Factor Selection and Factor Strength

An Application to U.S. Stock Market Return

Research Plan

Zhiyuan Jiang

I.D:28710967

Supervisor : Dr Natalia Bailey

Dr David Frazier

July 24, 2020

Contents

1 Introduction and Motivation

Capital Asset Pricing Model (CAPM) (Sharpe (1964), Lintner (1965), and Black (1972)) introduces a risk pricing paradigm. The model divided asset's risk into two parts: systematic risk and asset specified idiosyncratic risk. Researches (see Fama and French (1992), Carhart (1997), Kelly, Pruitt, and Su (2019)) has shown that, by adding new factors into the CAPM model, the multi-factor CAPM outperform the initial model in risk pricing. Because of this, factor identification becomes an important topic in financial. Harvey and Liu (2019) had collected over 500 factors from papers published in the top financial and economic journals, and they found the growth of new factors speed up since 2008.

But we should notice that not all factors can pass the significant test comfortably almost every time like factors in three-factor model (Fama & French, 1992). Pesaran and Smith (2019) provide a criteria called factor strength to measure such discrepancy. In general, if a factor can generate loading significantly different from zero for all assets, then we call such a factor a strong factor. And the less significant loading a factor can generate, the weaker the strength it has.

In his 2011 president address Cochrane emphasis the importance of finding factor which can provides independent information about average return and risk. With regard of this, a number of scholars had applied various methods to find such factor. Harvey and Liu (2017) provided a bootstrap methods to adjust the threshold of factor loading's significant test, trying to exclude some falsely significant factor caused by multiple-test problem. Some other scholars use machine learning methods to reduce the potential candidates, more precisely, a stream of them have used a shrinkage and subset selection method called Lasso (Tibshirani, 1996) and it's variations to find suitable factors. One example is Rapach, Strauss, and Zhou (2013). They applied the Lasso regression, trying to find some characteristics from a large group to predict the global stock market's return.

But an additional challenge is that factors, especially in the high-dimension, are commonly correlated. Kozak, Nagel, and Santosh (2020) point out that when facing a group of correlated factors, Lasso will only pick several highly correlated factors seemly randomly, and then ignore the other and shrink them to zero. In other word, Lasso fails to handle the correlated factor appropriately.

Therefore, the main empirical problem in this project is: how to select useful factors from a large group of highly correlated candidates. To answer this problem, we employed a two-step method.

First, we consider the selection of factor base on their strength. And then we will use another variable selection method called Elastic Net (Zou & Hastie, 2005) to select factors. With regard of the first stage, Bailey, Kapetanios, and Pesaran (2020) provides a consistent estimates method for the factor strength, and we will use such method to exam the strength of each candidate factors, and filter out those spurious factors, therefore,reduce the dimension of the number of potential factors. For the second part, elastic net fixes the problem of Lasso can not handle correlated variables by adding extra penalty term, which makes it suitable for our purpose.

For the rest of this plan, we will first go through some literatures relates with CAPM model and methods about selecting factors. Then we will give detailed description of factor strength and elastic net, as well as some preliminary discoveries about factor strength. Finally, we will talks about the further plan for this project.

2 Related Literature

This project is builds on papers devoted on risk pricing. Formularised d by Sharpe (1964), Lintner (1965), and Black (1972), the CAPM model only contains the market factor, which is denotes by the difference between market return and risk free return. Fama and French (1992) develop the model into three-factors, and then it been extend it into four (Carhart, 1997), and five (Fama & French, 2015). Recent research created a six-factors model and claim it outperform all other sparse factor model. (Kelly et al., 2019).

This project also connect with papers about involving factors has no or weak correlation with assets' return into CAPM model. Kan and Zhang (1999) found that the test-statistic of FM two-stage regression (Fama & MacBeth, 1973) will inflate when incorporating factors which are independent with the cross-section return. Kleibergen (2009) pointed out how a factor with small loading would deliver a spurious FM two-pass risk premia estimation. Gospodinov, Kan, and Robotti (2017) show how the involving of a spurious factor will distort the statistical inference of parameters. And, Anatolyev and Mikusheva (2018) studied the behaviours of the model with the presence of weak factors under asymptotic settings, find the regression will lead to an inconsistent risk premia result.

