Rutgers Business School
Newark and New Brunswick

MQF Research: Hedge Fund Alphas Test

Jiewei Wu

Professor: Yangru Wu

2021 Spring

Abstract

In light of the paper 'Thousands of Alpha Tests' [2] written by professor Yuan Liao, Dacheng Xiu, and Stefano Giglio, I'm interested in the various approaches and techniques applied and would like to replicate some of the results in the paper under the supervision and help of Professor Yangru Wu. The ultimate goal is to find reliable funds which could generate steady alpha compared to the benchmark while ruling out the data snooping bias. I'm going to focus on the empirical studies instead of the DGP, which are used by the authors to validate the powerness of their improvements.

Keywords: Multiple Testing, Asset Pricing Models, Risk Premium, Factor Analysis, Screening, B-H Method, False Discovery Rate, Matrix Completion, Wild-Bootstrap, Fama-Mecbeth Regression, Fama-French Factors

1. Introduction

When dealing with empirical finance problems, researchers will always encounter multiple testing. Rather then testing one or a small number of null hypothesis, situations like identifying among hundreds of alphas which add explanatory power to the existing models and telling which funds are able to generate significant positive alphas require multiple testing. However, problems come with multiple testing, closely related to data snooping [4], that when tests' number becomes large, some of the rejections from the hypothesis may be purely due to chance and construct a relative high portion of all rejections. This proportion is defined as "false discovery proportion"(FDP). This problem becomes even more serious if the patterns we are going to identify are rare compared to the full sample such that the FDP can even be 100% in extreme cases, which implies all rejections take place because of pure luck.

To solve the data snooping problem comes with multiple testing [2], I use the B-H procedure which will strictly control the "false discovery rate"(FDR), which is the expectation of FDP. FDR is associated with the underlying true data generating process and thus not observable. However, one can use the average of FDP in simulated DGP to estimate FDR. Rather than being too conservative at traditional B-H procedure in rejecting then null, I also use the screening procedure before B-H. This procedure will eliminate a set of extremely unskilled funds that are "deep in the null" and increases the proportion of funds selected.

In addition to this, in my work to identify the funds that could generate significant alphas, I make two more contributions. First, I adopt the iterative matrix completion algorithm to recover missing entries in the matrix of asset returns. An application of this algorithm is introduced to the famous "Netflix Prize" problem: Given a extremely sparse matrix in which columns and rows represent the users and movies and the matrix elements show ranking of user to the movies, the competitors are required to recover the full matrix. This problem can be solved based on the assumption that users have only a small number of reasons combined to make the ranking, which implies the low-rank characteristic of the matrix. In my work, I make the same assumption that the matrix is low-rank because hedge fund returns can be possibly explained by common factors and anomalies, the number of which is inferior to the matrix size.

Second, to prove significance of hedge fund performance we select, I use a wild-bootstrap for multiple testing of alphas, which can be proved robust to missing factors and missing entries. Instead of resampling the hedge fund return like the traditional bootstrap, I bootstrap the residuals in the asset pricing model. This can better account for the influence of common factors to the fund returns and restrict the uncertainties of simulations on the alphas alone. It can be proved using wild-bootstrap would control the FDR much better the traditional bootstrap. As shown from the simulations, the wild-bootstrap also outperforms the asymptotic approach.

2. Material and Methods

2.1 Data Cleaning

To illustrate the effectiveness of our methodology, I introduced the hedge fund monthly returns from Lipper hedge fund dataset, covering the time period from 1994 to 2020. I'm going to use the data from 1994 to 2018 to select the funds with significant alpha and data from 2019 to 2020 to simulate the returns of our selected funds, comparing with the average performance and benchmark performance. To get a better estimator of the risk premium and more robust matrix completion outcome, I only focus on the funds that had a relatively long continuous time series, which is equal or greater than 36 months. To avoid model distortion by outliers and improve robustness, I also dropped the funds which have returns greater than 100% and lower than minus 50%. I denote N as the number of funds we finally selected, which is 2000 and T as the number of periods in in-sample data, which is 300. The fund returns data are cleaned to a matrix with N by T matrix with approximate 70% missing entries.

