly value the assets of hedge funds may lead to sanctions by the authorities for valuation shortcomings. It is effective in continuously controlling risk, while hedge funds would unavoidably distract attention and require more investment in their operation and management in case of any violation of the AIFMD. The compliance cost from the valuation and depository is continuous and may constitute another important part of the total cost."*

As a sensitivity test, it would be useful to work with the dataset of hedge funds and verify/test whether they have experienced decreasing performance over time (consistent with your meta-analysis in the previous chapters).

Thank you for pointing this out. You are right that any development in hedge fund performance may have been caused by a secular trend that is independent of any regulatory changes. Nevertheless, we believe that the methodology that we use in Chapter 4 is specifically designed to address this issue. Specifically we use the DDD and PSM. These approaches compare the performance changes in "treated" hedge funds that are affected by the regulatory change relative to the corresponding changes in "control" hedge funds that are unaffected by the regulatory change. Furthermore, the PSM method ensures that the "treated" group is as comparable as possible to the "control" group. These approaches allow us to examine whether the performance decline in the "treated" group is more pronounced relative to any potential performance decline in the "control" group. Therefore, we would expect any secular trend affecting all hedge funds to be "differenced away" by using DDD and PSM. The observed effect should thus be solely attributable to the impact of the regulatory change that we analyze.

Referee: 3 – Professor Robert Reed

Introduction

I choose to provide my comments chapter by chapter.

Chapter 2: Is research on hedge fund performance published selectively? A quantitative survey

This chapter assembles a large database of published, primary studies that estimate abnormal returns for hedge funds. It then uses this database to investigate the presence of publication bias.

There are several reasons why one might expect to see publication bias. First, there is wide latitude in which hedge fund performance data is included in any given study. There is no common, publicly available database where hedge fund performance is curated. Hedge funds can choose which, if any, of their funds, to report to one or more private data providers.

Further, not all hedge funds report to all private data providers. Thus the hedge fund performance data included in any given study will differ across studies depending on the data providers the studies rely on.

Further, because there are no uniform reporting requirements, researchers are free to further select which hedge fund performance data to include in their study. This creates an opportunity to select hedge fund data that is supportive of the result they hope to show.

In addition, there is no commonly accepted specification for estimating abnormal returns.

Specifications can range from the simple one-factor, classical CAPM model to three-, four-, and even seven-factor models. This gives researchers freedom to select across multiple specifications for those that are most favorable to publication.

And even for a given specification, there are multiple estimation methods that provide alternatives in how standard errors are estimated and the extent to which clustering and endogeneity are addressed in estimating model coefficients.

Thus, the conditions are ripe to discover publication bias. However, despite a thorough search involving a wide variety of testing strategies, Yang does not find much evidence of publication bias. The only evidence consistent with publication bias is that publication bias may be present among the subsample of studies that use IV methods. This is consistent with Brodeur et al. (2020), who argued that IV methods create additional opportunities for researchers to "game the system" through the strategic inclusion of instruments.

(...)

Given that the empirical results are robust and the study has already been published in a good journal, I don't think anything else needs to be done.

Thank you for your positive evaluation of this paper.

Minor comments

1) I think it is useful to distinguish between publication selection bias caused by researchers and journals desiring to publish statistically significant results, and the selection bias generated by hedge funds selectively choosing which results to report. Presumably, they will choose to report the results that make the hedge funds look especially successful. This type of selection bias will not be a function of the standard error and thus will not be identified, or corrected, via the inclusion of a standard error variable

Thank you for your comment. We agree that the two biases you point out are conceptually different and they will be reflected in the data in different ways. The methods of detecting publication selection bias allow us to quantify an adjust for the former bias caused by researchers and journals desiring to publish statistically significant results because this bias likely introduces an association between the magnitude of the estimated performance coefficient and its standard error. In contrast the latter bias generated by hedge funds selectively choosing which results to report to the commercial databases is of different nature and it must be treated differently. We argue that the direction of this latter bias is not obvious. Unsuccessful hedge funds may choose not to report their performance to the commercial databases because their poor results are not likely to attract new investors. This would make the hedge funds look especially successful, as you might argue. Nevertheless, very successful hedge funds may also skip reporting their performance to commercial databases because they might be closed for new capital and therefore have no incentives to advertise their success.

We agree that the latter bias should not be a function of the standard error and thus will not be identified, or corrected, via the inclusion of a standard error variable. To quantify the effect of this bias we need to rely on the differences in the performance coefficients reported in primary studies that do and do not correct for these biases in their primary analysis. For example, the backfilling bias may be addressed by discarding initial months of performance data in the commercial databases. Furthermore, the survivorship bias may be addressed by using data from databases that do not purge the data on inactive funds or by using performance data for the funds of funds rather than individual hedge funds. When collecting our data set from the primary studies we code dummy variables that indicate whether or not a given estimate is obtained using one of these adjustment techniques. We are then able to use the differences between these estimates to evaluate the impact of a given bias.

We observe that while the publication selection bias does not significantly affect the inferences about the overall performance of hedge funds, the impact of the backfilling bias is indeed substantial. The difference between these two empirical results further underscores the importance of distinguishing between these two biases.

2) On page 35, Yang states: "*In the last two columns of Panel A of Table 2.2 we report our weighted least squares estimates of Equation 2.2. In the fifth column, we weigh the*

observations by the inverse of their standard error (WLS). This approach gives less weight to less precise estimates, which helps to adjust for potential heteroskedasticity in our observations. The sixth column shows our results from estimation when the observations are weighted by the inverse of the number of estimates reported in a given study (wNOBS)." There is an inconsistency here. Standard WLS uses "inverse variance" weights. So weighting by the inverse of the SE squared. Practically, this means dividing all the observations by the SE. I thought this is what Yang meant, but then she said Column 6 weighted by number of estimates. To be consistent with the WLS estimates, the Column 6 estimates should divide by the square root of the number of estimates. However, it would be better to use the appropriate language and say that Columns 5 and 6 weight by the inverse of the squared SE and the inverse of the number of estimates per study, respectively.

Thanks for drawing our attention to the point. Indeed, the formulation you refer to is imprecise. We have changed the expression about the issue in the text and notes:

> *"In the fifth column, we weigh the observations by the inverse of their* ***squared*** *standard error (WLS)."*

Chapter 3: Where Have All the Alphas Gone? A Meta-Analysis of Hedge Fund Performance

This study is a very well done meta-analysis! Having investigated the presence of publication bias in the previous chapter, it then proceeds to estimate the abnormal return associated with hedge funds. It reports several interesting results.

First, Yang finds that hedge funds have experienced declining abnormal returns over time. Second, while it is not possible to obtain proprietary information about fund fees, she is able to determine whether returns are gross or net of fees. This allows Yang to obtain an overall estimate of the size of hedge fund fees. Yang estimates that hedge funds charge their investors more than 5.0% annually.

Third, based on a best practice estimate, Yang concludes that "we cannot reject the null hypothesis that hedge funds currently generate no abnormal after-fee return for their investors." Actually, I would say it is worse than that. The point estimate she obtains is -0.079 (see Table 3.4). This implies an annual negative abnormal return of approximately negative 0.95 percent

The reason Yang emphasizes that she cannot reject the null of no abnormal returns is because of the large 95% confidence interval (-0.393 to 0.23, which equates to annual returns of -4.70% to 2.76%).

There are other interesting findings that I won't mention. Overall, this is a well-executed, interesting chapter

(...)

Thank you for your positive evaluation of this paper. We appreciate that you consider this chapter a very well done meta-analysis, you recognize the connection between Chapter 2 and Chapter 3, and you find several results in Chapter 3 interesting.

Concerning the negative sign of the hedge funds' current performance estimates that we present in Table 3.4. We agree that these results may be seen as suggesting that hedge funds' current performance is actually worse than "statistically indistinguishable from zero". In the test of Chapter 3, we point out that the negative sign of these estimates is consistent across different hedge fund strategies. At the same time, as you rightly point out in your report, the confidence intervals are rather wide. This prevents us from drawing stronger conclusions that would depict an even less favorable picture of the value hedge funds generate (or destroy) for their investors. While we opted to be conservative in interpreting these results, we expect the readers to be able to observe these numbers and make appropriate interpretations for themselves.

Minor comments

1) The Abstract is not related to the content of the chapter. The abstract is about banking regulation, but this chapter is about hedge fund performance.

Thank you for pointing out this issue. You are quite right. By mistake, we included the wrong abstract when compiling the .pdf file of this dissertation thesis. We have now included the correct version of the abstract, which we also copy below.

"We conduct a systematic meta-analysis of the factors influencing hedge fund performance estimates published between 2001 and 2021. Using a sample of 1,019 intercept terms from regressions of hedge fund returns on risk factors (the "alphas") collected from 74 studies, we document a strong downward trend in reported alphas that persists even after controlling for heterogeneity in hedge fund characteristics and research design choices in the underlying studies. Our best-practice estimates of current performance are not reliably different from zero for all common hedge fund strategies. In addition, we provide an estimate of the sizeable impact of management and performance fees charged by hedge funds. We also document how reported performance estimates vary with hedge fund and study characteristics. Our results suggest that while hedge funds have generated positive value for investors in the past, on average, they no longer do so."

2) Note the similarity between the BMA and OLS estimates in Table 3.3. This is not a coincidence. The BMA estimates are weighted averages of OLS estimates.

Thank you for pointing out this relationship. I absolutely agree that the similarity between the results based on the BMA and the OLS is by no means a coincidence. As you argue and as we also point out in the research design section of this chapter, the BMA results essentially are the weighted averages of OLS estimates under various regression specifications that differ in the inclusion of different combinations of control variables. Hence, I agree that it is not at all surprising that the two results are similar. Nevertheless, I still believe that despite this similarity, the OLS result is incrementally informative relative to the BMA-based result, which justifies reporting it in Table 3.3. First, while the BMA results are based on a multitude of various regression models, the OLS results are based on a single regression model. Thus, our BMA estimation can also be interpreted as a way of identifying the set of explanatory variables that are relevant to the OLS estimation. Second, while we consider the BMA results as our baseline findings and we rely on them in drawing our conclusions, we also report the OLS result as a robustness check. These auxiliary results are based on a regression estimation approach that most readers are likely familiar with and are readily comparable to results reported in prior studies. Reporting both sets of results also gives the readers an opportunity to choose from the estimates based on the frequentist approach or the Bayesian approach, whichever they prefer. We have newly included in the updated version of this dissertation thesis the quote below that explains this relationship between the two sets of results.

"To check the robustness of the BMA results, we also report in the right panel of Table 3.3 the ordinary least squares (OLS) estimates, which are based on a single regression model including a set of the most relevant variables identified by BMA. The frequentist OLS approach is better comparable to prior study results and thus may enhance the understanding of the effect of these relevant factors."

