In-sample and Out-of-sample Sharpe Ratios of Multi-factor Asset Pricing Models

RAYMOND KAN, XIAOLU WANG, and XINGHUA ZHENG*

This version: November 2020

*Kan is from the University of Toronto, Wang is from Iowa State University, and Zheng is from Hong Kong University of Science and Technology. We thank Svetlana Bryzgalova, Peter Christoffersen, Victor DeMiguel, Andrew Detzel, Junbo Wang, Guofu Zhou, seminar participants at Chinese University of Hong Kong, London Business School, Louisiana State University, University of Toronto, and conference participants at 2019 CFIRM Conference for helpful comments. Corresponding author: Raymond Kan, Joseph L. Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, Canada M5S 3E6; Tel: (416) 978-4291; Fax: (416) 978-5433; Email: `kan@chass.utoronto.ca`.

In-sample and Out-of-sample Sharpe Ratios of Multi-factor Asset Pricing Models

Abstract

For many multi-factor asset pricing models proposed in the recent literature, their implied tangency portfolios have substantially higher sample Sharpe ratios than that of the value-weighted market portfolio. In contrast, such high sample Sharpe ratio is rarely delivered by professional fund managers. This makes it difficult for us to justify using these asset pricing models for performance evaluation. In this paper, we explore if estimation risk can explain why the high sample Sharpe ratios of asset pricing models are difficult to realize in reality. In particular, we provide finite sample and asymptotic analyses of the joint distribution of in-sample and out-of-sample Sharpe ratios of a multi-factor asset pricing model. For an investor who does not know the mean and covariance matrix of the factors in a model, the out-of-sample Sharpe ratio of an asset pricing model is substantially worse than its in-sample Sharpe ratio. After taking into account of estimation risk, our analysis suggests that many of the newly proposed asset pricing models do not provide superior out-of-sample performance than the value-weighted market portfolio.

For performance evaluation, earlier literature typically uses the value-weighted market portfolio as the benchmark portfolio, as suggested by the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965). Over time, finance researchers find that the CAPM fails to completely explain the cross-section of expected returns of stocks, especially when they are grouped based on various firm characteristics (like size and book-to-market). As a result, many multi-factor asset pricing models were proposed to remedy the shortcomings of the CAPM and these new asset pricing models are now routinely used for performance evaluation.

In this paper, we first empirically examine the performence of the CAPM and seven popular multi-factor asset pricing models. The seven multi-factor models are: (1) Fama and French (1993) 3-factor model (FF-3), which adds size and book-to-market factor to the CAPM, (2) Carhart (1997) 4-factor model (Carhart-4), which adds the momentum factor to FF-3, (3) Betting against beta (BAB) of Frazzini and Pedersen (2014), which adds the return difference of low and high beta portfolios as an additional factor to the CAPM, (4) Fama and French (2015) 5-factor model (FF-5), which adds profitability and investment factors to FF-3, (5) Hou, Xue, and Zhang (2015) q-factor model (HXZ q), which adds size, investment and profitability factors to the market factor, (6) Barillas and Shanken (2018) 6-factor model (BS-6). The BS-6 model combines the market factor, size factor, momentum factor, HXZ's profitability and investment factors, and a monthly updated book-to-market factor from Asness and Frazzini (2013), and (7) Hou, Mo, Xue, and Zhang (2019) q^5 model (HMXZ q^5), which adds an expected growth factor to Hou, Xue, and Zhang (2015) q-factor model.

We notice one common feature for these multi-factor models, i.e., they all deliver very high sample Sharpe ratios. Over the period 1967/1–2018/12, the value-weighted market portfolio has a monthly sample Sharpe ratio of 0.117, and the seven multi-factor asset pricing models can produce a monthly sample Sharpe ratio from 0.195 for FF-3 to 0.634 for HMXZ q^5. In addition, it is interesting to note that the sample Sharpe ratios for asset pricing models are steadily increasing over time. It started with 0.117 for the CAPM of 1964, increased to 0.195 for the 3-factor model proposed in Fama and French (1993, FF-3), and finally reached 0.634 for the q^5 model recommended in Hou, Mo, Xue, and Zhang (2019, HMXZ q^5). The increase of sample Sharpe ratio for newer asset pricing model is not entirely surprising given the recent work of Barillas and Shanken (2017), in which they suggest that comparison of the performance of models with traded factors can be simply reduced to comparison of their

Sharpe ratios, and such comparison is independent of the choice of test assets. Naturally, superior asset pricing model that is uncovered over time should have higher Sharpe ratio than the Sharpe ratios of asset pricing models in the past. Continuing at this rate, we can expect future generation of asset pricing models to deliver even higher sample Sharpe ratios.

Given that so many of the recent asset pricing models can generate substantially higher Sharpe ratios than that of the value-weighted market portfolio, one would expect many of the professional investors, like mutual fund managers, ought to be able to do the same. Unfortunately, the empirical data tell us a different story. Over our sample period of 1993/1–2018/12, only 34.83% of the funds outperform the CAPM based on the before-fee sample Sharpe ratio.[1] Comparing with the HMXZ q^5 model, the proportion of the funds beating the benchmark decreases to barely 0.03%. It is entirely possible that mutual fund managers, as a group, are not very good investors when compared with the finance academics. So we shift our attention to the performance of Berkshire Hathaway Inc. (ticker symbol: BRK.A), which was managed by Warren Buffett, arguably the most illustrious investor of our generation. The sample Sharpe ratio of BRK.A over the period 1976/11–2018/12 is 0.228. It only outperformed the CAPM and FF-3 (with sample Sharpe ratios of 0.142 and 0.205, respectively) but underperformed the other six asset pricing models, which have sample Sharpe ratios ranging from 0.287 to 0.608 over the same period.[2]

The mutual fund results and the performance of Berkshire Hathaway Inc. are surprising. Given that we look at so many funds, one would like to think that even in the absence of ability, some funds should have done better than the benchmark simply because of luck. The fact that almost no mutual funds, and even not the best investor of our generation, were able to beat the benchmark suggests something might be wrong with the benchmark. There are many possible explanations as to why real world investors did not get Sharpe ratios that are nearly as large as what is suggested by the academic multi-factor models. One possibility is that investors are not mean-variance investors and they are not trying to hold the portfolio with maximum Sharpe ratio. For example, investors may have utility function that would make them care about higher order moments, and they may want to avoid the portfolio with maximum Sharpe ratio if it comes with some undesirable distributional properties.[3] A second

[1] We focus on the before-fee Sharpe ratio because it reflects what a mutual fund manager can actually accomplish.

[2] Frazzini, Kabiller, and Pedersen (2018) suggest that Buffett generated insignificant alpha once we control for the "betting against beta" factor and "quality minus junk" factor of Asness, Frazzini, and Pedersen (2019).

[3] While this is possible, it would have a hard time explaining why mutual fund managers are not interested in holding portfolios with high Sharpe ratios because they are often evaluated based on their Sharpe ratios.

possibility is that frictions in the markets (e.g., short-selling constraints, transaction costs, and taxes) prevent investors from realizing the observed returns in those factor portfolios, especially those that involve long-short portfolios (see Novy-Marx and Velikov (2016), Patton and Weller (2019), and Detzel, Novy-Marx, and Velikov (2019)). A third possibility is that the multi-factor asset pricing models are subject to the repeating testing problem (see Lo and MacKinlay (1990) and Harvey and Liu (2013)), and the surviving models today may have unusually high sample Sharpe ratios as compared with their population Sharpe ratios. This is particularly true for models that are motivated by anomalies, but models that are motivated by theories are not immune to this problem.

In this paper, we focus on the fourth possibility. That is, investors cannot get those sample Sharpe ratios because of estimation risk. The empirically obtained Sharpe ratios for multi-factor models are in-sample Sharpe ratios. The in-sample Sharpe ratio is computed based on the *ex post* tangency portfolio and it is not attainable by investors. While researchers typically use in-sample Sharpe ratio to estimate population Sharpe ratio, the population Sharpe ratio is also not directly relevant for investors unless they know the true mean and covariance matrix of the factors. If investors have only historical data to work with, they can get neither the in-sample Sharpe ratio nor the population Sharpe ratio as implied by the asset pricing model. What investors can get is the out-of-sample Sharpe ratio of the sample optimal portfolio, and this is subject to estimation risk.

Our empirical analysis suggests that the out-of-sample Sharpe ratio of a multi-factor model tends to be significantly inferior to its in-sample Sharpe ratio. Taking into account estimation errors, it is not entirely clear whether a multi-factor asset pricing model can deliver superior out-of-sample Sharpe ratio than the market portfolio (which has no estimation risk). Relative to the out-of-sample Sharpe ratios of various multi-factor asset pricing models, mutual fund performance looks much better. We see more mutual funds can beat various multi-factor asset pricing models out-of-sample.

However, one cannot draw definite conclusion based on our empirical results on the out-of-sample performance of asset pricing models because they are only point estimates and are subject to sampling fluctuations.[4] The out-of-sample performance is sensitive to the length of the estimation window as well as the choice of the out-of-sample period. Instead of relying only on the sample estimates, we need to understand the distributional property of

[4]Our sub-period empirical results confirm that the out-of-sample Sharpe ratio of the multi-factor asset pricing models can be very volatile.

the out-of-sample Sharpe ratio of a multi-factor asset pricing model. This calls for a serious theoretical analysis.

Under the assumption that returns of traded factors of an asset pricing model are i.i.d. multivariate normally distributed over time, we derive a simple stochastic representation of the in-sample and out-of-sample Sharpe ratios of an asset pricing model. Such representation enables us to obtain the finite sample marginal and joint distribution of the in-sample and out-of-sample Sharpe ratios. We show that the out-of-sample Sharpe ratio is always lower than the population Sharpe ratio, and that the in-sample Sharpe ratio is an upward biased estimator of the population Sharpe ratio. The gap between the in-sample and out-of-sample Sharpe ratios can be significant and as a result the in-sample Sharpe ratio of a multi-factor model does not give a reliable prediction of what an investor can obtain out of sample. However, at the same time, we find that, even though the factor returns are assumed to be i.i.d. over time, the in-sample and out-of-sample Sharpe ratios are positively correlated. Such dependence suggests that conditional on the realized in-sample Sharpe ratio, investors are able to make better inference about the out-of-sample Sharpe ratio. We show how investors can make inferences of the distribution of the out-of-sample Sharpe ratio of an asset pricing model based on our theoretical results. With the ability to make such an inference, investors will be in a better position to make a judgement of whether to invest in the factors suggested by an asset pricing model.

Although we provide the exact distribution of the in-sample and out-of-sample Sharpe ratios of an asset pricing model, researchers may opt to use the asymptotic distribution because asymptotic distribution is often simpler to use than the finite sample distribution. The simple stochastic representation of the in-sample and out-of-sample Sharpe ratios derived in this paper provides an easy way to obtain their limiting distributions. We consider the limiting distribution of the in-sample and out-of-sample Sharpe ratios under different assumptions of the number of factors (N) and number of periods (T). When N is fixed and $T \to \infty$, the limiting distribution of the in-sample Sharpe ratio is well known, but that of the out-of-sample Sharpe ratio is new. We also provide the limiting distribution for the case when $N \to \infty$ and $T \to \infty$ but $N/T \to \rho \in (0, 1)$. This limiting distribution is not easy to obtain and is currently unavailable in the literature.