This project also relates to some researches effort to identify useful factors from a group of potential factors. Harvey, Liu, and Zhu (2015) exam over 300 factors published on journals, presents that the traditional threshold for a significant test is too low for newly proposed factor, and they suggest to adjust the p-value threshold to around 3. Methods like a Bayesian procedure introduced by Barillas and Shanken (2018) were used to compare different factor models. Pukthuanthong, Roll, and Subrahmanyam (2019) defined several criteria for ”genuine risk factor”, and base on those criteria introduced a protocol to exam does a factor associated with the risk premium.

This project will attempt to address the factor selection problem by using machine learning techniques. Gu, Kelly, and Xiu (2020) elaborate the advantages of using emerging machine learning algorithms in asset pricing such as more accurate predict result, and superior efficiency. Various machine learning algorithms have been adopted on selecting factors for the factor model, especially in recent years. Lettau and Pelger (2020) applying Principle Components Analysis on investigating the latent factor of model. Lasso method, since it's ability to select features, is popular in the field of the factor selection. Feng, Giglio, and Xiu (2019) used the double-selected Lasso method (Belloni, Chernozhukov, & Hansen, 2014),and a grouped lasso method (Huang, Horowitz, & Wei, 2010) is used by Freyberger, Neuhierl, and Weber (2020) on picking factors from a group of candidates. Kozak et al. (2020) used a Bayesian-based method, combing with both Ridge and Lasso regression, argues that the sparse factor model is ultimately futile.

3 Methodology

3.1 Factor Strength

Capital Asset Pricing Model (CAPM) is the benchmark for pricing the systematic risk of a portfolio. Consider the following multi-factor models for n different assets and T observations with stochastic error term ε_{it}:

$$r_{it} - r_{ft} = a_i + \beta_{im}(r_{mt} - r_{ft}) + \sum_{j=1}^{k} \beta_{ij} f_{jt} + \varepsilon_{it} \tag{1}$$

In the left-hand side, we have r_{it} denotes the return of security i at time t, where $i = 1,2,3,\cdots,N$ and $t = 1,2,3,\cdots T$. r_{ft} denotes the risk free rate at time t. In the other hand, a_i is the constant term. r_{mt} is the market average return and therefore, $(r_{mt} - r_{ft})$ is the excess return of the market. Corresponding β_{im} is the lading of market excess return or market factor. f_{jt} of $j = 1,2,3\cdots k$ is potential risk factor under consideration. b_{ij} represents the factor loading for each k risk factors.

The factor strength of factor f_{jt} as α_j from Pesaran and Smith (2019), and Bailey et al. (2020) is defined as the pervasiveness of a factor.

If we run the OLS regression for equation (1) with only one facto f_{jt}, we will obtain n different factor loading $\hat{\beta}_{it}$. For each of the factor loading $\hat{\beta}_{ij}$, we can construct a t-test to test does the loading equals to zero. The test statistic will be $t_{jt} = \frac{\hat{\beta}_{ij}-0}{\hat{\sigma}_{jt}}$ where $\hat{\sigma}_i$ is the standard error of $\hat{\beta}_{ij}$. Then we defined π_{nT} as the proportion of significant factor's amount to the total factor loadings amount:

$$\hat{\pi}_{nT} = \frac{\sum_{i=1}^{n} \hat{\ell}_{i,nT}}{n} \tag{2}$$

$\ell_{i,nT}$ is an indicator function as: $\ell_{i,nT} := \mathbf{1}[|t_{jt}| > c(n)]$. If the t-statistic t_{jt} is greater than the critical value $c_p(n)$, $\hat{\ell}_{i,nT} = 1$. In other word, we will count one if the factor loading $\hat{\beta}_{ij}$ is significant. $c_p(n)$ represent the critical value of a test with test size p. The critical value is calculated by:

$$c_p(n) = \Phi^{-1}(1 - \frac{p}{2n^{\delta}}) \tag{3}$$

Here, $\Phi^{-1}(\cdot)$ is the inverse cumulative distribution function of a standard normal distribution, and δ is a non-negative value represent the critical value exponent. The traditional method to calculate critical value has not fixed the multiple testing problem. One of the most commonly used adjustment for multiple testing problem is Bonferroni correction. When n as sample size goes to infinity, however, the Bonferroni correction can not yield satisfying results since the $\frac{p}{2n^{\delta}} \to 0$ when $n \to \infty$. Therefore, Bailey, Kapetanios, and Pesaran (2016) provides another adjustment with additional exponent δ to constrain the behaviour of n.