Correspondingly, we have updated the way we comment on the OLS results presented in Table 3.3, for example, see the quote below.

"In the OLS regression, all of the nine variables included in the BMA model with the best fit are also significant at a better than 5% level."

3) I would suggest that Yang report I-squared. In particular, it would be interesting to note what percent of total variation can be explained by the included regressors. Alternatively, if she continues to base her main estimates on OLS estimation, she could report R-squared.

Thank you for pointing out this issue to us. We have considered reporting some measures of heterogeneity of the data we use for our meta-analysis, e.g., I-squared. We recognize that in some settings reporting information on the percentage of variation across studies that is due to heterogeneity rather than chance may have important implications for the interpretation of the results. Nevertheless, after some consideration, we have agreed with my co-authors not to report I-squared in this study. There are several reasons that led us to this decision. First, we argue that the relevance of heterogeneity measures varies across fields of study and the nature of how the underlying measures are typically estimated in the primary studies. We also believe that in economics sample heterogeneity is commonly considered substantial without explicitly reporting estimates of the heterogeneity measures. Second, there is no universally accepted consensus about the relevant cut of levels for the I-squared measure in different fields. According to Cochrane Review Guidelines, an I-squared above 0.75 is considerable heterogeneity in medicine, and according to Vivalt (2020), an I-squared of 0.9 is common in economics meta-analyses. This limits the usefulness of the measure in informing researchers about the underlying data characteristics and the consequences these characteristics should have for research design choices. Third, we also observe the controversy in the prior methodological literature about the problematic precision, with which these measures are typically investigated. Fourth, since the purpose of our study is to explain the heterogeneity in the hedge fund performance estimates reported in the primary studies, estimating I-squared is not the first-order consideration for our analysis. Given these limitations, we have doubts about whether I-squared can reasonably be considered a reliable measure of the inconsistency of studies' results and how informative would reporting the estimate be for the readers.

In contrast, in response to your suggestion, we have newly included the R-squared in Table 3.3 for the OLS regression as a measure of the proportion of total variation that can be explained by the included regressors. Correspondingly, we have updated the main body text of that chapter that refers to that newly included measure, e.g., see the quote below.

> *" Thus, our results identify nine key characteristics that are essential for explaining the heterogeneity in reported alphas. And the R-squared in the OLS regression indicates that these nine variables explain 23% of the variability."*

Chapter 4: The Impact of Regulatory Change on Hedge Fund Performance

This chapter investigates the effect of the Alternative Investment Fund Manager Directive (AIFMD) regulation on EU hedge fund performance. For data, it draws panel data on 26,862 hedge funds over the period 1994 to 2021 from Eureakhedge, a commercial hedge fund database. The AIFMD imposed tighter regulations and greater transparency on hedge funds operating in the EU. The AIFMD was passed in 2011 and became effective in 2013.

Yang proposes two hypotheses:

1. H1: EU hedge funds have lower risk-adjusted performance after the AIFMD was transposed into EU countries' national laws.

2. H2: EU Hedge funds domiciled in countries with strict regulations before the AIFMD would experience a less drop than the hedge funds domiciled in countries with lax regulations.

Yang uses two main empirical methods. The first is a triple difference-in-difference procedure that compared pre- and post- AIFMD hedge fund performance in funds with exemption and non-exemption, and in the EU and outside the EU. The second method is propensity score matching. She finds support for both hypotheses.

This is another nicely done study that combines careful institutional knowledge with a specialized dataset and employs appropriate empirical procedures.

(...)

Thank you for your positive evaluation of this paper.

Minor comments

1) On page 119 in H2, it would be better to say a "smaller drop" instead of a "less drop".

Thank you for pointing this out. I have changed the wording, as you suggested.

"H2: EU Hedge funds domiciled in countries with strict regulations before the AIFMD would experience a ***less drop*** *than the hedge funds domiciled in countries with lax regulations."*

2) On page 131, second and third lines from the top: Shouldn't "experienced a negatively significant influence from 0.08 to 0.15" be "0.03 to 0.15?

Thank you for pointing this out. Indeed, we overlooked this misleading formulation in the text. In response to your comment we have changed the text to the following:

> *"Our results, obtained from excess returns, three-factor model, and four-factor model as the measurements of hedge fund performance show that within the EU, the managers of larger funds and nonUCITS funds that are subject to AIFMD experienced a negatively significant influence from* ***0.03 to 0.15*** *compared with managers of funds with exemption while the one-factor model and FH model do not convey such significant result to support H1."*

3) On page 131, seven and eight lines from the bottom: Shouldn't "the EU funds without exemption experienced a drop varying from 0.17 to 0.29" be "0.06 to 0.29".

Thank you for pointing this out. Indeed, we overlooked this misleading formulation in the text. In response to your comment we have changed the text to the following:

> *"It shows, based on the alphas from risk factor models, the EU funds without exemption experienced a drop varying from* ***0.06 to 0.29*** *in performance after the rule compared with the corresponding NonEU funds of other areas."*

4) It has been a while since I have done PSM, but I think it is generally considered good practice to report a figure that shows the degree of overlap between the predicted probabilities of "treatment" between the treatment and control groups

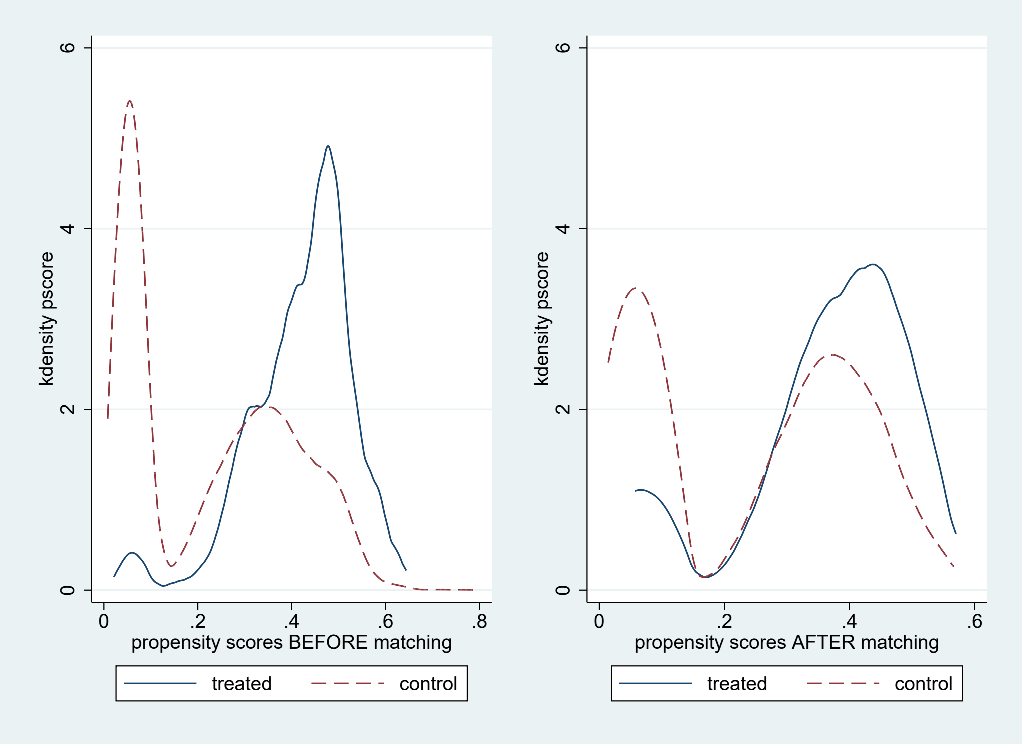
Thank you for the comment. I agree that it is indeed advisable to include such a figure. I added Figure 5.2 to show the distributions of propensity scores of the treated and control groups before and after matching. The figure clearly shows that the PSM increased the similarity of the characteristics of the treatment and control groups. Provided that my PSM-based results are qualitatively similar to the main results based on the DDD, this indicates that the empirical findings I present in this chapter are unlikely to be driven by fundamental differences between the hedge funds in my treatment group and those in the control group.

5) It would be nice to provide more details about the kind of matching that was done. Is it the nearest neighbor? With replacement? Without replacement? If with replacement, was any allowance made for correcting standard errors due to using the same controls?

Thank you for the good point. I have added the following explanation to provide more details on how I applied the PSM methodology in my research setting.

> *"First, we exactly match the date and investment strategy to make sure the significant factors, including the time period and strategy, do not influence the funds' performance. Then, we use the funds' characteristics to obtain their propensity score to match the treated observations with control variables using the nearest-neighbor method without replacement. To ensure the quality of matches, we set a caliper of 1%."*

Figure 5.2: Comparison of samples before and after PSM



6) In the interests of open science, Yang might think of making the data and code publicly available so that the results of the chapter are push-button replicable.

We agree that to facilitate future research in this area and to ensure replicability of the published results it is desirable to make the data and code publicly available. My coauthors routinely do so for the meta-analyses they publish. Nevertheless, this study is based on data from the Eureakhedge database, which we purchased specifically for this project with the use of funding kindly provided by the Charles University Grant Foundation (GAUK). The contract with the data provider allows me to use the data for research purposes and publish aggregated results, but it does not allow us to make the observations for individual hedge funds publicly available.

Bibliography

Ackermann, C., R. McEnally, & D. Ravenscraft (1999): "The performance of hedge funds: Risk, return, and incentives." *The Journal of Finance* **54(3)**: pp. 833–874.

van Aert, R. C. & M. van Assen (2020): "Correcting for publication bias in a meta-analysis with the p-uniform* method." *Working paper*, Tilburg University & Utrecht University, Netherllands.

Agarwal, V., Y. Arisoy, & N. Y. Naik (2017): "Volatility of aggregate volatility and hedge fund returns." *Journal of Financial Economics* **125(3)**: pp. 491–510.

Agarwal, V., G. Bakshi, & J. Huij (2009a): "Do higher-moment equity risks explain hedge fund returns?" *CFR Working Paper 10-07*, Centre for Financial Research (CFR), University of Cologne, Germany.

Agarwal, V., N. D. Daniel, & N. Y. Naik (2009b): "Role of managerial incentives and discretion in hedge fund performance." *The Journal of Finance* **64(5)**: pp. 2221–2256.

Agarwal, V., V. Fos, & W. Jiang (2013a): "Inferring reporting-related biases in hedge fund databases from hedge fund equity holdings." *Management Science* **59(6)**: pp. 1271–1289.

Agarwal, V., T. C. Green, & H. Ren (2018): "Alpha or beta in the eye of the beholder: What drives hedge fund flows?" *Journal of Financial Economics* **127(3)**: pp. 417–434.

Agarwal, V., W. Jiang, Y. Tang, & B. Yang (2013b): "Uncovering hedge fund skill from the portfolio holdings they hide." *The Journal of Finance* **68(2)**: pp. 739–783.