With our exact distribution results, we can evaluate the accuracy of the two asymptotic distributions. In approximating the exact distribution of the in-sample Sharpe ratio, the traditional fixed N asymptotic does not perform well but the fixed N/T asymptotic works very well, even when N is small. In approximating the exact distribution of the out-of-sample Sharpe ratio, the fixed N asymptotic works well for very small N. As N increases, the fixed N/T asymptotic does a better job than the fixed N asymptotic, but both approximations significantly deviate from the exact distribution. Therefore, unless N is very small, one is better off using the exact distribution to draw inference.

The remainder of the paper is organized as follows. Section I presents the empirical results that motivate our theoretical analysis. Section II outlines the problem and presents the finite sample analysis of in-sample and out-of-sample Sharpe ratios of an asset pricing model with multiple factors. Section III presents two different asymptotic distributions of the in-sample and out-of-sample Sharpe ratios of a multi-factor asset pricing model and evaluates the accuracy of the two asymptotic distributions. Section IV discusses methods for inferring the distribution of out-of-sample Sharpe ratio of a multi-factor asset pricing model based on our finite sample results. The final section concludes and the Appendix contains all the proofs of the paper.

I. Empirical Results

A. Factor Data and In-sample Sharpe Ratios

In our empirical exercise, we consider eight popular asset pricing models. They are (1) CAPM of Sharpe (1964) and Lintner (1965), (2) Fama-French (1993) 3-factor model (FF-3), (3) Carhart (1997) 4-factor model (Carhart-4), (4) Betting against beta (BAB) of Frazzini and Pedersen (2014), (5) Fama-French (2015) 5-factor model (FF-5), (6) Hou, Xue, and Zhang (2015) q-factor model (HXZ q), (7) Barillas and Shanken (2018) 6-factor model (BS-6), and (8) Hou, Mo, Xue, and Zhang (2019) q^5 model (HMXZ q^5). The sample period is 1967/1–2018/12. Monthly factor returns of CAPM, FF-3, Carhart-4, FF-5 are obtained from Ken French's website. Monthly returns of the betting-against-beta factor in BAB and the monthly updated value factor in BS-6 are available from AQR's website. We thank Kewei Hou for sharing the data for the q and q^5 factors with us.

Table I about here

In Table I, we report the maximum sample Sharpe ratio that one can obtain based on the factors from the eight asset pricing models using data from the full sample period (1967–

2018) as well as two subperiods (1967–1992 and 1993–2018). The year in which the model was first published is also included in the table. In addition, p-values based on the Gibbons-Ross-Shanken F-test, comparing the Sharpe ratio of a given model with that of the CAPM, are also reported in the table.

Over the period 1967/1–2018/12, the value-weighted market portfolio (i.e., CAPM) has a monthly sample Sharpe ratio of 0.117, and the other asset pricing models all produce significantly higher (at 1% level) Sharpe ratios than that of the value-weighted market portfolio. It is interesting to note that the sample Sharpe ratios for asset pricing models are steadily increasing with their dates of publication. It started with 0.117 for the CAPM of 1964, increased to 0.195 for FF-3 of 1993, and finally reached 0.634 for HMXZ q^5 of 2019. Continuing at this rate, we can expect future generation of asset pricing models to deliver even higher sample Sharpe ratios.

The same pattern holds across the two subperiods (1967/1–1992/12 and 1993/1–2018/12), although the sample Sharpe ratios of all the asset pricing models (other than the CAPM) are higher in the first subperiod. Note that except for the CAPM, all the other models are published in the second subperiod.

B. Mutual Fund Data

We obtain monthly mutual fund after-fee return data over the period of 1993/1–2018/12 from the CRSP Survivor-bias-free US Mutual Fund database. Before-fee returns are computed by adding back one-twelfth of the annual expense ratio to the after-fee returns. A mapping file from MFLINKs is used to aggregate returns of different classes of the same fund portfolio based on the total net assets (MTNA) at the beginning of the period.

We focus on active domestic equity funds. Domestic equity funds are identified when the first two letters of 'crsp_obj_cd' is 'ED' and the third letter is either 'C' or 'Y', but the third and the fourth letters are not 'YH' or 'YS' and 'si_obj_cd' is not 'OPI'. We exclude index funds from our sample. Index funds are identified by checking whether the name of the funds contain the word 'index' or 'idx'.

We examine fund performance in terms of their Sharpe ratios. For a given fund, we compute both before-fee and after-fee Sharpe ratios using monthly return data. Although the after-fee Sharpe ratios are more relevant to investors, the before-fee Sharpe ratios reflect what mutual fund managers can actually accomplish, which is our focus. We obtain fund

Sharpe ratios for the sample period 1993–2018 as well as two subperiods 1993–2005 and 2006–2018. For a fund to be included in our sample for a given period, we require the fund to have at least 60 non-missing monthly return data in the period. This results in 3,494 funds in the period of 1993–2018, and 2,107 and 2,513 funds in the subperiod of 1993–2005 and 2006–2018 respectively.

Table II about here

In Table II, we report the cross-sectional mean, median, first and third quartile of sample Sharpe ratios of the mutual funds in our sample, and compare the fund performance with the sample Sharpe ratios of various asset pricing models as documented in Table I. Table II shows that in terms of before-fee sample Sharpe ratio, only 34.83% of the funds outperformed the value-weighted market portfolio of the CAPM. For the FF-3 model, only 17.60% of the funds were able to deliver higher Sharpe ratio. The percentage of funds that can outperform the more recent asset pricing models is very small. Comparing with the HMXZ q^5 model, only one out of 3494 funds (i.e., 0.03%) can deliver higher before-fee Sharpe ratio. Similar results hold for both subperiods, but the underperformance of mutual funds relative to multi-factor asset pricing models is more notable in the first subperiod than in the second subperiod. For example, for the first subperiod of 1993/1–2005/12, only 1% of the mutual funds have higher before-fee Sharpe ratio than the FF-3 model; whereas for the second subperiod of 2006/1–2018/12, 16.99% of the mutual funds deliver higher before-fee Sharpe ratio than that of the FF-3 model. However, for both subperiods, there is only one fund that has higher before-fee Sharpe ratio than that of the HMXZ q^5 model.

C. Out-of-sample Sharpe Ratios

Note that the Sharpe ratios for the multi-factor models in Table I are in-sample Sharpe ratios. The in-sample Sharpe ratio is computed based on the *ex post* tangency portfolio and it is not attainable by investors. When investors do not know the true mean and covariance matrix of the traded factors and they have only historical data to work with, they cannot get the in-sample Sharpe ratio or the population Sharpe ratio as implied by the asset pricing model. What investors can get is the out-of-sample Sharpe ratio of the sample optimal portfolio, which is subject to estimation risk.

To get some idea of how estimation risk can affect the out-of-sample performance of multi-factor asset pricing models, we consider a situation where the investor has to estimate the optimal portfolio using historical data. We consider an out-of-sample period of 1993/1–2018/12 (which coincides with the sample period of the mutual fund data) as well as its two subperiods (1993/1–2005/12 and 2006/1–2018/12). For each one of these three out-of-sample periods, we assume the investor constructs a sample optimal portfolio using monthly data from 1967/1 to the beginning of the out-of-sample period. The investor then holds this sample optimal portfolio throughout the out-of-sample period. Out-of-sample Sharpe ratio (OS-SR) of the asset pricing model is computed using the portfolio returns in the out-of-sample period.[5] For comparison, we also obtain the in-sample Sharpe ratio (IS-SR) of the asset pricing model in the out-of-sample period. Table III reports the results for the three out-of-sample periods. We compare the Sharpe ratios of a given model with that of the CAPM, and report the significance in the table. The in-sample Sharpe ratio comparison is based on the Gibbons-Ross-Shanken test, whereas the out-of-sample Sharpe ratio comparison is based on a one-sided test using the asymptotic distribution.

Table III reveals that the out-of-sample Sharpe ratios of the multi-factor models are inferior to their in-sample Sharpe ratios, often by a substantial amount. For example, for the period 1993–2018, five multi-factor models have in-sample Sharpe ratios higher than the CAPM at the 1% level and one model outperforms the CAPM at the 5% level. The only model that does not have an in-sample Sharpe ratio significantly higher than that of the CAPM is FF-3. When the out-of-sample Sharpe ratio is considered, the only model that significantly outperforms the CAPM over 1993–2018 is HMXZ q^5.[6] Similar pattern can also be seen in both 1993–2005 and 2006–2018, though the multi-factor models all perform better in 1993–2005 than in 2006–2018 in terms of both in-sample and out-of-sample Sharpe ratios. Across both out-of-sample subperiods, only the HMXZ q^5 model can deliver a significantly higher out-of-sample Sharpe ratio than that of the CAPM.

$\boxed{\text{Table III about here}}$

In Table IV, we conduct similar exercise as in Table II by comparing the Sharpe ratios of mutual funds with those of the asset pricing models; but instead of in-sample Sharpe

[5]This empirical exercise is performed by Fama and French (2018, Table 3). However, instead of studying the properties of the out-of-sample Sharpe ratio, Fama and French (2018) used this out-of-sample Sharpe ratio to test whether competing models have equal population Sharpe ratio based on a bootstrap experiment.

[6]Given that there is no estimation risk involved when holding the market portfolio, the in-sample and out-of-sample Sharpe ratios of the CAPM are identical.

ratios, we use the out-of-sample Sharpe ratios of the asset pricing models for comparison in Table IV. We find that a larger percentage of mutual funds can beat the multi-factor asset pricing models in Table IV than in Table II. For example, 36.46% of mutual funds have higher before-fee sample Sharpe ratio than the out-of-sample Sharpe ratio of FF-3 in 1993–2018, whereas only 17.60% of the funds can beat the in-sample Sharpe ratio of FF-3 in the same period. The subperiod results (1993–2005 and 2006–2018) in Table IV suggest that the outperformance of the mutual funds relative to the out-of-sample Sharpe ratios of the multi-factor asset pricing models is more evident in the second subperiod 2006–2018. For example, in 2006–2018, 93.24% of the mutual funds outperform FF-3. The only exception is HMXZ q^5, for which we continue to see a very small percentage (0.08%) of the funds that can deliver a superior sample Sharpe ratio in the second subperiod.

Table IV about here

Our data analysis so far suggests that the out-of-sample Sharpe ratio of a multi-factor model tends to be inferior to its in-sample Sharpe ratio. Taking into account estimation errors, it is not entirely clear whether a multi-factor asset pricing model can deliver superior out-of-sample Sharpe ratio than the market portfolio (which has no estimation risk). The only exception is the newest model considered, i.e., HMXZ q^5, which presents robust performance even after taking into account of estimation risk.