After obtain the $\hat{\pi}_{nT}$, we can use the following formula to estimate our strength indicator α_j:

$$\hat{\alpha} = \begin{cases} 1 + \frac{\ln(\hat{\pi}_{nT})}{\ln n} & \text{if } \hat{\pi}_{nT} > 0, \\ 0, & \text{if } \hat{\pi}_{nT} = 0. \end{cases}$$

From the estimation, we can find out that $\hat{\alpha} \in [0, 1]$

$\hat{\alpha}$ represent the pervasiveness of a factor. Here we denote $[n^{\alpha}]$, $[\cdot]$ will take the integer part of number inside. For factor f_{jt}:

$$|f_{jt}| > c_p(n) \quad i = 1, 2, \ldots, [n^{\alpha_j}]$$

$$|f_{jt}| = 0 \quad i = [n^{\alpha_j}] + 1, [n^{\alpha_j}] + 2, \ldots, n$$

For a factor has strength $\alpha = 1$, factor loading will be significant for every assets at every time. The more observation the factor can significantly influence, the stronger the factor is, and vice versa. Therefore, we can use the factor strength to exclude those factor has only very limited pricing power, in other word, those factor can only generate significant loading on very small portion of assets.

3.2 Elastic Net

Elastic net is variable selection model that can be used for factor selection, introduced by Zou and Hastie (2005). Applying elastic net method to estimate the factor loading β_{ij} requires:

$$\hat{\beta}_{ij} = \underset{\beta_{ij}}{\arg\min}\{\sum_{i=1}^{n}[(r_{it} - r_{ft}) - \beta_{ij} f_{jt}]^2 + \lambda_2 \sum_{i=1}^{n} \beta_{ij}^2 + \lambda_1 \sum_{i=1}^{n} |\beta_{ij}|\} \tag{4}$$

Because the Lasso regression only contains L_1 penalty term $\sum_{i=1}^{n} |\beta_{ij}|$, it will shows no preference when selecting variables when they are highly correlated. So when Lasso regression will either randomly choose factors from highly correlated candidates, or eliminate them together as a whole. Elastic Net, however, by containing L_2 penalty term $\sum_{i=1}^{n} \beta_{ij}^2$, solves this problem. The L_2 penalty term tend to shrink the potential parameters when they does not provide enough explanatory power, but it will not remove redundant factors. Therefore, the elastic net method will shrink those parameters associated with the correlated factors and keep them, or drop them if they are redundant at pricing risk.

4 Preliminary Result

In current stage, we have studied the property of estimator of factor strength α under finite sample scenario. In purpose of this, we have designed and applied a Monte Carlo Simulation. The design details and result table can be find at the Appendix A and Appendix B

To measure the goodness of simulation, we calculate the difference between the estimated factor strength and assigned true factor strength and refer the difference as bias. Base on the bias, we also calculated the Mean Squared Error (MSE) for each setting.

The table shows that both the bias and MSE of $\hat{\alpha}$ for different value of the α, N and T.

From the table 1 and table 2 we can see that when the α is at a relatively low level, the estimator tends to overestimate the strength. For instance, under the setting of $T = 120, N = 500$ and $\alpha = 0.5$, the bias is over 0,2, indicates that the estimated strength $\hat{\alpha}$ is around 0.7. The overestimation, however decreased with the increase of the α. Under the same setting as above, if the strength is assigned to 0.9, the bias significantly reduced to only 0.023. When the α touched it's upper bound as $\alpha = 1$, the bias disappears. Therefore we can conclude that the error converge to zero when the strength α increases. Precision of this estimator improves as α increases toward unity for given T and N increases as well.

5 Further Plan

For the next step of this project, we will start the empirical analyse. Considering the variations of companies included, we will use companies return from Standard & Poors (S&P) 500 index as assets, and examine the strength of potential factors (150 in total) to be included in the final factor model specification. The time span will be 10 years, and in order to cover both recession and flourish, we use data from 2008 to 2018. The return will be calculated on monthly basis, so the time observation for each stock will be t = 120.