In my empirical test, I'm also trying to rule out the possible common factors that joinly affect the fund returns, either observable or latent. Thus the individual risk premium, which is alpha, is obtained as the cross-sectional regression intercept in the asset pricing model. As for as the observable factors, we follow the commonly used Fama-French five factor model and also add three factors in terms of reversal and momentum. Finally I selected eight observable factors, which is market, size, value/growth, profit, investment, short-term reversal, long-term reversal, and momentum. I reorganized the data in a matrix K by T, in which I denote K as the number of factors and T as the periods.

2.2 Calculate Factor Returns and Factor Loadings

To rule out the common factors which contributes to the factor returns, I introduce the Fama-Mecbeth two-step regression, for the purpose of accounting for the non-tradable factors. Even though the risk premium for the tradable factors can be estimated using the arithmetic mean, it's not applicable in our case with non-tradable factors. In addion to this, this process is separated in to two steps for both observable and latent factors. For the obersavle factors, the factors are given and I use time series regression to calculate the factor loadings. For the latent factors, I apply singular value decomposition(SVD) with iterative singular value threshold algorithm to solve the regularized matrix completion optimzation problem with regard to the residual matrix with missing entries. The factors and factor loadings for latent factors can thus be derived from two more regress steps to rule-out the missing value bias [1].

First, Obtain $\widehat{\beta}_o$ and the residual matrix $Z = (Z_{i,t})_{N*T}$:

$$\widehat{\beta}_{o,i} = (F_{o,i}F'_{o,i})^{-1}(F_{o,i}R_i) \tag{2.1}$$

$$z_{i,t} = r_{i,t} - \bar{r}_i - \widehat{\beta'_{o,i}}(f_{o,i} - \bar{f_{o,i}}) \tag{2.2}$$

where $F_{o,i}$ denotes the factor matrices, R_i is the fund returns with missing values, $\bar{r}_i = \frac{1}{T_i}\sum_{i\in\mathscr{T}_i} r_{i,t}$ is the average return of for fund i at its observable time points,$\bar{f_{o,i}} = \frac{1}{T_i}\sum_{i\in\mathscr{T}_i} f_{o,t}$ is the average of observable factors over time in which fund i is observed.

Next, let's formulate the matrix completion optimization problem. The goal is to recover an $N \times T$ low-rank matrix X. Z can be written as $X+E$, in which E represents the noise. Moreover, because Z is not fully observed, I use Ω as indicator for the missing entries. I employ the nuclear-norm penalized regression approach to recover X:

$$\widehat{X} = \arg\max_{X} \|(Z-X)\times\Omega\| + \lambda_{NT}\|X\| \tag{2.3}$$

where $\|X\|$ denotes the matrix nuclear norm and $\lambda_{NT} > 0$ is a tuning parameter.Then, I go iterative algorithm as follows to finish matrix completion.

1. Start with initial guess for missing values in Z
2. Soft threshold by some threshold a: apply SVD of Z giving $Z = UDV^t$, where D denotes the singular values by d_i. Define $S_a(Z) = UD_aV^t$, where D_a is the diagnal matrix with entries $max(d_i - a, 0)$
3. Calculate the difference between Z and $S_a(Z)$
4. Repeat until convergence

I first obtain β_0 from time series regressions using observable factors alone, and then obtain β_1 by applying SVD to the residual matrix from time series regressions. The estimated β_0 and β_1 are stacked together as β. As I have done in the matrix completion process, I have already got $\widehat{\beta}_0$ and a low rank matrix $\widehat{X}$. Then I estimate the latent factors and their loadings using $\widehat{X}$ as,

$$\widehat{v}_{l,t} = (\sum_{i\in\mathcal{N}_t} b_i b_i')^{-1} \sum_{i\in\mathcal{N}_t} b_i z_{it}, \quad t = 1,\dots,T \tag{2.4}$$

$$\widehat{\beta}_{l,i} = (\sum_{t\in\mathcal{T}_i} \widehat{v}_{l,t}\widehat{v}_{l,t}')^{-1} \sum_{t\in\mathcal{T}_i} \widehat{v}_{l,t} z_{it}, \quad i = 1,\dots,N, \tag{2.5}$$

where $(b_1,\dots,b_{K_l})$ is the top K_l left singular-vectors of $\widehat{X}$. Thus the resulting $\widehat{\beta}$ can be given by,