Given that investors do not know the mean and covariance matrix of the factors, they ought to be more interested in the out-of-sample Sharpe ratio of an asset pricing model. However, one should not draw definitive conclusion based on the results in Tables III and IV because these are only point estimates, and they are subject to sampling fluctuations. The out-of-sample performance can be sensitive to the length of the estimation window as well as the choice of the out-of-sample period. To provide additional insights, we conduct theoretical analysis of the distribution of both the in-sample and the out-of-sample Sharpe ratio of a multi-factor asset pricing model in the rest of the paper.

II. Finite Sample Distribution of In-sample and Out-of-sample Sharpe Ratios

We consider a multi-factor asset pricing model with N traded factors with $N \geq 2$. Let r_t be the excess returns of the N traded factors at time t. The elements of r_t can be returns of risky assets in excess of risk-free rate, or they can be return differences of two risky assets. The mean and covariance matrix of r_t are defined as $\mu = \mathbb{E}[r_t]$ and $\Sigma = \text{Var}[r_t]$, respectively. We assume μ is a nonzero vector and Σ is positive definite. For a mean-variance investor who wants to hold a portfolio with a target standard deviation of σ, it is easy to show that his optimal portfolio has weights of

$$w^* = \frac{\sigma}{\theta}\Sigma^{-1}\mu$$

in the N traded factors, where $\theta = \sqrt{\mu'\Sigma^{-1}\mu}$ is the maximum Sharpe ratio that one can obtain from the N factors. Obviously, the Sharpe ratio of the optimal portfolio is

$$\frac{w^{*\prime}\mu}{\sqrt{w^{*\prime}\Sigma w^*}} = \frac{\mu'\Sigma^{-1}\mu}{\sqrt{\mu'\Sigma^{-1}\mu}} = \sqrt{\mu'\Sigma^{-1}\mu} = \theta. \tag{1}$$

When the investor does not know the mean and covariance matrix of the factors, he has to estimate θ using historical data. Suppose he has data on r_t for $t = 1, \ldots, T$. The sample estimator of θ is

$$\hat{\theta} = \sqrt{\hat{\mu}'\hat{\Sigma}^{-1}\hat{\mu}}, \tag{2}$$

where $\hat{\mu}$ and $\hat{\Sigma}$ are the sample estimators of μ and Σ, and they are given by

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t, \tag{3}$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(r_t - \hat{\mu})(r_t - \hat{\mu})'. \tag{4}$$

We call $\hat{\theta}$ the *in-sample Sharpe ratio* of the factors (or asset pricing model), which is an *ex post* measure of performance and it is unattainable by investors.

For an investor who does not know μ and Σ, he needs to estimate the weights of the optimal portfolio, and therefore, θ is also unattainable out-of-sample. A natural estimator of w^* is

$$\hat{w} = \frac{\sigma}{\hat{\theta}}\hat{\Sigma}^{-1}\hat{\mu}. \tag{5}$$

The out-of-sample mean and variance of this estimated optimal portfolio are

$$\hat{w}'\mu = \frac{\sigma}{\hat{\theta}}\hat{\mu}'\hat{\Sigma}^{-1}\mu, \tag{6}$$

$$\hat{w}'\Sigma\hat{w} = \frac{\sigma^2}{\hat{\theta}^2}\hat{\mu}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\hat{\mu}. \tag{7}$$

The out-of-sample Sharpe ratio of the estimated optimal portfolio is then given by

$$\tilde{\theta} = \frac{\hat{w}'\mu}{\sqrt{\hat{w}'\Sigma\hat{w}}} = \frac{\hat{\mu}'\hat{\Sigma}^{-1}\mu}{(\hat{\mu}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\hat{\mu})^{\frac{1}{2}}}. \tag{8}$$

We call $\tilde{\theta}$ the *out-of-sample Sharpe ratio* of the factors (or asset pricing model). Unlike $\hat{\theta}$ or θ, $\tilde{\theta}$ is what an investor can actually obtain out-of-sample by holding the sample optimal portfolio $\hat{w}$.

Note that both $\hat{\theta}$ and $\tilde{\theta}$ are random variables because they depend on the realizations of $\hat{\mu}$ and $\hat{\Sigma}$. Our first task is to obtain the finite sample joint distribution of $(\hat{\theta}, \tilde{\theta})$. For this purpose, we assume r_t is identically and independently distributed (i.i.d.) as a multivariate normal distribution with mean μ and covariance Σ. Under this assumption, the following proposition provides a great simplification of this problem by providing a stochastic representation of $(\hat{\theta}, \tilde{\theta})$ that only depends on four univariate random variables (instead of $\hat{\mu}$ and $\hat{\Sigma}$).

Proposition 1: *Suppose* $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \Sigma)$ *and* $N \geq 2$. *Let* $u_1 \sim \chi^2_{T-N}$, $b \sim \text{Beta}((T - N + 1)/2, (N-1)/2)$, *and they are independent of each other. Conditional on* b, *let* $\tilde{z} \sim \mathcal{N}(\sqrt{b}\sqrt{T}\theta, 1)$ *and* $\tilde{u} \sim \chi^2_{N-1}((1-b)T\theta^2)$, *and they are independent of each other and* u_1, *where* $\chi^2_\nu(\delta)$ *stands for a noncentral chi-squared random variable with* ν *degrees of freedom and a noncentrality parameter of* δ. *We have*

$$\hat{\theta} \overset{d}{=} \frac{\sqrt{\tilde{z}^2 + \tilde{u}}}{\sqrt{u_1}}, \tag{9}$$

$$\tilde{\theta} \overset{d}{=} \frac{\theta\tilde{z}}{\sqrt{\tilde{z}^2 + \tilde{u}}}. \tag{10}$$

Besides providing a great simplification of the problem, Proposition 1 also reveals that the joint distribution of $(\hat{\theta}, \tilde{\theta})$ depends only on N, T, and θ. Once θ is known, we do not need to know the individual elements of μ and Σ to obtain the joint distribution of $(\hat{\theta}, \tilde{\theta})$. For asset pricing models with different μ and Σ (but same N), their joint distributions of $(\hat{\theta}, \tilde{\theta})$ are the same as long as they have the same θ.

With the stochastic representation of $(\hat{\theta}, \tilde{\theta})$ in Proposition 1, we can easily obtain the exact moments and joint moments of $\hat{\theta}$ and $\tilde{\theta}$. The following Lemma presents some low order moments of $\hat{\theta}$ and $\tilde{\theta}$.

Lemma 1. *Suppose* $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \Sigma)$ *and* $N \geq 2$. *We have*

$$\mathbb{E}[\hat{\theta}] = \frac{\Gamma\left(\frac{N+1}{2}\right)\Gamma\left(\frac{T-N-1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)\Gamma\left(\frac{T-N}{2}\right)} {}_1F_1\left(-\frac{1}{2}; \frac{N}{2}; -\frac{T\theta^2}{2}\right) \quad \text{for } T \geq N+2, \tag{11}$$

$$\mathbb{E}[\tilde{\theta}] = \frac{\theta^2\sqrt{T}\Gamma\left(\frac{N+1}{2}\right)\Gamma\left(\frac{T-N+2}{2}\right)\Gamma\left(\frac{T}{2}\right)}{\sqrt{2}\Gamma\left(\frac{N+2}{2}\right)\Gamma\left(\frac{T-N+1}{2}\right)\Gamma\left(\frac{T+1}{2}\right)} {}_1F_1\left(\frac{1}{2}; \frac{N+2}{2}; -\frac{T\theta^2}{2}\right) \quad \text{for } T \geq N+1, \tag{12}$$

$$\mathbb{E}[\hat{\theta}^2] = \frac{T\theta^2 + N}{T-N-2} \quad \text{for } T \geq N+3, \tag{13}$$

$$\mathbb{E}[\tilde{\theta}^2] = \theta^2\left[\frac{T-N+1}{T} - \frac{(N-1)(T-N)}{NT} {}_1F_1\left(1; \frac{N+2}{2}; -\frac{T\theta^2}{2}\right)\right] \quad \text{for } T \geq N+1, \tag{14}$$

$$\mathbb{E}[\hat{\theta}\tilde{\theta}] = \frac{\theta^2\sqrt{T}(T-N)\Gamma\left(\frac{T}{2}\right)}{\sqrt{2}(T-N-1)\Gamma\left(\frac{T+1}{2}\right)} \quad \text{for } T \geq N+2, \tag{15}$$

where $\Gamma(a)$ *is the gamma function and* ${}_1F_1(a; b; x)$ *is the confluent hypergeometric function.*

With the above expressions, we can then compute $\text{Var}[\hat{\theta}]$, $\text{Var}[\tilde{\theta}]$, and $\text{Cov}[\hat{\theta}, \tilde{\theta}]$. These results allow us to prove some important inequalities on $\mathbb{E}[\hat{\theta}]$, $\mathbb{E}[\tilde{\theta}]$ and $\text{Cov}[\hat{\theta}, \tilde{\theta}]$, which are given in the following lemma.

Lemma 2. *Suppose* $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \Sigma)$ *and* $N \geq 2$. *We have*

$$\mathbb{E}[\tilde{\theta}] < \theta < \mathbb{E}[\hat{\theta}], \tag{16}$$

$$\text{Cov}[\hat{\theta}, \tilde{\theta}] > 0. \tag{17}$$

The inequality $\mathbb{E}[\tilde{\theta}] < \theta$ is not surprising. Since $\hat{w}$ is estimated with errors, we expect the out-of-sample Sharpe ratio of $\hat{w}$ is not as good as the Sharpe ratio of the true optimal portfolio. In fact, the inequality can be strengthened to $\tilde{\theta} < \theta$ because unless $\hat{w}$ is proportional to w^* (and this event has probability zero), we must have $\tilde{\theta} < \theta$. The inequality $\theta < \mathbb{E}[\hat{\theta}]$ is also not surprising because the in-sample Sharpe ratio is computed after $\hat{\mu}$ and $\hat{\Sigma}$ are observed, and this look-ahead bias on average allows the sample optimal portfolio to have a better in-sample performance than the population Sharpe ratio. In Figure 1, we plot $\mathbb{E}[\tilde{\theta}]/\mathbb{E}[\hat{\theta}]$ as a function of T for various values of N (3 and 6) and θ (0.2 and 0.4). It can be seen that estimation errors significantly drive down the out-of-sample performance of a multi-factor asset pricing model. When $N = 6$ and $\theta = 0.2$, $\mathbb{E}[\tilde{\theta}]$ is only 22% of $\mathbb{E}[\hat{\theta}]$ when $T = 50$. The ratio $\mathbb{E}[\tilde{\theta}]/\mathbb{E}[\hat{\theta}]$ is increasing in T but even with $T = 300$, it is still only 0.6777. Therefore, in-sample Sharpe ratio does not give a reliable prediction of what an investor can expect to obtain in out-of-sample. For higher θ and smaller N, the signal-to-noise ratio is higher and hence $\mathbb{E}[\tilde{\theta}]/\mathbb{E}[\hat{\theta}]$ is higher, but the loss due to estimation risk is still quite significant, especially when the length of time series is short.

