Individual Stock-picking Skills in Active Mutual Funds

Yixin Chen*

January 28, 2019

I propose a new methodology to identify active mutual funds that can predictively outperform passive benchmarks due to stock-picking skills by applying a first-order stochastic dominance (FSD) condition. The FSD condition is implemented by extracting information from fund holdings to bootstrap counterfactual random portfolios with given factor loadings and degree of diversification. A simple factor model and simulation results show that the FSD condition complements conventional alpha-based metrics by developing robustness to heteroscedasticity and benchmark mis-specification problems. Empirically, the identified funds outperform passive benchmarks by large magnitudes out of sample. Findings on fund characteristics and flows further support the results.

*University of Rochester, Simon Business School. Contact: yixin.chen@simon.rochester.edu
This paper is based on the first chapter of my PhD dissertation. I owe an enormous intellectual debt to my advisors: Hui Chen, Leonid Kogan (Co-Chair) and Jonathan Parker (Co-Chair). I would also like to thank Zahi Ben-David, Randy Cohen, Simon Gervais, Dan Greenwald, Adam Guren, Marcin Kacperczyk, Mark Kritzman, Rich Lindsey, Andy Lo, Andrey Malenko, Paul Mende, Jun Pan, Eric So, Adrien Verdelhan, Jiang Wang, Andy Weisman, Haoxiang Zhu, and seminar participants at MIT, Goldman Sachs, University of Rochester, SAIF, CKGSB, CityU of HK, NTU, HKUST, University of Notre Dame, Bluemountain Capital, AQR, Fudan University and SHUFE for valuable comments. All errors are my own.

1 Introduction

It is well-known that skilled active mutual fund managers who can predictably outperform passive benchmarks are difficult to identify.[1] Current evaluation methods measure a fund manager's skills by comparing the time average of her returns[2] to that of an appropriately chosen benchmark. For example, fund manager i's skills can be measured as:

$$\hat{\alpha}_i = \frac{1}{T} \sum_t \left(r_{i,t} - r_{i,t}^b \right),$$

where $\{r_{i,t}\}_{t=1}^T$ are fund i's realized returns[3], $\left\{r_{i,t}^b\right\}_{t=1}^T$ are the returns of a passive benchmark representing fund i's exposure to systematic factors.

While this in-sample alpha is straightforward to construct, it has poor predictive power of the manager's out-of-sample alpha within the active mutual fund industry. Table 1 replicates the Carhart (1997) regression during a recent sample period. According to the table, none of the post-ranking alphas of the in-sample-alpha-ranked portfolios of funds is statistically significant at 5% confidence level. The predictive power of the in-sample Carhart four-factor alpha for future fund performance is weak and statistically insignificant.

[1]See, for example, Carhart (1997), Kosowski et al. (2006), Barras, Scaillet, and Wermers (2010), Fama and French (2010), etc.

[2]All performances are before fees unless specified otherwise.

[3]I don't make the distinction between fund managers and funds in this paper. I will use the term fund managers and funds interchangeably.

Table 1: The Weak Persistence of Alpha

Decile	α (in %)	mkt	smb	hml	umd	SR	IR
1	-1.37*	1.03***	0.35***	0.07***	0.01	0.48	-0.41
	[-1.88]	[63.2]	[13.4]	[2.81]	[0.79]		
2	-0.56	1.00***	0.21***	0.05**	0.00	0.52	-0.19
	[-0.84]	[62.2]	[9.13]	[2.17]	[-0.07]		
3	-0.26	0.99***	0.15***	0.07***	0.00	0.54	-0.10
	[-0.46]	[65.5]	[6.37]	[3.30]	[-0.31]		
4	0.08	0.98***	0.15***	0.05**	0.00	0.56	0.03
	[0.15]	[75.4]	[8.13]	[2.22]	[0.02]		
5	-0.16	1.00***	0.13***	0.06***	0.00	0.55	-0.07
	[-0.32]	[75.5]	[7.09]	[2.79]	[0.28]		
6	0.07	0.99***	0.12***	0.04**	0.00	0.56	0.03
	[0.14]	[83.0]	[6.03]	[2.02]	[0.27]		
7	-0.08	1.00***	0.15***	0.04*	0.01	0.55	-0.03
	[-0.16]	[74.4]	[6.27]	[1.79]	[0.87]		
8	0.36	1.00***	0.19***	0.02	0.02	0.58	0.14
	[0.64]	[75.0]	[9.27]	[0.69]	[1.31]		
9	1.04*	1.01***	0.23***	-0.02	0.02	0.61	0.36
	[1.80]	[67.6]	[8.86]	[-0.62]	[1.55]		
10	1.13	1.05***	0.40***	-0.14***	0.05**	0.59	0.30
	[1.46]	[52.9]	[14.2]	[-4.44]	[2.32]		

This table documents the before-fees performance of the trading strategy that sorts funds by their historical alpha. By the end of each quarter, the active equity mutual funds in the cross section are sorted into ten deciles based on the four-factor alpha computed from their proceeding 24 months' before-fees returns. The trading strategy is rebalanced every three months. The post-ranking annualized before-fees alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. The sample period is from January 1991 to December 2015.

The search for skilled managers based on in-sample alpha has poor out-of-sample performance because it has little power to distinguish skill from luck given the short fund performance histories.[4] The conventional positive alpha ($\alpha_i > 0$) condition requires the mean of the fund's return to be higher than the mean of the benchmark return, i.e. $\mathbb{E}\left(r_{i,t}\right) > \mathbb{E}\left(r^b_{i,t}\right)$. Thus, it suffers from the empirical problem that mean is difficult

[4] See Kosowski et al. (2006), Fama and French (2010) and Barras, Scaillet, and Wermers (2010) for ex post analysis.

to estimate in finite sample, and it is relatively easy for unskilled managers to achieve high in-sample alphas with excessive risk-taking or unobservable factor exposures[5].

In this paper, I focus on a subset of skilled fund managers who generate positive alpha by making profitable bets on firm specific risks (stock-picking), and show that a new first-order stochastic dominance (FSD) condition can be imposed to identify such skilled stock-pickers. The new FSD condition states that the return distribution of a skilled stock-picker should first-order stochastically dominate that of a counterfactual manager with similar investment style but no stock-picking skills, i.e. $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$. The FSD condition is a more stringent requirement than conditions focused on the mean alone such as $\alpha_i > 0$, because it uses information from the entire distribution of returns and excludes funds with heavy left tails in their return distributions due to excessive risk-taking or unobservable factor exposures. Intuitively, the new FSD condition is stronger than the conventional positive alpha condition but weaker than arbitrage. The positive alpha condition states that the fund manager's performance should be on average better than the passive benchmark, whereas arbitrage requires the manager's performance to exceed the benchmark in every single period. The FSD condition, on the other hand, identifies fund managers with high (low) probability to outperform (underperform) the unskilled counterfactual managers.

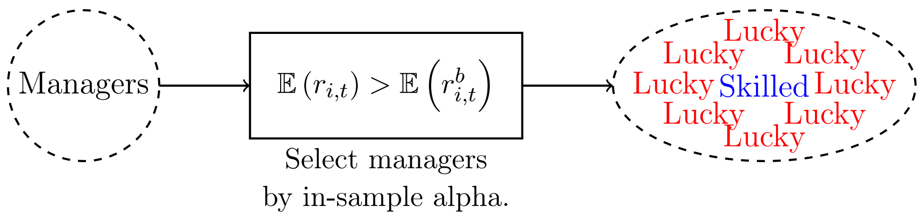
As a methodological contribution of this paper, I construct an FSD filter to select funds by testing the FSD condition. The test requires extending the benchmark from a **single return** to a **return distribution in each period**. Specifically, for each fund in each period, I draw a counterfactual return distribution from a bootstrap exercise by creating replica funds with random portfolios. The replica funds maintain the same portfolio weights as the original fund and invest in stocks with similar observ-

[5] See the empirical evidence documented by Chevalier and Ellison (1997) and a more recent structural estimation by Koijen (2014).

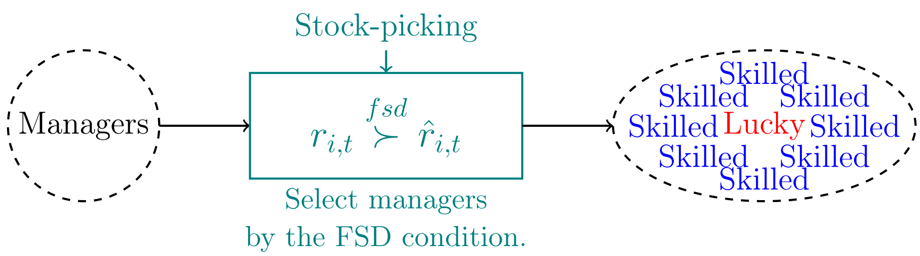
able characteristics, so that the replica funds resemble the original fund in the degree of diversification and loadings on observable factors.[6] However, the specific choices of stocks in a replica fund's portfolio are determined randomly. As a result, the replica funds inherit the characteristics of the original fund, meanwhile break the association between portfolio weights and stock choices, which reflects the stock-picking skills of the fund manager. The comparison between the original fund's return and the replica funds' return distribution then enables the econometrician to conduct a statistical test on the manager's stock-picking skills in each single period with only one observation. With repeated observations over time, the FSD filter selects funds by requiring the percentiles of the original fund's returns among the replica funds to first-order stochastically dominate a standard uniform distribution.

[6] I do not make the distinction between characteristics and factor loadings in this paper as discussed in Daniel and Titman (1997). However, the same methodology is applicable with either characteristics or factor loadings whenever the difference between the two is important.

Figure 1: Positive Alpha Filter VS FSD Filter



(a) Positive Alpha Filter



(b) FSD Filter

This figure illustrates the comparison between the positive alpha filter and the FSD filter. Panel (a) illustrates the ineffectiveness of the positive alpha filter in identifying skilled mutual fund managers; whereas Panel (b) proposes the application of the new FSD filter. The FSD filter selects managers whose return distributions first-order stochastically dominate counterfactual managers with similar investment styles but no stock-picking skills, i.e. $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$.

To understand the source of the additional statistical power that the new FSD condition is able to provide, I conduct simulations to compare the performance of a filter based on the conventional positive alpha condition with the performance of the FSD filter, as illustrated in Figure 1. Simulations show that the FSD filter outperforms the positive alpha filter in handling two statistical problems in finite samples – heteroscedasticity and benchmark mis-specification. The heteroscedasticity problem is defined as idiosyncratic volatility being time-varying and more volatile than the fund's true alpha. I show that the FSD filter has better performance than the positive alpha filter with heteroscedasticity calibrated to the real-world level because the positive alpha filter places equal weight on all observations regardless of idiosyncratic volatility; whereas the FSD filter places relatively higher (lower) weights on observations from high (low) signal-to-noise ratio periods. The benchmark mis-specification problem is defined as the situation that some managers might take on risk factors that are not observable to the econometrician. The performance of the positive alpha filter suffers due to the additional noise from the unobservable factors. The positive alpha filter tends to erroneously select mis-specified managers who take on unobservable factors with high in-sample realizations rather than the truly skilled managers who are able to deliver positive out-of-sample alphas. The FSD filter, on the other hand, is unaffected by this problem thanks to a detection mechanism. The unobservable factors taken by the mis-specified managers induce heavier left tails in their return distributions compared to the replica funds thereby violating the FSD condition. The FSD condition and the positive alpha condition are not mutually exclusive. The FSD requirement complements the alpha measure by providing robustness to heteroscedasticity and benchmark mis-specification in selecting fund managers who are good at picking stocks.

In the empirical part of this paper, I show that the FSD filter is indeed effective in selecting skilled stock-pickers. From January 1991 to December 2015, the FSD filter identifies a time-varying group of active mutual funds that are able to, on average, outperform the Carhart four-factor benchmark by 203 bps ($t = 2.78$) per year out of sample before management fees (78 bps ($t = 1.07$) per year after fees). More interestingly, even though there is only weak performance persistence among all funds in the cross section, among the funds selected by the FSD filter, the performance persistence is much stronger with in-sample alpha being significantly predictive of out-of-sample alpha. The combination of the FSD filter and the standard $\hat{\alpha}$ sort is especially powerful in identifying skilled stock-pickers, with the top quintile of funds in the second-stage sort by in-sample alpha outperforming the Carhart benchmark by as high as 371 bps ($t = 3.35$) per year before management fees (240 bps ($t = 2.17$) per year after fees). By investigating the fund return gaps[7], I further verify that about half of the Carhart four-factor alphas of the outperforming funds are resulted from profitable unobserved within-quarter trades. The finding lends additional support to the empirical success of the FSD filter in identifying skilled stock-pickers and is consistent with the view that profitable information is usually short-lived in a stock market that is largely liquid and efficient.

The investigation into the observable characteristics of the outperforming funds also produces interesting findings. The identified outperforming funds manifest characteristics that are distinctive from an average fund in the industry in the following aspects:

1. They have similar sizes as an average fund measured as asset under management, but they are able to charge higher fees.

2. They keep fewer stocks within their portfolios.

3. Controlling for realized in-sample alphas, funds that satisfy the FSD condition attract more flows.

The finding with the fee setting is partially consistent with the prediction by Berk and Green (2004) in the sense that more successful funds are able to extract higher rents, but inconsistent in the specific mechanism. In Berk and Green (2004), skilled managers demand compensation by growing the size of their funds meanwhile keeping the fees fixed. My finding, on the other hand, suggests that the outperforming managers are able to charge higher fees directly rather than growing the size of their funds. The finding with portfolio concentration is consistent with the theory by Van Nieuwerburgh

[7] See Kacperczyk, Sialm, and Zheng (2008).

and Veldkamp (2010). My finding verifies their prediction that informed investors could voluntarily keep under-diversified portfolios in order to become specialized when information acquisition is costly. The finding with the fund flows echoes the empirical work by Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016). Based on their arguments that fund flows reflect investors' evaluations of managers' skills, I show that investors infer the quality of funds from the properties of their return distributions that are beyond the mean or alpha. Yet, the positive out-of-sample alphas of the identified funds also suggest that the magnitudes of the fund flows are still insufficient to fully arbitrage away all the outperformances according to the logic proposed by Berk and Green (2004).

In short, the contribution of this paper is both empirical and methodological. Empirically, I provide evidence to show the existence of stock-picking skills in the mutual fund industry and document the characteristics of the outperforming funds. On the other hand, the methodology developed in this paper is not limited to mutual funds. The new performance evaluation strategy can be applied to any investor who aspires to generate profits by actively picking stocks.

The remaining of this article is organized as the following. Section 2 offers a review of the related literature. Section 3 provides details about the bootstrap exercise to construct the counterfactual return distribution in each single period. Section 4 describes the construction of the FSD filter using the counterfactual return distribution. Section 5 includes theoretical proofs and simulation exercises to illustrate the advantageous econometric properties of the FSD condition over the positive alpha condition. Section 6 describes the data and documents the empirical findings. Section 7 concludes.

2 Related Literature

Systematic academic research on the active mutual fund industry dates back to, at least, Treynor and Mazuy (1966) and Sharpe (1966). Early empirical work such as Jensen (1968) and Malkiel (1995) establish that active mutual funds, on average, cannot outperform the market index before fees, and significantly under-perform the passive index after fees. The findings are largely consistent with the efficient market hypothesis proposed by Malkiel and Fama (1970). Despite the mediocre average performance, researchers also investigate whether historical fund performances can be used to select funds that are able to deliver superior returns in the future. Early work by Hendricks, Patel, and Zeckhauser (1993), Goetzmann and Ibbotson (1994) and Brown and Goetzmann (1995) document the "hot-hand" effect that funds with good performances in the past also tend to outperform their peers going forward. However, the classic paper by Carhart (1997) demonstrates that much of the "hot-hand" effect can be attributed to the momentum of stock prices discovered by Jegadeesh and Titman (1993), and fund performance does not seem to persist once stock price momentum is adjusted for.

Since the seminal paper by Wermers (2000), researchers started to use survivorship-bias-free holdings data to better characterize fund styles and identify different types of investment skills. The influential paper by Daniel et al. (1997) divides the universe of stocks into size, value and momentum buckets, and characterizes funds' styles by their portfolio weights in different stock buckets. Moreover, they show that there is persistent stock-picking skills, but no significant market-timing skills in their sample period. Later papers then discover that various holdings characteristics can be used to infer the skills of fund managers and predict their future performances. For example, Cohen, Coval, and Pástor (2005) find that funds that have overlapping holdings with past successful funds tend to outperform others going forward; Kacperczyk, Sialm, and Zheng (2008) find that funds with profitable unobserved actions are also likely to generate trading profits in the future; Cremers and Petajisto (2009) show that funds with more active weights outperform the ones that are suspected to be closet indices. Other notable examples include: Grinblatt and Titman (1989), Grinblatt, Titman, and Wermers (1995), Chen, Jegadeesh, and Wermers (2000), Kacperczyk, Sialm, and Zheng (2005), Alexander, Cici, and Gibson (2006), Jiang, Yao, and Yu (2007), Kacperczyk and Seru (2007), Baker et al. (2010), Da, Gao, and Jagannathan (2010), Huang, Sialm, and Zhang (2011), Kacperczyk, van Nieuwerburgh, and Veldkamp (2014), Agarwal et al. (2015), etc. This paper contributes to this line of literature by proposing a new test on the manager's information advantage regarding the idiosyncratic risks of the securities she keeps in the portfolio. In the same spirit, Iskoz and Wang (2003) also propose a methodology to test whether a money manager incorporates private information in portfolio construction by investigating the connections between fund holdings and return distributions. The difference between this paper and their work is that they consider the relation between general types of private information and future stock return distributions; whereas this paper focuses on a particular type of private information on firm-specific risks and imposes a detailed restriction on fund return distributions – the FSD condition. Also, this paper discusses the empirical implementation of the FSD condition and documents the findings when such a methodology is applied to the active mutual fund industry.

The relation between luck and skill in the context of the active mutual fund industry was first empirically investigated by Kosowski et al. (2006). Their paper proposes a bootstrap exercise in the time-series to verify the existence of fund skills ex post. Fama and French (2010) employs a similar methodology and verifies the results of Kosowski et al. (2006) in a more recent sample period. Barras, Scaillet, and Wermers (2010) classifies funds into three categories: unskilled, zero-alpha and skilled, by implementing a novel statistical procedure to account for false discoveries. This paper is closely related to these three papers in that it also aims to account for luck in fund managers' performances. However, this paper contributes to this line of research in two ways. First, existing papers can only test fund skills in the time-series with repeated observations; whereas this paper shows that a statistical test on stock-picking skills can be conducted with even only one observation by carrying out a bootstrap procedure in the cross section rather than in the time series. Second, existing papers focus on the ex post identification of funds whose realized alphas are unlikely to be explained by luck; whereas this paper aims to identify skilled fund managers with short history so that profitable trading strategies can be formed. In a contemporaneous paper, Ren (2017) also proposes the application of stochastic dominance to adjust for luck and identify skilled mutual fund managers. However, in her paper, the stochastic dominance conditions are applied in the comparison between the performances of two actual funds; whereas, in this paper, I compare the performances between an actual fund and a cohort of counterfactual funds. My methodology differs from her in two key aspects. First, by constructing counterfactual funds, my methodology is able to explicitly control for fund characteristics or factor loadings. Second, by transforming the FSD condition from the return space to the ranking space, my methodology normalizes the null distribution to a standard uniform distribution, thereby avoiding the technical difficulties in implementing the FSD condition.