$$\widehat{\beta} = (\widehat{\beta}_o, \widehat{\beta}_l) \tag{2.6}$$

2.3 Estimate Factor Risk Premium and Alpha

In this step I would get the estimated alpha of each fund. As discussed in some previous works, I can estimate the factor risk premium using only the time series arithmetic mean only in the case in which all factors are excess returns. However, when the following benchmark model includes non-tradable factors, estimating the alphas of asset or factor returns requires two-pass Fama-MacBeth regressions based on the following equation. The first stage estimates β using time series regressions of individual fund returns onto the benchmark factors, and the second stage involves a cross-sectional regression of average returns onto the estimated β, where the residuals of this regression yield estimates of alpha, denoted as $\hat{\alpha}$.

$$r_t = \alpha + \beta\lambda + \beta\left(f_t - \mathbb{E}\left(f_t\right)\right) + u_t \tag{2.7}$$

where f_t is a $K*1$ vector of factors and u_t is the idiosyncratic component. The parameter λ is a $K*1$ vector of factor risk premium. Hence in this step, I first using time series regressions of individual fund returns onto the benchmark factors to estimate β.

Specifically, in addition to the observed factors in the benchmark model, some important factors are omitted, thus attributing to alpha what truly is just exposure to the omitted risk factors. Therefore, in general cases the above equation can be written as,

$$r_t = \alpha + \begin{bmatrix} \beta_o & \beta_l \end{bmatrix} \begin{bmatrix} \lambda_o \\ \lambda_l \end{bmatrix} + \begin{bmatrix} \beta_o & \beta_l \end{bmatrix} \begin{bmatrix} f_{o,t} - \mathbb{E} f_{o,t} \\ f_{l,t} - \mathbb{E} f_{l,t} \end{bmatrix} + u_t, \tag{2.8}$$

where $f_{0,t}$ is a the vector of observable factors, and $f_{1,t}$ is the vector of latent factors.

In order to deal with the latent factors, I need to adopt different methods to estimate β. Following the method proposed by [3], when there are only latent factors, I estimate $\hat{\beta}_l$ as following using the plain vanilla PCA,

$$\hat{\beta}_l = \sqrt{N}(b_1, \ldots, b_K) \tag{2.9}$$

where $b_1, \ldots, b_K$ are the largest K eigenvectors of the $N*N$ sample covariance matrix of the residual matrix of regression adopted on the linear model.

Furthermore, for dealing with unbalanced panels, I use a singular value decomposition (SVD) based matrix completion method to get a low rank matrix $\hat{X}$ by decomposing the residual matrix Z and use the singular vectors of X to estimate β, which I have introduce IN DETAIL in (2.4) and (2.5).

In the next step, we will run a cross-sectional regression of $\bar{r}$ to $\widehat{\beta}$, which is estimated by the last step, and a constant regressor 1_N to get the slope $\widehat{\lambda}$.

$$\widehat{\lambda} = (\widehat{\beta}' \mathbb{M}_{1_N} \widehat{\beta})^{-1} (\widehat{\beta}' \mathbb{M}_{1_N} \bar{r}) \tag{2.10}$$

With the slope $\widehat{\lambda}$, I could estimate α by subtracting the estimated risk premium from average returns by $\widehat{\alpha} = \bar{r} - \widehat{\beta}\widehat{\lambda}$. To be specific, in this estimation there would be some unbalanced panels bias since the residual matrix is not fully observed. For dealing with those panels, we should introduce an additional term $\widehat{A_i}$ to de-bias the estimated alphas due to unbalanced panels. By writing $\widehat{\xi}_i' = e_i' - \widehat{\beta}_i'(\widehat{\beta}'\mathbb{M}_{1_N}\widehat{\beta})^{-1}\widehat{\beta}'\mathbb{M}_{1_N}, e_i' = (0, \ldots 0, 1, 0, \ldots, 0), \widehat{g}_i = \frac{1}{T_i}\sum_{t \in \mathscr{T}_i} \widehat{v}_t' \widehat{\beta}_i, \widehat{H}_{o,i} = \widehat{V}_{l,i}\mathbb{M}_{1_{T_i}} F_{o,i}'(F_{o,i}\mathbb{M}_{1_{T_i}} F_{o,i}')^{-1}$, and $\widehat{H}_o = \widehat{V}_l \mathbb{M}_{1_T} F_o' (F_o \mathbb{M}_{1_T} F_o')^{-1}$, then calculate $\widehat{A_i}$ as following,

$$\widehat{A_i} = \widehat{\beta}_{l,i}' (\widehat{H}_{o,i} - \widehat{H}_o) \widehat{\lambda}_o - \widehat{\xi}_i' \widehat{g} \tag{2.11}$$