Agarwal, V., K. A. Mullally, N. Y. Naik *et al.* (2015): "The economics and finance of hedge funds: A review of the academic literature." *Foundations and Trends in Finance* **10(1)**: pp. 1–111.

Agarwal, V. & N. Naik (2002): "Characterizing systematic risk of hedge funds with buy-and-hold and option based strategies. london business school." *Technical report*, Working paper.

Agarwal, V. & N. Y. Naik (2000): "On taking the 'alternative' route: the risks, rewards, and performance persistence of hedge funds." *The Journal of Alternative Investments* **2(4)**: pp. 6–23.

Agarwal, V. & N. Y. Naik (2004): "Risks and portfolio decisions involving hedge funds." *The Review of Financial Studies* **17(1)**: pp. 63–98.

Aggarwal, R. K. & P. Jorion (2010): "Hidden survivorship in hedge fund returns." *Financial Analysts Journal* **66(2)**: pp. 69–74.

Ahoniemi, K. & P. Jylha (2014): "Flows, Price Pressure, and Hedge Fund Returns." *Financial Analysts Journal* **70(5)**: pp. 73–93.

Aiken, A., C. P. Clifford, & J. Ellis (2013): "Out of the Dark: Hedge Fund Reporting Biases and Commercial Databases." *Review of Financial Studies* **26(1)**: pp. 208–243.

Amin, G. S. & H. M. Kat (2003): "Hedge fund performance 1990–2000: Do the 'money machines' really add value?" *Journal of Financial and Quantitative Analysis* **38(2)**: pp. 251–274.

Ammann, M. & P. Moerth (2005): "Impact of fund size on hedge fund performance." *Journal of Asset Management* **6(3)**: pp. 219–238.

Ammann, M. & P. Moerth (2008a): "Impact of Fund Size and Fund Flows on Hedge Fund Performance." *The Journal of Alternative Investments* **11(1)**: pp. 78–96.

Ammann, M. & P. Moerth (2008b): "Performance of Funds of Hedge Funds." *The Journal of Wealth Management* **11(1)**: pp. 46–63.

Andrews, I. (2018): "Valid Two-Step Identification-Robust Confidence Sets for GMM." *The Review of Economics and Statistics* **100(2)**: pp. 337–348.

Andrews, I. & M. Kasy (2019): "Identification of and correction for publication bias." *American Economic Review* **109(8)**: pp. 2766–2794.

Aragon, G. O. (2007): "Share restrictions and asset pricing: Evidence from the hedge fund industry." *Journal of Financial Economics* **83(1)**: pp. 33–58.

Aragon, G. O., M. Hertzel, & Z. Shi (2013): "Why do hedge funds avoid disclosure? evidence from confidential 13f filings." **48(5)**: pp. 1499–1518. Publisher: Cambridge University Press.

Aragon, G. O. & J. Spencer Martin (2012): "A unique view of hedge fund derivatives usage: Safeguard or speculation?" *Journal of Financial Economics* **105(2)**: pp. 436–456.

Asness, C., R. Krail, & J. Liew (2001): "Do Hedge Funds Hedge?" *The Journal of Portfolio Management* **28(1)**: pp. 6–19.

Association of the Luxembourg Fund Industry (2014): "Luxembourg hedge funds."

Astakhov, A., T. Havranek, & J. Novak (2019): "Firm size and stock returns: A quantitative survey." *Journal of Economic Surveys* **33(5)**: pp. 1463–1492.

Avramov, D., L. Barras, & R. Kosowski (2013): "Hedge fund return predictability under the magnifying glass." *Journal of Financial and Quantitative Analysis* **48(4)**: pp. 1057–1083.

Bajzik, J., T. Havranek, Z. Irsova, & J. Schwarz (2020): "Estimating the Armington Elasticity: The Importance of Data Choice and Publication Bias." *Journal of International Economics* **127(C)**: p. 103383.

Bali, T. G., S. Brown, & K. O. Demirtas (2013): "Do Hedge Funds Outperform Stocks and Bonds?" *Management Science* **59(8)**: pp. 1887–1903.

Bali, T. G., S. J. Brown, & M. O. Caglayan (2011): "Do hedge funds' exposures to risk factors predict their future returns?" *Journal of Financial Economics* **101(1)**: pp. 36–68.

Bali, T. G., S. J. Brown, & M. O. Caglayan (2014): "Macroeconomic risk and hedge fund returns." *Journal of Financial Economics* **114(1)**: pp. 1–19.

Baquero, G., J. Ter Horst, & M. Verbeek (2005): "Survival, look-ahead bias, and persistence in hedge fund performance." *Journal of Financial and Quantitative analysis* **40(3)**: pp. 493–517.

Barth, D., J. Joenvaara, M. Kauppila, & R. Wermers (2020): "The hedge fund industry is bigger (and has performed better) than you think." *OFR Working Paper 20–01*, Office of Financial Research (OFR), U.S. Department of the Treasury, USA: Washington, DC.

Ben-David, I., J. Birru, & A. Rossi (2020): "The performance of hedge fund performance fees." *NBER Working Paper 27454*, National Bureau of Economic Research.

Berk, J. B. & R. C. Green (2004): "Mutual fund flows and performance in rational markets." *Journal of political economy* **112(6)**: pp. 1269–1295.

Bhardwaj, G., G. B. Gorton, & K. G. Rouwenhorst (2014): "Fooling Some of the People All of the Time: The Inefficient Performance and Persistence of Commodity Trading Advisors." *Review of Financial Studies* **27(11)**: pp. 3099–3132.

Bianchi, R. J. & M. E. Drew (2010): "Hedge fund regulation and systemic risk." **19(1)**: pp. 6–29. Publisher: Routledge _eprint: https://doi.org/10.1080/10854666.2010.10854666.

Billio, M., L. Frattarolo, & L. Pelizzon (2014): "A time-varying performance evaluation of hedge fund strategies through aggregation." *Bankers, Markets & Investors* **(129)**: pp. 40–58.

Black, F. (1972): "Capital market equilibrium with restricted borrowing." *The Journal of Business* **45(3)**: pp. 444–455.

Blanco-Perez, C. & A. Brodeur (2020): "Publication Bias and Editorial Statement on Negative Findings." *The Economic Journal* **130(629)**: pp. 1226–1247.

Blitz, D. (2018): "Are Hedge Funds on the Other Side of the Low-Volatility Trade?" *The Journal of Alternative Investments* **21(1)**: pp. 17–26.

Bollen, N. P., J. Joenväärä, & M. Kauppila (2021): "Hedge fund performance: End of an era?" *Financial Analysts Journal* **77(3)**: pp. 109–132.

Bollen, N. P. & V. K. Pool (2009): "Do hedge fund managers misreport returns? Evidence from the pooled distribution." *The Journal of Finance* **64(5)**: pp. 2257–2288.

Bollen, N. P. B. & R. E. Whaley (2009): "Hedge Fund Risk Dynamics: Implications for Performance Appraisal." *The Journal of Finance* **64(2)**: pp. 985–1035.

Bom, P. R. D. & H. Rachinger (2019): "A kinked meta-regression model for publication bias correction." *Research Synthesis Methods* **10(4)**: pp. 497–514.

Boyson, N., C. W. Stahel, & R. M. Stulz (2006): "Is there hedge fund contagion?" *NBER Working Paper 12090*, National Bureau of Economic Research.

Brav, A., W. Jiang, F. Partnoy, & R. Thomas (2008): "Hedge fund activism, corporate governance, and firm performance." *The Journal of Finance* **63(4)**: pp. 1729–1775.

Breuer, M. (2021): "How Does Financial-Reporting Regulation Affect Industry-Wide Resource Allocation?" *Journal of Accounting Research* **59(1)**: pp. 59–110. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-679X.12345.

Brodeur, A., N. Cook, & A. Heyes (2020): "Methods matter: P-hacking and publication bias in causal analysis in economics." *American Economic Review* **110(11)**: pp. 3634–3660.

Brodeur, A., M. Le, M. Sangnier, & Y. Zylberberg (2016): "Star Wars: The Empirics Strike Back." *American Economic Journal: Applied Economics* **8(1)**: pp. 1–32.

Brown, R. (2012): "Framework for hedge fund return and risk attribution." *The Journal of Investing* **21(4)**: pp. 8–23.

Brown, S., W. Goetzmann, B. Liang, & C. Schwarz (2008): "Mandatory disclosure and operational risk: Evidence from hedge fund registration." **63(6)**: pp. 2785–2815. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2008.01413.x.

Brown, S., Y. Lu, S. Ray, & M. Teo (2018): "Sensation seeking and hedge funds." *The Journal of Finance* **73(6)**: pp. 2871–2914.

Brown, S. J., W. N. Goetzmann, R. G. Ibbotson *et al.* (1999): "Offshore hedge funds: Survival and performance, 1989–95." *The Journal of Business* **72(1)**: pp. 91–117.

Brown, S. J., W. N. Goetzmann, & B. Liang (2004): "Fees on fees in funds of funds." *Journal of Investment Management* **2(C)**: pp. 9–56.

Bruns, S. B. & J. P. A. Ioannidis (2016): "p-Curve and p-Hacking in Observational Research." *PloS ONE* **11(2)**: p. e0149144.

Buller, J. & N. Lindstrom (2013): "Hedging its bets: The UK and the politics of european financial services regulation." **18(3)**: pp. 391–409. Publisher: Routledge _eprint: https://doi.org/10.1080/13563467.2012.700924.

Buraschi, A., R. Kosowski, & F. Trojani (2014): "When there is no place to hide: Correlation risk and the cross-section of hedge fund returns." *The Review of Financial Studies* **27(2)**: pp. 581–616.

Cao, C., B. Goldie, B. Liang, & L. L. Petrasek (2016): "What is the nature of hedge fund manager skills? Evidence from the risk-arbitrage strategy." *Journal of Financial and Quantitative Analysis* **51(3)**: pp. 929–957.

Capocci, D., A. Corhay, & G. Hubner (2005): "Hedge fund performance and persistence in bull and bear markets." *The European Journal of Finance* **11(5)**: pp. 361–392.

Capocci, D. & G. Hubner (2004): "Analysis of hedge fund performance." *Journal of Empirical Finance* **11(1)**: pp. 55–89.

Carhart, M. M. (1997): "On persistence in mutual fund performance." *The Journal of finance* **52(1)**: pp. 57–82.

Carter, E. C., F. D. Schonbrodt, W. M. Gervais, & J. Hilgard (2019): "Correcting for bias in psychology: A comparison of meta-analytic methods." *Advances in Methods and Practices in Psychological Science* **2(2)**: pp. 115–144.

Cassar, G. & J. Gerakos (2011): "Hedge funds: Pricing controls and the smoothing of self-reported returns." *The Review of Financial Studies* **24(5)**: pp. 1698–1734.