My findings on the observable characteristics of the outperforming funds are related to a number of earlier findings in the literature. The finding that the outperforming funds have the same size as the industry average but are able to charge higher fees is partially consistent with Berk and Green (2004) that skilled managers are able to extract higher rents from fund investors, although the specific mechanism is different. The finding that a large portion of the alphas of the identified outperforming funds are due to their unobserved within-quarter trades is consistent with Kacperczyk, Sialm, and Zheng

(2008), where they show that return gap is indicative of the skills of a fund manager. The finding that the outperforming funds tend to keep more concentrated portfolios verifies the theoretical prediction by Van Nieuwerburgh and Veldkamp (2010) that informed investors can voluntarily choose to become under-diversified when information acquisition is endogenous. The finding that among the funds selected by the FSD filter, the ones with larger alphas also have more trading is related to Pástor, Stambaugh, and Taylor (2017), where they show that skilled managers are able to make more profits when they trade more. Finally, the finding that controlling for realized alpha, funds that satisfy the FSD condition attract more flows than others echoes the work by Barber, Huang, and Odean (2016) and Berk and van Binsbergen (2016), where they argue that fund flows reflect investors' evaluations on fund managers' skills.

3 Benchmark Extension

Under current performance evaluation methodologies, the return of a fund $r_{i,t}$ is compared to a benchmark return $r^b_{i,t}$ in every period. The single-period fund outperformance is then computed as the difference between these two returns, $r_{i,t} - r^b_{i,t}$. This approach provides a point estimate of the fund's outperformance in this single period, yet offers no information about its statistical significance. The extension from the benchmark return $r^b_{i,t}$ to the counterfactual return distribution $\langle \hat{r}_{i,t} \rangle$ allows the econometrician to obtain both the point estimate and the statistical significance of the single-period fund outperformance due to stock-picking by comparing $r_{i,t}$ to $\langle \hat{r}_{i,t} \rangle$. The additional distributional information can then be used to implement the FSD condition as elaborated in Section 4.

The construction of the counterfactual return distribution is a bootstrap exercise that mimics the fund's portfolio by investing in stocks of similar characteristics with the same portfolio weights meanwhile randomizes the specific stock choices. Specifically, the procedure can be summarized with the following 4 steps:

1. Retrieve the most recent portfolio of the fund that is available.

2. Create a replica portfolio by replacing each stock in the original portfolio with a new stock[8] of similar characteristics that is randomly chosen, meanwhile keeping the portfolio weights unchanged.

3. Compute the hypothetical return of the replica portfolio by taking the inner product between the portfolio weights and the returns of the replaced stocks.

4. Repeat Step 2 and Step 3 to generate a distribution of counterfactual portfolio returns.

[8]The new stock can be the same as the original stock.

In order to find stocks of similar characteristics with a given stock as required in Step 2, I follow and extend the approach proposed by Daniel et al. (1997) and Wermers (2003). For US stocks that are traded on AMEX, NYSE and Nasdaq, I first sort them into 5 size buckets by their market capitalization.[9] Within each size bucket, I further divide the stocks into 5 value buckets by their book-to-market ratio. Then I repeat the same procedure and divide the stocks within each value bucket into 5 momentum buckets by their preceding one-year return. Lastly, I divide the stocks in each momentum bucket further into 5 volatility buckets by their return volatility. The procedure thus categorizes all stocks into $5 \times 5 \times 5 \times 5 = 625$ non-overlapping buckets and is repeated once in a year by the end of June. For each stock within the original portfolio, Step 2 is carried out by finding a random replacing stock within the same bucket as the original stock. The weights of the replica portfolio are kept unchanged as the original portfolio. Table 2 offers an example to illustrate the bootstrap procedure. Panel (a) is a snapshot of the portfolio of Longleaf Partners Fund by the end of 2012/12. Panel (b) is a simulated replica portfolio.

In Step 3, the hypothetical return of fund i's replica in period t is computed as:

$$\hat{r}_{i,t} = \sum_{j} w_{i,j,t-1} \cdot \tilde{r}_{\hat{j},t}$$

where $w_{i,j,t-1}$ denotes the portfolio weight of stock j within the portfolio of fund i by the end of period $t-1$; $\hat{j}$ denotes the random replacement of stock j in the replica portfolio; $\tilde{r}_{\hat{j},t}$ denotes the return of the replacing stock in period t.

A few comments are in order regarding the bootstrap procedure. First, by controlling for stock characteristics in the replica portfolios, I ignore the fact that the choice of stock characteristics might also reflect the fund manager's skills. In other words,

[9] The breakpoints of the size buckets are defined by NYSE stocks only.

by comparing the real fund with replica funds of similar factor loadings and degree of diversification, the bootstrap exercise only measures the stock-picking skills of the manager and is silent about potential factor-timing skills[10]. Therefore, my methodology only focuses on the search of fund managers with a particular set of skills and is not a general diagnosis on all potential skills that a manager might have. Secondly, I extend the Daniel et al. (1997) stock classification to include volatility as an additional dimension. The matching of volatility serves two purposes. On the one hand, recent literature has documented that stock-level volatility might represent systematic risks that are priced in the cross section of stocks.[11] On the other hand, I match the stock-level idiosyncratic volatility in the replica funds to the real fund so that the portfolio-level idiosyncratic volatility of the replica funds would also be comparable to that of the real fund. The purpose of the matching of idiosyncratic volatility will be further discussed in Section 4. Thirdly, the holdings information of the real fund is only employed to extract the weight distribution and stock characteristics of the fund's investment. The specific choices of stocks in the real portfolio are not used. Therefore, even though the holdings information is only empirically available at quarterly frequency, the counterfactual return distribution can be constructed at higher frequencies, such as monthly or daily frequencies, by interpolating portfolio characteristics. Finally, the extension from a single benchmark return to the counterfactual return distribution extracts additional information about the distribution from the data so that a statistical test on stock-picking skills can be formed in every period. Outperformance computed under current evaluation methodologies as the difference between a fund's return and the benchmark return can be regarded as a point estimator of the manager's skills in a single period; whereas the comparison between the fund's return and the counterfactual return distri-

[10]See Grinblatt and Titman (1989) and Daniel et al. (1997) for the definition of the two types of investing skills

[11]See, for example, Ang et al. (2009), Fu (2009), etc.

bution provides both the point estimate and the statistical significance of the manager's stock-picking skills in each period.

4 FSD Implementation

This section demonstrates how to implement an FSD filter with the time-series of the counterfactual return distributions $\{\langle \hat{r}_{i,t} \rangle\}_{t=1}^{T}$. A test statistic for the FSD condition is proposed, and its finite-sample distribution is computed with bootstrap simulation.

4.1 FSD Motivation

I provide heuristic arguments here to demonstrate the type of investment skills that the FSD condition ($r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$) is able to identify. I defer formal proofs to Appendix A.

Consider a frictionless financial market with a factor structure. There are J factors (denoted as $\{F_{j,t}\}_{j=1}^{J}$) that are observable to both fund managers and the econometrician; and L factors (denoted as $\{f_{l,t}\}_{l=1}^{L}$) that are only observable to fund managers but not to the econometrician. There are K stocks traded in the market. The excess return (relative to the risk-free rate) of any stock within this market can be decomposed into three parts: the exposure to the J observable factors, the exposure to the L unobservable factors, and the idiosyncratic component:

$$\tilde{r}_{k,t} = r_f + \sum_j \beta_{k,j} F_{j,t} + \sum_l \gamma_{k,l} f_{l,t} + \epsilon_{k,t}$$

where $\tilde{r}_{k,t}$ denotes the return of stock k at time t; $F_{j,t}$ ($f_{l,t}$) is the realization of the observable (unobservable) factor j (l) at time t; $\beta_{k,j}$ ($\gamma_{k,l}$) denotes of the loading of stock k on factor j (l); $\epsilon_{k,t}$ is the idiosyncratic shock in stock k's return. I assume that the random variables on the right-hand side of this equation are independent to each other.

The return of fund i at time t is:

$$
\begin{aligned}
r_{i,t} &= \sum_k w_{i,k,t-1}\tilde{r}_{k,t} \\
&= \sum_k w_{i,k,t-1}\left(r_f + \sum_j \beta_{k,j}F_{j,t} + \sum_l \gamma_{k.l}f_{l,t} + \epsilon_{k,t}\right) \\
&= r_f + \sum_j \left(\sum_k w_{i,k,t-1}\beta_{k,j}\right)F_{j,t} + \sum_l \left(\sum_k w_{i,k,t-1}\gamma_{k,l}\right)f_{l,t} + \sum_k w_{i,k,t-1}\epsilon_{k,t} \\
&\equiv r_f + \sum_j \beta_{i,j,t}F_{j,t} + \sum_l \gamma_{i,l,t}f_{l,t} + \sum_k w_{i,k,t-1}\epsilon_{k,t}
\end{aligned}
$$

For each real fund in each period, replica portfolios are constructed according to the procedure in Section 3. During the construction, each stock within the real fund's portfolio is replaced randomly with another stock in the same bucket. The buckets are defined by the econometrician according to the J observable factors. Therefore, the return of a replica portfolio of fund i is:

$$
\begin{aligned}
\hat{r}_{i,t} &= \sum_k w_{i,k,t-1}\tilde{r}_{\hat{k},t} \\
&= \sum_k w_{i,k,t-1}\left(r_f + \sum_j \beta_{\hat{k},j}F_{j,t} + \sum_l \gamma_{\hat{k},l}f_{l,t} + \epsilon_{\hat{k},t}\right) \\
&= r_f + \sum_j \left(\sum_k w_{i,k,t-1}\beta_{\hat{k},j}\right)F_{j,t} + \sum_l \left(\sum_k w_{i,k,t-1}\gamma_{\hat{k},l}\right)f_{l,t} + \sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} \\
&\equiv r_f + \sum_j \hat{\beta}_{i,j,t}F_{j,t} + \sum_l \hat{\gamma}_{i,l,t}f_{l,t} + \sum_k w_{i,k,t-1}\epsilon_{\hat{k},t}
\end{aligned}
$$

I assume that the buckets are fine enough so that any two stocks in the same bucket have the same loadings on the observable factors, i.e. $\forall k, j,\ \beta_{k,j} = \beta_{\hat{k},j}$. Since the replica fund adopts the same weights as the original fund, it is easy to see that $\beta_{i,j,t} = \hat{\beta}_{i,j,t}$.

Appendix A provides formal arguments to show that the return of fund i first-order stochastically dominates its replica ($r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$) if the following three conditions are satisfied:

1. The manager of fund i has superior information about firm-specific risks so that

$$\mathbb{E}\left(\sum_k w_{i,k,t-1}\epsilon_{k,t}\right) > \mathbb{E}\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t}\right) = 0.$$

2. Fund i is sufficiently diversified so that the idiosyncratic component in fund return $\sum_k w_{i,k,t-1}\epsilon_{k,t}$ asymptotically follows a normal distribution.

3. Fund i is not biased towards unobservable factors so that $\forall l,\ \gamma_{i,l,t} = \hat{\gamma}_{i,l,t}$.

Intuitively, the FSD condition ($r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$) identifies fund managers whose factor exposures are well-captured by the loadings on observable factors, and who are able to generate profits by making bets on firm-specific risks in a well-diversified portfolio.

4.2 Ranking FSD

To empirically implement the FSD condition, the key step is to transform the return FSD condition to a ranking FSD condition.

Proposition 1. *A real fund's return being first-order stochastically dominant to a replica fund's return is equivalent to the condition that the ranking of the real fund's return among the cohort of replica funds being first-order stochastically dominant to the ranking of the replica fund's return among the cohort of replica funds. Moreover, the ranking of a replica fund's return among the cohort of replica funds follows a standard uniform distribution. That is*

$$r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t} \iff Pct\left(r_{i,t}, \langle\hat{r}_{i,t}\rangle\right) \overset{fsd}{\succ} Pct\left(\hat{r}_{i,t}, \langle\hat{r}_{i,t}\rangle\right) \sim Unif\left(0,1\right)$$

where $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ *(*$Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$*) denotes the percentile of the real (replica) fund's return in the companion return distribution;* $Unif\left(0,1\right)$ *denotes the uniform distribution with support* $[0,1]$*.*

Proof. Since all replica funds are constructed randomly in the bootstrap procedure, $Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)\sim Unif\left(0,1\right)$ is obvious.

The equivalence condition is immediate from $Pct\left(\cdot,\langle\hat{r}_{i,t}\rangle\right)$ being monotonically increasing.

Denote $F_{t-1}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)(F_{t-1}^{Pct(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle)}(x))$ as the conditional CDF of the ranking $r_{i,t}(\hat{r}_{i,t})$; $\langle\hat{r}_{i,t}\rangle[x]$ as the value at the x percentile of $\langle\hat{r}_{i,t}\rangle$.

$$
\begin{aligned}
F_{t-1}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x) &\equiv Prob_{t-1}\left(Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)\leq x\right)\\
&= Prob_{t-1}\left(r_{i,t}\leq\langle\hat{r}_{i,t}\rangle[x]\right)\\
&\equiv F_{t-1}^{r_{i,t}}\left(\langle\hat{r}_{i,t}\rangle[x]\right)\\
&< F_{t-1}^{\hat{r}_{i,t}}\left(\langle\hat{r}_{i,t}\rangle[x]\right)\\
&= Prob_{t-1}\left(\hat{r}_{i,t}\leq\langle\hat{r}_{i,t}\rangle[x]\right)\\
&= Prob_{t-1}\left(Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)\leq x\right)\\
&\equiv F_{t-1}^{Pct(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)
\end{aligned}
$$

□

Proposition 2. $r_{i,t}\overset{fsd}{\succ}\hat{r}_{i,t}\iff F_{t-1}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)<F_{t-1}^{Pct(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)=x$*, where* F_{t-1} *denotes the conditional CDF.*

Proof. Immediate from Proposition 1. □

Figure 2 offers a graphical illustration of the FSD condition. Panel (a) is a demonstration regarding the relation between $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ and $Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$. The dashed line plots the PDF of $Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$, which is a flat horizontal line constant at 1 since $Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ follows a standard uniform distribution. The solid line is an example of $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$. Since $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)\overset{fsd}{\succ}Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$, the solid line has a smaller left tail compared to the dashed line, but a larger right tail in the PDF plot. Panel (b) offers the illustration on the same relation with CDF plots. The $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)\overset{fsd}{\succ}Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ condition is reflected in the plot as the solid curve strictly lies below the dashed line.

Proposition 2 establishes the relation between $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ and $Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ conditional on the information set by the end of $t-1$. Conditional relations between distributions are not empirically observable. However, Proposition 2 also indicates that the conditional CDF of $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ always lies below the 45 degree line, which is an appealing feature to facilitate time aggregation.

Proposition 3. $r_{i,t}\overset{fsd}{\succ}\hat{r}_{i,t},\ \forall t\Rightarrow F^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)<F^{Pct(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)=x$, *where F denotes the unconditional CDF.*

Proof. $F^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)=\mathbb{E}\left[F_{t-1}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)\right]<\mathbb{E}\left[F_{t-1}^{Pct(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)\right]=\mathbb{E}(x)=x.$ □

Note that the unconditional CDF of the ranking of a manager among replica funds is an empirically observable object. I therefore, employ the empirical counterpart of $F^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$, i.e. $\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$, to implement the FSD condition.

4.3 FSD Test Statistic

Figure 3 illustrates the construction of the FSD test statistic. Specifically, for each fund, I construct its empirical ranking CDF $\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$ with its historical returns.[12] The FSD test statistic $\hat{\theta}_i\in[0,1]$ is then defined as the measure of the region where

[12] I use the 24 proceeding monthly returns in the empirical exercise.

$\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$ lies below the 45 degree line. The FSD condition is perfectly satisfied in sample if $\hat{\theta}_i = 1$, and a higher $\hat{\theta}_i$ indicates a better fit of the FSD condition.

Figure 4 demonstrates the simulated finite-sample distribution of the FSD test statistic $\hat{\theta}$ constructed from 24 observations under the null hypothesis that $Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle) \sim Unif(0,1)$. The null distribution of $\hat{\theta}$ seems to follow a standard uniform distribution itself. According to the simulation, a test size of 10% (5%) corresponds to the critical value of 0.90 (0.95). Asymptotically, $\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x) - x$ follows a Brownian bridge under the null hypothesis that $Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle) \sim Unif(0,1)$. And it can be verified that the FSD test statistic $\hat{\theta}$ indeed follows a standard uniform distribution asymptotically.[13]

In general, the test statistic $\hat{\theta}$ can only be used to evaluate of the goodness of fit of the FSD condition, but it is unable to measure the magnitude of the fund's true α. Therefore, in practice, the FSD filter is better used in combination with the standard $\hat{\alpha}$ sort as demonstrated in Section 6. The FSD filter serves to rule out potential false positives; whereas the $\hat{\alpha}$ sort measures the magnitude of the potential stock-picking skills.

[13]Alternatively, the FSD test statistic can be constructed as $\hat{\theta}_i \equiv \max_{x\in[0,1]}\left(\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x) - x\right)$, which is a variant of the Kolmogorov–Smirnov (KS) statistic ($\max_{x\in[0,1]}\left|\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x) - x\right|$). This alternative test statistic does not have a tractable asymptotic null distribution. Empirically, the alternative test statistic generates similar findings as the test statistic defined in this section. The results are discussed in Appendix B.

5 Finite Sample Robustness

This section describes two specific mechanisms through which the FSD condition is able to improve the power of the conventional positive alpha condition. The two mechanisms correspond to two statistical problems that might be present in the data: heteroscedasticity and benchmark mis-specification. The heteroscedasticity problem is defined as the situation where idiosyncratic volatility is time-varying and more volatile than the fund's true alpha; whereas the benchmark mis-specification problem is defined as some managers taking on factors that are unobservable to the econometrician. The FSD condition possesses superior econometric properties compared to the positive alpha condition because:

1. Heteroscedasticity:

   (a) Positive alpha condition: Assigns equal weight to all observations regardless of the level of idiosyncratic volatility.

   (b) FSD condition: Weights observations differently according to the signal-to-noise ratio in different periods.

2. Benchmark Mis-specification:

   (a) Positive alpha condition: Tends to mistakenly select mis-specified managers who take on unobservable factors with high in-sample realizations as being skilled.

   (b) FSD condition: Offers a detection mechanism to rule out mis-specified managers by checking the left tails of their return distributions.

I provide both theoretical arguments and simulation results to illustrate the superiority of the FSD condition to the positive alpha condition in these aspects.