Thus, α could be estimated by,

$$\widehat{\alpha}_i = \bar{r}_i - \widehat{\beta}_i' \widehat{\lambda} + \widehat{A}_i, \quad i = 1, \ldots, N \tag{2.12}$$

2.4 Construct Inference Statistics

After we estimated the alpha, we need to construct inference statistics and this can be done in two ways.First is the asymptotic methods: the normalized alpha under the estimated standard deviation will prove to be follow standard normal distribution. [2] The second approach is the through bootstrap. After bootstrap the residual returns after asset pricing regression, we'll construct N, which is the bootstrap size, alphas thus I can use to construct inference results.

2.4.1 Asymptotic method

Regarding the Asymptotic method, we need rely on the the results of regression steps. The detailed construction process is presented as follows:

The standard error was calculated as follows:

$$se(\widehat{\alpha}_i) = \frac{1}{\sqrt{T_i}} \sigma_i \widehat{\sigma}_i \tag{2.13}$$

The estimated σ was given by:

$$\widehat{\sigma_i^2} = \frac{1}{T_i} \sum_{t \in \mathcal{T}_i} \widehat{u_{i,t}}^2 (1 - \widehat{v}_t' \widehat{\sum_f}^{-1} \widehat{\lambda})^2 \tag{2.14}$$

where $\widehat{u}_{i,t} = r_{i,t} - \bar{r}_i - \widehat{\beta}_i' \widehat{v}_t$ is the residual, and $\widehat{\sum_f} = \frac{1}{T} \sum_{t=1}^{T} \widehat{v}_t \widehat{v}_t'$.

What comes natural is: the t-statistics and p-values can be constructed as below formulas:

$$t_i = \frac{\widehat{\alpha}_i}{se(\widehat{\alpha}_i)} \tag{2.15}$$

$$p_i = 1 - \Phi(t_i), \quad i = 1, \ldots, N, \tag{2.16}$$

where$\Phi(\cdot)$ is the Gaussian cumulative distribution function.

This method is quite straightforward, which means that it mainly based on the results of regression steps and the $_i$ here could also be replaced by the two sided one that is $p_i = 2(1 - \Phi(|t_i|))$. What's more, this method could be applicable in both unbalanced panel and balanced panel. However, when it comes to the scenario that a large amount of data are missing, the asymptotic approach may lose its power.

More specifically, the conclusion derived from the asymptotic approach based on two main lemmas:

$$\sigma_{i,NT}^{-1} (\widehat{\alpha}_i - \alpha_i) \xrightarrow{d} N(0,1), \quad \sigma_{i,NT}^2 = \frac{1}{T} Var\left(u_{i,t}\left(1 - v_t^{-1} \textstyle\sum_f^{-2} \lambda\right)\right) + \frac{1}{N} Var(\alpha_i) \frac{1}{N} \beta_i' S_\beta^{-1} \beta_i, \tag{2.17}$$

When $N, T \to \infty$. However as the author proved, the asymptotic inference is not robust to a lot of missing data. On the contrary, the wild bootstrap can handle missing data well while also maintaining strong performance and validity.

2.4.2 Wild-Bootstrap

In this part, we introduce Bootstrapping p-values procedure [2], which can be used a method to effectively deal a large amount of data missing scenario.

We resample the residuals of fund return, that is essentially we generate bootstrap sample of fund residulas after the previous regression steps:

$$r_{i,t}^{\star} = \widehat{\beta}_i' \widehat{\lambda} + \widehat{\beta}_i' \widehat{v}_t + \widehat{u_{i,t}^{\star}}, \quad \widehat{u_{i,t}^{\star}} = \widehat{u_{i,t}} w_{i,t}, \quad for \quad t \in T_i \tag{2.18}$$

where the$w_{i,t}$ is the i.i.d. random variables, such that $\mathbb{E} w_{i,t} = 0$ and $Var(w_{i,t}) = 1$.