Cazachevici, A., T. Havranek, & R. Horvath (2020): "Remittances and economic growth: A meta-analysis." *World Development* **134**: p. 105021.

Chen, Y., M. Cliff, & H. Zhao (2017): "Hedge funds: The good, the bad, and the lucky." *Journal of Financial and Quantitative Analysis* **52(3)**: pp. 1081–1109.

Chen, Y. & B. Liang (2007): "Do market timing hedge funds time the market?" *Journal of Financial and Quantitative Analysis* **42(4)**: pp. 827–856.

Chincarini, L. & A. Nakao (2011): "Measuring hedge fund timing ability across factors." *The Journal of Investing* **20(4)**: pp. 50–70.

Christensen, G. & E. Miguel (2018): "Transparency, reproducibility, and the credibility of economics research." *Journal of Economic Literature* **56(3)**: pp. 920–980.

Clark, K. A. & K. D. Winkelmann (2004): "Active risk budgeting in action." *The Journal of Alternative Investments* **7(3)**: pp. 35–46.

Clarkson, D., S. Jaecklin, & K. Kaczmarski (2014): "DOMICILES OF ALTERNATIVE INVESTMENT FUNDS." .

Coggin, T., F. Fabozzi, & S. Rahman (1993): "The Investment Performance of U.S. Equity Pension Fund Managers: An Empirical Investigation." *The Journal of Finance* **48(3)**: pp. 1039–1055. 1039.

Connor, G. & M. Woo (2004): "An introduction to hedge funds." *Discussion paper 477*, Financial Markets Group, The London School of Economics and Political Science, UK: London.

Cumming, D. & N. Dai (2010): "A law and finance analysis of hedge funds." *Financial Management* **39(3)**: pp. 997–1026.

Cumming, D., N. Dai, & S. Johan (2020): "Dodd-franking the hedge funds." *Journal of banking & finance* **119**: p. 105216.

Dagenais, M. G. & D. L. Dagenais (1997): "Higher moment estimators for linear regression models with errors in the variables." *Journal of Econometrics* **76(1–2)**: pp. 193–221.

Dichev, I. D. & G. Yu (2011): "Higher risk, lower returns: What hedge fund investors really earn." *Journal of Financial Economics* **100(2)**: pp. 248–263.

Diez De Los Rios, A. & R. Garcia (2011): "Assessing and valuing the nonlinear structure of hedge fund returns." *Journal of Applied Econometrics* **26(2)**: pp. 193–212.

Ding, B. & H. A. Shawky (2007): "The Performance of Hedge Fund Strategies and the Asymmetry of Return Distributions." *European Financial Management* **13(2)**: pp. 309–331.

Ding, B., H. A. Shawky, & J. Tian (2009): "Liquidity shocks, size and the relative performance of hedge fund strategies." *Journal of Banking & Finance* **33(5)**: pp. 883–891.

Do, V., R. Faff, & J. Wickramanayake (2005): "An empirical analysis of hedge fund performance: The case of Australian hedge funds industry." *Journal of Multinational Financial Management* **15(4)**: pp. 377–393.

Duarte, J., F. A. Longstaff, & F. Yu (2007): "Risk and Return in Fixed-Income Arbitrage: Nickels in Front of a Steamroller?" *The Review of Financial Studies* **20(3)**: pp. 769–811.

Durbin, J. (1954): "Errors in variables." *Revue de l'institut International de Statistique* **22(1–3)**: pp. 23–32.

Easley, D., D. Michayluk, M. O'Hara, & T. J. Putniņš (2021): "The active world of passive investing." *Review of Finance* **25(5)**: pp. 1433–1471.

Edelman, D., W. Fung, D. Hsieh, & N. N. Naik (2012): "Funds of hedge funds: Performance, risk and capital formation 2005 to 2010." *Financial Markets and Portfolio Management* **26(1)**: pp. 87–108.

Edelman, D., W. Fung, & D. A. Hsieh (2013): "Exploring uncharted territories of the hedge fund Industry: Empirical characteristics of mega hedge fund firms." *Journal of Financial Economics* **109(3)**: pp. 734–758.

Edwards, F. R. & M. O. Caglayan (2001): "Hedge Fund Performance and Manager Skill." *Journal of Futures Markets* **21(11)**: pp. 1003–1028.

Egger, M., G. D. Smith, M. Schneider, & C. Minder (1997): "Bias in meta-analysis detected by a simple, graphical test." *BMJ: British Medical Journal* **315(7109)**: pp. 629–634.

Ehrenbergerova, D., J. Bajzik, & T. Havranek (2023): "When Does Monetary Policy Sway House Prices? A Meta-Analysis." *IMF Economic Review* **(forthcoming)**.

Eicher, T. S., C. Papageorgiou, & A. E. Raftery (2011): "Default priors and predictive performance in Bayesian model averaging, with application to growth determinants." *Journal of Applied Econometrics* **26(1)**: pp. 30–55.

Eling, M. & R. Faust (2010): "The performance of hedge funds and mutual funds in emerging markets." *Journal of Banking & Finance* **34(8)**: pp. 1993–2009.

Elliott, G., N. Kudrin, & K. Wuthrich (2022): "Detecting p-hacking." *Econometrica* **90(2)**: pp. 887–906.

Elton, E. J., M. J. Gruber, & C. R. Blake (2003): "Marginal stockholder tax effects and ex-dividend day behavior-thirty-two years later." *Conference paper*, Presented at the European Finance Association Meeting held on Aug 20-23, 2003, Scotland: Glasgow, and The American Finance Association Meeting held on Jan 3-5, 2004, California: San Diego.

Engert, A. (2010): "Transnational hedge fund regulation." **11(3)**: pp. 329–378.

European Securities and Markets Authority (2019): *ESMA annual statistical report: EU alternative investment funds 2020.* Publications Office.

Fagetan, A. M. (2020): *The Regulation of Hedge Funds: A Global Perspective.* Springer Nature.

Fama, E. F. (1970): "Efficient capital markets: A review of theory and empirical work." *The journal of Finance* **25(2)**: pp. 383–417.

Fama, E. F. & K. R. French (1992): "The cross-section of expected stock returns." *the Journal of Finance* **47(2)**: pp. 427–465.

Fama, E. F. & K. R. French (1995): "Size and book-to-market factors in earnings and returns." *Journal of Finance* **50(1)**: pp. 131–155.

Fama, E. F. & K. R. French (1996): "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance* **51(1)**: pp. 55–84.

Fama, E. F. & K. R. French (1998): "Value versus growth: The international evidence." *The journal of finance* **53(6)**: pp. 1975–1999.

Fernandez, C., E. Ley, & M. F. J. Steel (2001): "Benchmark priors for Bayesian Model Averaging." *Journal of Econometrics* **100(2)**: pp. 381–427.

Ferran, E. (2011): "After the crisis: The regulation of hedge funds and private equity in the EU." **12(3)**: pp. 379–414.

Fioretos, O. (2010): "Capitalist diversity and the international regulation of hedge funds." **17(4)**: pp. 696–723.

Frumkin, D. & D. Vandegrift (2009): "The effect of size, age, beta and disclosure requirements on hedge fund performance." *Journal of Derivatives & Hedge Funds* **15**: pp. 241–251.

Frydenberg, S., K. Hrafnkelsson, V. S. Bromseth, & S. Westgaard (2017): "Hedge Fund Strategies and Time-Varying Alphas and Betas." *The Journal of Wealth Management* **19(4)**: pp. 44–60.

Fung, H.-G., X. E. Xu, & J. Yau (2002): "Global Hedge Funds: Risk, Return, and Market Timing." *Financial Analysts Journal* **58(6)**: pp. 19–30.

Fung, W. & D. A. Hsieh (1997): "Empirical characteristics of dynamic trading strategies: The case of hedge funds." *The Review of Financial Studies* **10(2)**: pp. 275–302.

Fung, W. & D. A. Hsieh (1999): "A primer on hedge funds." *Journal of Empirical Finance* **6(3)**: pp. 309–331.

Fung, W. & D. A. Hsieh (2000): "Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases." *Journal of Financial and Quantitative Analysis* **35(3)**: pp. 291–307.

Fung, W. & D. A. Hsieh (2001): "The risk in hedge fund strategies: Theory and evidence from trend followers." *The Review of Financial Studies* **14(2)**: pp. 313–341.

Fung, W. & D. A. Hsieh (2002): "Hedge-fund benchmarks: Information content and biases." *Financial Analysts Journal* **58(1)**: pp. 22–34.

Fung, W. & D. A. Hsieh (2004a): "Hedge fund benchmarks: A risk-based approach." *Financial Analysts Journal* **60(5)**: pp. 65–80.

Fung, W. & D. A. Hsieh (2004b): "Extracting portable alphas from equity long-short hedge funds." *Journal of Investment Management* **2(4)**: pp. 57–75.

Fung, W., D. A. Hsieh, N. Y. Naik, & T. Ramadorai (2008): "Hedge funds: Performance, risk, and capital formation." *The Journal of Finance* **63(4)**: pp. 1777–1803.

Fung, W. K., D. A. Hsieh *et al.* (2006): "Hedge funds: An industry in its adolescence." *Federal Reserve Bank of Atlanta Economic Review* **91(4th quarter)**: pp. 1–34.

Furukawa, C. (2020): "Publication bias under aggregation frictions: Theory, evidence, and a new correction method." *MIT Working paper*, MIT.

Gârleanu, N. & L. H. Pedersen (2022): "Active and passive investing: Understanding samuelson's dictum." *The Review of Asset Pricing Studies* **12(2)**: pp. 389–446.

Gechert, S., T. Havranek, Z. Irsova, & D. Kolcunova (2022): "Measuring capital-labor substitution: The importance of method choices and publication bias." *Review of Economic Dynamics* **45(C)**: pp. 55–82.

George, E. I. (2010): "Dilution priors: Compensating for model space redundancy." In "IMS Collections Borrowing Strength: Theory Powering Applications Â– A Festschrift for Lawrence D. Brown," volume 6, p. 158Â–165. Institute of Mathematical Statistics.

Getmansky, M., P. A. Lee, & A. W. Lo (2015): "Hedge funds: A dynamic industry in transition." *Annual Review of Financial Economics* **7**: pp. 483–577.

Goetzmann, W. N., J. E. Ingersoll Jr, & S. A. Ross (2003): "High-water marks and hedge fund management contracts." *The Journal of Finance* **58(4)**: pp. 1685–1718.

Griffin, J. M. & J. Xu (2009): "How smart are the smart guys? A unique view from hedge fund stock holdings." *The Review of Financial Studies* **22(7)**: pp. 2531–2570.

Guasoni, P. & J. Obłój (2016): "The incentives of hedge fund fees and high-water marks." *Mathematical Finance* **26(2)**: pp. 269–295.

Gupta, B., B. Cerrahoglu, & A. Daglioglu (2003): "Evaluating hedge fund performance: Traditional versus conditional approaches." *The Journal of Alternative Investments* **6(3)**: pp. 7–24.