References

Anatolyev, S., & Mikusheva, A. (2018, 7). Factor models with many assets: strong factors, weak factors, and the two-pass procedure. *CESifo Working Paper Series*. Retrieved from http://arxiv.org/abs/1807.04094

Bailey, N., Kapetanios, G., & Pesaran, M. H. (2016, 9). Exponent of cross-sectional dependence: Estimation and inference. *Journal of Applied Econometrics*, *31*, 929-960. Retrieved from http://doi.wiley.com/10.1002/jae.2476 doi: 10.1002/jae.2476

Bailey, N., Kapetanios, G., & Pesaran, M. H. (2020). Measurement of factor strength: Theory and practice. *CESifo Working Paper*.

Barillas, F., & Shanken, J. (2018, 4). Comparing asset pricing models. *The Journal of Finance*, *73*, 715-754. Retrieved from http://doi.wiley.com/10.1111/jofi.12607 doi: 10.1111/jofi.12607

Belloni, A., Chernozhukov, V., & Hansen, C. (2014, 4). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, *81*, 608-650. doi: 10.1093/restud/rdt044

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business*, *45*, 444-455. Retrieved from www.jstor.org/stable/2351499

Carhart, M. M. (1997, 3). On persistence in mutual fund performance. *The Journal of Finance*, *52*, 57-82. Retrieved from http://doi.wiley.com/10.1111/j.1540-6261.1997.tb03808.x doi: 10.1111/j.1540-6261.1997.tb03808.x

Cochrane, J. H. (2011, 8). Presidential address: Discount rates. *The Journal of Finance*, *66*, 1047-1108. Retrieved from http://doi.wiley.com/10.1111/j.1540-6261.2011.01671.x doi: 10.1111/j.1540-6261.2011.01671.x

Fama, E. F., & French, K. R. (1992, 6). The cross-section of expected stock returns. *The Journal of Finance*, *47*, 427-465. Retrieved from http://doi.wiley.com/10.1111/j.1540-6261.1992.tb04398.x doi: 10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (2015, 4). A five-factor asset pricing model. *Journal of Financial*

Economics, *116*, 1-22. doi: 10.1016/j.jfineco.2014.10.010

Fama, E. F., & MacBeth, J. D. (1973, 5). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*, 607-636. doi: 10.1086/260061

Feng, G., Giglio, S., & Xiu, D. (2019, 1). *Taming the factor zoo: A test of new factors.* Retrieved from `http://www.nber.org/papers/w25481.pdf` doi: 10.3386/w25481

Freyberger, J., Neuhierl, A., & Weber, M. (2020, 4). Dissecting characteristics nonparametrically. *The Review of Financial Studies*, *33*, 2326-2377. Retrieved from `https://doi.org/10.1093/rfs/hhz123` doi: 10.1093/rfs/hhz123

Gospodinov, N., Kan, R., & Robotti, C. (2017, 9). Spurious inference in reduced-rank asset-pricing models. *Econometrica*, *85*, 1613-1628. doi: 10.3982/ecta13750

Gu, S., Kelly, B., & Xiu, D. (2020, 2). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*, 2223-2273. Retrieved from `https://doi.org/10.1093/rfs/hhaa009` doi: 10.1093/rfs/hhaa009

Harvey, C. R., & Liu, Y. (2017, 12). False (and missed) discoveries in financial economics. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3073799

Harvey, C. R., & Liu, Y. (2019, 3). A census of the factor zoo. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3341728

Harvey, C. R., Liu, Y., & Zhu, H. (2015, 10). … and the cross-section of expected returns. *The Review of Financial Studies*, *29*, 5-68. Retrieved from `https://doi.org/10.1093/rfs/hhv059` doi: 10.1093/rfs/hhv059

Huang, J., Horowitz, J. L., & Wei, F. (2010, 8). Variable selection in nonparametric additive models. *Annals of Statistics*, *38*, 2282-2313. doi: 10.1214/09-AOS781

Kan, R., & Zhang, C. (1999, 2). Two-pass tests of asset pricing models with useless factors. *The Journal of Finance*, *54*, 203-235. Retrieved from `http://doi.wiley.com/10.1111/0022-1082.00102` doi: 10.1111/0022-1082.00102

Kelly, B. T., Pruitt, S., & Su, Y. (2019, 12). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, *134*, 501-524. doi: 10.1016/j.jfineco.2019.05.001

Kleibergen, F. (2009, 4). Tests of risk premia in linear factor models. *Journal of Econometrics*, *149*, 149-173. doi: 10.1016/j.jeconom.2009.01.013

Kozak, S., Nagel, S., & Santosh, S. (2020, 2). Shrinking the cross-section. *Journal of Financial Economics*, *135*, 271-292. doi: 10.1016/j.jfineco.2019.06.008

Lettau, M., & Pelger, M. (2020, 2). Estimating latent asset-pricing factors. *Journal of Econometrics*. doi: 10.1016/j.jeconom.2019.08.012

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*, 13-37. doi: 10.2307/1924119

Pesaran, M. H., & Smith, R. P. (2019). The role of factor strength and pricing errors for estimation and inference in asset pricing models. *CESifo Working Paper Series*.