$\boxed{\text{Figure 1 about here}}$

The result that $\hat{\theta}$ and $\tilde{\theta}$ are positively correlated is somewhat surprising. Given that returns are i.i.d., we do not expect measure of past performance ($\hat{\theta}$) to predict future performance ($\tilde{\theta}$). The reason that $\hat{\theta}$ and $\tilde{\theta}$ are positively correlated is because both $\hat{\theta}$ and $\tilde{\theta}$ are functions of $\hat{w}$. In Figure 2, we present the correlation between $\hat{\theta}$ and $\tilde{\theta}$, $\rho(\hat{\theta}, \tilde{\theta})$, as a function of T for various values of N (3 and 6) and θ (0.2 and 0.4). Figure 2 shows that for all cases considered, there is non-trivial positive correlation between $\hat{\theta}$ and $\tilde{\theta}$. The correlation is higher for smaller θ. For some cases, the correlation can be as high as 0.333. As T increases, the correlation decreases gradually.

$\boxed{\text{Figure 2 about here}}$

Using the results of Proposition 1, we can obtain explicit expressions of the marginal distributions of $\hat{\theta}$ and $\tilde{\theta}$ as well as the joint distribution of $(\hat{\theta}, \tilde{\theta})$. For the marginal distribution of $\hat{\theta}$, we can write $\hat{\theta} \stackrel{d}{=} \sqrt{u_3}/\sqrt{u_1}$, where $u_3 = \tilde{z}^2 + \tilde{u} \sim \chi^2_N(T\theta^2)$. This implies that $\hat{\theta}^2 \stackrel{d}{=} u_3/u_1$ is proportional to a noncentral F-distribution with degrees of freedom N and $T - N$ and a noncentrality parameter $T\theta^2$, a well known result in the literature (see Gibbons, Ross, and Shanken (1989) and Kan and Robotti (2016)). Let $F^{\delta}_{m,n}(y)$ stand for the cumulative distribution function of a noncentral F random variable with m and n degrees of freedom and a noncentrality parameter δ. The distribution and density functions of $\hat{\theta}$ are given by

$$\mathbb{P}[\hat{\theta} < c] = \mathbb{P}[\hat{\theta}^2 < c^2] = F^{T\theta^2}_{N,T-N}\left(\frac{(T-N)c^2}{N}\right), \tag{18}$$

$$f_{\hat{\theta}}(c) = \frac{2c^{N-1}e^{-\frac{T\theta^2}{2}}}{(1+c^2)^{\frac{T}{2}}\mathrm{B}\left(\frac{N}{2}, \frac{T-N}{2}\right)}\, {}_1F_1\left(\frac{T}{2}; \frac{N}{2}; \frac{T\theta^2c^2}{2(1+c^2)}\right), \tag{19}$$

where $\mathrm{B}(a, b)$ is the beta function. In Figure 3, we plot the density function of $\hat{\theta}/\theta$ for two different values of N (3 and 6) and θ (0.2 and 0.4) with $T = 120$. It can be seen that $\hat{\theta}$ is quite volatile when $T = 120$. In addition, as suggested by Lemma 2, $\hat{\theta}$ is an upward biased estimator of θ. The relative bias of $\hat{\theta}$ increases with N and decreases with θ. Therefore, we can expect an asset pricing model with more factors to have a higher in-sample Sharpe ratio on average, even though the population Sharpe ratio may not improve from having more factors.

Figure 3 about here

For the marginal distribution of $\tilde{\theta}$, we use Proposition 1 and define $q = \tilde{z}/\sqrt{\tilde{u}}$ to obtain

$$\tilde{\theta} \stackrel{d}{=} \frac{\theta \tilde{z}}{\sqrt{\tilde{z}^2 + \tilde{u}}} = \frac{\theta q}{\sqrt{1+q^2}}. \tag{20}$$

Expression (20) suggests that $\tilde{\theta}$ is a monotonic increasing function of q. It also reveals that $-\theta \le \tilde{\theta} \le \theta$, so $\tilde{\theta}$ has a bounded support. Conditional on b, q is proportional to a doubly noncentral t-distribution, so we can compute the cumulative distribution function of $\tilde{\theta}$ using

$$\begin{aligned}
\mathbb{P}[\tilde{\theta} < c] &= \mathbb{P}\left[q < \frac{\frac{c}{\theta}}{\sqrt{1 - \frac{c^2}{\theta^2}}} \right] \\
&= \mathbb{P}\left[\tilde{z} < \frac{c\sqrt{\tilde{u}}}{\sqrt{\theta^2 - c^2}} \right] \\
&= \int_0^1 \int_0^\infty \Phi\left(\frac{c\sqrt{u}}{\sqrt{\theta^2 - c^2}} - \sqrt{b}\sqrt{T}\theta \right) f_{\tilde{u}}(u) f_b(b) \mathrm{d}u \mathrm{d}b \quad \text{for } -\theta < c < \theta,
\end{aligned} \tag{21}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal, $f_{\tilde{u}}(u)$ is the density function of $\chi^2_{N-1}((1-b)T\theta^2)$, and $f_b(b)$ is the density function of Beta$((T-N+1)/2, (N-1)/2)$. Taking derivative yields the density function of $\tilde{\theta}$ as follows

$$f_{\tilde{\theta}}(c) = \frac{\theta^2}{(\theta^2 - c^2)^{\frac{3}{2}}} \int_0^1 \int_0^\infty \phi\left(\frac{c\sqrt{u}}{\sqrt{\theta^2 - c^2}} - \sqrt{b}\sqrt{T}\theta \right) \sqrt{u} f_{\tilde{u}}(u) f_b(b) \mathrm{d}u \mathrm{d}b \quad \text{for } -\theta < c < \theta, \tag{22}$$

where $\phi(\cdot)$ is the density function of a standard normal.

In Figure 4, we plot the density function of $\tilde{\theta}/\theta$ for two different values of N (3 and 6) and θ (0.2 and 0.4) with $T = 120$.[7] It can be seen that $\tilde{\theta}$ is quite volatile when $T = 120$, especially when N is large and θ is small. For those cases, we can expect a substantial deterioration in the out-of-sample performance when an investor holds the sample optimal

[7] It can be shown that when $N > 3$, $\lim_{c \to \theta^-} f_{\tilde{\theta}}(c) = 0$ and $\lim_{c \to (-\theta)^+} f_{\tilde{\theta}}(c) = 0$. When $N = 3$, these two limits are finite, and when $N = 2$, these two limits are infinity. Proof of these results as well as the explicit expressions of of the two limits for the case of $N = 3$ are available upon request.

portfolio. For example, when $N = 6$ and $\theta = 0.2$, we have $\mathbb{P}[\tilde{\theta}/\theta < 0.8] = 0.7027$, so there is more than 70% probability that an investor will lose more than 20% of the Sharpe ratio due to estimation risk. However, if $\theta = 0.4$, then this probability drops down to 19%.

Figure 4 about here

We now turn our attention to the joint distribution of $(\hat{\theta}, \tilde{\theta})$. As it turns out, the joint cumulative distribution of $(\hat{\theta}, \tilde{\theta})$ can be written as a triple integral, whereas the joint density of $(\hat{\theta}, \tilde{\theta})$ can be written as a double integral. These expressions are summarized in the following Proposition.

Proposition 2: *Suppose* $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \Sigma)$ *and* $N \geq 2$. *When* $c_1 \geq 0$ *and* $0 < c_2 \leq \theta$, *we have*

$$\begin{aligned}
&\mathbb{P}[\hat{\theta} < c_1, \tilde{\theta} < c_2] \\
&= \int_0^\infty \int_0^1 \int_0^{c_1^2 v} \left[\Phi\left(\min\left[\frac{c_2\sqrt{u}}{\sqrt{\theta^2 - c_2^2}}, \sqrt{c_1^2 v - u}\right] - \sqrt{T}\theta\sqrt{b}\right) - \Phi\left(-\sqrt{c_1^2 v - u} - \sqrt{T}\theta\sqrt{b}\right)\right] \\
&\quad \times f_{\tilde{u}}(u) f_b(b) f_{u_1}(v) \mathrm{d}u \mathrm{d}b \mathrm{d}v, \qquad (23)
\end{aligned}$$

where $f_{\tilde{u}}(u)$ *is the density function of* $\chi^2_{N-1}((1-b)T\theta^2)$, $f_b(b)$ *is the density function of* $\mathrm{Beta}((T-N+1)/2, (N-1)/2)$, *and* $f_{u_1}(v)$ *is the density function of* χ^2_{T-N}. *When* $c_1 \geq 0$ *and* $-\theta \leq c_2 \leq 0$, *we have*

$$\begin{aligned}
&\mathbb{P}[\hat{\theta} < c_1, \tilde{\theta} < c_2] \\
&= \int_0^\infty \int_0^1 \int_0^{\frac{c_1^2(\theta^2 - c_2^2)v}{\theta^2}} \left[\Phi\left(\frac{c_2\sqrt{u}}{\sqrt{\theta^2 - c_2^2}} - \sqrt{T}\theta\sqrt{b}\right) - \Phi\left(-\sqrt{c_1^2 v - u} - \sqrt{T}\theta\sqrt{b}\right)\right] \\
&\quad \times f_{\tilde{u}}(u) f_b(b) f_{u_1}(v) \mathrm{d}u \mathrm{d}b \mathrm{d}v. \qquad (24)
\end{aligned}$$

The joint density of $(\hat{\theta}, \tilde{\theta})$ *is given by*[8]

$$f_{\hat{\theta},\tilde{\theta}}(c_1, c_2) = \int_0^1 \int_0^\infty f_{\tilde{u}}\left(\frac{c_1^2(\theta^2 - c_2^2)v}{\theta^2}\right) \phi\left(\frac{c_1 c_2 \sqrt{v}}{\theta} - \sqrt{T}\theta\sqrt{b}\right) \frac{2c_1^2 v^{\frac{3}{2}}}{\theta} f_{u_1}(v) f_b(b) \mathrm{d}v \mathrm{d}b \quad (25)$$

for $c_1 > 0$ *and* $-\theta < c_2 < \theta$.

[8]When N is even, the inner integrals of $f_{\tilde{\theta}}(c)$ and $f_{\hat{\theta},\tilde{\theta}}(c_1, c_2)$ can be solved analytically, so $f_{\tilde{\theta}}(c)$ and $f_{\hat{\theta},\tilde{\theta}}(c_1, c_2)$ can be evaluated using a single rather than a double integral. These results are available upon request.

In Figure 5, we plot the joint density of $(\hat{\theta}/\theta, \tilde{\theta}/\theta)$ for two different values of N (3 and 6) and θ (0.2 and 0.4) with $T = 120$. It can be seen that there is quite a bit of volatility in both $\hat{\theta}$ and $\tilde{\theta}$, especially when N is large and θ is small. Comparing the graphs of $N = 3$ with $N = 6$, we see that the joint density is an increasing function of $\tilde{\theta}$ when $N = 3$, and it reaches the maximum when $\tilde{\theta}/\theta = 1$. However, for $N = 6$, the joint density reaches an interior maximum at $\tilde{\theta}/\theta < 1$. This is consistent with the result in Figure 4, where we also see the marginal density of $\tilde{\theta}$ is an increasing function of $\tilde{\theta}$ when $N = 3$ but has an interior maximum when $N = 6$.