The loading of factors $\widehat{\beta^\star} = (\widehat{\beta_1^\star}, \ldots, \widehat{\beta_N^\star})$is obtained in the same way as before.

$$\widehat{\beta^\star} = (\widehat{V}_i \mathbb{M}_{1,T_i} \widehat{V}_i') (\widehat{V}_i \mathbb{M}_{1,T_i} \widehat{R}_i^{\star}) \tag{2.19}$$

The bootstrap risk premium $\lambda^\star$ can be obtained by doing the same process stated in matrix completion part.

In this step, we are going to estimate alpha and de-bias the results we get. The de-bias process is also the same in (2.12):

$$\widehat{\alpha_i^\star} = \bar{r}_i^\star - \widehat{\beta_i^\star}'\widehat{\lambda^\star} - \widehat{\xi_i^\star}'\widehat{g^\star}, \quad i = 1,\ldots,N, \tag{2.20}$$

$$\widehat{\xi_i^\star}' = e_i' - \widehat{\beta_i^\star}'\left(\widehat{\beta^\star}'\mathbb{M}_{1N}\widehat{\beta^\star}\right)^{-1}\widehat{\beta^\star}'\mathbb{M}_{1N} \tag{2.21}$$

$$e_i' = (0,\ldots,1,0,\ldots) \tag{2.22}$$

$$\widehat{g_i^\star} = \frac{1}{T_i}\sum_{t\in T_i}\widehat{v_t}'\widehat{\beta_i^\star} \tag{2.23}$$

Based on above steps, we can calculate the critical p-value in an empirical way. For each fund i,we repeat K times generating bootstrap alpha value for that fund,that is $\widehat{\alpha_{i,n}^\star}$ and$1 \leq n \leq K$.Therefore, the p-value could be gauged in Wild-Bootstrap way as:

$$p_i = \frac{1}{K}\sum_{n=1}^{K} 1\left\{\widehat{\alpha_{i,n}^\star} > \widehat{\alpha}_i\right\}, \quad i = 1,\ldots,K \tag{2.24}$$

where the $\widehat{\alpha}_i$ is given by (2.12).

2.5 Alpha Screen B-H Procedure

Through steps in section 2.3 and 2.4, we finally come to the stage of selecting funds. After obtaining the inference statistics through either the asymptotic or bootstrap method, I got a series of p-value for each fund. How am I going to decide which fund generates significant alpha? It's an easy question for one unique fund. In the case of multiple testing scenarios, we need to take into account the data snooping problem, which can be properly address by the traditional B-H procedure through controlling the FDR. In addition to this, we also take the idea of alpha screen into practice to refuse considering the p-values that are deep in the null so that a less conservative rejection result can be obtained.

2.5.1 Alpha Screen

The idea is that when some of the alphas are ”overwhelmingly negative”(which we call ”deep in the null”), it's safe to simply accept the null hypothesis with tiny probability of make the type II error. This would reduce the denominators in the B-H process and thus increase the reject horizon for p-values, leading more rejections. Based on this idea, I select the set of funds in the following sets:

$$\widehat{\mathscr{I}} = \left\{i \leq N : t_i > -\log(\log T)\sqrt{\log N}\right\} \tag{2.25}$$

2.5.2 B-H procedure

B-H procedure [2] can be proved to be the solution to identify the critical p-value and control False Discovery Rate. B-H procedure is shown below:

Step 1. Sort in ascending order the collection of p-values,$\{p_i : i = 1, ..., N\}$, of the individual test statistics $\{t_i\}$. Denote $p_{(1)} \leq ... \leq p_{(N)}$ as the sorted p-values.

Step 2. For $i = 1, ..., N$, reject $\mathbb{H}_0^i$ if $p_i \leq p_{(\widehat{k})}$, where $\hat{k}$ = max$\{i \in I : p_{(i)} \leq \tau i/N\}$.

Accept all other $\mathbb{H}_0^i$.