Harvey, C. R. (2017): "Presidential Address: The Scientific Outlook in Financial Economics." *The Journal of Finance* **72(4)**: pp. 1399–1440.

Harvey, C. R., Y. Liu, & H. Zhu (2016): ". . . and the cross-section of expected returns." *The Review of Financial Studies* **29(1)**: pp. 5–68.

Havranek, T. (2010): "Rose effect and the euro: Is the magic gone?" *Review of World Economics* **146(2)**: pp. 241–261.

Havranek, T. (2015): "Measuring intertemporal substitution: The importance of method choices and selective reporting." *Journal of the European Economic Association* **13(6)**: pp. 1180–1204.

Havranek, T. & Z. Irsova (2010): "Measuring bank efficiency: A meta-regression analysis." *Prague Economic Papers* **2010(4)**: pp. 307–328.

Havranek, T. & Z. Irsova (2017): "Do borders really slash trade? A meta-analysis." *IMF Economic Review* **65(2)**: pp. 365–396.

Havranek, T., Z. Irsova, L. Laslopova, & O. Zeynalova (2023): "Publication and Attenuation Biases in Measuring Skill Substitution." *The Review of Economics and Statistics* **forthcoming**.

Havranek, T., Z. Irsova, & T. Vlach (2018a): "Measuring the income elasticity of water demand: The importance of publication and endogeneity biases." *Land Economics* **94(2)**: pp. 259–283.

Havranek, T., Z. Irsova, & O. Zeynalova (2018b): "Tuition fees and university enrolment: A meta-regression analysis." *Oxford Bulletin of Economics and Statistics* **80(6)**: pp. 1145–1184.

Havranek, T. & M. Rusnak (2013): "Transmission Lags of Monetary Policy: A Meta-Analysis." *International Journal of Central Banking* **9(4)**: pp. 39–76.

Havranek, T., T. D. Stanley, H. Doucouliagos, P. Bom, J. Geyer-Klingeberg, I. Iwasaki, W. R. Reed, K. Rost, & R. C. M. van Aert (2020): "Reporting Guidelines for Meta-Analysis in Economics." *Journal of Economic Surveys* **34(3)**: pp. 469–475.

Hodder, J. E., J. C. Jackwerth, & O. Kolokolova (2014): "Recovering delisting returns of hedge funds." *Journal of Financial and Quantitative Analysis* **49(3)**: pp. 797–815.

Hong, X. (2014): "The dynamics of hedge fund share restrictions." *Journal of Banking & Finance* **49(C)**: pp. 82–99.

Huang, Y. S., C. R. Chen, & I. Kato (2017): "Different strokes by different folks: The dynamics of hedge fund systematic risk exposure and performance." *International Review of Economics & Finance* **48(C)**: pp. 367–388.

Ibbotson, R. G., P. Chen, & K. X. Zhu (2011): "The ABCs of hedge funds: Alphas, Betas, and Costs." *Financial Analysts Journal* **67(1)**: pp. 15–25.

Ioannidis, J. P., T. D. Stanley, & H. Doucouliagos (2017): "The Power of Bias in Economics Research." *The Economic Journal* **127(605)**: pp. F236–F265.

Irsova, Z., P. R. D. Bom, & H. Rachinger (2023): "Spurious Precision in Meta-Analysis." *CEPR Discussion Papers 17927*, C.E.P.R. Discussion Papers.

Irsova, Z., T. Havranek, & D. Herman (2018): "Does daylight saving save energy? A meta-analysis." *The Energy Journal* **39(2)**: pp. 35–61.

JAME, R. (2018): "Liquidity Provision and the Cross Section of Hedge Fund Returns." *Management Science* **64(7)**: pp. 3288–3312.

JEFFREYS, H. (1961): *Theory of Probability.* Oxford Classic Texts in the Physical Sciences. Oxford: Oxford University Press, third edition.

JENSEN, M. C. (1968): "The performance of mutual funds in the period 1945–1964." *The Journal of Finance* **23(2)**: pp. 389–416.

JOBMAN, D. R. (2002): *The handbook of alternative investments.* John Wiley & Sons, USA: New York.

JOENVAARA, J., M. KAUPILA, R. KOSOWSKI, & P. TOLONEN (2019a): "Hedge fund performance: What do we know?" *CEPR Working paper*, CEPR.

JOENVAARA, J. & R. KOSOWSKI (2021): "The Effect of Regulatory Constraints on Fund Performance: New Evidence from UCITS Hedge Funds." *Review of Finance* **25(1)**: pp. 189–233.

JOENVAARA, J., R. KOSOWSKI, & P. TOLONEN (2019b): "The Effect of Investment Constraints on Hedge Fund Investor Returns." *Journal of Financial and Quantitative Analysis* **54(4)**: pp. 1539–1571.

JOHNSTON, A. (2015): "Regulating hedge funds for systemic stability: The EU's approach." **21(6)**: pp. 758–786. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/eulj.12159.

JORDAN, A. & P. SIMLAI (2011): "Risk characterization, stale pricing and the attributes of hedge funds performance." *Journal of Derivatives & Hedge Funds* **17(C)**: pp. 16–33.

JORION, P. & C. SCHWARZ (2014): "The strategic listing decisions of hedge funds." *Journal of Financial and Quantitative Analysis* **49(3)**: pp. 773–796.

JYLHA, P., K. RINNE, & M. SUOMINEN (2014): "Do Hedge Funds Supply or Demand Liquidity?" *Review of Finance* **18(4)**: pp. 1259–1298.

KAAL, W. A., B. LUPPI, & S. PATERLINI (2014): "Did the dodd-frank act impact hedge fund performance?" *U of St. Thomas (Minnesota) Legal Studies Research Paper* **(14-09)**.

KAAL, W. A. & D. A. OESTERLE (2016): "The history of hedge fund regulation in the united states." *Legal History eJournal* .

KAMAL, S. (2012): "Hedge fund regulation and the eu directive." *Available at SSRN 2050411* .

Kanuri, S. (2020): “Hedge Fund Performance in Japan.” *Review of Pacific Basin Financial Markets and Policies* **23(03)**: p. 2050023.

Kehoe, E. (2013): “Hedge fund regulation for systemic risk: Largely possible.” **14(1)**: pp. 35–68.

Kim, E. (2014): “Rethinking hedge fund regulation: Focusing on the u.s., the u.k., and korea.” .

Klein, P., D. Purdy, I. Schweigert, & A. Vedrashko (2015): “The Canadian hedge fund industry: Performance and market timing.” *International Review of Finance* **15(3)**: pp. 283–320.

Kooli, M. & I. Stetsyuk (2021): “Are hedge fund managers skilled?” *Global Finance Journal* **49(C)**: p. 100574.

Kosowski, R., N. Y. Naik, & M. Teo (2007): “Do hedge funds deliver alpha? A Bayesian and bootstrap analysis.” *Journal of Financial Economics* **84(1)**: pp. 229–264.

Kotkatvuori-Ornberg, J., J. Nikkinen, & J. Peltomaki (2011): “Geographical focus in emerging markets and hedge fund performance.” *Emerging Markets Review* **12(4)**: pp. 309–320.

Kouwenberg, R. & W. T. Ziemba (2007): “Incentives and risk taking in hedge funds.” *Journal of Banking & Finance* **31(11)**: pp. 3291–3310.

Kranz, S. & P. Putz (2022): “Methods matter: p-Hacking and publication bias in causal analysis in economics: Comment.” *American Economic Review* **112(9)**: pp. 3124–3136.

Ley, E. & M. F. Steel (2009): “On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression.” *Journal of Applied Econometrics* **24(4)**: pp. 651–674.

Liang, B. (1999): “On the performance of hedge funds.” *Financial Analysts Journal* **55(4)**: pp. 72–85.

Liang, B. (2000): “Hedge funds: The living and the dead.” *Journal of Financial and Quantitative analysis* **35(3)**: pp. 309–326.

Liang, B. (2003): “The accuracy of hedge fund returns.” *The Journal of Portfolio Management* **29(3)**: pp. 111–122.

Liang, B. (2004): “Alternative investments: CTAs, hedge funds, and funds-of-funds.” *Journal of Investment Management* **2(4)**: pp. 76–93.

Lim, J., B. A. Sensoy, & M. S. Weisbach (2016): "Indirect incentives of hedge fund managers." *The Journal of Finance* **71(2)**: pp. 871–918.

Ling, Y., J. Yao, & W. Liu (2015): "Chinese Hedge Funds - Performance and Risk Exposures." *The Chinese Economy* **48(5)**: pp. 330–350.

Lintner, J. (1965): "Security prices, risk, and maximal gains from diversification." *The Journal of Finance* **20(4)**: pp. 587–615.

Lo, A. W. (2001): "Risk Management for Hedge Funds: Introduction and Overview." *Financial Analysts Journal* **57(6)**: pp. 16–33.

Lukaj, M. S. & G. M. Healy (2007): "Hedge fund regulation: current trends in the industry." **8(1)**: pp. 4–12.

Malkiel, B. G. & A. Saha (2005): "Hedge funds: Risk and return." *Financial Analysts Journal* **61(6)**: pp. 80–88.

Malladi, R. (2020): "Luck versus skill in evaluating hedge fund managers' performance." *Journal of Business and Management* **26(1)**: pp. 22–39.

Markowitz, H. (1952): "Portfolio selection." *Journal of Finance* **7(1)**: pp. 77–91.

Matousek, J., T. Havranek, & Z. Irsova (2022): "Individual discount rates: a meta-analysis of experimental evidence." *Experimental Economics* **25(1)**: pp. 318–358.

Meligkotsidou, L. & I. D. Vrontos (2008): "Detecting structural breaks and identifying risk factors in hedge fund returns: A Bayesian approach." *Journal of Banking & Finance* **32(11)**: pp. 2471–2481.

Mitchell, M. & T. Pulvino (2001): "Characteristics of Risk and Return in Risk Arbitrage." *The Journal of Finance* **56(6)**: pp. 2135–2175.

Mladina, P. (2015): "Illuminating Hedge Fund Returns to Improve Portfolio Construction." *The Journal of Portfolio Management* **41(3)**: pp. 127–139.

Molyboga, M. & C. L'Ahelec (2016): "A simulation-based methodology for evaluating hedge fund investments." *Journal of Asset Management* **17(6)**: pp. 434–452.

Moschella, M. (2011): "Getting hedge funds regulation into the eu agenda: the constraints of agenda dynamics." *Journal of European Integration* **33(3)**: pp. 251–266.

Mossin, J. (1966): "Equilibrium in a capital asset market." *Econometrica* **34(4)**: pp. 768–783.

Mozes, H. A. (2013): "Decomposing Hedge Fund Returns: What Hedge Funds Got Right for the Past 20 Years." *The Journal of Investing* **22(3)**: pp. 9–20.