Figure 5 about here

While it is interesting to obtain the unconditional distribution of the out-of-sample Sharpe ratio of an asset pricing model, it is equally important for us to understand the distribution of the out-of-sample Sharpe ratio of an asset pricing model conditional on its observed in-sample Sharpe ratio. Given that $\tilde{\theta}$ is not independent of $\hat{\theta}$, the conditional distribution of $\tilde{\theta}$ provides better information for an investor to decide whether to hold a sample optimal portfolio implied by a given asset pricing model. With the expression of the joint density of $(\hat{\theta}, \tilde{\theta})$ and the marginal density of $\hat{\theta}$, we can compute the density of $\tilde{\theta}$ conditional on $\hat{\theta}$ using

$$f_{\tilde{\theta}|\hat{\theta}}(c_2|c_1) = \frac{f_{\hat{\theta},\tilde{\theta}}(c_1, c_2)}{f_{\hat{\theta}}(c_1)}. \tag{26}$$

The conditional distribution of $\tilde{\theta}$ when $\hat{\theta} = c_1$ can be obtained by integrating the conditional density above. In Figure 6, we plot the conditional density of $\tilde{\theta}/\theta$ for two different values of N (3 and 6) and θ (0.2 and 0.4) with $T = 120$. The plot shows the conditional density of $\tilde{\theta}/\theta$ when conditional on three different values of $\hat{\theta}$, the first one is at the 10th percentile of $\hat{\theta}$ (solid line), the second one is at the 50th percentile of $\hat{\theta}$ (dotted line), and the last one is at the 90th percentile of $\hat{\theta}$ (dashed line). It can be seen that the conditional density of $\tilde{\theta}$ can be quite sensitive to the value of $\hat{\theta}$. In particular, the distribution of $\tilde{\theta}$ when conditional on a high value of $\hat{\theta}$ tends to dominate the distribution of $\tilde{\theta}$ when conditional on a low value of $\hat{\theta}$. This is consistent with the result in Lemma 2, in which we show that $\hat{\theta}$ and $\tilde{\theta}$ are positively correlated.

Figure 6 about here

The conditional moments of $\tilde{\theta}$ can also be derived. In the following lemma, we present explicit expressions for the first two moments of $\tilde{\theta}$ when conditional on a given value of $\hat{\theta}$.

Lemma 3. *Suppose* $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \Sigma)$ *and* $N \geq 2$. *We have*

$$\mathbb{E}[\tilde{\theta}|\hat{\theta}] = \frac{2\theta\sqrt{y}}{N} \frac{\Gamma\left(\frac{T-N+2}{2}\right) {}_1F_1\left(\frac{T+1}{2}; \frac{N+2}{2}; y\right)}{\Gamma\left(\frac{T-N+1}{2}\right) {}_1F_1\left(\frac{T}{2}; \frac{N}{2}; y\right)}, \tag{27}$$

$$\mathbb{E}[\tilde{\theta}^2|\hat{\theta}] = \theta^2 \left[\frac{T-N+1}{T} - \frac{(N-1)(T-N)}{NT} \frac{{}_1F_1\left(\frac{T}{2}; \frac{N+2}{2}; y\right)}{{}_1F_1\left(\frac{T}{2}; \frac{N}{2}; y\right)}\right] \tag{28}$$

for $T \geq N+1$, *where* $y = T\theta^2\hat{\theta}^2/[2(1+\hat{\theta}^2)]$.

In Figure 7, we plot $\mathbb{E}[\tilde{\theta}|\hat{\theta}]/\theta$ as a function of $\hat{\theta}/\theta$ for two different values of N (3 and 6) and θ (0.2 and 0.4) with $T = 120$. As expected, $\mathbb{E}[\tilde{\theta}|\hat{\theta}]/\theta$ is a monotonic increasing function of $\hat{\theta}/\theta$. It suggests that keeping N and θ constant, one would prefer an asset pricing model that has a higher in-sample Sharpe ratio. The sample optimal portfolio from such a model tends to also have a higher expected out-of-sample Sharpe ratio. However, Figure 7 also reveals that the relation between $\mathbb{E}[\tilde{\theta}|\hat{\theta}]/\theta$ and $\hat{\theta}/\theta$ also depends on N and θ. When N is small and θ is large, $\mathbb{E}[\tilde{\theta}|\hat{\theta}]/\theta$ is big but it does not vary much with $\hat{\theta}$. In contrast, when N is large and θ is small, the magnitude of $\mathbb{E}[\tilde{\theta}|\hat{\theta}]/\theta$ is small, but one can expect the out-of-sample performance of an asset pricing model to be greatly dependent on its past in-sample performance.

Figure 7 about here

III. Asymptotic Distributions of In-sample and Out-of-sample Sharpe Ratios

Although we provide the exact distribution of $(\hat{\theta}, \tilde{\theta})$ in Section II, researchers may opt to use the asymptotic distribution of $(\hat{\theta}, \tilde{\theta})$ instead. This is because asymptotic distribution is often simpler to use than finite sample distribution. This advantage will be particularly appealing when the approximation error of the asymptotic distribution is small. In this section, we present the limiting distribution of $(\hat{\theta}, \tilde{\theta})$ and evaluate the accuracy of the limiting distribution using the finite sample results from Section II. As it turns out, there are two different limiting distributions for $(\hat{\theta}, \tilde{\theta})$ depending on whether N is fixed or $N \to \infty$ as $T \to \infty$.

Our first limiting result is based on the traditional asymptotic analysis, which assumes N is fixed when $T \to \infty$. The following Proposition presents the limiting distribution of $(\hat{\theta}, \tilde{\theta})$ under this assumption.

Proposition 3: *Suppose* $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \Sigma)$ *and* $N \geq 2$*. When* N *is fixed and* $T \to \infty$*, we have*

$$\left[\begin{array}{c} \sqrt{T}(\hat{\theta} - \theta) \\ T(\tilde{\theta} - \theta) \end{array} \right] \xrightarrow{d} \left[\begin{array}{c} X \\ Y \end{array} \right], \tag{29}$$

where $X \sim \mathcal{N}\left(0, 1 + \frac{\theta^2}{2}\right)$, $Y \sim -(1+\theta^2)/(2\theta)\chi^2_{N-1}$*, and they are independent of each other.*

The limiting distribution of $\sqrt{T}(\hat{\theta}-\theta)$ is well known and it is easy to derive (see, for example, Barillas, Kan, Robotti, and Shanken (2019)). Interestingly, the asymptotic distribution of $\hat{\theta}$ is the same as the one for the single asset case as discussed in Lo (2002). This suggests that using $\hat{w}$ instead of w^* has no impact on the asymptotic distribution of $\hat{\theta}$, so estimation errors of $\hat{w}$ has only a second order impact on the asymptotic distribution of $\hat{\theta}$.

The result of the limiting distribution of $T(\tilde{\theta} - \theta)$ is new. There are two points to note here. First, unlike $\hat{\theta}$ which converges to θ at a rate of $1/\sqrt{T}$, $\tilde{\theta}$ converges to θ at a rate of $1/T$. Second, the limiting distribution of $T(\tilde{\theta} - \theta)$ is not a normal distribution, but instead it is a negative random variable which is proportional to χ^2_{N-1}. This is because while $\tilde{\theta}$ converges to θ, it is always less than θ in a finite sample, so the limiting distribution of $T(\tilde{\theta} - \theta)$ has to be a negative random variable.

Instead of the traditional asymptotic analysis which assumes fixed N, the following Proposition presents the limiting distribution assuming both N and T converge to infinity, but $N/T \to \rho \in (0, 1)$.

Proposition 4: *Suppose* $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \Sigma)$ *and* $N \geq 2$*. When* $N \to \infty$*,* $T \to \infty$*,* $N/T \to \rho \in (0, 1)$ *and* θ *is fixed, we have*

$$\left[\begin{array}{c} \sqrt{T}(\hat{\theta} - \bar{\theta}) \\ \sqrt{T}(\tilde{\theta} - \underline{\theta}) \end{array} \right] \xrightarrow{d} \mathcal{N}\left(0_2, \left[\begin{array}{cc} \frac{\theta^4+2\theta^2+\rho}{2(1-\rho)^2(\theta^2+\rho)} & \frac{\rho\theta^2}{2(\theta^2+\rho)^2} \\ \frac{\rho\theta^2}{2(\theta^2+\rho)^2} & \frac{\rho\theta^2}{2(\theta^2+\rho)}\left[\frac{(1-\rho)(2\rho+\theta^2)}{(\theta^2+\rho)^2} + 2 + \theta^2\right] \end{array} \right]\right), \tag{30}$$

where

$$\bar{\theta} = \frac{\sqrt{\theta^2 + \rho}}{\sqrt{1 - \rho}}, \tag{31}$$

$$\underline{\theta} = \frac{\theta^2\sqrt{1 - \rho}}{\sqrt{\theta^2 + \rho}}. \tag{32}$$

It can be readily shown that $\bar{\theta} > \theta$ and $\underline{\theta} < \theta$. The result that $\hat{\theta} \xrightarrow{p} \bar{\theta}$ under the assumption that $N/T \to \rho$ can be easily obtained from Theorem 4.6 of El Karoui (2010). The result that $\tilde{\theta} \xrightarrow{p} \underline{\theta}$ when $N/T \to \rho$ was shown in Ao, Li, and Zheng (2019). However, the limiting distribution of $(\hat{\theta}, \tilde{\theta})$ under the assumption that $N/T \to \rho$ is not easy to obtain and it is unavailable in the literature. We are able to obtain the limiting distribution in Proposition 4 because we have derived the exact distribution of $(\hat{\theta}, \tilde{\theta})$ in Proposition 1, and this allows us to obtain the limiting distribution of $(\hat{\theta}, \tilde{\theta})$ by taking the appropriate limit.

The bivariate normality result in Proposition 4 also allows us to obtain an approximation of the conditional distribution of $\tilde{\theta}$ when conditional on $\hat{\theta}$. In particular, we can approximate the conditional distribution of $\tilde{\theta}$ by using a normal distribution with conditional mean

$$\begin{aligned} \mathbb{E}[(\tilde{\theta} - \underline{\theta})|\hat{\theta}] &\approx \frac{\rho(1-\rho)^2\theta^2}{(\theta^2+\rho)(\theta^4+2\theta^2+\rho)}(\hat{\theta} - \bar{\theta}) \\ \Rightarrow \quad \mathbb{E}[\tilde{\theta}|\hat{\theta}] &\approx \frac{\sqrt{1-\rho}\,\theta^2(\theta^4+2\theta^2+\rho^2)}{\sqrt{\theta^2+\rho}\,(\theta^4+2\theta^2+\rho)} + \frac{\rho(1-\rho)^2\theta^2}{(\theta^2+\rho)(\theta^4+2\theta^2+\rho)}\hat{\theta}. \end{aligned} \tag{33}$$

In addition, the conditional variance of $\tilde{\theta}$ can be approximated by using

$$\mathrm{Var}[\tilde{\theta}|\hat{\theta}] \approx \frac{\rho\theta^2(1+\theta^2)^2(2+\theta^2)}{2T(\theta^2+\rho)(\theta^4+2\theta^2+\rho)}, \tag{34}$$

which is independent of $\hat{\theta}$.