2.5.3 B-H procedure with Alpha Screening

We can also combine the B-H procedure with Alpha Screening by replacing the set of funds to be test with the filtered set of funds by Alpha Screening. Suppose I is the set of filtered funds. Then we have:

Let $\left|\widehat{I}\right|$ denote the number of elements in I:

Step 1. Sort the p-values,$\{p_i : i = 1, ..., N\}$ for $\{p_i : i \in I\}$

Step 2. For $i \in I$, reject $\mathbb{H}_0^i$ if $p_i \leq p_{(\widehat{k})}$, where = max$\{i \in I : p_{(i)} \leq \tau i/|\widehat{I}|\}$.

Accept all other $\mathbb{H}_0^i$.

2.5.4 Realization

In this project, I use both procedures to select funds without missing the chance to analyze the effectiveness of alpha screening in practice (since the theoretical effectiveness has been proved [2]). After I calculate the p-value under Bootstrap method or Asymptotic method,I rank the p-values in an ascending order and find the critical p-value by using the prespecified value τ (I use $\tau = 0.05$ for all tests). From the first fund to the last(N^{th}) one in the ascending order, the i^{th} fund with p-value less than $i\tau/N$ will be rejected by null hypothesis and also be noted as "1" in both methods. Finally, there will be two series of 0/1 series, which represent the result of my test.

3. Results and Discussion

3.1 Hedge Fund Return Factor Structure

The left plot of Figure 3.1 shows the average monthly return of the 2000 full hedge fund sample. It can be seen that the histogram is slightly skewed to the right, which shows a greater percentage of the hedge funds could generate profits than those who lose money.

The right plot reports the first 15 eigenvalues of the co-variance matrix of excess returns. The ”latent” shows the eigenvalues of the raw returns while ”8 factor” outlines the case after time series regression to rule out the eight factors. Unlike in the paper where the eigenvalues shrink substantially after accounting for the impact of market return using CAPM model, we show a relatively insignificant influence of observable factors on the cross sectional return [2].

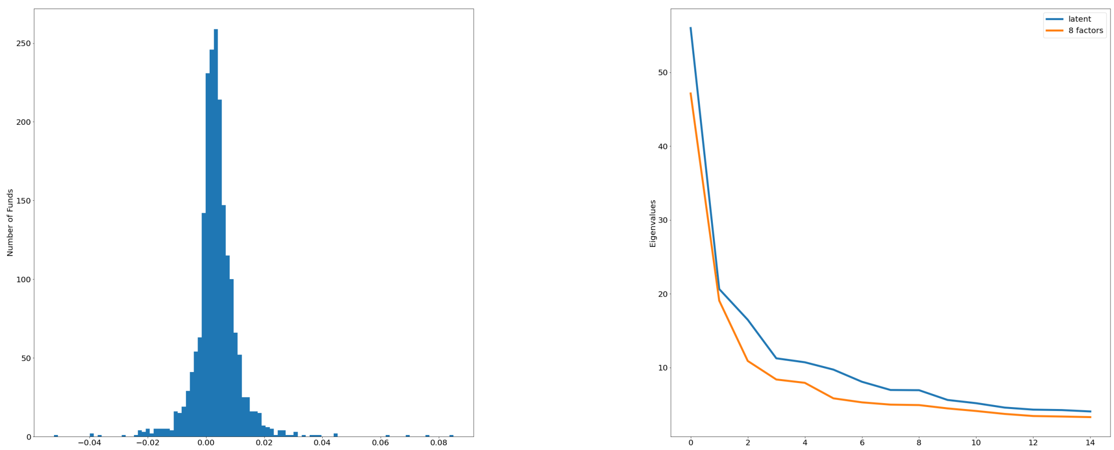

Figure 3.1: Properties of Hedge Fund Excess Returns

3.2 In sample Analysis

In this section, I aim to compare the funds selected by the adopted FDR controlled method with those selected by some other different methods. The following Table 3.1 shows the results of funds selection. To be specific, the first row reports the average alpha of selected funds and the second row shows the max alpha. Besides, the third row reports the t-stat of those alphas, while the last row shows the fractions of selected funds in the whole set of funds.

As for the columns, different levels of comparisons are also shown in the table. Overall, I want to compare the performance of selected funds after one dimension of technique changed while keeping others the same. The bootstrap and asymptotic results is compared under the mixed FDR control level. Moreover, I also make comparisons with regard to the number of observable and latent factors selected and their impact on the overall results of our fund selection. In addition to this, the results after selection processes with or without screening before the B-H procedure are also shown in the data for us to compare.