Pukthuanthong, K., Roll, R., & Subrahmanyam, A. (2019, 8). A protocol for factor identification. *Review of Financial Studies*, *32*, 1573-1607. Retrieved from `https://doi.org/10.1093/rfs/hhy093` doi: 10.1093/rfs/hhy093

Rapach, D. E., Strauss, J. K., & Zhou, G. (2013, 8). International stock return predictability: What is the role of the united states? *The Journal of Finance*, *68*, 1633-1662. Retrieved from `http://doi.wiley.com/10.1111/jofi.12041` doi: 10.1111/jofi.12041

Sharpe, W. F. (1964, 9). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*, 425-442. Retrieved from `http://doi.wiley.com/10.1111/j.1540-6261.1964.tb02865.x` doi: 10.1111/j.1540-6261.1964.tb02865.x

Tibshirani, R. (1996, 1). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*, 267-288. Retrieved from `http://doi.wiley.com/10.1111/j.2517-6161.1996.tb02080.x` doi: 10.1111/j.2517-6161.1996.tb02080.x

Zou, H., & Hastie, T. (2005, 4). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*, 301-320. Retrieved from `http://doi.wiley.com/10.1111/j.1467-9868.2005.00503.x` doi: 10.1111/j.1467-9868.2005.00503.x

A Monte Carlo Design

In this section, I will introduce the baseline design setting of the Monte Carlo Simulation and for the preliminary result please see section (4)

A.1 Monte Carlo Design

First, we will discuss the general design.

Consider the following model with stochastic error:

$$r_{it} - r_{ft} = f_1(r_{mt} - r_f) + f_2(\beta_{ij} f_{jt}) + \varepsilon_{it}$$

In this Monte Carlo simulation, we consider a dataset has $i = 1, 2, \ldots, n$ different assets, with $t = 1, 2, \ldots, T$ different observations. $j = 1, 2, \ldots, k$ different risk factors and one market factors are also included.

$f_1(\cdot)$ and $f_2(\cdot)$ are two different functions represent the unknown mechanism of market factor and other factors in pricing asset risk. $(r_{mt} - r_f)$ is the market return, calculated by market return r_{mt} minus risk free return r_{ft}. r_{it} is the stock return, f_{jt} denotes factors other than market factors and β_{ij} is the corresponding factor loading. ε_{it} is random error with structure can be defined in different designs. Notice that the β_{ij} will be influenced by each factor's strength α_j, where we have α as defined in section (3.1) And for each factor, we assume they follow a multinomial distribution with mean zero and a $k \times k$ variance-covariance matrix Σ. The diagonal of matrix Σ indicates the variance of each factor, and the rest represent the correlation among all k factors. In this model, we can control several parts to investigates different scenarios of the simulation:

A.2 Baseline Design

Follow the general model above, we assume both $f_1(a)$ and $f_2(a)$ are linear function:

$$f_1(r_{mt} - r_f) = a_i + \beta_{im}(r_{mt} - r_{ft})$$

$$f_2(\beta_{ij} f_{jt}) = \sum_{j=1}^{k} \beta_{ij} f_{ij}$$

Therefore, the model with single factor can be write as:

$$r_{it} = a_i + \sum_{j=1}^{k+1} \beta_{ij} f_{ij} + \varepsilon_{it}$$

The constant a_i is generated from a uniform distribution $U[-0.5, 0.5]$. β_{ij} is the factor loading, and f_{ij} is factor with strength α_j. Notice that here we have included the market factor r_{mi-r_f} as f_{i1}, which has strength equals to 1. To generate factors loading, we employed a two steps strategy. First we generate a whole factor loadings vector $\beta_{\mathbf{i}} = (\beta_{i1}, \beta_{i2} \cdots, \beta_{ik+1})$, All elements of the vector follows $IIDU(\mu_\beta - 0.2, \mu_\beta + 0.2)$. The μ_β has been equalled to 0.71 to ensure all values apart from zero. After generating the vector, we randomly selected $[n^{\alpha_j}]$ elements from $\beta_{\mathbf{i}}$ to keep their value and set the other elements value to zero. This step ensures the loading reflects the strength of each factor. For the stochastic error term, in this baseline design, we assume it follows a Standard Gaussian distribution, but we can easily extend it into a more complex form.