With the exact distribution result in Section II, we can evaluate the accuracy of the two different asymptotic distributions of $\hat{\theta}$ and $\tilde{\theta}$. In Figure 8, we plot the exact density of $\hat{\theta}/\theta$ (solid line) vs. its two approximations, one is based on the fixed N asymptotic (dashed line) and the other is based on the $N/T \to \rho$ asymptotic (dotted line). We consider two different values of N (3 and 6) and θ (0.2 and 0.4) with $T = 120$. It can be seen that the approximate distribution based on the traditional fixed N asymptotic does not perform well in all cases. This is because $\hat{\theta}$ has a bias, which the fixed N asymptotic distribution of $\hat{\theta}$ ignores. In contrast, the approximation based on the fixed N/T asymptotic works very well in all cases in approximating the exact distribution of $\hat{\theta}$. Therefore, if one would like to use an asymptotic distribution to approximate the exact distribution of $\hat{\theta}$, there is a compelling reason to use the fixed N/T asymptotic one, even when N is small and T is large.

Figure 8 about here

We now turn our attention to the density of $\tilde{\theta}$. In Figure 9, we plot the exact density of $\tilde{\theta}/\theta$ (solid line) vs. its two approximations, one is based on the fixed N asymptotic (dashed line) and the other is based on the $N/T \to \rho$ asymptotic (dotted line). We consider two different values of N (3 and 6) and θ (0.2 and 0.4) with $T = 120$. It can be seen that when $N = 3$, the fixed N asymptotic approximation works very well, especially when $\theta = 0.4$. In contrast, the fixed N/T asymptotic approximation of $\tilde{\theta}$, which assumes it has a limiting normal distribution, does a poor job in approximating the exact distribution of $\tilde{\theta}$. When $N = 6$, the fixed N asymptotic approximation starts to deviate significantly from the exact distribution of $\tilde{\theta}$, especially for the case with $\theta = 0.2$. The fixed N/T asymptotic approximation improves with an increase in N, but it still deviates quite a bit from the exact distribution for $N = 6$. Therefore, unless N is very small, both asymptotic approximations do not provide reliable approximations of the exact distribution of $\tilde{\theta}$, and one is better off using the exact distribution of $\tilde{\theta}$ to draw inference.[9]

Figure 9 about here

With the expressions of conditional mean and conditional variance of $\tilde{\theta}$ in (33) and (34) and Proposition 4, we are also able to obtain the limiting distribution of $\tilde{\theta}$ conditional on $\hat{\theta}$, which is a normal distribution.[10] As shown in Figure 6, the conditional density based on the finite sample results clearly deviates from normality. Therefore, the asymptotic conditional distribution of $\tilde{\theta}$ also does not provide an accurate approximation of the exact conditional distribution of $\tilde{\theta}$ in finite samples.

IV. Making Inference of Out-of-Sample Sharpe Ratio

We now proceed to answer the following question. Given an asset pricing model and some historical data on its factor returns, how should an investor forecast the distribution of the out-of-sample Sharpe ratio of the sample optimal portfolio in the following period. Given that the asymptotic distributions of $\tilde{\theta}$ do not provide accurate approximation in the finite sample as shown in Section III, we focus on making inference using our finite sample results on in-sample and out-of-sample Sharpe ratio. Based on the frequentist approach, we present a procedure to construct the confidence interval of the distribution of the out-of-sample

[9]When N is larger, like $N \geq 10$, we find that the fixed N/T asymptotic approximation works remarkably well in approximating the exact distribution of $\tilde{\theta}$.

[10]In the case that N is fixed and $T \to \infty$, the limiting distribution of $\hat{\theta}$ and $\tilde{\theta}$ are independent of each other. Therefore, the limiting conditional distribution of $\tilde{\theta}$ is the same as the unconditional one in Proposition 3.

Sharpe ratio. In addition, we also show how to obtain the posterior predictive distribution of the out-of-sample Sharpe ratio using a Bayesian approach.

A. Confidence Interval of the Distribution of Out-of-Sample Sharpe Ratio

In Section II, we already derive the finite sample distribution of $\tilde{\theta}$ conditional on $\hat{\theta}$ in (26). This distribution, however, depends on the population Sharpe ratio, which is unknown to investors. Thus, we first consider the problem of constructing a confidence interval for the population Sharpe ratio of an asset pricing model. This problem is well studied in the literature. Under the assumption that $r_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \Sigma)$, we know

$$\frac{(T-N)\hat{\theta}^2}{N} \sim \mathcal{F}_{N,T-N}^{T\theta^2}, \tag{35}$$

where $\mathcal{F}_{N,T-N}^{T\theta^2}$ is a noncentral F-distribution with N and $T-N$ degrees of freedom, and a noncentrality parameter of $T\theta^2$. Since the noncentral F-distribution is decreasing in its noncentrality parameter, we can use the statistical method (see, for example, Casella and Berger (1990, Section 9.2.3)) to construct a $100(1-\alpha)\%$ confidence interval for θ. Using this methodology, we first plot the $100(\alpha/2)$ and $100(1-\alpha/2)$ percentiles of the distribution of $\hat{\theta}^2$ for different values of θ. We then draw a horizontal line at the observed value of $\hat{\theta}^2$. This horizontal line will first intersect the $100(1-\alpha/2)$ percentile line and then the $100(\alpha/2)$ percentile line of $\hat{\theta}^2$. The interval between these two intersection points gives us a $100(1-\alpha)\%$ confidence interval for θ. Mathematically, θ_L and θ_U are implicitly determined by the following two equations

$$F_{N,T-N}^{T\theta_L^2}(x) = 1 - \frac{\alpha}{2}, \tag{36}$$

$$F_{N,T-N}^{T\theta_U^2}(x) = \frac{\alpha}{2}, \tag{37}$$

where $x = (T-N)\hat{\theta}^2/N$. Note that since $F_{N,T-N}^{\delta}(x)$ is decreasing in the noncentrality parameter δ, (36) will not have a solution for θ_L when $F_{N,T-N}^{0}(x) < 1-\alpha/2$. In this case, we set $\theta_L = 0$. Similarly, if $F_{N,T-N}^{0}(x) < \alpha/2$, we cannot find a solution for θ_U in (37) and we set $\theta_U = 0$.

While we can obtain a confidence interval for θ, this information would be of little interest to an investor except for the case of a single factor model (like the CAPM). For multi-factor models, investors do not know the true optimal portfolio implied by the model, and hence the population Sharpe ratio is not directly relevant. Instead, investors are more interested in what they will receive out-of-sample from holding the sample optimal portfolio suggested by the asset pricing model. Therefore, we would like to make inference on the distribution of the out-of-sample Sharpe ratio of an asset pricing model.

With the confidence interval for θ, we are able to construct the confidence interval of the exact distribution of $\tilde{\theta}$ conditional on $\hat{\theta}$ using (26). Specifically, we can first use θ_L to compute the conditional distribution of $\tilde{\theta}$ and then use θ_U to compute another conditional distribution of $\tilde{\theta}$. This gives us a method for constructing a confidence interval for any percentile of the conditional distribution of $\tilde{\theta}$. We could also use this method to construct a confidence interval for $\mathbb{E}[\tilde{\theta}|\hat{\theta}]$.

In Table V, we consider an investor standing at the end of 2018. For three different estimation windows, $h = 60$, 120, and 240 months, we report $\hat{\theta}$ for the eight different asset pricing models. For each model, we report a 95% confidence interval for its θ, i.e., (θ_L, θ_U). It can be seen that there is a lot of uncertainty about θ. Even with $h = 240$ months, quite a few models have $\theta_L = 0$ (e.g., CAPM, FF-3, and Carhart-4). For $h = 240$ months, HMXZ q^5 has the highest in-sample Sharpe ratio of 0.469 but the 95% confidence interval for its θ is (0.308, 0.501), which is quite wide. Although the 95% confidence interval for θ of HMXZ q^5 dominates the 95% confidence interval for θ of the CAPM, which is (0, 0.229), this does not mean an investor ought to invest in the factors of the HMXZ q^5 model. This is because the population Sharpe ratio is unattainable to an investor who does not know the true mean and covariance matrix of the factors.

Table V about here

We now turn our attention to constructing confidence intervals for the 10th, 50th, and 90th percentiles of the conditional distribution of $\tilde{\theta}$ (when conditional on $\hat{\theta}$) for the multi-factor models. Just like the confidence interval for θ, the confidence intervals for different percentiles of the conditional distribution of $\tilde{\theta}$ are also quite wide, even if the length of estimation window is 240 months. When $h = 240$ months, the 95% confidence intervals for the median of $\tilde{\theta}$ of most mulit-factor models overlap with the confidence interval for θ of the CAPM. For example, the 95% confidence interval for the median of $\tilde{\theta}$ for the FF-5, HXZ q, and BS-6 models are (0.135, 0.408), (0.088, 0.362), and (0.074, 0.356), respectively, and it is not entirely clear that it dominates the CAPM, which has a 95% confidence interval of (0, 0.229) for θ. The only exception is the HMXZ q^5 model, which has a 95% confidence interval of (0.291, 0.562) for the median of $\tilde{\theta}$.[11]

An advantage of holding the value-weighted market portfolio is that we know we will get its population Sharpe ratio in out-of-sample. This is not the case for multi-factor models because due to estimation risk, its out-of-sample Sharpe ratio has a distribution. In addition, the uncertainty of the percentiles of the conditional distribution of $\tilde{\theta}$ also presents an additional source of risk to an investor. Taken as a whole, we do not find strong evidence that the recent multi-factor models can deliver superior out-of-sample performance than the value-weighted market portfolio, with the exception of the most recent HMXZ q^5 model.

B. Bayesian Approach

We next illustrate how to make inference of $\tilde{\theta}$ under the Bayesian approach by specifying a prior on θ. There are many possible choices for the prior distribution of θ. We consider a prior on θ that has a bounded support, say $0 \leq \theta \leq c$, because it is unlikely that θ of any asset pricing model can go to infinity. Specifically, we set the upper bound for the θ to $c = 0.6$, which is about four to five times the in-sample Sharpe ratio of the CAPM that we observe in our sample period, and it is about the level of the highest in-sample Sharpe ratio for the asset pricing models that we consider. The conventional wisdom in finance is that high Sharpe ratios are good deals and they are unlikely to survive, and a Sharpe ratio that is twice that of the market portfolio is already considered very high (e.g., Ross (1976), MacKinlay (1995), Cochrane and Saá-Requejo (2000)). Nevertheless, we would like to give the benefit of doubt to the multi-factor asset pricing models by allowing them to have a potentially much higher Sharpe ratio than that of the market portfolio.