Table 3.1: In-sample results

	Mixed FDR		Only observable			No screen	No screen	Only latent
	Bootstrap	Aver.	FDR	p<0.05	p>0.05	FDR	observ. FDR	FDR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average alpha	11.4	2.5	2.7	1.8	0.4	2.4	2.6	2.6
Max alpha	21.1	7.3	14.2	14.2	7.3	7.7	14.1	8.6
Average t-stat	1.03	1.60	1.38	1.06	0.47	1.49	1.37	1.43
Fraction selected	37.3	15.3	37.3	38.3	61.7	25.3	24.2	29.7

Note: The table reports the results of the multiple alpha tests for the 1,761 hedge funds in our sample, using different methodologies. The first row reports the average alpha of the funds selected; the second row the average t-stat of the funds selected; the third row the fraction of funds selected; the fourth row the p-value of the test that the average alpha is equal to zero; the fifth row the p-value of the test that the average alpha is equal to the one obtained using the full FDR methodology (first column). Each column corresponds to a different selection procedure: (1) Our FDR with 5 observable factors (FH7) and 3 latent factors, and bootstrap standard errors; (2) The same as (1) using asymptotic standard errors; (3) FDR with only 8 observable factors (FH7); (4) Funds with individual p-value below 0.05, using only observable factors; (5) Funds with individual p-value above 0.05, using only observable factors and alpha-screening; (6) Our FDR, without alpha-screening; (7) Standard FDR (with only observable factors) without alpha-screening; (8) Our FDR, with 8 latent factors and no observable factors.

In the table above, compare (1) and (2), we can see a clear outperformance of the bootstrap inference than the asymptotic inference. However, I also found this is because several of the selected funds have significant alphas than the others that impacted the overall result.

From the comparison in (3)(4)(5), we can clearly see the effectiveness of our FDR control process. Through the plain vanilla selection of p-values that are smaller than 0.05, the average t-statistics is only 1.06, which is not significant. The average alpha of those filtered through greater than 0.05 p-values is only 0.4, substantially less than the FDR control.

After comparing (6)(7) with the previous, the alpha screen procedure proved its validity because the fraction selected in (6)(7) is substantially less than the (1) while keeping the similar performance. Alpha screen provides a more robust and less conservative selection process on the B-H procedure by punishing the denominator as we show in the Material and Methods.

The performance of method(7)(8), generally speaking unlike what the author proved in his thesis, are proven to be not that important in my analysis. The ruling out of observable or latent factors have similar performance in the final outcome. This is partially because the latent factors captures most of the observable factor patterns in a implicit way.

3.3 Out sample Analysis

After I have selected the funds I need through the FDR control method, it's naturally think how my selected funds would behave in the outsample analysis. The tail 24 months of data is used as the outsample data, which is decided in the very beginning.After plotting the return distribution of each fund on each day, I then have a rough understanding of the quality of my selection.

After selecting funds by our criteria from our in sample set, we check the performance of these selected funds in the next 24 consecutive months. Below, I draw the average performance of these funds through the time and the cross sectional distribution of these funds. The orange