Naik, N. Y., T. Ramadorai, & M. Stromqvist (2007): "Capacity constraints and hedge fund strategy returns." *European Financial Management* **13(2)**: pp. 239–256.

Olden, A. & J. MĂ¸en (2022): "The triple difference estimator." **25(3)**: pp. 531–553.

Pal, M. (1980): "Consistent moment estimators of regression coefficients in the presence of errors in variables." *Journal of Econometrics* **14(3)**: pp. 349–364.

Patton, A. J. (2009): "Are 'market neutral' hedge funds really market neutral?" *The Review of Financial Studies* **22(7)**: pp. 2495–2530.

Patton, A. J. & T. Ramadorai (2013): "On the High-Frequency Dynamics of Hedge Fund Risk Exposures." *The Journal of Finance* **68(2)**: pp. 597–635.

Patton, A. J., T. Ramadorai, & M. Streatfield (2015): "Change you can believe in? Hedge fund data revisions." *The Journal of Finance* **70(3)**: pp. 963–999.

Posthuma, N. & P. J. Van der Sluis (2003): "A reality check on hedge funds returns." *Technical report*, ABP Investments (Netherlands: Schiphol) and Department of Finance at Vrije Universiteit Amsterdam (Netherlands: Amsterdam).

Quaglia, L. (2011): "The 'old' and 'new' political economy of hedge fund regulation in the european union." **34(4)**: pp. 665–682. Publisher: Routledge _eprint: https://doi.org/10.1080/01402382.2011.572383.

Racicot, F.-E. & R. Théoret (2009): "Integrating volatility factors in the analysis of the hedge fund alpha puzzle." *Journal of Asset Management* **10**: pp. 37–62.

Racicot, F.-E. & R. Théoret (2013): "The procyclicality of hedge fund alpha and beta." *Journal of Derivatives & Hedge Funds* **19(C)**: pp. 109–128.

Racicot, F.-E. & R. Théoret (2014): "Cumulant instrument estimators for hedge fund return models with errors in variables." *Applied Economics* **46(10)**: pp. 1134–1149.

Raftery, A. E. (1995): "Bayesian Model Selection in Social Research." *Sociological Methodology* **25**: pp. 111–163.

Ranaldo, A. & L. Favre (2005): "Hedge fund performance and higher-moment market models." *The Journal of Alternative Investments* **8(3)**: pp. 37–51.

Rathner, S. (2012): "The performance of socially responsible investment funds: A meta-analysis." *Working Paper 2012-03*, Working Papers in Economics and Finance.

RESTREPO, F. (2020): "Hedge Fund Regulation, Performance, and Risk-Taking: Re-Examining the Effect of the Dodd-Frank Act." *SSRN Electronic Journal* .

REVELLI, C. & J.-L. VIVIANI (2015): "Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis." *Business Ethics: A European Review* **24(2)**: pp. 158–185.

ROODMAN, D., J. G. MACKINNON, M. O. NIELSEN, & M. D. WEBB (2018): "Fast and wild: Bootstrap inference in Stata using boottest." *Queen's Economics Department Working Paper 1406*, Department of Economics, Queen's University, Canada: Kingston.

RÖSCH, D. M., A. SUBRAHMANYAM, & M. A. VAN DIJK (2017): "The dynamics of market efficiency." *The review of financial studies* **30(4)**: pp. 1151–1187.

RUSNAK, M., T. HAVRANEK, & R. HORVATH (2013): "How to solve the price puzzle? a meta-analysis." *Journal of Money, Credit and Banking* **45(1)**: pp. 37–70.

RZAKHANOV, Z. & G. JETLEY (2019): "Competition, scale and hedge fund performance: Evidence from merger arbitrage." *Journal of Economics and Business* **105(C)**: p. 105841.

SABBAGHI, O. (2012): "Hedge fund return volatility and comovement: recent evidence." *Managerial Finance* **38(1)**: pp. 101–119.

SADKA, R. (2010): "Liquidity risk and the cross-section of hedge-fund returns." *Journal of Financial Economics* **98(1)**: pp. 54–71.

SADKA, R. (2012): "Hedge-fund performance and liquidity risk." *Journal of Investment Management* **10(C)**: pp. 60–72.

SAMI, G. (2009): "A comparative analysis of hedge fund regulation in the united states and europe." p. 35.

SANDVIK, S. H., S. FRYDENBERG, S. WESTGAARD, & R. K. HEITMANN (2011): "Hedge Fund Performance in Bull and Bear Markets: Alpha Creation and Risk Exposure." *The Journal of Investing* **20(1)**: pp. 52–77.

SCHNEEWEIS, T. (1998): "Dealing with myths of hedge fund investment." *The Journal of Alternative Investments* **1(3)**: pp. 11–15.

SHARPE, W. F. (1966): "Mutual fund performance." *The Journal of Business* **39(1)**: pp. 119–138.

Shi, Z. (2017): "The impact of portfolio disclosure on hedge fund performance." *Journal of Financial Economics* **126(1)**: pp. 36–53.

Stafylas, D., K. Anderson, & M. Uddin (2018): "Hedge fund performance attribution under various market conditions." *International Review of Financial Analysis* **56(C)**: pp. 221–237.

Stafylas, D. & A. Andrikopoulos (2020): "Determinants of hedge fund performance during 'good' and 'bad' economic periods." *Research in International Business and Finance* **52(C)**: p. 101130.

Stambaugh, R. F. (2014): "Presidential Address: Investment Noise and Trends." *The Journal of Finance* **69(4)**.

Stanley, T. & H. Doucouliagos (2010): "Picture This: A Simple Graph That Reveals Much Ado About Research." *Journal of Economic Surveys* **24(1)**: pp. 170–191.

Stanley, T. D. (2001): "Wheat from Chaff: Meta-analysis as Quantitative Literature Review." *Journal of Economic Perspectives* **15(3)**: pp. 131–150.

Stanley, T. D. (2005): "Beyond Publication Bias." *Journal of Economic Surveys* **19(3)**: pp. 309–345.

Stanley, T. D. & H. Doucouliagos (2017): "Neither fixed nor random: weighted least squares meta-regression." *Research Synthesis Methods* **8(1)**: pp. 19–42.

Stanley, T. D., S. B. Jarrell, & H. Doucouliagos (2010): "Could It Be Better to Discard 90% of the Data? A Statistical Paradox." *The American Statistician* **64(1)**: pp. 70–77.

Stoforos, C. E., S. Degiannakis, & T. B. Palaskas (2017): "Hedge fund returns under crisis scenarios: A holistic approach." *Research in International Business and Finance* **42(C)**: pp. 1196–1207.

Stulz, R. M. (2007): "Hedge funds: Past, present, and future." *Journal of Economic Perspectives* **21(2)**: pp. 175–194.

Sullivan, R. N. (2021): "Hedge Fund Alpha: Cycle or Sunset?" *The Journal of Alternative Investments* **23(3)**: pp. 55–79.

Sun, L. (2018): "Implementing valid two-step identification-robust confidence sets for linear instrumental-variables models." *Stata Journal* **18(4)**: pp. 803–825.

Sun, Z., A. Wang, & L. Zheng (2012): "The Road Less Traveled: Strategy Distinctiveness and Hedge Fund Performance." *The Review of Financial Studies* **25(1)**: pp. 96–143.

Teo, M. (2009): "The Geography of Hedge Funds." *The Review of Financial Studies* **22(9)**: pp. 3531–3561.

The European Parliament and the Council of the European Union (2011): "Directive 2011/61/EU of the european parliament and of the council of 8 june 2011 on alternative investment fund managers and amending directives 2003/41/EC and 2009/65/EC and regulations (EC) no 1060/2009 and (EU) no 1095/2010 text with EEA relevance."

Vivalt, E. (2020): "How much can we generalize from impact evaluations?" *Journal of the European Economic Association* **18(6)**: pp. 3045–3089.

Vrontos, S. D., I. D. Vrontos, & D. Giamouridis (2008): "Hedge fund pricing and model uncertainty." *Journal of Banking & Finance* **32(5)**: pp. 741–753.

Woll, C. (2012): *The Defense of Economic Interests in the European Union: The Case of Hedge Fund Regulation*. Campus Verlag. Pages: 197-212 Publication Title: Crisis and Control: Institutional Change in Financial Market Regulation.

Woll, C. (2013): "Lobbying under pressure: The effect of salience on e uropean u nion hedge fund regulation." *JCMS: Journal of Common Market Studies* **51(3)**: pp. 555–572.

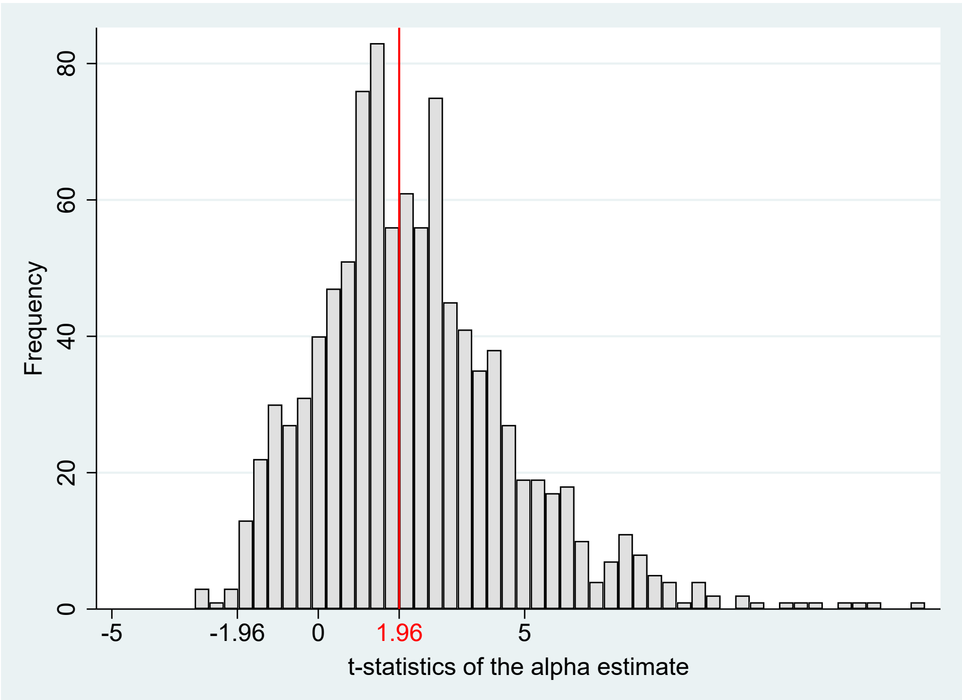
Yang, F., T. Havranek, Z. Irsova, & J. Novak (2023): "Is research on hedge fund performance published selectively? a quantitative survey." *Journal of Economic Surveys* .