A simple prior distribution for θ is the beta distribution. Specifically, we assume that

$$\theta \sim c\text{Beta}(\alpha, \beta), \tag{38}$$

where α and β are parameters to be specified. The mean of the prior distribution of θ is given by

$$\mathbb{E}[\theta] = \frac{\alpha c}{\alpha + \beta}. \tag{39}$$

[11] In untabulated results, we also obtain the 95% confidence interval for the unconditional distribution of $\tilde{\theta}$. As expected, the unconditional distribution of $\tilde{\theta}$ is more volatile than the conditional distribution of $\tilde{\theta}$, and the 95% confidence intervals for the 10th, 50th, and 90th percentiles of the unconditional $\tilde{\theta}$ are wider than those for the conditional $\tilde{\theta}$.

If we assume the prior for the CAPM has a mean of 0.15, then we need $\alpha/(\alpha+\beta) = 1/4$, which can be accomplished by setting $\alpha = 2$ and $\beta = 6$. With these parameter values, the prior distribution of θ is right skewed and has a mode at 1/6. For an asset pricing model with more factors, we may choose a lower β so that the mean of the prior distribution of θ is higher.

Given the prior distribution of θ and the in-sample Sharpe ratio ($\hat{\theta}$), we can then obtain the posterior predictive distribution of $\tilde{\theta}$. The posterior predictive density of $\tilde{\theta}$ can be computed using

$$f(\tilde{\theta}|\hat{\theta}) = \int_0^c f(\tilde{\theta}|\theta,\hat{\theta})f(\theta|\hat{\theta})\mathrm{d}\theta = \frac{\int_0^\infty f(\tilde{\theta}|\theta,\hat{\theta})f(\theta,\hat{\theta})\mathrm{d}\theta}{f(\hat{\theta})} = \frac{\int_0^c f(\tilde{\theta},\hat{\theta}|\theta)f(\theta)\mathrm{d}\theta}{\int_0^c f(\hat{\theta}|\theta)f(\theta)\mathrm{d}\theta} \tag{40}$$

for $-c < \tilde{\theta} < c$, and the k-th moment of the posterior predictive distribution of $\tilde{\theta}$ is given by

$$\mathbb{E}[\tilde{\theta}^k|\hat{\theta}] = \int_{-c}^{c} \tilde{\theta}^k f(\tilde{\theta}|\hat{\theta})\mathrm{d}\tilde{\theta}. \tag{41}$$

Following the proof of Lemma 3, it can be readily shown that the first two moments can be expressed as single integrals.

In Table VI, we consider an investor standing at the end of 2018 and apply our results to compute the posterior predictive distribution of $\tilde{\theta}$. For three different estimation windows, h = 60, 120, and 240 months, we first compute the in-sample Sharpe ratio ($\hat{\theta}$) for the eight asset pricing models. Given the prior distribution of θ, we then obtain the posterior predictive mean and standard deviation of $\tilde{\theta}$. For the prior distribution of θ, we set $\alpha = 2$ and consider three different values of β (6, 4, and 2).

Table VI about here

At the end of 2018, all the multi-factor models have a higher in-sample Sharpe ratio than that of the CAPM, regardless of the length of estimation window. For example, when h = 60 months, the CAPM has $\hat{\theta} = 0.201$, whereas some of the multi-factor models have much higher $\hat{\theta}$ (e.g., 0.555 for BAB and 0.567 for HMXZ q^5). However, a Bayesian investor will incorporate his prior when deciding whether to invest in the sample optimal portfolio associated with an asset pricing model. In general, the posterior mean of θ will be some weighted average of prior mean of θ and $\hat{\theta}$. For example, if investor has a prior of θ with $\beta = 6$ (which implies a prior mean of 0.15 for θ), then the posterior mean of θ for the

CAPM (which is also its $\tilde{\theta}$) will only be 0.159, and it is less than its $\hat{\theta}$ of 0.201. For multi-factor models, the Bayesian investor will also need to incorporate the uncertainty due to the estimation risk associated with investing in the sample optimal portfolio. For example, even with the impressive $\hat{\theta}$ of 0.555 for BAB and 0.567 for HMXZ q^5, the posterior predictive mean of $\tilde{\theta}$ for these two models are only 0.253 and 0.160, respectively. In addition, the posterior predictive standard deviation of $\tilde{\theta}$ for these two models are quite a bit higher (0.094 for BAB and 0.099 for HMXZ q^5) than that of the CAPM. For all the other multi-factor models, they show a lower posterior predictive mean of $\tilde{\theta}$ than that for the CAPM. For priors with $\beta = 4$ and 2, the corresponding prior mean for θ are 0.2 and 0.3, respectively. For investors who hold a strong belief that the multi-factor models have very high Sharpe ratios, then the posterior predictive mean of $\tilde{\theta}$ will tend to be higher. However, even with $\beta = 2$, the posterior predictive mean of $\tilde{\theta}$ for BAB and HMXZ q^5 are still only 0.396 and 0.308, which is far lower than their in-sample Sharpe ratios of 0.555 and 0.567. For all the other asset pricing models, their posterior predictive mean of $\tilde{\theta}$ are not higher than that of the CAPM.

With a longer estimation window, the estimation risk associated with $\tilde{\theta}$ is reduced. However, even for $h = 120$ and 240 months, the estimation risk is still quite substantial, and the posterior predictive mean of $\tilde{\theta}$ for the multi-factor models are still quite a bit lower than their $\hat{\theta}$. For example, when $h = 120$ months and with $\beta = 6$, we still see only BAB and HMXZ q^5 to have a higher posterior predictive mean of $\tilde{\theta}$ than that of the CAPM. Finally, with $h = 240$ months, we now see more mulit-factor pricing models delivering a higher posterior predictive mean of $\tilde{\theta}$ than that of the CAPM, but they come with a higher posterior predictive standard deviation. Nevertheless, FF-3 and Carhart-4 continue to have a lower posterior predictive mean of $\tilde{\theta}$ than that of the CAPM even with $h = 240$ months. In summary, for a Bayesian investor, he should not be simply investing in the multi-factor model with the highest $\hat{\theta}$. When it comes to selecting models, he needs to incorporate his prior on θ and to take into account the estimation risk associated with the in-sample optimal portfolio. In our empirical example, we find that even though many of the popular multi-factor models have higher in-sample Sharpe ratio than that of the CAPM at the end of the sample period, a Bayesian investor does not always favor these models over the CAPM. The decision crucially depends on his prior as well as the length of estimation window.

V. Conclusion

Academic asset pricing models have produced increasingly large sample Sharpe ratios over time. Starting with the value-weighted market portfolio of the CAPM, which only produced a sample Sharpe ratio of 0.117, we now have multi-factor model that produced sample Sharpe ratio of 0.634, more than five times larger than that of the market portfolio. At the same time, we do not see any real world investor who can generate Sharpe ratio that is anywhere close to what is suggested by these popular asset pricing models. This presents a serious problem because these multi-factor models are often used as benchmarks to evaluate the performance of mutual fund managers, or to determine the cost of capital for capital budgeting.

The high sample Sharpe ratios of the popular multi-factor asset pricing models are also at odd with a long-standing belief in finance that high Sharpe ratios are good deals and they are unlikely to survive. For example, Ross (1976) assumed that no portfolio can have Sharpe ratio that is twice as large as that of the market portfolio. MacKinlay (1995) considered Fama-French 3-factor model has unreasonably high sample Sharpe ratio, even after taking into account of sampling variability. Cochrane and Saá-Requejo (2000) assumed that no asset should have a Sharpe ratio that is twice that of the S&P500 (which they assumed to have an annual value of 0.5, or a monthly value of 0.1443) and use this assumption to derive bounds on option prices.

While there are a number of possible reasons of why the recent asset pricing models produce high sample Sharpe ratios that seem unattainable by real world investors, our paper focuses on one possible explanation, i.e., estimation risk. For multi-factor models, investors need to know how to optimally allocate the weights into the various factors of the model. When the mean and covariance matrix of the factors are not known, investor will not be able to hold the true optimal portfolio. Instead, investors need to estimate the optimal weights and this will lead to deteriorated out-of-sample performance for the sample optimal portfolio as compared with the true optimal portfolio.

In this paper, we provide an analysis of the finite sample joint distribution of the in-sample and out-of-sample Sharpe ratios of the sample optimal portfolio. This analysis allows us to understand the uncertainty that investors face when they invest in the sample optimal portfolio of a multi-factor asset pricing model. Our analysis also allows us to easily obtain the limiting joint distributions of the in-sample and out-of-sample Sharpe ratios under various assumptions. More importantly, our analysis allows us to predict the out-of-sample performance of a sample optimal portfolio constructed based on the factors of an asset pricing model. When estimation risk is taken into account, we do not find very strong evidence that many of the multi-factor asset pricing models can deliver superior out-of-sample performance than the value-weighted market portfolio, with the lone exception of the most recent HMXZ q^5 model.

One of the limitations of our analysis is that it is based on the i.i.d. multivariate normality assumption for the returns of the traded factors. With fat-tailed distributions, it is quite conceivable that the problem with estimation risk is more severe than in the normality case. So one should take our result as a lower bound on the impact of estimation risk on the out-of-sample performance of multi-factor models. In addition, if there is a concern that parameters in these multi-factor models are not constant over time, then there is an additional source of risk that hampers the out-of-sample performance of the sample optimal portfolio based on a multi-factor model.