5.1 Simulation Environment

In order to illustrate the arguments, I construct an artificial economy with the following specifications. I simulate 1000 fund managers, among whom 20 are skilled with α being 25 bps per month. The percentage of skilled managers and the magnitude of their α is determined according to the findings documented in Fama and French (2010). Except for the case studying benchmark mis-specification, I assume a one-factor structure, e.g. the market factor $r_{m,t}$. Without loss of generality, I assume that all funds have unit loading on the single factor, and the risk-free rate is zero. I also assume that the single factor follows normal distribution: $r_{m,t} \sim N(0, 0.06^2)$, i.e. the factor has 0 mean and 6% monthly ($6\% \times \sqrt{12} = 21\%$ annualized) volatility. Therefore, the return of an unskilled manager is

$$r_{i,t}^{unskilled} = r_{m,t} + u_{i,t}, \quad \mathbb{E}(u_{i,t}) = 0,\ r_{m,t} \sim N\left(0, 0.06^2\right)$$

and the return of a skilled manager is

$$r_{i,t}^{skilled} = 25bps + r_{m,t} + u_{i,t}, \quad \mathbb{E}(u_{i,t}) = 0,\ r_{m,t} \sim N\left(0, 0.06^2\right).$$

The performances of the replica funds in the construction of the FSD test statistic have the same properties as the unskilled managers. That is

$$\hat{r}_{i,t} = r_{m,t} + \hat{u}_{i,t}, \quad \hat{u}_{i,t} \sim u_{i,t},\ r_{m,t} \sim N\left(0, 0.06^2\right)$$

where $\hat{u}_{i,t}$ and $u_{i,t}$ have the same probability distribution but a mutually independent.

I then select 20 best performing managers in the simulated data using $\hat{\alpha}$ and $\hat{\theta}$, i.e. the test statistics of the positive alpha condition and the FSD condition, respectively; and compare the accuracy of these two filters in identifying skilled fund managers.

5.2 Robustness to Heteroscedasticity

In order to study the influence of return heteroscedasticity on the performance measures, I assume that the volatility of idiosyncratic returns follows the following (almost) AR1 process:

$$\sigma_{i,t} = \max\left\{\sigma_{i,t-1} + \rho\left(\bar{\sigma} - \sigma_{i,t-1}\right) + \zeta\epsilon_t,\ 0\right\}.$$

where ρ determines the speed of mean-reversion of the volatility process, $\bar{\sigma}$ is the long-run volatility, ζ is the volatility of the volatility process, and ϵ_t is a standard normal shock, i.e. $\epsilon_t \sim N(0,1)$.

For simplicity, I assume that the idiosyncratic return components are also normal:

$$u_{i,t} \sim \hat{u}_{i,t} \sim N\left(0, \sigma_{i,t}^2\right).$$

In the simulations, I specify: $\rho = 0.9^*$, $\bar{\sigma} = 0.01^*$; and I consider three different degrees of heteroscedasticity: $\zeta_L = 0.0012$, $\zeta_M = 0.0018$ and $\zeta_H = 0.0024^*$.[14] The return processes for the managers in this heteroscedastic economy are thus fully specified.

In this economy, the positive alpha filter should have poor performance when heteroscedasticity is high because it assigns equal weight to all observations so that lucky shocks from high volatility periods are of large magnitudes and are difficult to be cancelled out by shocks in other periods.

The FSD filter, on the other hand, naturally does not suffer from this problem because idiosyncratic volatility is adjusted period by period when the ranking of the real fund return among replica funds $\left(Pct\left(r_{i,t}, \langle\hat{r}_{i,t}\rangle\right)\right)$ is taken. Conceptually, this adjustment mechanism is similar to a weighted least square (WLS) regression. But the WLS regression requires the specification of the volatility process of the residual, whereas the FSD filter accounts for heteroscedasticity via bootstrap simulations and

[14]Values with the "*" sign are calibrated with the real data.

does not make structural assumptions about the volatility process.

Figure 5 compares the effectiveness of the positive alpha filter versus the FSD filter in the heteroscedastic economy. The x-axis is the formation period from which the performance measures are constructed. The y-axis is the average number of skilled managers that the corresponding filter is able to identify over 500 simulation paths. The black solid (blue dotted, red dashed) line represents the level of heteroscedasticity with $\zeta = 0.0012$ ($\zeta = 0.0018$, $\zeta = 0.0024^{*}$). Panel (a) plots the effectiveness of the positive alpha filter; whereas Panel (b) plots the effectiveness of the FSD filter.

From the figure, the real-world level of heteroscedasticity is mild enough so that the positive alpha filter is virtually unaffected. Interestingly, Panel (b) of the figure shows that the effectiveness of the FSD filter improves with heteroscedasticity. The intuition of this result is that idiosyncratic volatility gets adjusted in each period. High volatility periods are assigned with low weights and low volatility periods are assigned with high weights. Therefore, the FSD filter improves with heteroscedasticity because it is able to take advantage of the high signal-to-noise ratio in periods with low idiosyncratic volatility.

5.3 Benchmark Mis-specification Detection

Another difficult problem encountered by the positive alpha condition is that the passive index that is used to benchmark a fund's performance might be inappropriately chosen and does not fully account for the fund's exposure to systematic factors so that the outperformance in each period as well as the overall α might be measured with error.

To fix ideas, I modified the aforementioned simulation environment to introduce a third type of fund managers – the mis-specified managers. The return processes of the three types of managers are as follows:

The unskilled managers:

$$r_{i,t}^{unskilled} = r_{m,t} + u_{i,t}, \quad u_{i,t} \sim N\left(0, 0.01^2\right), \; r_{m,t} \sim N\left(0, 0.06^2\right).$$

The skilled managers:

$$r_{i,t}^{skilled} = 25bps + r_{m,t} + u_{i,t}, \quad u_{i,t} \sim N\left(0, 0.01^2\right), \; r_{m,t} \sim N\left(0, 0.06^2\right).$$

The mis-specified managers:

$$r_{i,t}^{mis-spec} = r_{m,t} + f_{i,t} + u_{i,t}, \quad u_{i,t} \sim N\left(0, 0.01^2\right), \; r_{m,t} \sim N\left(0, 0.06^2\right), \; f_{i,t} \sim N\left(0, \sigma_f^2\right).$$

The returns of the replica funds have the same properties for all three types of managers. That is

$$\hat{r}_{i,t} = r_{m,t} + \hat{u}_{i,t}, \quad \hat{u}_{i,t} \sim u_{i,t}, \; r_{m,t} \sim N\left(0, 0.06^2\right).$$

Notice that for the mis-specified managers, the replica funds have the same exposure to the observable factor $r_{m,t}$, but do not load on the unobservable factor $f_{i,t}$.

The mis-specified managers have no stock-picking skills so that they generate zero α. The difference between the mis-specified and the unskilled managers is that a mis-specified manager takes on factor risk $f_{i,t}$ that is not observable to the econometrician. Thus, $f_{i,t}$ is not controlled for in the replica funds. The volatility of the uncontrolled factor σ_f is a measure of the severity of mis-specification in this economy.

I assume idiosyncratic volatility to be constant in this case so that $u_{i,t} \sim \hat{u}_{i,t} \sim N\left(0, 0.01^2\right)$. I consider three levels of uncontrolled factor volatility with $\sigma_f = 0.01$ ($\sigma_f = 0.03$, $\sigma_f = 0.05$) representing the case of mild (moderate, severe) mis-specification. In the most severe case, the volatility of the missing factor is comparable to the volatility of the market so that this specification is empirically plausible. For simplicity, the factor is assumed to follow normal distribution: $f_{i,t} \sim N\left(0, \sigma_f^2\right)$, and the uncontrolled factors for two managers are uncorrelated. There are 1000 managers in total in the economy. Among them, 20 managers are skilled, 100 managers are mis-specified, and the remaining 880 managers are unskilled.

The existence of the mis-specified managers might severely compromise the effectiveness of the positive alpha filter. To see that, the $\hat{\alpha}$ of a mis-specified manager is

$$\begin{aligned}\hat{\alpha}_i &= \frac{1}{T}\sum_t \left(r_{i,t} - r_{m,t}\right) \\ &= \frac{1}{T}\sum_t f_{i,t} + \frac{1}{T}\sum_t u_{i,t}.\end{aligned}$$

Thus the existence of unobservable factors might obscure the measurement of skill because $\hat{\alpha}_i$ can be dominated by the realization of $\frac{1}{T}\sum_t f_{i,t}$ especially when sample is short (T is small) or mis-specification is severe (σ_f is large).

The FSD condition is able to alleviate this problem by offering a detection mechanism. Suppose the fund has significant loading on a factor that is not controlled in the bootstrap process during the construction of the replica portfolios, then the first-order stochastic dominance condition is likely to be violated. Indeed, during periods in which the uncontrolled factor has large positive realizations, the manager would rank highly compared to the replica funds, and vice versa during periods when the uncontrolled factor has large negative realizations. As as result, the PDF of the ranking of the manager compared to the replica funds ($Pct\left(r_{i,t}, \langle\hat{r}_{i,t}\rangle\right)$) shall have both large left and right tails, violating the first-order stochastic dominance condition. Figure 6 offers a graphical illustration of the detection mechanism.

The following proposition provides a more general statement for this argument when the uncontrolled factor is allowed to have non-zero risk premium.

Proposition 4. *Consider a mis-specified manager's return process:*

$$
\begin{aligned}
r_{i,t} &= r_f + \sum_j \beta_{i,j} F_{j,t} + \sum_l \gamma_{i,l} f_{l,t} + e_{i,t} \\
&\equiv r_f + \sum_j \beta_{i,j} F_{j,t} + \tilde{f}_{i,t} + e_{i,t} \\
\tilde{f}_{i,t} &\sim N\left(\mu_f, \sigma_f^2\right), \quad e_{i,t} \sim N\left(0, \sigma_i^2\right), \quad \tilde{f}_{i,t} \perp e_{i,t}.
\end{aligned}
$$

where $\{F_{j,t}\}_{j=1}^{J}$ *are the observable factors,* $\{f_{l,t}\}_{l=1}^{L}$ *are the unobservable factors,* $e_{i,t}$ *is the idiosyncratic component.*

A corresponding replica fund has the return process:

$$
\begin{aligned}
\hat{r}_{i,t} &= r_f + \sum_j \beta_{i,j} F_{j,t} + \hat{e}_{i,t} \\
\hat{e}_{i,t} &\sim N\left(0, \sigma_i^2\right).
\end{aligned}
$$

The first-order stochastic dominance condition $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ *is violated as long as* $\sigma_f > 0$.

Proof. $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t} \iff \tilde{f}_{i,t} + e_{i,t} \overset{fsd}{\succ} \hat{e}_{i,t}$.

Since $\tilde{f}_{i,t} \perp e_{i,t}$, $\tilde{f}_{i,t} + e_{i,t} \sim N\left(\mu_f, \sigma_f^2 + \sigma_i^2\right)$.

Denote the PDF of $\tilde{f}_{i,t} + e_{i,t}$ as $\phi(x) \equiv \frac{1}{\sqrt{2\pi\left(\sigma_f^2+\sigma_i^2\right)}} \exp\left[-\frac{\left(x-\mu_f\right)^2}{2\left(\sigma_f^2+\sigma_i^2\right)}\right]$; and the PDF of $\hat{e}_{i,t}$ as $\hat{\phi}(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{x^2}{2\sigma_i^2}\right]$.

Define

$$
\begin{aligned}
L(x) &\equiv \frac{\phi(x)}{\hat{\phi}(x)} \\
&= \sqrt{\frac{\sigma_i^2}{\left(\sigma_f^2+\sigma_i^2\right)}} \exp\left[\frac{x^2}{2\sigma_i^2} - \frac{\left(x-\mu_f\right)^2}{2\left(\sigma_f^2+\sigma_i^2\right)}\right] \\
&= \sqrt{\frac{\sigma_i^2}{\left(\sigma_f^2+\sigma_i^2\right)}} \exp\left[\frac{\left(\sigma_f^2+\sigma_i^2\right)x^2 - \sigma_i^2\left(x-\mu_f\right)^2}{2\sigma_i^2\left(\sigma_f^2+\sigma_i^2\right)}\right] \\
&= \sqrt{\frac{\sigma_i^2}{\left(\sigma_f^2+\sigma_i^2\right)}} \exp\left[\frac{\sigma_f^2 x^2 + 2\sigma_i^2\mu_f x - \sigma_i^2\mu_f^2}{2\sigma_i^2\left(\sigma_f^2+\sigma_i^2\right)}\right].
\end{aligned}
$$

If $\sigma_f > 0$, then $\lim_{x\to-\infty} L(x) = +\infty$. Therefore, $\tilde{f}_{i,t} + e_{i,t}$ has a larger left tail compared to $\hat{e}_{i,t}$, and the first-order stochastic dominance condition is violated. $\square$

Of course, the proposition only holds under the special condition that both the uncontrolled factors and the idiosyncratic risk are normal, and are independent to each other. The conclusion of the proposition can be violated if one considers alternative distribution specifications of the uncontrolled factors. However, the proposition conveys the intuition that the FSD condition is able to detect the existence of the uncontrolled factors because the mis-specified managers are likely to have a larger left tail in their return distributions compared to the replica funds when they take on factors that are not controlled by the econometrician, as demonstrated in Figure 7.

Figure 8 and Figure 9 compare the effectiveness of the positive alpha filter versus the FSD filter in simulation. Figure 8 demonstrates the two filters' ability to identify skilled managers; whereas Figure 9 illustrates their tendencies to select mis-specified managers. Specifically, the x-axis is the formation period from which the performance measures are constructed. For Figure 8, the y-axis is the average number of skilled managers that the corresponding filter is able to identify over 500 simulation paths; whereas for Figure 9, the y-axis is the average number of mis-specified managers that the corresponding filter erroneously selects over 500 simulation paths. The black solid (blue dotted, red dashed) line represents the situation where benchmark mis-specification is mild $\sigma_f = 0.01$ (moderate $\sigma_f = 0.03$, severe $\sigma_f = 0.05$).

From the figures, the positive alpha filter identifies fewer skilled managers when benchmark mis-specification becomes more severe because it tends to erroneously select mis-specified managers whose uncontrolled factors have high in-sample realizations. The FSD filter, however, is unaffected by benchmark mis-specification because the FSD condition excludes the mis-specified managers as their rankings relative to replica funds would have both larger left and right tails compared to a uniform distribution.

6 Empirical Findings

6.1 Data

I obtain monthly after-fees fund returns along with other fund characteristics such as fund size, age, name, expense ratio, etc. from CRSP Survivor-Bias-Free US Mutual Fund Database. I compute the before-fees returns by adding back the expense ratio to the after-fees fund returns. I obtain fund holdings from Thomson Reuters Mutual Fund Holdings (s12), formerly known as the CDA/Spectrum Mutual Fund Holdings Database. Both databases are standard in this line of research. Their popularity arose largely due to their efforts to eliminate survivorship bias by making an attempt to include all funds that have ever existed in the US market. In fact, Linnainmaa (2013) raised the concern of a potential reverse survivorship bias by using these databases as funds hit by a series of unlucky negative shocks tend to exit the market, leaving behind trajectories of poor performances without the chances to "clear their names". Therefore, my finding of superior out-of-sample performances is unlikely to be caused potential survivorship bias. I follow the standard approach to link these two databases with the MFLINKS database constructed by Prof. Russ Wermers, and I obtain stock prices and returns from the CRSP Monthly Stock File.

I limit my focus to domestic, open-end, actively managed, US equity funds. I employ the investment objectives code (crsp_obj_cd) that has been recently introduced by CRSP as my screening variable to identify such funds.[15] Doshi, Elkamhi, and Simutin (2015) shows that the funds identified with the crsp_obj_cd are almost identical to the funds identified with the investment objectives codes from other data vendors that

[15] I include funds with crsp_obj_cd that begins with "EDC" or "EDY"; exclude funds with crsp_obj_cd being "EDYH" or "EDYS"; and exclude option income funds with Strategic Insight Objectives code being "OPI". I then eliminate index funds by screening fund names.

have been used in earlier literature.[16] To reduce the impact from very small funds, I require the funds in my sample to have at least \$10 million under management and hold at least 10 stocks in their portfolios. I aggregate funds with multiple share classes into a single class as these different share classes have the same portfolio composition. In order to have enough funds for this project, I take the sample period from January 1991 to December 2015. I have 2693 distinctive funds in my sample and 227,710 fund-month observations. Table 3 documents the summary statistics of the funds that are included in my sample.

6.2 Out-of-sample Performances

Table 4 documents the out-of-sample performances of the funds identified by the FSD condition. Specifically, by the end of each quarter, I compute the empirical CDF of the percentile of each fund in the counterfactual return distribution ($\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$). I then construct the FSD test statistic $\hat{\theta}$ for each fund. I install the FSD filter to select funds with $\hat{\theta} \geq 0.90$, which corresponds to a 10% test size. The FSD filter alone is able to identify a group of fund managers who are able to outperform the Carhart benchmark by 203 bps per year before fees (78 bps per year after fees) out of sample.

If the FSD filter is indeed able to identify skilled fund managers, then one should expect performance persistence among the FSD selected funds. To verify this conjecture, I further sort the FSD identified funds into 5 quintiles by their realized Carhart four-factor alphas during the proceeding 24 months. Table 4 shows that for those FSD satisfying funds, historical alphas do predict future performances. Specifically, the average out-of-sample alphas of the funds increase monotonically with historical realized alphas. The funds in the top quintile are able to, on average, outperform the Carhart four-factor benchmark by as much as 371 bps per year before fees (240 bps per year

[16] I thank the authors for sharing their SAS code online.

after fees). The finding of alpha persistence among the identified funds is consistent with the arguments that the FSD condition is able to identify a group of fund managers who are potentially skilled at stock-picking.

Figure 10 plots the time-series of the before-fees performances of the selected mutual funds. Panel (a) plots the time-series of the fund performances for all funds selected by the FSD filter; whereas Panel (b) plots the top quintile of the funds with the highest historical alpha within the funds selected by the FSD filter. The blue dashed (red solid) line is the cumulative before-fees return of the selected mutual funds (the market). Figure 11 plots the time-series of the before-fees outperformances of the selected mutual funds. The outperformance is defined as the cumulative return of a trading strategy that longs the portfolio of identified funds and shorts the market. From the figure, fund outperformances seem to be most pronounced during the dot-com bubble periods, but the outperformances are in general consistent over the sample period.

Figure 12 plots the histograms of the before-fees excess returns of the selected mutual funds. Panel (a) plots the histogram of the returns in excess of the Carhart four-factor benchmark for all funds selected by the FSD filter; whereas Panel (b) plots the histogram of the returns in excess of the Carhart four-factor benchmark for the top quintile of the funds with the highest historical alpha within the funds selected by the FSD filter. From the figure, it is obvious that the identified funds are more likely to realize positive excess returns than negative excess returns compared to the Carhart four-factor benchmark. And also, the excess return distributions feature heavier right tails than left tails, which is consistent with the requirement of the FSD condition.

6.3 Comparison to Alternative Performance Measures

How does the FSD filter compare to other performance measures? Table 5 and Table 6 compare the out-of-sample performance of the FSD filter to two conventional performance measures: alpha and information ratio (IR). To make the results comparable, I repeat the search exercise in the previous subsection but replace the FSD test statistic with the alternative performance measures. Specifically, I use the alternative performance measures (alpha or IR) to select the top 10% of the funds in the first stage. Then among these funds, I further sort them into 5 quintiles based on their in-sample Carhart four-factor alphas.