Zigraiova, D., T. Havranek, Z. Irsova, & J. Novak (2021): "How puzzling is the forward premium puzzle? A meta-analysis." *European Economic Review* **134(C)**: p. 103714.

Appendix A

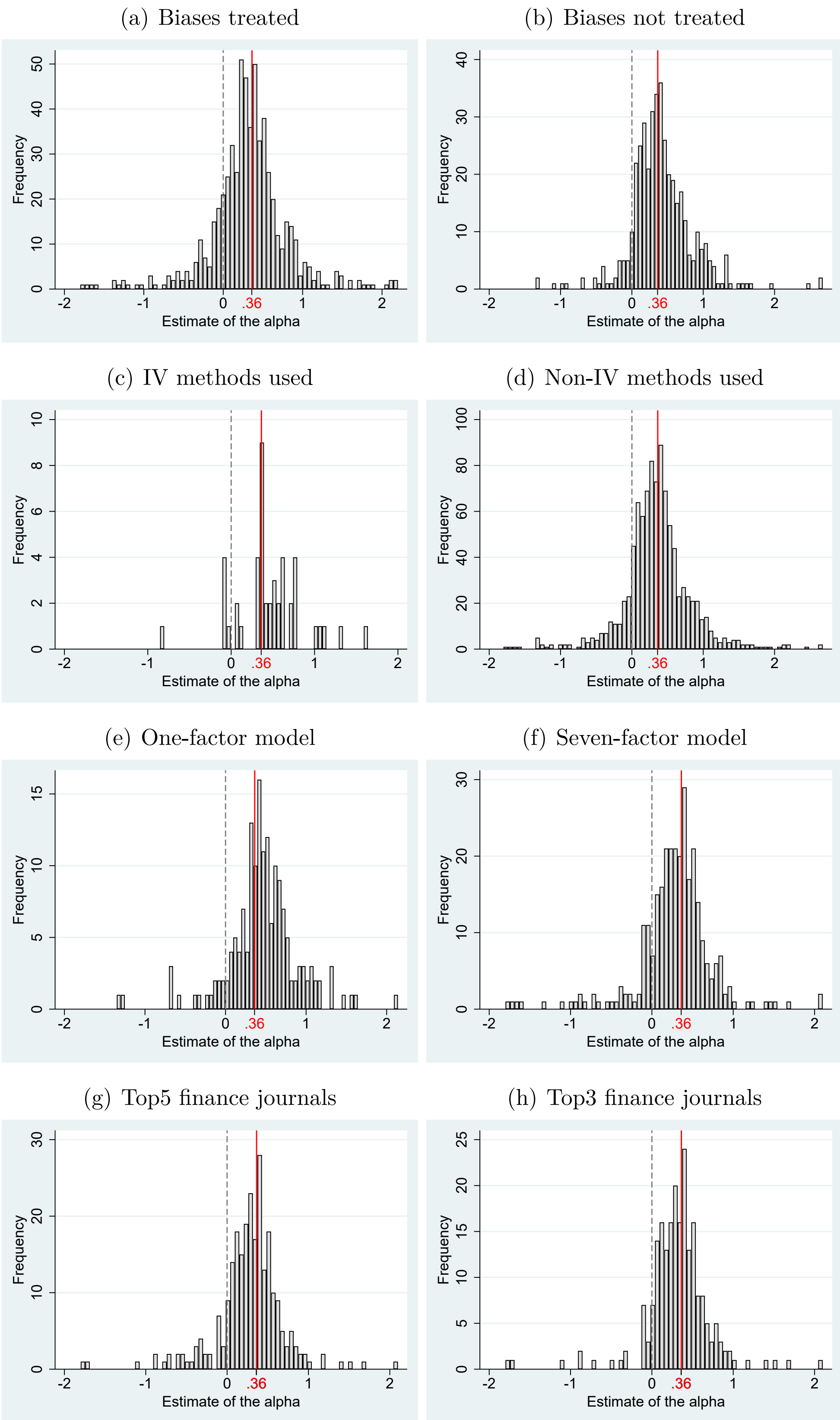
Appendix to Chapter 2

Figure A1: Distribution of t-statistics



Note: The figure represents the distribution of t-statistics of the reported estimates of the alpha. Red lines represents critical value of 1.96 associated with significance at the 5% level and the value of 0 associated with changing the sign of the estimate. We exclude estimates with large t-statistics from the figure for ease of exposition but include them in statistical tests.

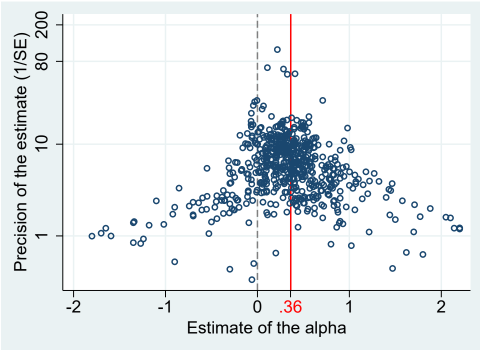
Figure A2: Histograms for subsamples



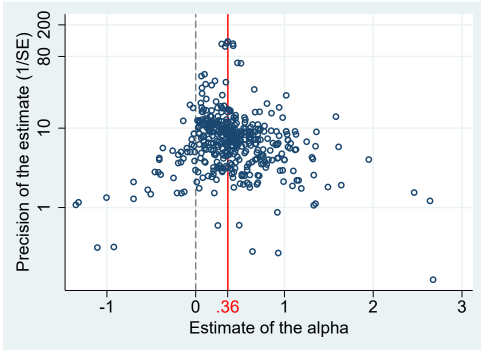
Notes: The figure depicts funnel plots of the reported alphas divided based on the treatment of biases, the implementation of methods (IV = instrumental variables) and models, and the quality of journals in finance. The solid vertical line denotes the sample mean; the dashed vertical line denotes the null alpha.

Figure A3: Funnel plots for subsamples

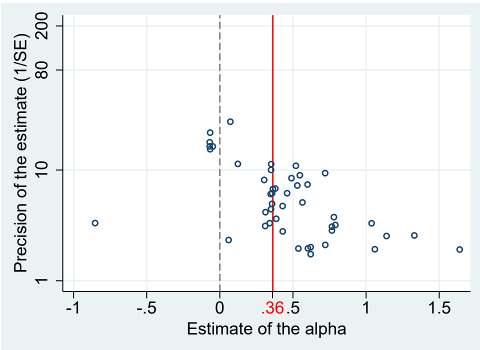
(a) Biases treated



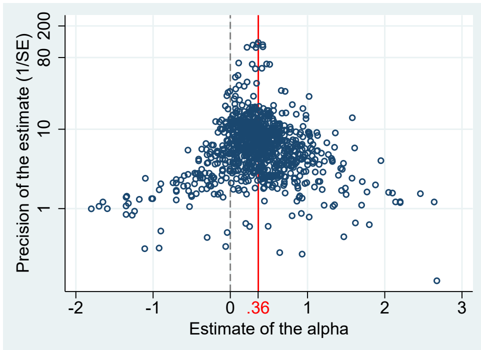
(b) Biases not treated



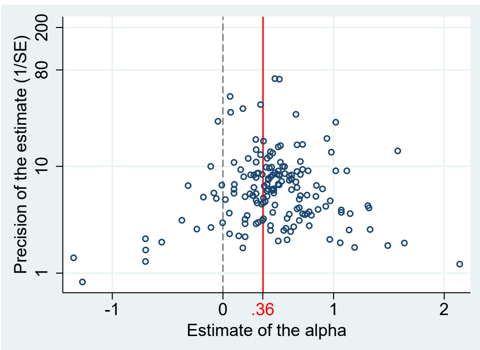
(c) IV methods used



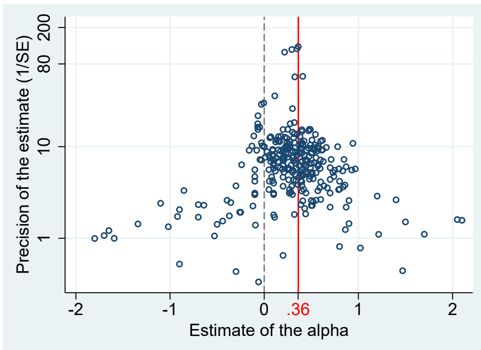
(d) Non-IV methods used



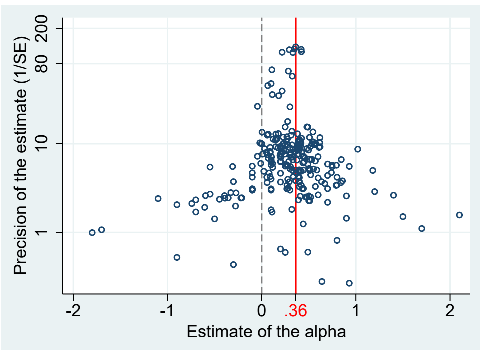
(e) One-factor model



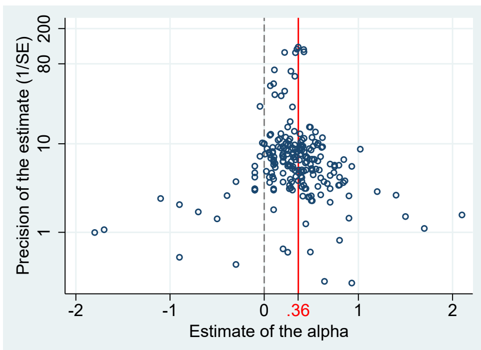
(f) Seven-factor model



(g) Top5 finance journals



(h) Top3 finance journals



Notes: The figure depicts funnel plots of the reported alphas divided based on the treatment of biases, the implementation of methods (IV = instrumental variables) and models, and the quality of journals in finance. The solid vertical line denotes the sample mean; the dashed vertical line denotes the null alpha.

Table A1: Specification test of Andrews & Kasy (2019)

	All estimates	Bias treated	Bias not treated	IV estimates	Non-IV estimates
Correlation	0.330 [0.264, 0.392]	0.359 [0.275, 0.430]	0.318 [0.191, 0.431]	0.389 [-0.031, 0.769]	0.322 [0.243, 0.379]
Observations	1,019	605	414	46	973
		One-factor model	Seven-factor model	Top 5 journals	Top 3 journals
Correlation		0.096 [-0.091, 0.255]	0.287 [0.166, 0.407]	0.262 [0.155, 0.375]	0.275 [0.159, 0.393]
Observations		167	298	256	218

Notes: The table shows the inverse publication-probability-weighted correlations between $log(\alpha)$ and $log(SE(\alpha))$, tests developed by Kranz & Putz (2022) for viability of Andrews & Kasy (2019) publication bias test. If all the assumptions of the selection model hold, the correlation should be zero. Bootstrapped standard errors in parentheses.