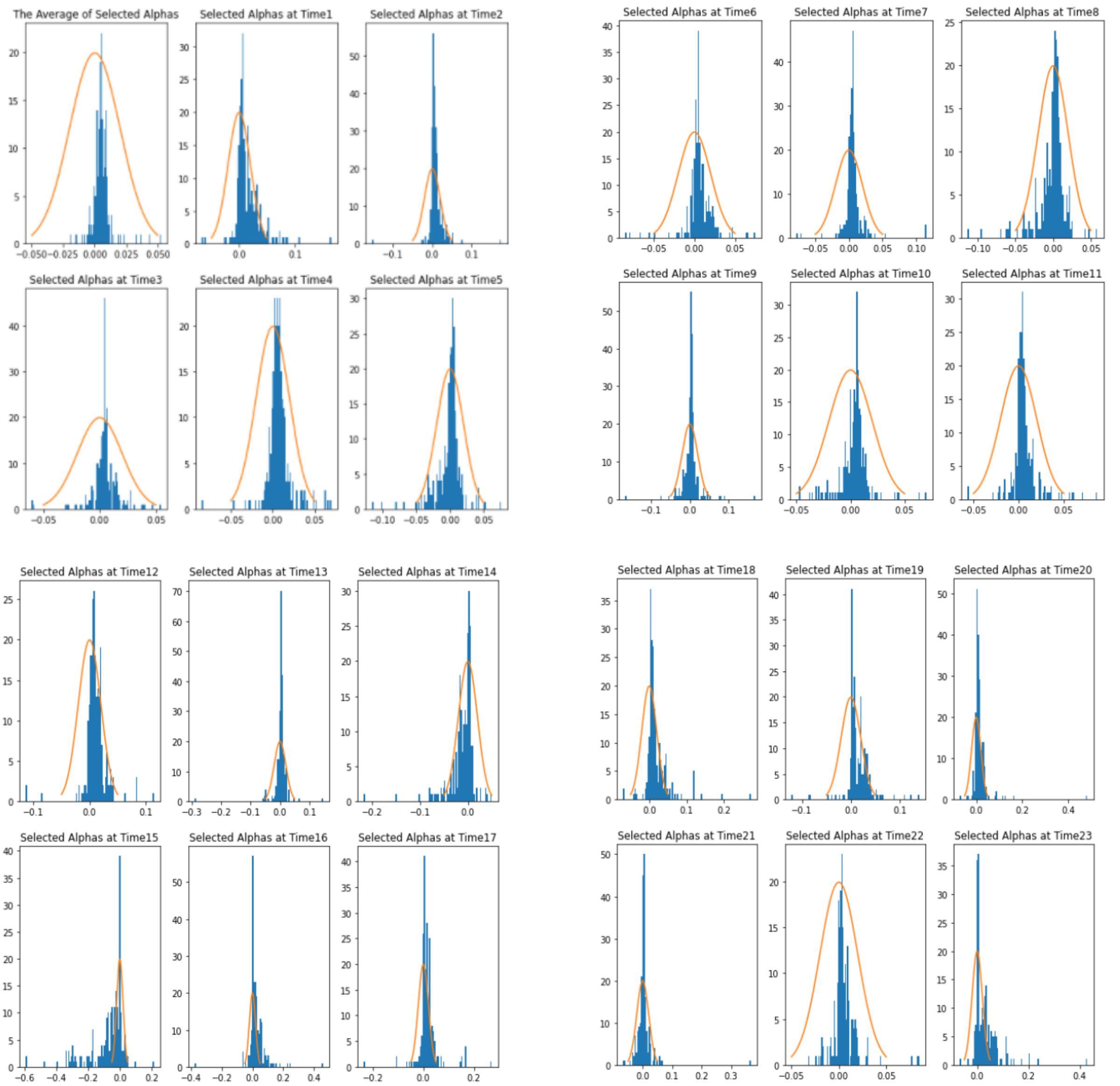

Figure 3.2: Performance of Selected Alphas in Out Sample Set

line show s the standard normal distribution which is aligned as a scale of our outsample performance.

Out of 24 plots, only 5 of the plots show a left-skewd distribution, while the others are all shown to generate a positive return. However the return given is still highly unstable, which shows our selection process is still not robust in the out sample.Nonetheless, I still think this a very big step for the researcher to step inside in to the mechanism of hedge funds to generate significant alphas. The last few years have seen a burgeoning strand of literature on the applications of machine learning techniques to high dimensional problems in asset pricing, in which data snooping leads to potentially numerous false discoveries. Here I replicates a way in the thesis to rigorously account for the data snooping bias, taking into account explicitly the specific properties of the finance context to which it is applied.

As for the future work, there remain many other settings in which our high-dimensional multiple-testing framework can be applied: for example, the evaluation of multiple potential new factors against an existing asset pricing model. .

Bibliography

[1] COCHRANE, J. Asset pricing: Revised edition.

[2] GIGLIO, S., LIAO, Y., AND XIU, D. Thousands of alpha tests. *Chicago Booth Research Paper*, 18-09 (2020), 2018–16.

[3] GIGLIO, S., AND XIU, D. Asset pricing with omitted factors. *Chicago Booth Research Paper Series* (2019).

[4] HALBERT, W. A reality check for data snooping. *Econometrica*, 5 (2000), 1097–1126.