Consistent with Carhart (1997), Table 5 shows that the in-sample alpha measure is unable to select funds that are able to significantly outperform the Carhart four-factor benchmark out of sample. The first-stage funds are only able to outperform the Carhart four-factor benchmark by 113 bps before fee, which is not statistically significant. After fees, these funds underperform the Carhart four-factor benchmark by 16 bps. Moreover, there is also lack of performance persistence among the first-stage funds. The out-of-sample alphas of the fund portfolios in the second stage are not monotonic with respect to in-sample alphas.

Table 6 shows that the search based on information ratio works better than alpha. The funds in the first stage are able to generate an out-of-sample four-factor alpha of 169 bps with a t-statistic of 2.81. After fee, the out-of-sample alpha drops to 49 bps and is no long statistically significant. In the second stage, the high quintile funds also perform better than the search based on alpha. However, the out-of-sample alphas are still non-monotonic with respect to in-sample alphas.

In sum, Table 5 and Table 6 show that the FSD filter offers better out-of-sample performance than the search based on in-sample alpha or information ratio.

6.4 Fund Characteristics

Table 7 compares the observable characteristics of the identified funds with the cross-sectional average of all funds in the sample. From the table, the funds identified by the

FSD filter do not differ in size compared to an average fund in the industry. However, the funds in higher second-stage quintiles charge more fees. The finding is partially consistent with Berk and Green (2004) in the sense that more skilled managers are able to extract higher rents in equilibrium, although the specific mechanism is different. In Berk and Green (2004), skilled managers receive compensation by growing the size of their funds, leaving the fees unchanged. My finding, on the other hand, suggests that they demand higher fees directly. Funds in higher second-stage quintiles also tend to keep fewer stocks in their portfolios, thereby being more concentrated. This finding is consistent with the theory proposed by Van Nieuwerburgh and Veldkamp (2010) that informed investors can choose to be specialized when information acquisition is costly. The finding that funds in higher quintiles also trade more is related to the finding by Pástor, Stambaugh, and Taylor (2017) that trades by active mutual fund managers tend to be profitable.

6.5 The Return Gap

One potential concern regarding the outperformances of the identified funds is that instead of possessing stock-picking skills, those funds might be loading on momentum factors that the Carhart benchmark does not perfectly control for. In order to rule out such possibility, I study the return gaps of the identified funds. The return gap is defined as the difference between a fund's actual return from the hypothetical return that the fund might have earned by keeping the portfolio weights at the beginning of the quarter unchanged throughout the entire quarter:

$$rgap_{i,t} \equiv r_{i,t} - \sum_j w_{i,j,\underline{t}} \tilde{r}_{j,t}$$

where $w_{i,j,\underline{t}}$ denotes the portfolio weight of stock j of fund i at the most recent quarter-end of month t; $\tilde{r}_{j,t}$ denotes the return of stock j during month t.

The return gap measures the profitability of the unobserved within-quarter actions conducted by a fund manager. I regress the return gaps of the identified funds against the Carhart four-factor benchmark:

$$rgap_{i,t} = \alpha_i^{rgap} + \beta_{m,i}^{rgap}(r_{m,t} - r_f) + \beta_{smb,i}^{rgap} smb_t + \beta_{smb,i}^{rgap} smb_t + \beta_{hml,i}^{rgap} hml_t + \beta_{umd,i}^{rgap} umd_t + \epsilon_{i,t}^{rgap}.$$

The results are documented in Table 8. The table shows that the out-of-sample alphas resulting from the return gaps also increase monotonically with historical alphas for funds selected by the FSD filter. Moreover, the return gaps account for about half of the total out-of-sample alphas for all quintiles. The finding is consistent with the results of Kacperczyk, Sialm, and Zheng (2008) that the return gap is indicative of manager skills. The profitability of the return gap offers strong support that the identified managers are skilled because they are able to make profitable within-quarter trades. It rules out the concern that the outperformances of the identified managers are entirely driven by their loadings on some uncontrolled momentum factors.

6.6 The Mimicking Strategy

The return gap analysis suggests that the identified fund managers are able to generate profits from their unobserved within-quarter actions. Therefore, it should also suggest that it would be difficult for an out-sider to free-ride on those managers stock-picking endeavors. Indeed, Table 9 documents the performance of the trading strategy that aims to mimic the performances of the selected outperforming mutual fund managers. To ensure implementability, by the end of each quarter, the stock holdings from the end of the previous quarter are retrieved for the managers who have been identified by the FSD filter. The trading strategy then invests in the stocks that the managers were holding as of the end of the previous quarter. The portfolios are rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics.

Consistent with the analysis on the return gap, the mimicking strategy loses about half of the profitability compare to the managers' total before-fees performances. Interestingly, the profitability of the mimicking strategy is comparable to that of the after-fees returns that investors are able to earn by investing in the funds. The finding suggests that the fees of the identified funds might be set rationally in equilibrium. This result is consistent with the findings documented by Frank et al. (2004) for a limited sample of high-expense funds.

6.7 Fund Flow Responses

According to Berk and van Binsbergen (2015) and Barber, Huang, and Odean (2016), fund flows contain information about fund investors' evaluations of managers' investment skills. In order to understand whether fund investors infer managers' skills using signals correlated with the FSD condition, I run the following Fama-Macbeth regression:

$$Flow_{i,t} = Const + \delta_{\mathbf{0}} \times FSD_{i,t-1} + (\beta + \delta_{\mathbf{1}} \times FSD_{i,t-1}) \times \hat{\alpha}_i^{[t-1-T,t-1]} + X_i + \epsilon_{i,t},$$

where X_i represents control variables including fund age, log fund size, fees, and the number of stocks in the portfolio; $\hat{\alpha}_i^{[t-1-T,t-1]}$ is the trailing in-sample realized Carhart four-factor alpha; $FSD_{i,t-1}$ is a dummy variable that equals to one for funds with a FSD test statistic $\hat{\theta}$ higher than 0.90, and zero otherwise.

The parameters of interest are δ_0 and δ_1. A positive δ_0 indicates that controlling for realized in-sample alphas, the funds that satisfy the FSD condition attract more flows than other funds. A positive δ_1 indicates that flows are more sensitive to realized in-sample alphas for funds that satisfy the FSD condition.

Table 10 presents the regression results. Both δ_0 and δ_1 are highly significantly positive. The finding suggests that fund investors reveal preferences towards fund return distribution properties beyond the first moment (mean/alpha). They appreciate funds with return distributions satisfying the FSD condition more than other funds controlling for realized alpha. On the other hand, the still positive out-of-sample alphas of the FSD identified funds suggest the presence of certain informational frictions in the fund market so that the fund outperformances are not fully arbitraged away as in the Berk and Green (2004) equilibrium.

7 Conclusion

Due to the strong influence of luck in fund managers' performances, the search for skilled managers with predictable outperformances is a challenging task. Existing alpha based evaluation methods have poor out-of-sample performances because alpha is related to the mean of the return distribution and mean is difficult to estimate in short samples. I show that, by limiting the search scope to a specific subset of skilled managers – the skilled stock-pickers, a new first-order stochastic dominance condition can be imposed to improve the effectiveness of the search. The new FSD filter complements the conventional $\hat{\alpha}$ sort because it is robust to finite-sample problems such as heteroscedasticity and benchmark mis-specification. The empirical part of this paper demonstrates the superior performance of the combination of the new FSD filter and the standard $\hat{\alpha}$ sort in identifying outperforming stock-pickers. My findings confirm various theoretical and empirical results discussed earlier in the literature and are also able to shed new light on our understanding about the active mutual fund industry.

References

Agarwal, Vikas, Kevin A Mullally, Yuehua Tang, and Baozhong Yang. 2015. "Mandatory portfolio disclosure, stock liquidity, and mutual fund performance." *The Journal of Finance* 70 (6):2733–2776.

Alexander, Gordon J, Gjergji Cici, and Scott Gibson. 2006. "Does motivation matter when assessing trade performance? An analysis of mutual funds." *The Review of Financial Studies* 20 (1):125–150.

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2009. "High idiosyncratic volatility and low returns: International and further US evidence." *Journal of Financial Economics* 91 (1):1–23.

Baker, Malcolm, Lubomir Litov, Jessica A Wachter, and Jeffrey Wurgler. 2010. "Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements." *Journal of Financial and Quantitative Analysis* 45 (5):1111–1131.

Barber, Brad M, Xing Huang, and Terrance Odean. 2016. "Which factors matter to investors? Evidence from mutual fund flows." *The Review of Financial Studies* 29 (10):2600–2642.

Barras, Laurent, Olivier Scaillet, and Russ Wermers. 2010. "False discoveries in mutual fund performance: Measuring luck in estimated alphas." *The journal of finance* 65 (1):179–216.

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

Berk, Jonathan B and Jules H van Binsbergen. 2015. "Measuring skill in the mutual fund industry." *Journal of Financial Economics* 118 (1):1–20.

———. 2016. “Assessing asset pricing models using revealed preference.” *Journal of Financial Economics* 119 (1):1–23.

Brown, Stephen J and William N Goetzmann. 1995. “Performance persistence.” *The Journal of finance* 50 (2):679–698.

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

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers. 2000. “The value of active mutual fund management: An examination of the stockholdings and trades of fund managers.” *Journal of Financial and quantitative Analysis* 35 (3):343–368.

Chevalier, Judith and Glenn Ellison. 1997. “Risk taking by mutual funds as a response to incentives.” *Journal of Political Economy* 105 (6):1167–1200.

Cohen, Randolph B, Joshua D Coval, and L’uboš Pástor. 2005. “Judging fund managers by the company they keep.” *The Journal of Finance* 60 (3):1057–1096.

Cremers, KJ Martijn and Antti Petajisto. 2009. “How active is your fund manager? A new measure that predicts performance.” *Review of Financial Studies* 22 (9):3329–3365.

Da, Zhi, Pengjie Gao, and Ravi Jagannathan. 2010. “Impatient trading, liquidity provision, and stock selection by mutual funds.” *The Review of Financial Studies* 24 (3):675–720.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers. 1997. “Measuring mutual fund performance with characteristic-based benchmarks.” *The Journal of finance* 52 (3):1035–1058.

Daniel, Kent and Sheridan Titman. 1997. "Evidence on the characteristics of cross sectional variation in stock returns." *the Journal of Finance* 52 (1):1–33.

Doshi, Hitesh, Redouane Elkamhi, and Mikhail Simutin. 2015. "Managerial activeness and mutual fund performance." *Review of Asset Pricing Studies* 5 (2):156–184.

Fama, Eugene F and Kenneth R French. 2010. "Luck versus skill in the cross-section of mutual fund returns." *The journal of finance* 65 (5):1915–1947.

Frank, Mary Margaret, James M Poterba, Douglas A Shackelford, and John B Shoven. 2004. "Copycat funds: Information disclosure regulation and the returns to active management in the mutual fund industry." *The Journal of Law and Economics* 47 (2):515–541.

Fu, Fangjian. 2009. "Idiosyncratic risk and the cross-section of expected stock returns." *Journal of Financial Economics* 91 (1):24–37.

Goetzmann, William N and Roger G Ibbotson. 1994. "Do winners repeat?" *The Journal of Portfolio Management* 20 (2):9–18.

Grinblatt, Mark and Sheridan Titman. 1989. "Mutual fund performance: An analysis of quarterly portfolio holdings." *Journal of business* :393–416.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers. 1995. "Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior." *The American economic review* :1088–1105.

Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser. 1993. "Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988." *The Journal of finance* 48 (1):93–130.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang. 2011. "Risk shifting and mutual fund performance." *The Review of Financial Studies* 24 (8):2575–2616.

Iskoz, Sergey and Jiang Wang. 2003. "How to Tell If a Money Manager Knows More?" Tech. rep., National Bureau of Economic Research.

Jegadeesh, Narasimhan and Sheridan Titman. 1993. "Returns to buying winners and selling losers: Implications for stock market efficiency." *The Journal of finance* 48 (1):65–91.

Jensen, Michael C. 1968. "The performance of mutual funds in the period 1945–1964." *The Journal of finance* 23 (2):389–416.

Jiang, George J, Tong Yao, and Tong Yu. 2007. "Do mutual funds time the market? Evidence from portfolio holdings." *Journal of Financial Economics* 86 (3):724–758.

Kacperczyk, Marcin and Amit Seru. 2007. "Fund manager use of public information: New evidence on managerial skills." *The Journal of Finance* 62 (2):485–528.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. 2005. "On the industry concentration of actively managed equity mutual funds." *The Journal of Finance* 60 (4):1983–2011.

———. 2008. "Unobserved actions of mutual funds." *Review of Financial Studies* 21 (6):2379–2416.

Kacperczyk, Marcin, Stijn van Nieuwerburgh, and Laura Veldkamp. 2014. "Time-varying fund manager skill." *The Journal of Finance* 69 (4):1455–1484.

Koijen, Ralph SJ. 2014. "The Cross-Section of Managerial Ability, Incentives, and Risk Preferences." *The Journal of Finance* 69 (3):1051–1098.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White. 2006. "Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis." *The Journal of finance* 61 (6):2551–2595.

Linnainmaa, Juhani T. 2013. "Reverse survivorship bias." *The Journal of Finance* 68 (3):789–813.

Malkiel, Burton G. 1995. "Returns from investing in equity mutual funds 1971 to 1991." *The Journal of finance* 50 (2):549–572.

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

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor. 2017. "Do funds make more when they trade more?" *The Journal of Finance* 72 (4):1483–1528.

Ren, Jue. 2017. "Mutual Fund Style Analysis: A Stochastic Dominance Approach." *working paper* .

Ross, Stephen A. 1976. "The arbitrage theory of capital asset pricing." *Journal of economic theory* 13 (3):341–360.

Sharpe, William F. 1966. "Mutual fund performance." *The Journal of business* 39 (1):119–138.

Treynor, Jack and Kay Mazuy. 1966. "Can mutual funds outguess the market." *Harvard business review* 44 (4):131–136.

Van Nieuwerburgh, Stijn and Laura Veldkamp. 2010. "Information acquisition and under-diversification." *The Review of Economic Studies* 77 (2):779–805.

Wermers, Russ. 2000. "Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses." *The Journal of Finance* 55 (4):1655–1695.

———. 2003. "Is money really'smart'? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence." .

Appendix

A. The Factor Model

I propose a factor model in this section to formalize the definition of stock-picking skills. The model serves to clarify the specific assumptions required to establish the FSD condition. The analysis shows that one set of sufficient conditions to impose the FSD condition is to limit the search scope to fund managers who are: 1) skilled at stock-picking; 2) unbiased towards unobservable factors; 3) sufficiently diversified.

A.1 The Economy

The economy considered in this section is a frictionless financial market with a factor structure. There are J factors (denoted as $\{F_{j,t}\}_{j=1}^{J}$) that are observable to both fund managers and the econometrician; and L factors (denoted as $\{f_{l,t}\}_{l=1}^{L}$) that are only observable to fund managers but not to the econometrician. There are K stocks traded in the market. The excess return (relative to the risk-free rate) of any stock within this market can be decomposed into three parts: the exposure to the J observable factors, the exposure to the L unobservable factors, and the idiosyncratic component:

$$\tilde{r}_{k,t} = r_f + \sum_j \beta_{k,j} F_{j,t} + \sum_l \gamma_{k,l} f_{l,t} + \epsilon_{k,t}$$

where $\tilde{r}_{k,t}$ denotes the return of stock k at time t; $F_{j,t}$ ($f_{l,t}$) is the realization of the observable (unobservable) factor j (l) at time t; $\beta_{k,j}$ ($\gamma_{k,l}$) denotes of the loading of stock k on factor j (l); $\epsilon_{k,t}$ is the idiosyncratic shock in stock k's return.

The factors and the idiosyncratic shocks represent different sources of risks and are assumed to be mutually independent. For an economic interpretation, the factors

$\{F_{j,t}\}_{j=1}^J$ and $\{f_{l,t}\}_{l=1}^L$ can be regarded as $J+L$ different types of market-wide risks; whereas the idiosyncratic shocks $\{\epsilon_{k,t}\}_{k=1}^K$ represent firm-specific risks.

Assumption 1. *The factors* $\{F_{j,t}\}_{j=1}^J$ *and* $\{f_{l,t}\}_{l=1}^L$*, and the idiosyncratic shocks* $\{\epsilon_{k,t}\}_{k=1}^K$ *are mutually independent. That is* $\forall j,\ j' \quad F_{j,t} \perp F_{j',t}$, $\forall l,\ l' \quad f_{l,t} \perp f_{l',t}$, $\forall k,\ k' \quad \epsilon_{k,t} \perp \epsilon_{k',t}$, $\forall j,\ l \quad F_{j,t} \perp f_{l,t}$, $\forall j,\ k \quad F_{j,t} \perp \epsilon_{k,t}$, $\forall l,\ k \quad f_{l,t} \perp \epsilon_{k,t}$, *where* $\perp$ *denotes that two random variables are independent to each other.*

Regarding the idiosyncratic shocks, they have zero expectation under the econometrician's information set so that there is no asymptotic arbitrage in this economy according to Ross (1976).

Assumption 2. *Idiosyncratic shocks have zero expectation under the econometrician's information set:* $\forall k, \quad \mathbb{E}\left(\epsilon_{k,t}\right) = 0$.

A.2 Connection between Real and Replica Fund Returns

For each real fund in each period, replica portfolios are constructed according to the procedure in Section 3. During the construction, each stock within the real fund's portfolio is replaced randomly with another stock in the same bucket. In this economy, the buckets are defined by the econometrician according to the J observable factors.