Table A2: Clustered for author teams

	OLS	FE	BE	IV	WLS	wNOBS
Publication bias (λ)	-0.0152	-0.0265	0.0602	0.178	0.324	0.0497
	(0.171)	(0.206)	(0.131)	(0.359)	(0.310)	(0.124)
	[-0.488, 0.365]			[-0.617, 1.114]	[-0.377, 1.097]	[-0.313, 0.458]
				{-0.639, 1.067}		
Effect beyond bias (κ)	0.366***	0.369***	0.350***	0.316***	0.301***	0.353***
	(0.0453)	(0.0516)	(0.0474)	(0.0852)	(0.0440)	(0.0397)
	[0.265, 0.470]			[0.150, 0.474]	[0.183, 0.416]	[0.274, 0.434]
First-stage robust F-stat				11.71		
Studies	74	74	74	73	74	74
Observations	1,019	1,019	1,019	979	1,019	1,019

Notes: The table reports the results of the regression $\alpha_{ij} = \kappa + \lambda \cdot SE(\alpha_{ij}) + \epsilon_{ij}$, where α_{ij} denotes the i-th alpha coefficient estimated in the j-th study, and $SE(\alpha_{ij})$ denotes its standard error. FE: study-level fixed effects, BE: study-level between effects, IV: the inverse of the square root of the number of observations is used as an instrument for the standard error, WLS: model is weighted by the inverse of the squared standard error of an estimate, wNOBS: model is weighted by the inverse of the number of estimates per study. Standard errors, clustered at the level of authors, are reported in parentheses. 95% confidence intervals from wild bootstrap in square brackets (Roodman *et al.* 2018). In curly brackets we show the two-step weak-instrument-robust 95% confidence interval based on Andrews (2018) and Sun (2018).* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Tests of p-hacking

	20 bins	15 bins	10 bins
Test for non-increasingness	0.469	0.179	0.403
Test for monotonicity and bounds	0.242	0.223	0.481
Observations ($p <= 0.15$)	663	663	663
Total observations	1,019	1,019	1,019

Notes: Results of p-hacking tests based on Elliott *et al.* (2022) for the whole sample.

Appendix B

Appendix to Chapter 3

Table B1: Definition and Descriptive Statistics of Explanatory Variables

Variable	Description	Mean	SD	WM
Alpha	The estimate of the alpha (response variable).	0.362	0.477	0.365
Standard error (SE)	Standard error of the alpha. The variable is important for gauging publication bias.	0.251	0.285	0.240
SE * IV method	The interaction term between the standard error and IV method capturing the publication bias among IV estimates.	0.011	0.060	0.004
Dependent variable				
Individual funds	= 1 if the dependent variables is defined as returns of individual funds and 0 otherwise.	0.172	0.377	0.264
Equal-weighted funds	= 1 if the dependent variable is defined as equal-weighted returns and 0 otherwise.	0.494	0.500	0.517
Value-weighted funds	= 1 if the dependent variable is defined as value-weighted returns and 0 otherwise (reference category).	0.335	0.472	0.220
Net-of-fee returns	= 1 if the dependent variable is defined in net-of fees returns and 0 otherwise.	0.966	0.182	0.939
Gross returns	= 1 if the dependent variable is defined in gross returns including fees and 0 otherwise (reference category).	0.034	0.182	0.061
Data characteristics				
Cross-section data	= 1 if cross-sectional data are used to estimate the effect.	0.839	0.368	0.735
Longitudinal data	= 1 if longitudinal data are used to estimate the effect (reference category).	0.161	0.368	0.265
Data year	The logarithm of the mean year of the data used minus the earliest average year in our data (base = 1990).	2.405	0.603	2.485
Database: default	= 1 if the estimates are based on the data provided by either TASS, HFR, BarclayHedge, or EurekaHedge databases and 0 otherwise.	0.527	0.500	0.626
Database: CST	= 1 if the estimates are based on the data provided by Credit Suisse/Tremont/Dow Jones Credit Suisse database and 0 otherwise.	0.251	0.434	0.154

Continued on next page

Table B1: Definition and Descriptive Statistics of Explanatory Variables (continued)

Variable	Description	Mean	SD	WM
Database: CISDM	= 1 if the estimates are based on the data provided by CISDM database and 0 otherwise.	0.175	0.380	0.206
Database: hand-collected	= 1 if the estimates are based on the data collected by hand and 0 otherwise.	0.022	0.145	0.038
Database: other	= 1 if the estimates are based on other than aforementioned databases.	0.164	0.370	0.172
Number of databases	Total number of databases used to estimate alpha.	1.366	1.048	1.480
Structural variation				
Developed markets	= 1 if the estimates are based on the data of developed market economies (IMF classification).	0.137	0.344	0.134
World markets	= 1 if the estimates are based on the data of global markets (reference category for geographical location).	0.863	0.344	0.866
Bull market	= 1 if the estimates are relevant to bull market conditions.	0.038	0.192	0.016
Bear market	= 1 if the estimates are relevant to bear market conditions.	0.038	0.192	0.016
Hedge fund strategy				
Strategy: all funds	= 1 if the estimates are based on the data of all funds and 0 otherwise.	0.238	0.426	0.345
Strategy: equity hedge	= 1 if the estimates are based on the data of equity hedge funds and 0 otherwise.	0.225	0.418	0.186
Strategy: event driven	= 1 if the estimates are based on the data of event driven funds (merger arbitrage, distressed securitities) and 0 otherwise.	0.111	0.314	0.102
Strategy: relative value	= 1 if the estimates are based on the data of relative value strategy funds (fixed income arbitrage, convertible arbitrage) and 0 otherwise.	0.092	0.290	0.085
Strategy: global	= 1 if the estimates are based on the global hedge funds and 0 otherwise.	0.153	0.360	0.105
Strategy: fund of funds	= 1 if the estimates are based on the data of funds of hedge funds and 0 otherwise.	0.066	0.248	0.080
Strategy: multi	= 1 if the estimates are based on the data of multistrategy funds and 0 otherwise.	0.039	0.194	0.020
Strategy: other	= 1 if other strategy of hedge funds is used for estimation (reference category for the group of strategies).	0.076	0.264	0.078
Estimation technique				
IV method	= 1 if instrumental variable approach (such as GMM and 2SLS) is used for estimation.	0.045	0.208	0.017
non-IV method	= 1 if other than IV method is used for estimation (reference category for methods).	0.955	0.208	0.983
1-factor model	= 1 if one-factor model or its modifications are used to estimate the alpha.	0.164	0.370	0.139
3-factor model	= 1 if three-factor model or its modifications are used to estimate the alpha.	0.070	0.255	0.081
4-factor model	= 1 if four-factor model or its modifications are used to estimate the alpha.	0.201	0.401	0.161

Continued on next page

Table B1: Definition and Descriptive Statistics of Explanatory Variables (continued)

Variable	Description	Mean	SD	WM
7-factor model	= 1 if seven-factor model or its modifications are used to estimate the alpha.	0.292	0.455	0.363
Modeling model uncertainty	= 1 if methods dealing with model uncertainty (such as stepwise regression or model averaging) are used to estimate the alpha.	0.139	0.346	0.112
Asset-based model	= 1 if asset-based models are used to estimate the alpha.	0.079	0.269	0.095
Other model	= 1 if other (sophisticated) models are used (reference category for the group of models).	0.055	0.228	0.049
Survivorship treated	= 1 if the survisorship bias is dealt with and 0 otherwise.	0.576	0.494	0.616
Backfilling treated	= 1 if the backfilling bias is dealt with and 0 otherwise.	0.301	0.459	0.343
Publication characteristics				
Publication year	The logarithm of the year when the study appeared in Google Scholar normalized by the year of the earliest publication in our sample.	2.260	0.713	2.252
Citations	The logarithm of the number of per-year citations of the study in Google Scholar.	1.748	1.041	1.773
Impact factor	The discounted recursive RePEc impact factor of the outlet.	3.650	5.081	4.034

Notes: The table provides the definition, the (unweighted) mean value (Mean), the standard deviation (SD), and the mean weighted by the inverse of the number of estimates reported per study (WM) for the explanatory variables that we use in our regression analysis. GMM denotes the generalized method of moments, and 2SLS denotes two-stage least squares method.

Appendix C

Appendix to Chapter 4

Table C1: Variable Description

Variable	Description
Dependent variable	
Alpha_FH model	Alpha obtained by Fung& Hsieh (2008) eight-factor model.
Alpha_1 factor	Alpha obtained by CAPM model.
Alpha_3 factors	Alpha obtained by Fama & French (1998) three-factor model.
Alpha_4 factors	Alpha obtained by Carhart (1997) four-factor model.
Excess Return	The performance is measured by the difference between excess return and risk-free return.
Control variable	
aum_euros	The lagged AUM in euros.
lnaum	The log terms of fund i's assets in euros at the end of period t-1.
aumsq	The quadratic terms of fund i's assets in euros at the end of period t-1.
Independent variable (DiD and DDD)	
EU	Indicator variable equal to 1 if the fund is located or marketing in European Union and 0 otherwise.
Size	Indicator variable equal to 0 if aum_euros≤100 or 100<aum_euros≤500 & Leverage="No" & lockup is at least five years (meet exemption requirements) and 1 otherwise.
Post	Indicator variable equal to 1 if the period is after the AIFMD has be imposed and 0 otherwise.
EU_Post	Variable equal to EU*Post.
NonUCITS	Indicator variable equal to 0 if the fund is subject to UCITS and 1 otherwise. Status in 2021 used for the entire period.
Nonexemption	Variable equal to Size*Nonucits.
Nonexemption_Post	Variable equal to Nonexemption*Post.
EU_Nonexemption	Variable equal to EU*Nonexemption.
EU_Nonexemption_Post	Variable equal to EU*Nonexemption*Post.
Disclosure	Variable equal to the business extent of the disclosure index from World Bank Open Data.
Disclosure_Post	Variable equal to Disclosure*Post.

Continued on next page

Table C1: Variable Description (continued)

Variable	Description
Disclosure_Nonexemption	Variable equal to Disclosure*Nonexemption.
Disclosure_Nonexemption_Post	Variable equal to Disclosure*Nonexemption*Post.
PSM variable	
treated_post	Indicator variable equal to 1 if the fund is authorized and is within the authorized periods and 0 otherwise.
Hurdle Rate	Indicator variable equal to 1 if the fund has hurdle rate and 0 otherwise.
Flow_period	The flows during the 2 year window.
Watermark	Indicator variable equal to 1 if the fund uses watermark to evaluate managers' performance and 0 otherwise.
Leverage	Indicator variable equal to 1 if the fund use the leverage and 0 otherwise.
Management Fee	The management fee.
Performance Fee	The performance fee.
Lockup	Indicator variable equal to 1 if the fund has the lockup requirement and 0 otherwise.

Notes: The table provides the variable list and their descriptions.