Follow the same idea, we also construct a two factor model:

$$r_{it} = a_i + \beta_{im}(r_{mt} - r_{ft}) + \sum_{j=1}^{k} \beta_{ij} f_{jt} + \varepsilon_{it}$$

Here the $r_{mt} - r_{ft}$ is the market factor which assumably has strength $\alpha_m = 1$. β_m is the market factor loading as a vector with all elements different from zero.

For each of the those different models, we consider the $T = \{120, 240, 360\}$, $n = \{100, 300, 500\}$. The market factor will have strength $\alpha_m = 1$ all the time, and the strength of the other factor in two factor model will be $\alpha_x = \{0.5, 0.7, 0.9, 1\}$. For every setting, we will replicate 500 times independently, all the constant a_i and loading β_i will be re-generated for each replication. To exam the goodness of estimation, we calculate the bias between our true underneath factor strength α and the estimated strength $\hat{\alpha}$ as $bias = |\alpha - \hat{\alpha}|$. We also use the bias to calculate the Mean Square Error (MSE) $MSE = \frac{1}{n} \sum_{i=1}^{n} (bias_i)^2$

B Simulation Result Table

Table 1: Simulation result of single factor model

	Single Factor					
	Bias			MSE		
	$\alpha = 0.5$					
n \ T	120	240	360	120	240	360
100	0.194	0.188	0.199	0.050	0.047	0.053
300	0.224	0.224	0.226	0.062	0.062	0.062
500	0.229	0.237	0.225	0.064	0.067	0.062
	$\alpha = 0.7$					
100	0.093	0.090	0.092	0.013	0.012	0.013
300	0.101	0.098	0.101	0.014	0.008	0.014
500	0.101	0.107	0.100	0.015	0.015	0.014
	$\alpha = 0.9$					
100	0.023	0.022	0.023	0.001	0.001	0.001
300	0.023	0.023	0.024	0.001	0.001	0.001
500	0.023	0.023	0.024	0.001	0.001	0.001
	$\alpha = 1.0$					
100	0.000	0.000	0.000	0.000	0.000	0.000
300	0.000	0.000	0.000	0.000	0.000	0.000
500	0.000	0.000	0.000	0.000	0.000	0.000

This table shows the result of one risk factor model. We simulated scenarios of factor strength equals to 0.5, 0.7, 0.9, and 1 with different time, assets size combination. The replication times is 500

Table 2: Simulation result of two factor model

	Two Factor					
	Bias			MSE		
$\alpha_j = 0.5$, $\alpha_m = 1.0$						
T / n	120	240	360	120	240	360
100	0.221	0.219	0.221	0.050	0.049	0.050
300	0.253	0.253	0.253	0.042	0.064	0.065
500	0.268	0.266	0.269	0.072	0.071	0.071
$\alpha_j = 0.7$, $\alpha_m = 1.0$						
100	0.100	0.101	0.100	0.010	0.010	0.010
300	0.113	0.113	0.112	0.013	0.013	0.013
500	0.118	0.118	0.119	0.014	0.014	0.014
$\alpha_j = 0.9$, $\alpha_m = 1.0$						
100	0.024	0.023	0.024	0.001	0.001	0.001
300	0.025	0.025	0.025	0.001	0.001	0.001
500	0.026	0.025	0.025	0.001	0.001	0.001
$\alpha_j = 1.0$, $\alpha_m = 1.0$						
100	0.000	0.000	0.000	0.000	0.000	0.000
300	0.000	0.000	0.000	0.000	0.000	0.000
500	0.000	0.000	0.000	0.000	0.000	0.000

This table shows the result of two factor model, with one market factor and one risk factor. We simulated scenarios of factor strength equals to 0.5, 0.7, 0.9, and 1 with different time, assets size combination. The replication times is 500