Assumption 3. *For each stock within the original portfolio, its replacement in the replica portfolio has the same exposure to observable factors, i.e.*

$$\forall k,\ j \quad \beta_{k,j} = \beta_{\hat{k},j}$$

where $\beta_{k,j}$ *is stock* k*'s exposure to observable factor* j*;* $\hat{k}$ *labels the replacing stock of stock* k*;* $\beta_{\hat{k},j}$ *is the exposure to observable factor* j *of stock* k*'s replacing stock.*

For expositional clarity, I adopt the following notation. I denote the return of fund i in period t as

$$
\begin{aligned}
r_{i,t} &= \sum_k w_{i,k,t-1} \tilde{r}_{k,t} \\
&= \sum_k w_{i,k,t-1} \left(r_f + \sum_j \beta_{k,j} F_{j,t} + \sum_l \gamma_{k.l} f_{l,t} + \epsilon_{k,t} \right) \\
&= r_f + \sum_k w_{i,k,t-1} \left(\sum_j \beta_{k,j} F_{j,t} \right) + \sum_k w_{i,k,t-1} \left(\sum_l \gamma_{k.l} f_{l,t} \right) + \sum_k w_{i,k,t-1} \epsilon_{k,t} \\
&= r_f + \sum_j \left(\sum_k w_{i,k,t-1} \beta_{k,j} \right) F_{j,t} + \sum_l \left(\sum_k w_{i,k,t-1} \gamma_{k,l} \right) f_{l,t} + \sum_k w_{i,k,t-1} \epsilon_{k,t} \\
&\equiv r_f + \sum_j \beta_{i,j,t} F_{j,t} + \underbrace{\underbrace{\sum_l \gamma_{i,l,t} f_{l,t}}_{v_{i,t}} + \underbrace{\sum_k w_{i,k,t-1} \epsilon_{k,t}}_{e_{i,t}}}_{u_{i,t}} \\
&\equiv r_f + \sum_j \beta_{i,j,t} F_{j,t} + u_{i,t}.
\end{aligned}
$$

Likewise, the return of fund i's replica is written as

$$\begin{aligned}
\hat{r}_{i,t} &= \sum_k w_{i,k,t-1}\tilde{r}_{\hat{k},t} \\
&= \sum_k w_{i,k,t-1}\left(r_f + \sum_j \beta_{\hat{k},j}F_{j,t} + \sum_l \gamma_{\hat{k},l}f_{l,t} + \epsilon_{\hat{k},t}\right) \\
&= r_f + \sum_k w_{i,k,t-1}\left(\sum_j \beta_{\hat{k},j}F_{j,t}\right) + \sum_k w_{i,k,t-1}\left(\sum_l \gamma_{\hat{k},l}f_{l,t}\right) + \sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} \\
&= r_f + \sum_j\left(\sum_k w_{i,k,t-1}\beta_{\hat{k},j}\right)F_{j,t} + \sum_l\left(\sum_k w_{i,k,t-1}\gamma_{\hat{k},l}\right)f_{l,t} + \sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} \\
&\equiv r_f + \sum_j \hat{\beta}_{i,j,t}F_{j,t} + \underbrace{\underbrace{\sum_l \hat{\gamma}_{i,l,t}f_{l,t}}_{\hat{v}_{i,t}} + \underbrace{\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t}}_{\hat{e}_{i,t}}}_{\hat{u}_{i,t}} \\
&\equiv r_f + \sum_j \hat{\beta}_{i,j,t}F_{j,t} + \hat{u}_{i,t}
\end{aligned}$$

The excess return of a real (replica) fund can be decomposed into three parts: the exposure to the observable factors $\sum_j \beta_{i,j,t}F_{j,t}$ $(\sum_j \hat{\beta}_{i,j,t}F_{j,t})$, the exposure to the unobservable factors $v_{i,t} \equiv \sum_l \gamma_{i,l,t}f_{l,t}$ $(\hat{v}_{i,t} \equiv \sum_l \hat{\gamma}_{i,l,t}f_{l,t})$, and the exposure to the idiosyncratic shocks $e_{i,t} \equiv \sum_k w_{i,k,t-1}\epsilon_{k,t}$ $(\hat{e}_{i,t} \equiv \sum_k w_{i,k,t-1}\epsilon_{\hat{k},t})$.

Note that fund i and its replica have the same portfolio weights by construction. A direct outcome of this and Assumption 3 is that the real fund and the replica have the same exposure to observable factors.

Proposition 5. *For observable factors, the original portfolio and the replica portfolio have the same loadings, i.e.* $\forall i,\ j,\ t, \quad \beta_{i,j,t} = \hat{\beta}_{i,j,t}$.

Proof. According to Assumption 3, $\forall k,\ j \quad \beta_{k,j} = \beta_{\hat{k},j}$. Therefore, $\beta_{i,j,t} = \sum_K w_{i,k,t-1}\beta_{k,j} = \sum_K w_{i,k,t-1}\beta_{\hat{k},j} = \hat{\beta}_{i,j,t}$, i.e. the original portfolio and the replica portfolio have the same loadings on observable factors. □

By construction, the econometrician ensures that the replica fund has the same loadings on observable factors as the original fund. As for the unobservable factors, the replica fund's loadings should be unbiased since the stocks within the replica fund are picked randomly once the observable factor loadings are match.

Proposition 6. *For well-diversified portfolio weight distribution* $\{w_{i,k,t-1}\}$*, the replica fund's loadings on the unobservable factors are unbiased. That is,* $\forall i,\ l,\quad \hat{\gamma}_{i,l,t} = \sum_K w_{i,k,t-1}\gamma_{\hat{k},l} \approx \sum_K w_{i,k,t-1}\bar{\gamma}_{k,l}$*, where* $\bar{\gamma}_{k,l} \equiv \mathbb{E}\left(\gamma_{k.l} | \{\beta_{k,j}\}\right)$ *is the average loading on factor* $f_{l,t}$ *for stocks in the bucket identified by* $\{\beta_{k,j}\}$*.*

Proof. Denote

$$\begin{aligned}
\delta_{\hat{k},l} &\equiv \gamma_{\hat{k}.l} - \mathbb{E}\left(\gamma_{\hat{k}.l} | \left\{\beta_{\hat{k},j}\right\}\right) \\
&= \gamma_{\hat{k}.l} - \mathbb{E}\left(\gamma_{k.l} | \{\beta_{k,j}\}\right) \\
&= \gamma_{\hat{k}.l} - \bar{\gamma}_{k,l}
\end{aligned}$$

$\delta_{\hat{k},l}$ is the deviation of stock $\hat{k}$'s loading on factor l from the average loading on factor l of the stocks in the same bucket as $\hat{k}$. Since stock $\hat{k}$ is chosen randomly within the bucket, $\mathbb{E}\left(\delta_{\hat{k},l} | \{\beta_{k,j}\}\right)$[17].

The second line comes from the fact that $\beta_{\hat{k},j} = \beta_{k,j}$.

$$\begin{aligned}
\hat{\gamma}_{i,l,t} &= \sum_k w_{i,k,t-1}\gamma_{\hat{k},l} \\
&= \sum_k w_{i,k,t-1}\left(\bar{\gamma}_{k,l} + \delta_{\hat{k},l}\right) \\
&= \sum_k w_{i,k,t-1}\bar{\gamma}_{k,l} + \sum_k w_{i,k,t-1}\delta_{\hat{k},l} \\
&\overset{LLN}{\approx} \sum_k w_{i,k,t-1}\bar{\gamma}_{k,l}
\end{aligned}$$

[17]Here, the expectation is taken before the fund is constructed.

The last line is given by Law of Large Number. It holds when the portfolio is sufficiently diversified. □

Proposition 6 shows that the replica fund's loadings on the unobservable factors are unbiased because the stocks within the replica fund's portfolio are chosen randomly. On the other hand, if the loadings on the observable factors can already sufficiently characterize the style of the original fund, then the original fund's loadings on the unobservable factors should also be unbiased.

Definition 1. Fund i's style is well-specified by the observable factors iff

$$
\begin{aligned}
&\sum_k w_{i,k,t-1} \left(\gamma_{k.l} - \mathbb{E}\left(\gamma_{k.l} \middle| \{\beta_{k,j}\}\right)\right) \\
\equiv &\sum_k w_{i,k,t-1} \delta_{k.l} \\
\approx &0
\end{aligned}
$$

Equivalently,

$$
\gamma_{i,l,t} \approx \sum_k w_{i,k,t-1} \bar{\gamma}_{k,l}
$$

As mentioned in Section 3, the matching of style between the original fund and the replica funds indicates that only stock-picking skills rather than factor-timing skills can be measured in this exercise. A fund manager manifests stock-picking skills in his ability to predict idiosyncratic stock returns $\{\epsilon_{k,t}\}_{k=1}^K$. The stock-picking skills of a fund manager can be defined more formally as follows.

Definition 2. Define skilled stock-pickers as the managers with superior information about $\{\epsilon_{k,t}\}_{k=1}^K$ and satisfy:

1. Better firm-specific information, not contingent on factor realizations:

$$\mathbb{E}\left(e_{i,t}\middle|\left\{w_{i,k,t-1}\right\}\right) = \mathbb{E}\left(\sum_k w_{i,k,t-1}\epsilon_{k,t}\middle|\left\{w_{i,k,t-1}\right\}\right) \equiv \alpha_{i,t}$$
$$> \mathbb{E}\left(\hat{e}_{i,t}\middle|\left\{w_{i,k,t-1}\right\}\right) = \mathbb{E}\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t}\middle|\left\{w_{i,k,t-1}\right\}\right) = 0$$

and

$$\epsilon_{k,t} \perp F_{j,t}\middle|\left\{w_{i,k,t-1}\right\}, \quad \forall k,\ j$$

$$\epsilon_{k,t} \perp f_{l,t}\middle|\left\{w_{i,k,t-1}\right\}, \quad \forall k,\ l$$

$$\epsilon_{k,t} \perp \epsilon_{k',t}\middle|\left\{w_{i,k,t-1}\right\}, \quad \forall k,\ k'$$

2. Sufficiently diversified, so that Proposition 6 holds and Central Limit Theorem applies:

$$\sum_k w_{i,k,t-1}\epsilon_{k,t}\middle|\left\{w_{i,k,t-1}\right\} \sim N\left(\alpha_{i,t}, Var\left(\sum_k w_{i,k,t-1}\epsilon_{k,t}\middle|\left\{w_{i,k,t-1}\right\}\right)\right)$$
$$\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t}\middle|\left\{w_{i,k,t-1}\right\} \sim N\left(0, Var\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t}\middle|\left\{w_{i,\hat{k},t-1}\right\}\right)\right)$$

3. Style well-specified by observable factors, no bias towards unobservable factors:

$$\sum_k w_{i,k,t-1}\left(\gamma_{k.l} - \mathbb{E}\left(\gamma_{k.l}\middle|\left\{\beta_{k,j}\right\}\right)\right)$$
$$\equiv \sum_k w_{i,k,t-1}\delta_{k.l}$$
$$\approx 0$$

Equivalently,

$$\gamma_{i,l,t} \approx \sum_k w_{i,k,t-1}\bar{\gamma}_{k,l} \approx \hat{\gamma}_{i,l,t}$$

An important clarification is warranted. Definition 2 does not aim to exclusively define all types of stock-picking skills under common sense. Instead, the definition only serves to draw the boundary of the empirical search. In other words, a skilled stock-picker under common sense might be excluded by Definition 2 for being significantly under-diversified or biased towards certain unobservable factors to the econometrician. The purpose of this project is to identify a large enough subset of all skilled stock-pickers who are able to deliver positive alpha out of sample.

The rest of this subsection intends to show that, the return distribution of a skilled stock-picker defined above is a mean shift from the return distribution of a corresponding replica fund. That is

$$\begin{aligned}
& r_{i,t} - \alpha_{i,t} \sim \hat{r}_{i,t} \\
\iff & \sum_j \beta_{i,j,t} F_{j,t} + \sum_l \gamma_{i,l,t} f_{l,t} + \sum_k w_{i,k,t-1}\epsilon_{k,t} - \alpha_{i,t} \\
& \sim \sum_j \hat{\beta}_{i,j,t} F_{j,t} + \sum_l \hat{\gamma}_{i,l,t} f_{l,t} + \sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} \\
\iff & \sum_j \beta_{i,j,t} F_{j,t} + v_{i,t} + e_{i,t} - \alpha_{i,t} \sim \sum_j \hat{\beta}_{i,j,t} F_{j,t} + \hat{v}_{i,t} + \hat{e}_{i,t}.
\end{aligned}$$

So far, the following have already been established:

1. $\sum_j \beta_{i,j,t} F_{j,t} = \sum_j \hat{\beta}_{i,j,t} F_{j,t}$.

2. $v_{i,t} = \sum_l \gamma_{i,l,t} f_{l,t} \approx \hat{v}_{i,t} = \sum_l \hat{\gamma}_{i,l,t} f_{l,t}$.

3. $e_{i,t} \perp v_{i,t}$, $e_{i,t} \perp \sum_j \beta_{i,j,t} F_{j,t}$, $\hat{e}_{i,t} \perp \hat{v}_{i,t}$ and $\hat{e}_{i,t} \perp \sum_j \hat{\beta}_{i,j,t} F_{j,t}$.

The next step is to prove that $e_{i,t} - \alpha_{i,t} \sim \hat{e}_{i,t}$ for a skilled stock-picker.

Proposition 7. *For a skilled stock-picker defined in Definition 2, the residual risk of the fund portfolio is a mean shift to that of the corresponding replica fund. That is,*

$$e_{i,t} - \alpha_{i,t} \sim \hat{e}_{i,t}.$$

Proof. According to Central Limit Theorem,

$$\begin{aligned} e_{i,t} | \{w_{i,k,t-1}\} &\sim N\left(\alpha_{i,t}, Var\left(\sum_k w_{i,k,t-1}\epsilon_{k,t} | \{w_{i,k,t-1}\}\right)\right) \\ \hat{e}_{i,t} | \{w_{i,k,t-1}\} &\sim N\left(0, Var\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} | \left\{w_{i,\hat{k},t-1}\right\}\right)\right) \end{aligned}$$

All we need to show is that $Var\left(\sum_k w_{i,k,t-1}\epsilon_{k,t} | \{w_{i,k,t-1}\}\right) \approx Var\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} | \left\{w_{i,\hat{k},t-1}\right\}\right)$. Notice that the real fund and the replica fund are picking stocks from the same stock volatility buckets, so that

$$\begin{aligned} & Var\left(\tilde{r}_{k,t}\right) = Var\left(\tilde{r}_{\hat{k},t}\right) \\ \Longleftrightarrow \ & Var\left(\sum_l \gamma_{k,l} f_{l,t} + \epsilon_{k,t}\right) = Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t} + \epsilon_{\hat{k},t}\right) \end{aligned}$$

On the other hand $\gamma_{k,l} = \bar{\gamma}_{k,l} + \delta_{k,l}$ and $\gamma_{\hat{k},l} = \bar{\gamma}_{k,l} + \delta_{\hat{k},l}$, so that the difference between

$\gamma_{k,l}$ and $\gamma_{\hat{k},l}$ could be diversified away. Therefore,

$$
\begin{aligned}
& Var\left(\sum_k w_{i,k,t-1}\epsilon_{k,t} | \left\{w_{i,k,t-1}\right\}\right) \\
=& \sum_k w_{i,k,t-1}^2 Var\left(\epsilon_{k,t}\right) \\
=& \sum_k w_{i,k,t-1}^2 \left(Var\left(\sum_l \gamma_{k,l} f_{l,t} + \epsilon_{k,t}\right) - Var\left(\sum_l \gamma_{k,l} f_{l,t}\right)\right) \\
=& \sum_k w_{i,k,t-1}^2 \left(Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t} + \epsilon_{\hat{k},t}\right) - Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t}\right)\right) \\
& + \left(Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t}\right) - Var\left(\sum_l \gamma_{k,l} f_{l,t}\right)\right) \\
=& Var\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} | \left\{w_{i,\hat{k},t-1}\right\}\right) + \sum_k w_{i,k,t-1}^2 \left(Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t}\right) - Var\left(\sum_l \gamma_{k,l} f_{l,t}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
& \sum_k w_{i,k,t-1}^2 \left(Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t}\right) - Var\left(\sum_l \gamma_{k,l} f_{l,t}\right)\right) \\
=& \sum_k w_{i,k,t-1}^2 \left(\sum_l \left(\gamma_{\hat{k},l}^2 - \gamma_{k,l}^2\right) Var\left(f_{l,t}\right)\right) \\
\approx& \sum_k w_{i,k,t-1}^2 \left(\sum_l 2\bar{\gamma}_{k,l}\left(\delta_{\hat{k},l} - \delta_{k,l}\right) Var\left(f_{l,t}\right)\right) \\
=& 2\sum_l \left(\sum_k w_{i,k,t-1}^2 \bar{\gamma}_{k,l}\left(\delta_{\hat{k},l} - \delta_{k,l}\right)\right) Var\left(f_{l,t}\right) \\
\approx& \frac{2}{K}\sum_l \left(\sum_k w_{i,k,t-1} \bar{\gamma}_{k,l}\left(\delta_{\hat{k},l} - \delta_{k,l}\right)\right) Var\left(f_{l,t}\right) \\
\overset{LLN}{\approx}& 0
\end{aligned}
$$

□

Proposition 8. *The return of a skilled stock-picker defined in Definition 2 is approxi-*

mately a mean shift to the return of the replica fund. That is,

$$r_{i,t} - \alpha_{i,t} \sim \hat{r}_{i,t}.$$

Proof. Directly from the fact that:

1. $\sum_j \beta_{i,j,t} F_{j,t} = \sum_j \hat{\beta}_{i,j,t} F_{j,t}$.

2. $v_{i,t} = \sum_l \gamma_{i,l,t} f_{l,t} \approx \hat{v}_{i,t} = \sum_l \hat{\gamma}_{i,l,t} f_{l,t}$.

3. $e_{i,t} \perp v_{i,t}$, $e_{i,t} \perp \sum_j \beta_{i,j,t} F_{j,t}$, $\hat{e}_{i,t} \perp \hat{v}_{i,t}$ and $\hat{e}_{i,t} \perp \sum_j \hat{\beta}_{i,j,t} F_{j,t}$.

4. $e_{i,t} - \alpha_{i,t} \sim \hat{e}_{i,t}$.

□

The intuition of this result is straightforward. According to Definition 2, the real fund and the replica fund have the same exposures to all systematic risks. If the real fund's manager is skilled at picking stocks, then the idiosyncratic component of the real fund's return has higher mean compared to the replica fund's return. On the other hand, the distribution of the idiosyncratic component of the real fund has the same shape as the replica fund due to the Central Limit Theorem and the fact that the volatility of the stocks in both funds are matched by construction. The result then follows because of the independence between the systematic and idiosyncratic components in fund returns.

Proposition 9. *The return of a skilled stock-picker defined in Definition 2 first-order stochastically dominates the return of the replica fund. That is,*

$$r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}.$$

Proof. The result is directly from Proposition 8 and the fact that $\alpha_{i,t} > 0$ according to Definition 2. □

B. The Alternative FSD Test Statistic

As discussed in the footnote of Section 4.3, the FSD test statistic can be also constructed as a variant of the Kolmogorov–Smirnov (KS) statistic. The original KS statistic is defined as

$$KS \equiv \max_{x\in[0,1]} \left| \hat{F}^{Pct(r_{i,t},\langle \hat{r}_{i,t}\rangle)}(x) - x \right|,$$

which detects whether the distribution of $Pct(r_{i,t}, \langle \hat{r}_{i,t}\rangle)$ differs from the standard uniform distribution ($Unif(0,1)$).

The original Kolmogorov–Smirnov test is a two-sided test. However, the FSD condition is a one-sided restriction on the CDF of the distribution. Therefore, one can also construct an FSD test statistic by modifying the KS statistic:

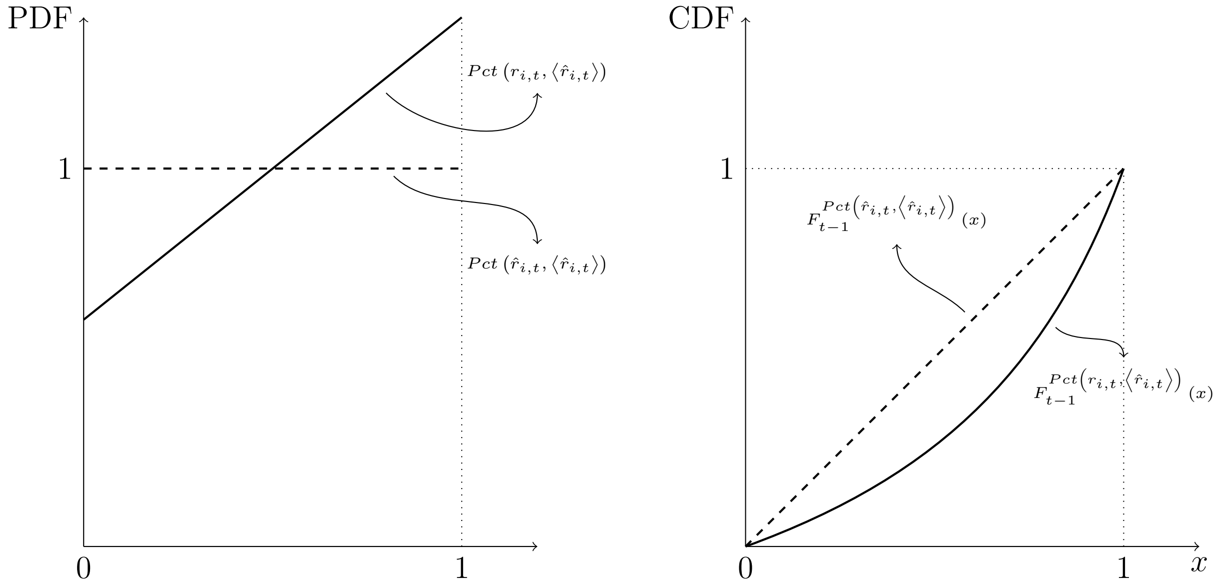
$$\hat{\theta} \equiv \max_{x\in[0,1]} \left(\hat{F}^{Pct(r_{i,t},\langle \hat{r}_{i,t}\rangle)}(x) - x \right).$$

Figure 13 illustrates the construction of the alternative FSD test statistic. From the figure, it is clear that the FSD condition is better satisfied with a smaller test statistic. Unlike the FSD test statistic constructed in Section 4, the alternative FSD test statistic does not have a tractable asymptotic null distribution. Figure 14 plots the simulated finite-sample null distribution of the alternative FSD test statistic. According to the simulation, a test size of 10% (5%) corresponds to a critical value of 0.039 (0.027).

Table 11 documents the out-of-sample performance of the search based on the alternative FSD test statistic. Comparing Table 11 to Table 4, one can see that the alternative FSD test statistic generates similar empirical findings as the main FSD test statistic. Alas, the out-of-sample alphas in the second stage are not perfectly monotonic to in-sample alphas, with the alpha in Quintile 4 being smaller than Quintile 3 and 5.

Figures

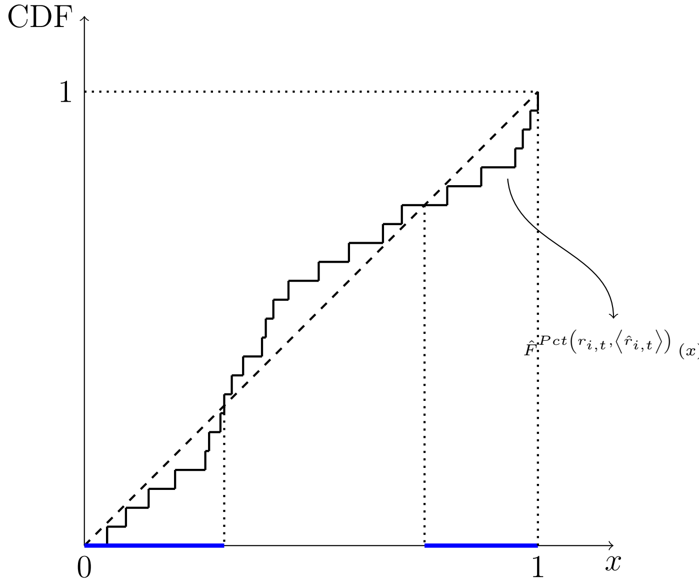
Figure 2: Graphical Illustration of the FSD Condition



(a) (b)

This figure offers a graphical illustration of the FSD condition. Panel (a) is a demonstration regarding the relation between $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ and $Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$. The dashed line plots the PDF of $Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$, which is a flat horizontal line constant at 1 since $Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ follows a standard uniform distribution. The solid line is an example of $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$. Since $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right) \overset{fsd}{\succ} Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$, the solid line has a smaller left tail compared to the dashed line, but a larger right tail in the PDF plot. Panel (b) offers the illustration on the same relation with CDF plots. The $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right) \overset{fsd}{\succ} Pct\left(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle\right)$ condition is reflected in the plot as the solid curve strictly lies below the dashed line.

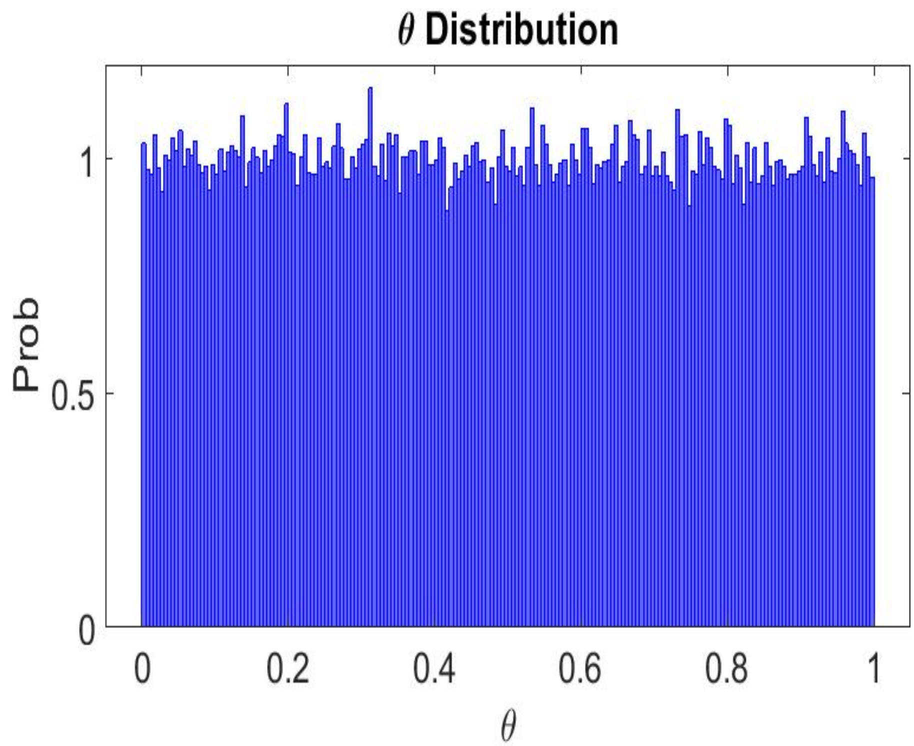
Figure 3: FSD Test Statistic $\hat{\theta}$



Test Statistic: $\hat{\theta}$ = length of —

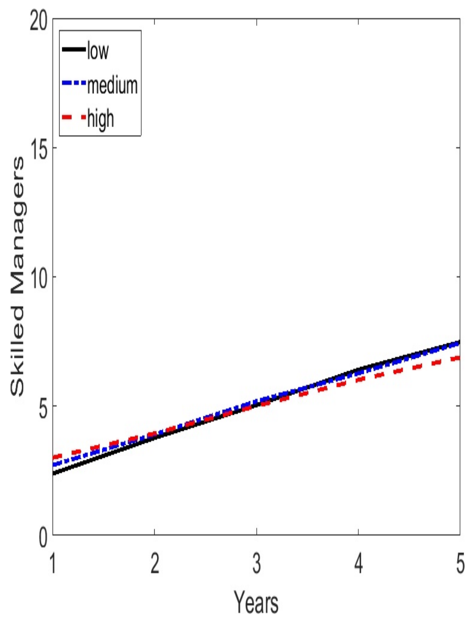
This figure illustrates the construction of the FSD test statistic $\hat{\theta}$. The solid step function is an illustration of the empirical CDF of $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$. The test statistic $\hat{\theta}$ is constructed as the measure of the region where the empirical CDF falls below the dashed 45 degree line.

Figure 4: Simulated Null Distribution of the FSD Test Statistic

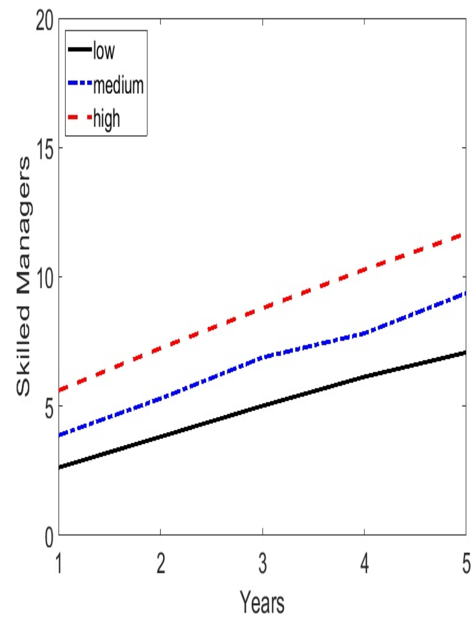


This figure plots finite-sample distribution of the FSD test statistic $\hat{\theta}$ constructed from 24 observations under the null hypothesis that $Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$ follows a standard uniform distribution.

Figure 5: The Heteroscedastic Case: Filter Accuracy



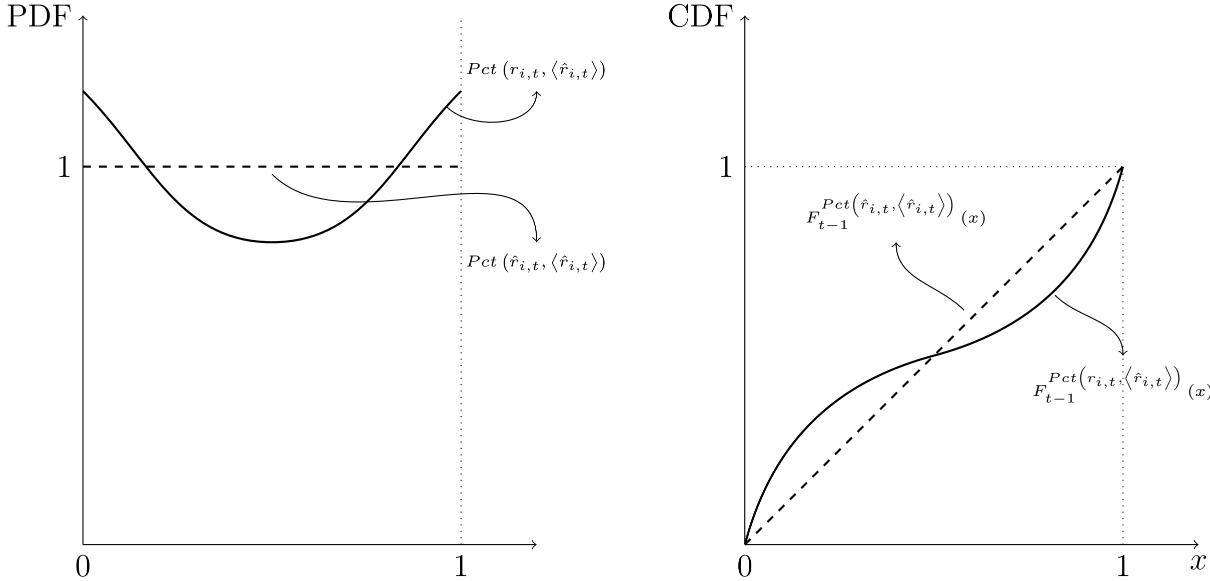
(a) The $\alpha_i > 0$ Filter



(b) The $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ Filter

The plot compares the accuracy of the $\alpha_i > 0$ and the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filters when idiosyncratic volatility is time-varying. The x-axis is the formation period from which the performance measures are constructed. The y-axis is the average number of skilled managers that the corresponding performance measure is able to identify over 500 simulation paths. The black solid (blue dotted, red dashed) line represents the level of heteroscedasticity with $\zeta = 0.0012$ ($\zeta = 0.0018$, $\zeta = 0.0024^*$). Panel (a) plots the effectiveness of the $\alpha_i > 0$ filter; whereas Panel (b) plots the effectiveness of the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filter.

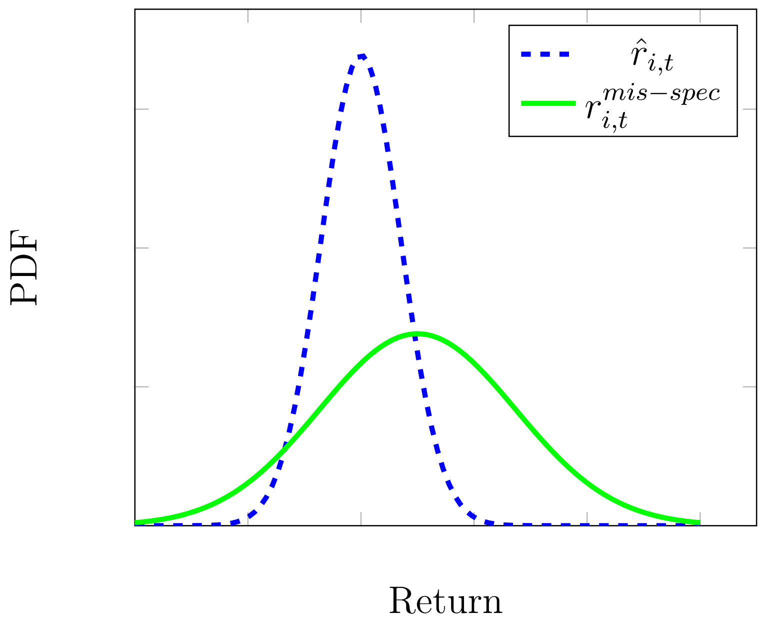
Figure 6: Detection for Uncontrolled Factors



(a) (b)

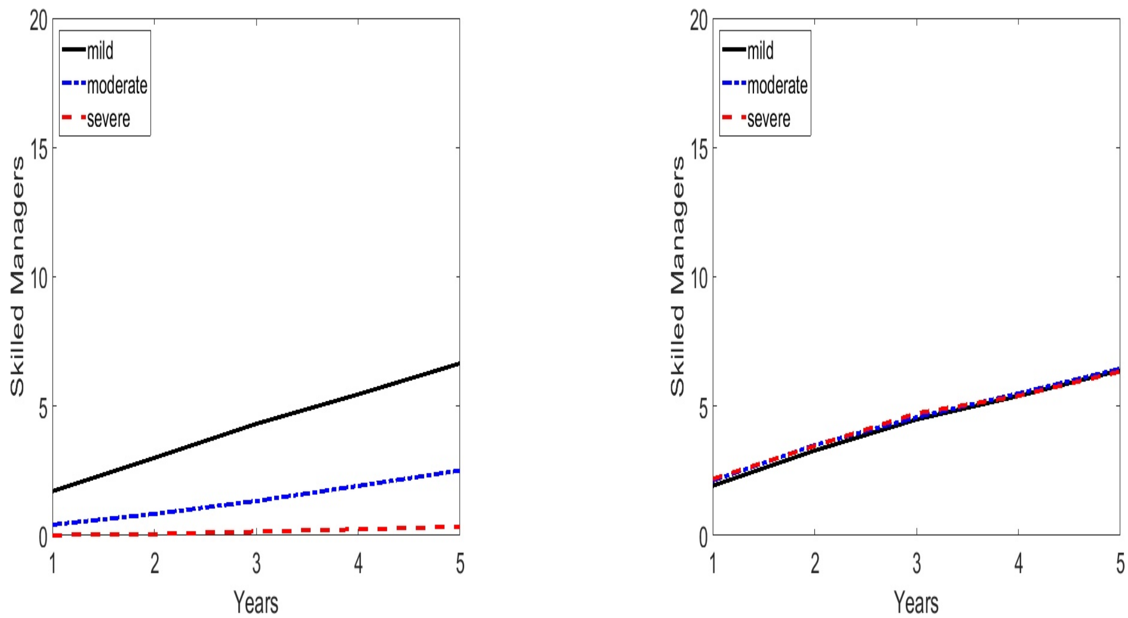
This figure offers a graphical illustration of the detection mechanism of the FSD condition to exclude mis-specified managers. Panel (a) shows that the PDF of the mis-specified manager's ranking $(Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right))$ has both larger left and right tails compared to the standard uniform distribution due to the uncontrolled risk factor; Panel (b) shows that the CDF of the mis-specified manager's ranking $\left(Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right)\right)$ goes above the 45 degree line for some region, thereby violating the FSD condition.

Figure 7: Mis-specified Return Distribution



This figure illustrates that the return distribution of a mis-specified manager carries a heavier left tail compared to the return distribution of the replica fund, thereby violating the FSD condition.

Figure 8: The Mis-specification Case: Filter Accuracy

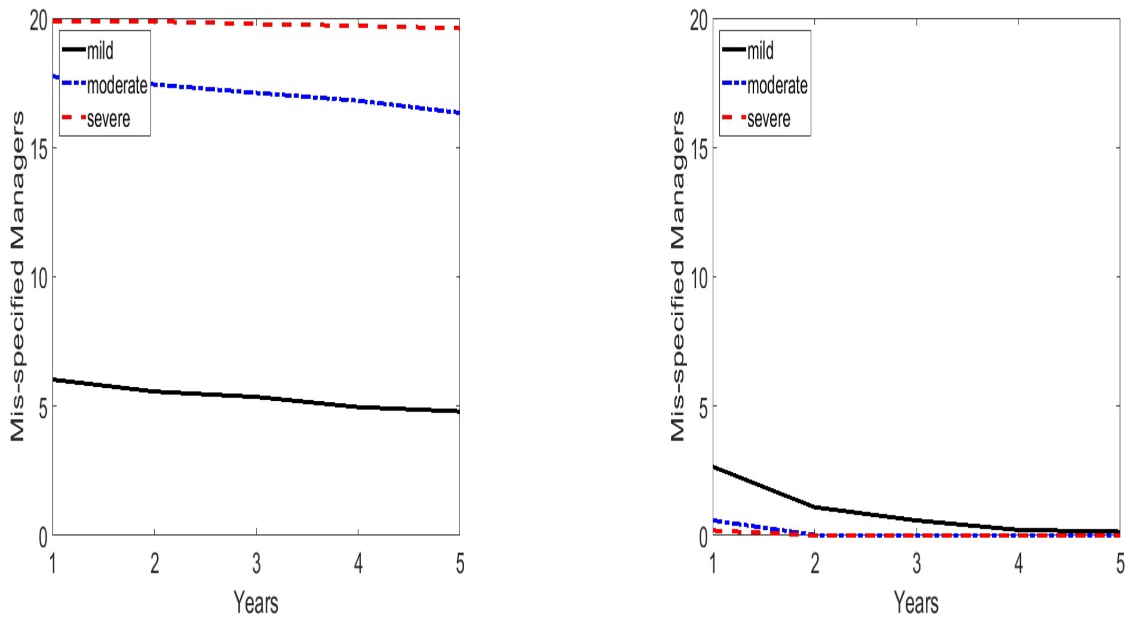


(a) The $\alpha_i > 0$ Filter

(b)The $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ Filter

The plot compares the accuracy of the $\alpha_i > 0$ and the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filters in the presence of benchmark mis-specification. The x-axis is the formation period from which the performance measures are constructed. The y-axis is the average number of skilled managers that the corresponding performance measure is able to identify over 500 simulation paths. The black solid (blue dotted, red dashed) line represents the situation where benchmark mis-specification is mild $\sigma_f = 0.01$ (moderate $\sigma_f = 0.03$, severe $\sigma_f = 0.05$). Panel (a) plots the effectiveness of the $\alpha_i > 0$ filter; whereas Panel (b) plots the effectiveness of the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filter.

Figure 9: The Mis-specification Case: Filter Mistake

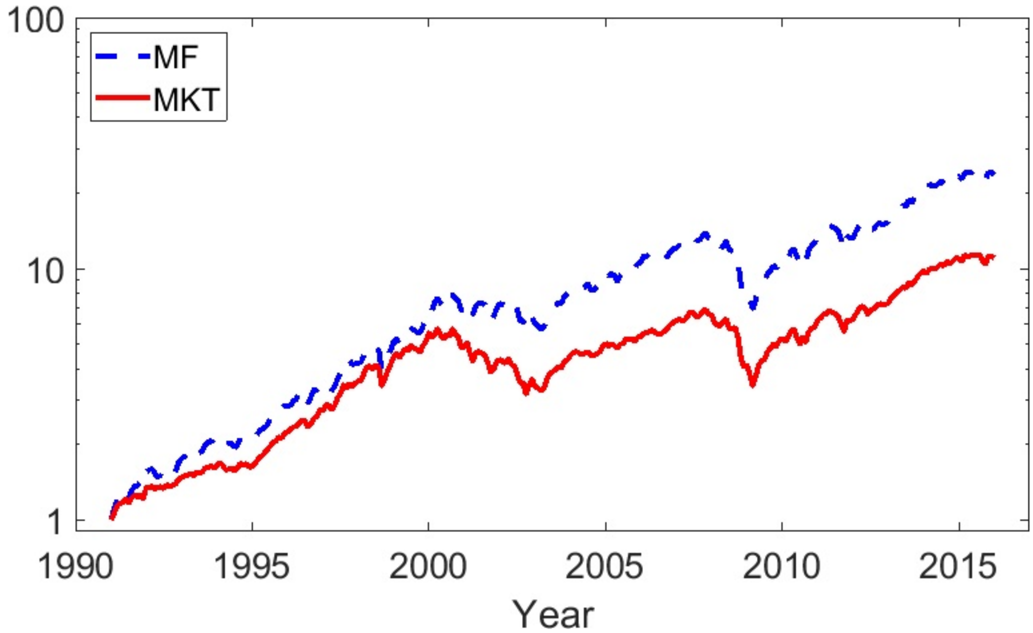


(a) The $\alpha_i > 0$ Filter

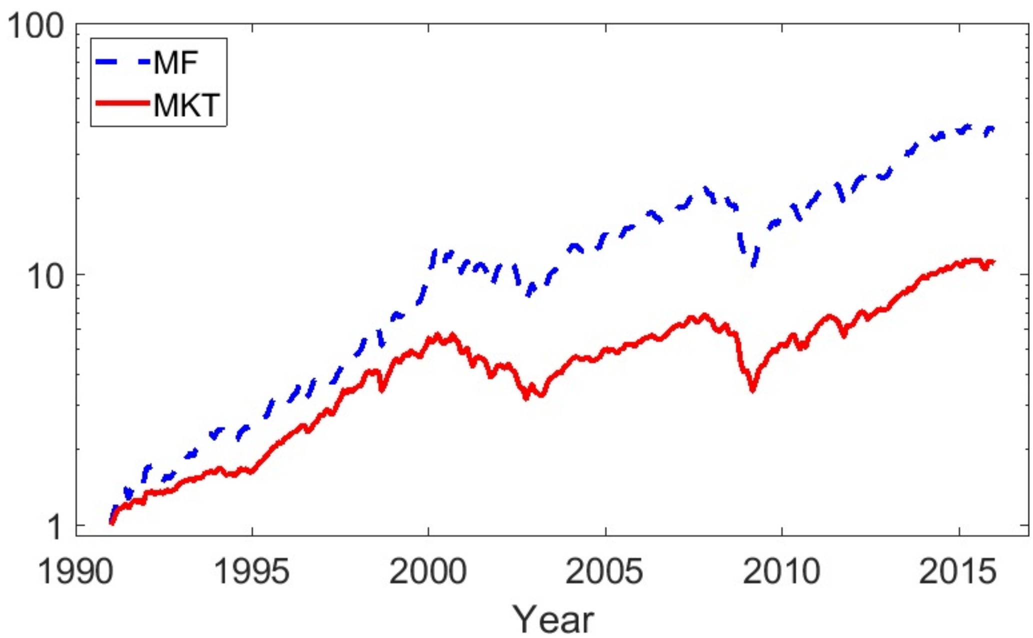
(b)The $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ Filter

The plot compares the tendency to select mis-specified managers of the $\alpha_i > 0$ and the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filters in the presence of benchmark mis-specification. The x-axis is the formation period from which the performance measures are constructed. The y-axis is the average number of mis-specified managers that the corresponding performance measure erroneously selects over 500 simulation paths. The black solid (blue dotted, red dashed) line represents the situation where benchmark mis-specification is mild $\sigma_f = 0.01$ (moderate $\sigma_f = 0.03$, severe $\sigma_f = 0.05$). Panel (a) plots the performance of the $\alpha_i > 0$ filter; whereas Panel (b) plots the performance of the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filter.

Figure 10: Out-of-sample Fund Performances: Time Series



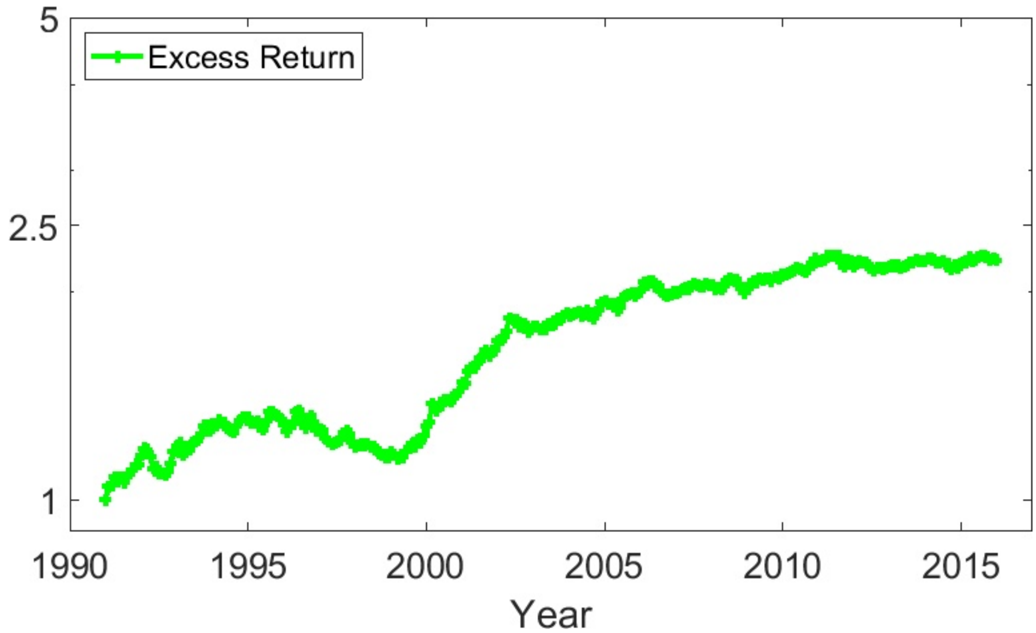
(a) 1st Stage



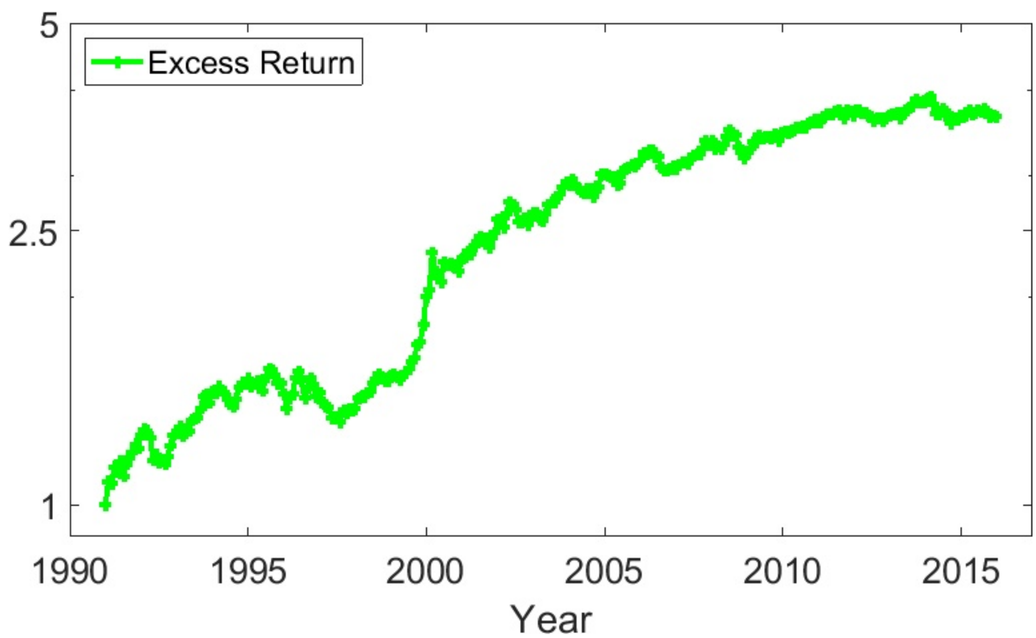
(b) Top α Quintile, 2nd Stage

This panel of plots documents the time-series of the before-fees performances of the selected mutual funds. Panel (a) plots the time-series of the fund performances for all funds satisfying the FSD condition; whereas Panel (b) plots the top quintile of the funds with the highest historical α within the funds satisfying the FSD condition. The blue dashed line is the cumulative before-fees return of the selected mutual funds, and the red solid line is cumulative return of the market.

Figure 11: Out-of-sample Fund Outperformances: Time Series



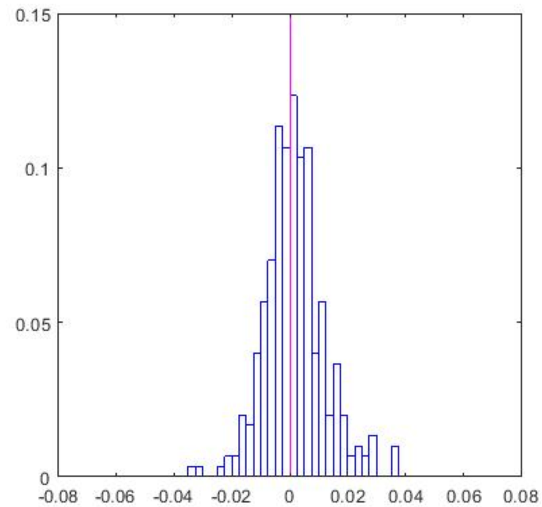
(a) 1st Stage



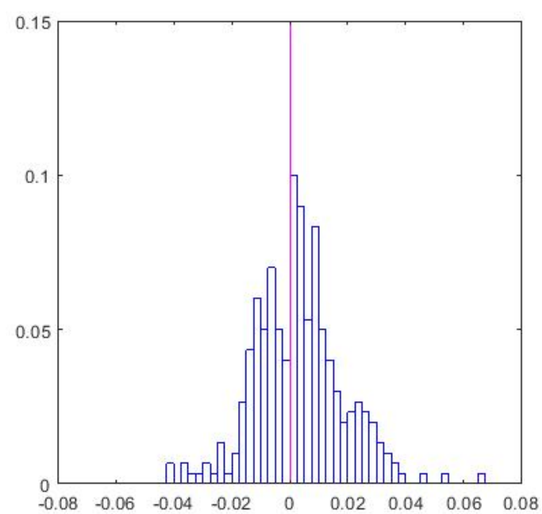
(b) Top α Quintile, 2nd Stage

This panel of plots documents the time-series of the before-fees out-performances of the selected mutual funds. The out-performance is defined as the cumulative return of longing the portfolio of the identified funds and shorting the market. Panel (a) plots the time-series of the fund out-performances for all funds satisfying the FSD condition; whereas Panel (b) plots the top quintile of the funds with the highest historical α within the funds satisfying the FSD condition.

Figure 12: Out-of-sample Fund Excess Returns: Histogram



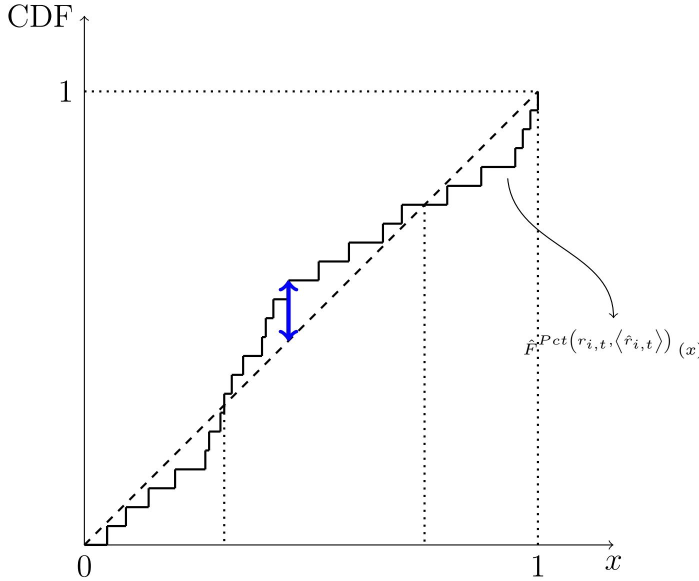
(a) 1st Stage



(b) Top $\hat{\alpha}$ Quintile, 2nd Stage

This panel of plots documents the histograms of the before-fees Carhart excess returns of the selected mutual funds. Panel (a) plots the histogram for all funds satisfying the FSD condition; whereas Panel (b) plots the histogram of the top quintile of the funds with the highest historical α within the funds satisfying the FSD condition.

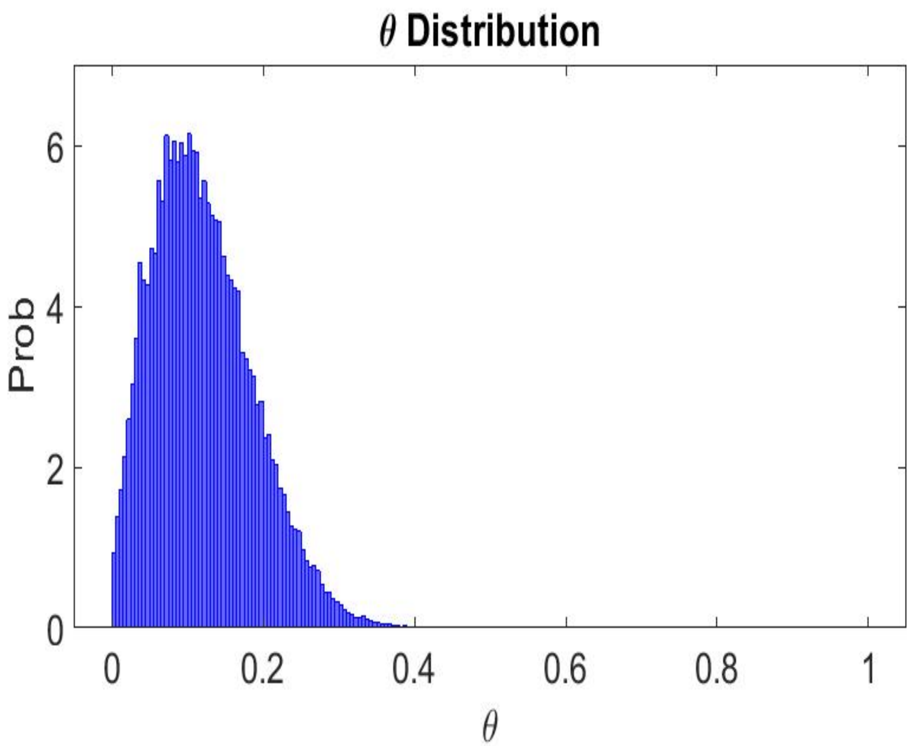
Figure 13: Alternative FSD Test Statistic



Test Statistic: $\hat{\theta}$ = length of $\longleftrightarrow$

This figure illustrates the construction of the alternative FSD test statistic. The solid step function is an illustration of the empirical CDF of $Pct\left(r_{i,t},\langle\hat{r}_{i,t}\rangle\right)$. The alternative FSD test statistic is defined as $\hat{\theta}_i \equiv \max\left(\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x) - x\right)$.

Figure 14: Simulated Null Distribution of the Alternative FSD Test Statistic



This figure plots finite-sample distribution of the alternative FSD test statistic ($\hat{\theta} \equiv \max\left(\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x) - x\right)$) constructed from 24 observations under the null hypothesis that $Pct\left(r_{i,t}, \langle\hat{r}_{i,t}\rangle\right)$ follows a standard uniform distribution.

Tables

Table 2: Bootstrap Example

Stock Ticker	Bucket No.	Weight(%)
ABT	597	5.24
BEN	552	4.68
BK	576	6.73
BRK	561	4.65
CHK	603	8.07
CNX	428	7.16
DELL	529	5.55
DIS	598	5.70
DTV	502	8.19
FDX	534	8.00
L	581	10.00
LVLT	405	6.15
MDLZ	592	5.39
PHG	356	1.26
TRV	568	6.65
VMC	489	6.58

(a) The Real Fund

Stock Ticker	Bucket No.	Weight(%)
MO	597	5.24
DOV	552	4.68
SYK	576	6.73
PG	561	4.65
C	603	8.07
BBY	428	7.16
A	529	5.55
DIS	598	5.70
WU	502	8.19
FDX	534	8.00
MDT	581	10.00
LVLT	405	6.15
MDLZ	592	5.39
PHG	356	1.26
FISV	568	6.65
ARW	489	6.58

(b) The Replica Fund

This table offers an example in order to illustrate the bootstrap procedure that constructs replica portfolios and the counterfactual return distribution. Panel (a) is a snapshot of the portfolio of Longleaf Partners Fund by the end of 2012/12. Panel (b) is a simulated replica portfolio. The replica portfolio is created by replacing each stock in the real portfolio with another stock that is randomly chosen in the same bucket as the original stock. As an extension of Daniel et al. (1997), the buckets are defined by four stock characteristics: size, book-to-market, momentum and volatility. The portfolio weights of the replica fund are identical as the real fund.

Table 3: Fund Characteristics Summary

Year	# of Funds	Age	TNA (in 10^6 \$)	# of Stocks	Fees (in bps)
1991	178	20	583	89	110
1992	213	19	684	92	128
1993	250	17	776	98	131
1994	282	17	778	95	128
1995	299	18	1175	102	127
1996	312	18	1560	110	125
1997	387	17	1783	110	120
1998	473	16	1941	104	119
1999	589	15	2312	111	123
2000	642	15	1871	105	123
2001	705	14	1472	115	126
2002	847	14	1071	107	134
2003	948	14	1367	109	136
2004	1043	14	1462	107	133
2005	1130	14	1470	115	131
2006	1195	14	1537	116	127
2007	1233	15	1643	117	122
2008	1260	15	913	121	123
2009	1275	16	1227	125	125
2010	1235	16	1406	121	122
2011	1268	16	1293	108	117
2012	1272	17	1408	110	116
2013	1254	17	1925	114	113
2014	1197	18	2033	111	111
2015	1163	19	1984	113	110

This table documents the characteristics of the funds in my sample. The documented characteristics include total number of funds in the sample, the average fund age, the average total net assets in million dollars, the average number of stocks that a fund holds, the average fee that a fund charges in basis points per year.

Table 4: Out-of-sample Performances of the FSD Selected Funds

Quintile	Sample Share	α (in %)	mkt	smb	hml	umd	SR	IR
				Before Fees				
1	1.83%	0.57	1.03***	0.29***	0.03	0.05**	0.60	0.13
		[0.62]	[47.0]	[9.29]	[0.79]	[2.53]		
2	1.94%	1.48*	1.01***	0.20***	0.07**	0.02	0.65	0.36
		[1.77]	[56.9]	[5.87]	[2.32]	[1.11]		
3	1.96%	2.24**	1.00***	0.23***	0.09**	0.05**	0.72	0.50
		[2.41]	[43.4]	[6.77]	[2.11]	[2.53]		
4	1.94%	2.38**	1.03***	0.27***	0.01	0.06**	0.70	0.49
		[2.47]	[46.7]	[6.98]	[0.25]	[2.47]		
5	1.88%	3.71***	1.07***	0.44***	−0.10***	0.08***	0.74	0.70
		[3.35]	[38.9]	[13.2]	[−2.83]	[3.28]		
1st Stage	9.55%	2.03***	1.03***	0.28***	0.02	0.05***	0.69	0.56
		[2.78]	[61.1]	[10.5]	[0.70]	[2.98]		
All Funds	100%	0.04	1.00***	0.21***	0.02	0.01	0.56	0.02
		[0.07]	[81.4]	[10.7]	[1.17]	[0.95]		
				After Fees				
1	1.83%	−0.69	1.03***	0.29***	0.03	0.05**	0.52	−0.15
		[−0.74]	[47.0]	[9.25]	[0.79]	[2.55]		
2	1.94%	0.29	1.01***	0.20***	0.07**	0.02	0.58	0.07
		[0.35]	[57.2]	[5.86]	[2.30]	[1.09]		
3	1.96%	1.02	1.00***	0.23***	0.09**	0.05**	0.64	0.23
		[1.09]	[43.4]	[6.74]	[2.09]	[2.55]		
4	1.94%	1.10	1.03***	0.27***	0.01	0.06**	0.63	0.22
		[1.14]	[46.7]	[6.98]	[0.26]	[2.49]		
5	1.88%	2.40**	1.07***	0.44***	−0.10***	0.08***	0.66	0.45
		[2.17]	[39.2]	[13.3]	[−2.86]	[3.30]		
1st Stage	9.55%	0.78	1.03***	0.28***	0.02	0.05***	0.62	0.22
		[1.07]	[61.2]	[10.5]	[0.68]	[2.99]		
All Funds	100%	−1.18**	1.01***	0.21***	0.02	0.01	0.48	−0.53
		[−2.39]	[81.9]	[10.7]	[1.17]	[1.00]		

This table documents the out-of-sample performances of the funds whose returns first order stochastically dominated the returns of the replica funds during the 24 months prior to portfolio formation. Specifically, by the end of each quarter, the empirical CDF of the percentile of each fund in the counterfactual return distribution $(F^{Pct(r_{i,t},\langle \hat{r}_{i,t}\rangle)}(x))$ is computed. The FSD test statistic $\hat{\theta}$ is then constructed for each fund. The first stage selects funds with $\hat{\theta} \geq 0.90$, which corresponds to a test size of 10%. The selected funds are then sorted into 5 quintiles based on their proceeding 24 months' four-factor alpha. The trading strategy is rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

Table 5: Out-of-sample Performance of Alternative Performance Measure: Alpha

Quintile	Sample Share	α (in %)	mkt	smb	hml	umd	SR	IR
		Before Fees						
1	1.90%	0.02	1.00***	0.35***	−0.03	0.04*	0.55	0.00
		[0.02]	[55.2]	[8.95]	[−1.00]	[1.77]		
2	2.02%	0.71	1.03***	0.34***	−0.07**	0.08***	0.60	0.19
		[0.91]	[57.3]	[11.8]	[−2.33]	[3.86]		
3	2.02%	0.81	1.06***	0.36***	−0.16***	0.06**	0.56	0.17
		[0.87]	[42.8]	[12.6]	[−3.52]	[2.44]		
4	2.02%	2.80***	1.04***	0.44***	−0.20***	0.04	0.64	0.53
		[2.64]	[38.0]	[10.1]	[−4.76]	[1.41]		
5	1.95%	1.25	1.09***	0.49***	−0.22***	0.05	0.55	0.22
		[1.03]	[35.1]	[12.8]	[−6.12]	[1.35]		
1st Stage	9.89%	1.13	1.05***	0.40***	−0.14***	0.05**	0.59	0.30
		[1.46]	[52.9]	[14.2]	[−4.44]	[2.32]		
All Funds	100%	0.04	1.00***	0.21***	0.02	0.01	0.56	0.02
		[0.07]	[81.4]	[10.7]	[1.17]	[0.95]		
		After Fees						
1	1.90%	−1.18	1.00***	0.35***	−0.03	0.04*	0.48	−0.31
		[−1.57]	[55.4]	[8.99]	[−1.01]	[1.78]		
2	2.02%	−0.57	1.04***	0.34***	−0.07**	0.08***	0.52	−0.15
		[−0.73]	[57.2]	[11.8]	[−2.31]	[3.88]		
3	2.02%	−0.50	1.07***	0.36***	−0.16***	0.06**	0.49	−0.11
		[−0.54]	[43.9]	[12.6]	[−3.54]	[2.47]		
4	2.02%	1.53	1.05***	0.44***	−0.20***	0.04	0.57	0.29
		[1.44]	[38.3]	[10.2]	[−4.78]	[1.41]		
5	1.95%	−0.08	1.09***	0.49***	−0.23***	0.05	0.48	−0.02
		[−0.07]	[35.1]	[12.9]	[−6.10]	[1.33]		
1st Stage	9.89%	−0.16	1.05***	0.40***	−0.14***	0.06**	0.52	−0.04
		[−0.21]	[53.4]	[14.2]	[−4.44]	[2.33]		
All Funds	100%	−1.18**	1.01***	0.21***	0.02	0.01	0.48	−0.53
		[−2.39]	[81.9]	[10.7]	[1.17]	[1.00]		

This table documents the out-of-sample performances of the funds selected by the four-factor alpha during the 24 months prior to portfolio formation. Specifically, by the end of each quarter, the top 10% funds with highest historical alpha are selected in the first stage. The selected funds are then sorted into 5 quintiles based on their proceeding 24 months' four-factor alpha. The trading strategy is rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

Table 6: Out-of-sample Performance of Alternative Performance Measure: Information Ratio (IR)

Quintile	Sample Share	α (in %)	mkt	smb	hml	umd	SR	IR
		Before Fees						
1	1.90%	0.71	0.94***	−0.04*	0.09***	0.01	0.61	0.27
		[1.29]	[75.7]	[−1.65]	[3.95]	[0.95]		
2	2.02%	1.78**	0.94***	0.13***	0.10***	0.01	0.68	0.48
		[2.29]	[48.9]	[4.49]	[3.00]	[0.47]		
3	2.02%	1.53**	0.97***	0.31***	0.02	0.07***	0.67	0.44
		[2.25]	[67.7]	[12.6]	[0.56]	[3.63]		
4	2.02%	2.54**	1.04***	0.35***	−0.16***	0.05*	0.65	0.51
		[2.55]	[38.8]	[8.64]	[−3.61]	[1.78]		
5	1.95%	1.98*	1.06***	0.48***	−0.19***	0.06*	0.61	0.39
		[1.82]	[37.0]	[14.1]	[−5.43]	[1.95]		
1st Stage	9.89%	1.69***	0.99***	0.25***	−0.03	0.04**	0.66	0.57
		[2.81]	[68.2]	[10.7]	[−1.30]	[2.27]		
All Funds	100%	0.04	1.00***	0.21***	0.02	0.01	0.56	0.02
		[0.07]	[81.4]	[10.7]	[1.17]	[0.95]		
		After Fees						
1	1.90%	−0.29	0.95***	−0.04*	0.09***	0.01	0.54	−0.11
		[−0.52]	[74.8]	[−1.67]	[3.81]	[0.80]		
2	2.02%	0.56	0.94***	0.13***	0.10***	0.01	0.60	0.15
		[0.72]	[49.7]	[4.45]	[3.01]	[0.47]		
3	2.02%	0.31	0.98***	0.31***	0.02	0.07***	0.60	0.09
		[0.45]	[67.9]	[12.6]	[0.59]	[3.71]		
4	2.02%	1.32	1.04***	0.35***	−0.16***	0.05*	0.58	0.26
		[1.32]	[39.1]	[8.65]	[−3.63]	[1.78]		
5	1.95%	0.67	1.07***	0.48***	−0.19***	0.06*	0.54	0.13
		[0.61]	[37.0]	[14.2]	[−5.41]	[1.92]		
1st Stage	9.89%	0.49	0.99***	0.25***	−0.03	0.04**	0.58	0.16
		[0.81]	[69.0]	[10.7]	[−1.30]	[2.26]		
All Funds	100%	−1.18**	1.01***	0.21***	0.02	0.01	0.48	−0.53
		[−2.39]	[81.9]	[10.7]	[1.17]	[1.00]		

This table documents the out-of-sample performances of the funds selected by the information ratio (IR) and four-factor alpha during the 24 months prior to portfolio formation. Specifically, by the end of each quarter, the top 10% funds with highest historical IR are selected in the first stage. The selected funds are then sorted into 5 quintiles based on their proceeding 24 months' four-factor alpha. The trading strategy is rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

Table 7: Characteristics of the FSD Selected Funds

Quintile	Sample Share	Age	Age Norm.	TNA Norm.	# of Stocks Norm.	Fees (in bps)	Fees Norm.	Turnover Ratio	Turnover Norm.
1	1.83%	15.16	0.93	0.66	1.05	129.36	1.05	0.84	1.04
2	1.94%	16.27	1.00	0.99	1.09	118.01	0.96	0.70	0.87
3	1.96%	15.08	0.93	0.88	1.02	120.78	0.98	0.72	0.89
4	1.94%	15.65	0.96	0.95	0.96	126.99	1.03	0.75	0.93
5	1.88%	15.79	0.97	0.88	0.79	136.46	1.10	0.83	1.03
1st Stage	9.55%	15.58	0.96	0.88	0.98	126.06	1.02	0.77	0.95
All Funds	100%	16.24	1.00	1.00	1.00	123.40	1.00	0.81	1.00

This table documents the characteristics of the funds selected by the FSD filter and compare them with the cross-sectional average. “Norm.” denotes the normalization procedure that takes the ratio of the corresponding variable with the cross-sectional average when all funds in the sample are included. “Sample Share” is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

Table 8: The Return Gaps

Quintile	Sample Share	α^{rgap} (in %)	mkt	smb	hml	umd
1	1.83%	0.78**	0.03***	0.03*	−0.01	0.02***
		[2.35]	[3.63]	[1.88]	[−0.92]	[3.64]
2	1.94%	0.79***	0.03***	0.02	0.00	0.02***
		[3.66]	[5.21]	[1.60]	[0.21]	[3.68]
3	1.96%	0.96***	0.04***	0.01	0.00	0.01***
		[3.57]	[5.41]	[0.61]	[0.20]	[2.96]
4	1.94%	0.72***	0.03***	0.01	0.00	0.01**
		[2.68]	[4.09]	[1.32]	[−0.10]	[2.20]
5	1.88%	1.63***	0.03**	−0.02	−0.03	0.03***
		[3.15]	[2.26]	[−0.64]	[−1.27]	[2.61]
1st Stage	9.55%	0.98***	0.03***	0.02	0.00	0.02***
		[4.75]	[5.17]	[1.22]	[−0.31]	[4.12]
All Funds	100%	0.18	0.01**	0.02	0.00	0.02***
		[0.93]	[2.17]	[1.62]	[0.58]	[4.03]

This table documents the return gaps of the FSD selected mutual funds. The return gap is defined as the difference between a mutual fund's actual return versus the hypothetical return generated by keeping the holdings within the mutual fund's portfolio by the end of the proceeding quarter. The time-series of the return gaps of different funds are then averaged within the corresponding quintiles and regressed against the Carhart four factors. The portfolios are rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

Table 9: Performance of the Mimicking Strategy

Quintile	Sample Share	α (in %)	mkt	smb	hml	umd	SR	IR
1	1.83%	0.38	1.06***	0.26***	0.06	−0.04*	0.54	0.07
		[0.34]	[38.0]	[6.26]	[1.46]	[−1.94]		
2	1.94%	1.25	1.07***	0.17***	0.06*	−0.06***	0.59	0.28
		[1.42]	[54.8]	[3.73]	[1.76]	[−3.14]		
3	1.96%	1.63	1.06***	0.24***	0.11**	0.00	0.64	0.32
		[1.58]	[39.9]	[5.15]	[2.19]	[−0.06]		
4	1.94%	1.50	1.09***	0.29***	−0.01	−0.01	0.60	0.28
		[1.45]	[45.0]	[6.02]	[−0.20]	[−0.31]		
5	1.88%	2.24**	1.15***	0.43***	−0.11***	0.00	0.61	0.40
		[1.97]	[37.5]	[9.88]	[−3.00]	[0.04]		
1st Stage	9.55%	1.33	1.09***	0.27***	0.02	−0.02	0.60	0.32
		[1.64]	[53.6]	[6.87]	[0.68]	[−1.30]		
All Funds	100%	0.28	1.09***	0.18***	0.00	−0.08***	0.52	0.10
		[0.49]	[66.5]	[5.44]	[−0.03]	[−6.69]		

This table documents the performance of the trading strategy that aims to mimic the performances of the FSD selected mutual funds. To ensure implementability, by the end of each quarter, the stock holdings from the end of the previous quarter are retrieved for the managers who have been identified as skilled. The trading strategy then invests in the stocks that the managers were holding as of the end of the previous quarter. The portfolios are rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

Table 10: Flow Responses

	$Flow_{i,t}$	$Flow_{i,t}$	$Flow_{i,t}$	$Flow_{i,t}$
$\hat{\alpha}_i^{[t-1-T,t-1]}$	2.58***	2.38***	2.28***	2.27***
	[40.0]	[37.7]	[37.5]	[37.4]
$FSD_{i,t-1}$ ($\times$100)		0.94***		0.46***
		[17.8]		[7.41]
$FSD_{i,t-1} \times \hat{\alpha}_i^{[t-1-T,t-1]}$			1.85***	1.12***
			[16.5]	[7.96]
Controls	Y	Y	Y	Y
# of Obs	227710	227710	227710	227710

This table documents the Fama-Macbeth regression results of

$$Flow_{i,t} = Const + \delta_{\mathbf{0}} \times FSD_{i,t-1} + (\beta + \delta_{\mathbf{1}} \times FSD_{i,t-1}) \times \hat{\alpha}_i^{[t-1-T,t-1]} + X_i + \epsilon_{i,t}.$$

The dependent variable is the flow of each fund in every month, and the independent variables include the trailing in-sample alpha, the dummy variable corresponding to the FSD condition, and the interaction between the two. The control variables include fund age, log fund size, fees, and the number of stocks in the portfolio. The sample period is from January 1991 to December 2015.

Table 11: Out-of-sample Performance of the Alternative FSD test statistic

Quintile	Sample Share	α (in %)	mkt	smb	hml	umd	SR	IR
		Before Fees						
1	1.74%	0.97	1.02***	0.24***	0.05	0.04*	0.62	0.22
		[1.06]	[49.2]	[7.81]	[1.26]	[1.96]		
2	1.86%	1.30	1.02***	0.20***	0.05*	0.02	0.63	0.30
		[1.48]	[54.1]	[5.76]	[1.73]	[1.08]		
3	1.85%	2.73***	0.98***	0.17***	0.11**	0.03	0.75	0.60
		[2.93]	[38.5]	[4.23]	[2.37]	[1.40]		
4	1.86%	1.30	1.02***	0.28***	0.04	0.05**	0.64	0.26
		[1.33]	[46.4]	[7.73]	[1.03]	[1.99]		
5	1.78%	3.53***	1.06***	0.42***	−0.10***	0.07***	0.72	0.68
		[3.26]	[38.7]	[12.5]	[−2.97]	[2.81]		
1st Stage	9.08%	1.97***	1.02***	0.26***	0.03	0.04**	0.69	0.53
		[2.62]	[57.0]	[9.31]	[0.95]	[2.42]		
All Funds	100%	0.04	1.00***	0.21***	0.02	0.01	0.56	0.02
		[0.07]	[81.4]	[10.7]	[1.17]	[0.95]		
		After Fees						
1	1.74%	−0.30	1.03***	0.24***	0.05	0.04**	0.55	−0.07
		[−0.32]	[49.0]	[7.77]	[1.24]	[1.99]		
2	1.86%	0.06	1.03***	0.20***	0.05*	0.02	0.56	0.01
		[0.07]	[54.4]	[5.75]	[1.71]	[1.05]		
3	1.85%	1.49	0.99***	0.17***	0.11**	0.03	0.67	0.33
		[1.59]	[38.5]	[4.22]	[2.36]	[1.44]		
4	1.86%	0.02	1.02***	0.28***	0.04	0.05**	0.56	0.00
		[0.02]	[46.4]	[7.71]	[1.02]	[2.01]		
5	1.78%	2.21**	1.07***	0.42***	−0.10***	0.07***	0.65	0.43
		[2.04]	[38.9]	[12.5]	[−2.99]	[2.84]		
1st Stage	9.08%	0.70	1.03***	0.26***	0.03	0.04**	0.61	0.19
		[0.93]	[57.1]	[9.28]	[0.92]	[2.43]		
All Funds	100%	−1.18**	1.01***	0.21***	0.02	0.01	0.48	−0.53
		[−2.39]	[81.9]	[10.7]	[1.17]	[1.00]		

This table documents the out-of-sample performance of the filter constructed with the alternative FSD test statistic ($\max_{x\in[0,1]}\left(\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)-x\right)$). Specifically, by the end of each quarter, the empirical CDF of the percentile of each fund in the counterfactual return distribution ($F^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$) is computed. The alternative FSD test statistic $\hat{\theta}$ is then constructed for each fund. The first stage selects funds with $\hat{\theta} \leq 0.039$, which corresponds to a test size of 10%. The selected funds are then sorted into 5 quintiles based on their proceeding 24 months' four-factor alpha. The trading strategy is rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.