Appendix

Proof of Proposition 1: Under the multivariate normality assumption, it is well known that $\hat{\mu}$ and $\hat{\Sigma}$ are independent of each other and have the following distributions:

$$\hat{\mu} \sim \mathcal{N}(\mu, \Sigma/T), \tag{A1}$$

$$\hat{\Sigma} \sim \mathcal{W}_N(T-1, \Sigma/T), \tag{A2}$$

where $\mathcal{W}_N(T-1, \Sigma/T)$ is a Wishart distribution with $T-1$ degrees of freedom and covariance matrix Σ/T. Define $\eta = \Sigma^{-\frac{1}{2}}\mu/\theta$, we have $\eta'\eta = 1$. Let P be an $N \times N$ orthonormal matrix with its first column is equal to η. By defining

$$z = \sqrt{T}P'\Sigma^{-\frac{1}{2}}\hat{\mu} \sim \mathcal{N}\left(\left[\begin{array}{c} \sqrt{T}\theta \\ 0_{N-1} \end{array}\right], I_N\right), \tag{A3}$$

$$W = TP'\Sigma^{-\frac{1}{2}}\hat{\Sigma}\Sigma^{-\frac{1}{2}}P \sim \mathcal{W}_N(T-1, I_N), \tag{A4}$$

we can write

$$\hat{\theta} = (\hat{\mu}'\hat{\Sigma}^{-1}\hat{\mu})^{\frac{1}{2}} = (z'W^{-1}z)^{\frac{1}{2}}, \tag{A5}$$

$$\tilde{\theta} = \frac{\mu'\hat{\Sigma}^{-1}\hat{\mu}}{(\hat{\mu}'\hat{\Sigma}^{-1}\Sigma\hat{\Sigma}^{-1}\hat{\mu})^{\frac{1}{2}}} = \frac{\sqrt{T}\theta e_1'W^{-1}z}{(Tz'W^{-2}z)^{\frac{1}{2}}} = \frac{\theta e_1'W^{-1}z}{(z'W^{-2}z)^{\frac{1}{2}}}, \tag{A6}$$

where $e_1 = [1,\ 0_{N-1}']'$. Define an $N \times N$ orthonormal matrix $Q = [\tilde{z},\ Q_1]$ with its first column is equal to $\tilde{z} \equiv z/(z'z)^{\frac{1}{2}}$. Let

$$A = (Q'W^{-1}Q)^{-1} = \left[\begin{array}{cc} \tilde{z}'W^{-1}\tilde{z} & \tilde{z}'W^{-1}Q_1 \\ Q_1'W^{-1}\tilde{z} & Q_1'W^{-1}Q_1 \end{array}\right]^{-1} \equiv \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right] \sim \mathcal{W}_N(T-1, I_N), \tag{A7}$$

where A_{11} is the $(1,1)$ element of A. Theorem 3.2.10 of Muirhead (1982) suggests that

$$u_1 \equiv A_{11\cdot 2} = A_{11} - A_{12}A_{22}^{-1}A_{21} \sim \chi^2_{T-N}, \tag{A8}$$

and it is independent of A_{12} and A_{22}. In addition, using the result of Dickey (1967), we can show that

$$-A_{22}^{-1}A_{21} \sim \frac{x}{\sqrt{u_2}}, \tag{A9}$$

where $x \sim \mathcal{N}(0_{N-1}, I_{N-1})$, $u_2 \sim \chi^2_{T-N+1}$, and they are independent of each other and u_1. Since the distribution of A is independent of z, x, u_1 and u_2 are also independent of z. Using the formula for the inverse of a partitioned matrix, we can easily verify that

$$\tilde{z}'W^{-1}\tilde{z} = A_{11\cdot 2}^{-1} = \frac{1}{u_1}, \tag{A10}$$

$$Q_1' W^{-1} \tilde{z} = -A_{22}^{-1} A_{21} A_{11\cdot 2}^{-1} = \frac{x}{u_1 \sqrt{u_2}}. \tag{A11}$$

With these identities, we can write

$$z' W^{-2} z = z' W^{-1} \left(\tilde{z} \tilde{z}' + Q_1 Q_1' \right) W^{-1} z = (z'z) \left(\frac{1}{u_1^2} + \frac{x'x}{u_1^2 u_2} \right). \tag{A12}$$

Let $z_1 \sim \mathcal{N}(\sqrt{T}\theta, 1)$ and $x_1 \sim \mathcal{N}(0,1)$ be the first element of z and x, respectively. We can write $z'z = z_1^2 + u$ and $x'x = x_1^2 + u_3$, where $u \sim \chi^2_{N-1}$ and $u_3 \sim \chi^2_{N-2}$. It follows that

$$z' W^{-2} z = \frac{(z_1^2 + u)}{u_1^2} \left(1 + \frac{x_1^2 + u_3}{u_2} \right). \tag{A13}$$

Without loss of generality, let the first column of Q_1 be

$$\xi = \frac{(I_N - \tilde{z}\tilde{z}')e_1}{[e_1'(I_N - \tilde{z}\tilde{z}')e_1]^{\frac{1}{2}}} = \frac{(I_N - \tilde{z}\tilde{z}')e_1}{\sqrt{1 - \frac{z_1^2}{z'z}}}. \tag{A14}$$

From (A11), we know that

$$\frac{x_1}{u_1\sqrt{u_2}} = \xi' W^{-1} \tilde{z} = \frac{e_1' W^{-1} \tilde{z} - \frac{e_1' \tilde{z}}{u_1}}{\sqrt{1 - \frac{z_1^2}{z'z}}} = \frac{e_1' W^{-1} z - \frac{z_1}{u_1}}{\sqrt{u}}, \tag{A15}$$

and hence

$$e_1' W^{-1} z = \frac{z_1}{u_1} + \frac{x_1 \sqrt{u}}{u_1 \sqrt{u_2}} = \frac{1}{u_1} \left(z_1 + \frac{x_1 \sqrt{u}}{\sqrt{u_2}} \right). \tag{A16}$$

Define $q_1 = x_1/\sqrt{x'x}$ and $q_2 = z_2/\sqrt{u}$, where $z_2 \sim \mathcal{N}(0,1)$ is the second element of z. It is well known that q_1 is independent of $x'x$ and q_2 is independent of u (see for example, Theorem 1.5.6 of Muirhead (1982)). Since x is independent of u, q_1 and q_2 are independent of both u and $x'x$. In addition, q_1 and q_2 have the same distribution, so we can replace q_1 with q_2 and write

$$x_1 \sqrt{u} = \frac{x_1}{\sqrt{x'x}} \sqrt{x'x} \sqrt{u} \overset{d}{=} \frac{z_2}{\sqrt{u}} \sqrt{x'x} \sqrt{u} = z_2 \sqrt{x'x}. \tag{A17}$$

Letting $g = x'x/u_2$, we can write

$$e_1' W^{-1} z = \frac{z_1 + \sqrt{g} z_2}{u_1}, \tag{A18}$$

$$z' W^{-2} z = \frac{(z'z)(1+g)}{u_1^2}. \tag{A19}$$

Let

$$b = \frac{1}{1+g} = \frac{u_2}{x'x + u_2} \sim \text{Beta}\left(\frac{T-N+1}{2}, \frac{N-1}{2}\right) \tag{A20}$$

and $z'z = z_1^2 + z_2^2 + u_0$, where $u_0 \sim \chi^2_{N-2}$ and it is independent of z_1 and z_2, we have

$$\hat{\theta} = (z'W^{-1}z)^{\frac{1}{2}} = (z'z)^{\frac{1}{2}}(\tilde{z}'W^{-1}\tilde{z})^{\frac{1}{2}} \overset{d}{=} \frac{\sqrt{z_1^2 + z_2^2 + u_0}}{\sqrt{u_1}}, \tag{A21}$$

$$\tilde{\theta} = \frac{\theta e_1' W^{-1} z}{\sqrt{z'W^{-2}z}} = \frac{\theta(\sqrt{b}z_1 + \sqrt{1-b}z_2)}{(z'z)^{\frac{1}{2}}} \overset{d}{=} \frac{\theta(\sqrt{b}z_1 + \sqrt{1-b}z_2)}{\sqrt{z_1^2 + z_2^2 + u_0}}. \tag{A22}$$

Finally, let $\tilde{z} = \sqrt{b}z_1 + \sqrt{1-b}z_2 \sim \mathcal{N}(\sqrt{b}\sqrt{T}\theta, 1)$ and $\tilde{u} = z_1^2 + z_2^2 + u_0 - \tilde{z}^2 = z'z - \tilde{z}^2 \sim \chi^2_{N-1}((1-b)T\theta^2)$, and conditional on b, $\tilde{z}$ and $\tilde{u}$ are independent of each other. Therefore, we can write

$$\hat{\theta} \overset{d}{=} \frac{\sqrt{\tilde{z}^2 + \tilde{u}}}{\sqrt{u_1}}, \tag{A23}$$

$$\tilde{\theta} \overset{d}{=} \frac{\theta \tilde{z}}{\sqrt{\tilde{z}^2 + \tilde{u}}}. \tag{A24}$$

This completes the proof.

Proof of Lemma 1: We first cite some explicit expressions of moments of noncentral chi-squared and beta random variables. Suppose $X \sim \chi^2_\nu(\lambda)$ and $B \sim \text{Beta}(\nu_1, \nu_2/2)$. We have

$$\mathbb{E}[X^r] = \frac{2^r \Gamma\left(\frac{\nu}{2} + r\right)}{\Gamma\left(\frac{\nu}{2}\right)} {}_1F_1\left(-r; \frac{\nu}{2}; -\frac{\lambda}{2}\right) \quad \text{for } r > -\tfrac{\nu}{2}, \tag{A25}$$

$$\mathbb{E}[B^r] = \frac{\mathrm{B}(\nu_1 + r, \nu_2)}{\mathrm{B}(\nu_1, \nu_2)} \quad \text{for } r > -\nu_1, \tag{A26}$$

where $\mathrm{B}(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. (A25) is given in Krishnan (1967), and (A26) is obtained by direct integration. Using (A25) and the fact that $\tilde{z}^2 + \tilde{u} \sim \chi^2_N(T\theta^2)$ and it is independent of u_1, we can obtain $\mathbb{E}[\hat{\theta}]$ and $\mathbb{E}[\hat{\theta}^2]$ as

$$\begin{aligned} E[\hat{\theta}] &= E[(\tilde{z}^2 + \tilde{u})^{\frac{1}{2}}]E[u_1^{-\frac{1}{2}}] \\ &= \frac{\Gamma\left(\frac{N+1}{2}\right)\Gamma\left(\frac{T-N-1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)\Gamma\left(\frac{T-N}{2}\right)} {}_1F_1\left(-\frac{1}{2}; \frac{N}{2}; -\frac{T\theta^2}{2}\right) \quad \text{for } T \geq N+2, \end{aligned} \tag{A27}$$

$$\begin{aligned} \mathbb{E}[\hat{\theta}^2] &= E[\tilde{z}^2 + \tilde{u}]E[u_1^{-1}] \\ &= \frac{N + T\theta^2}{T-N-2} \quad \text{for } T \geq N+3, \end{aligned} \tag{A28}$$

For $\tilde{\theta}$, we use independence between b and (z_1, z_2, u_0) in (A22) and apply (A26) to obtain

$$\mathbb{E}[\tilde{\theta}] = \theta\mathbb{E}[b^{\frac{1}{2}}]\mathbb{E}\left[\frac{z_1}{\sqrt{z_1^2+z_2^2+u_0}}\right] = \theta\frac{\mathrm{B}\left(\frac{T-N+2}{2}, \frac{N-1}{2}\right)}{\mathrm{B}\left(\frac{T-N+1}{2}, \frac{N-1}{2}\right)}\mathbb{E}\left[\frac{z_1}{\sqrt{z_1^2+z_2^2+u_0}}\right]. \tag{A29}$$

By using the symmetry argument, the term $\sqrt{1-b}z_2$ drops out because $z_2 \sim \mathcal{N}(0,1)$. For the last expectation, we use a lemma in Kan and Wang (2019) to show that

$$\mathbb{E}\left[\frac{z_1}{\sqrt{z_1^2+z_2^2+u_0}}\right] = \sqrt{T}\theta\mathbb{E}\left[\frac{1}{\sqrt{y}}\right], \tag{A30}$$

where $y \sim \chi^2_{N+2}(T\theta^2)$. Then using (A25), we obtain

$$\mathbb{E}[\tilde{\theta}] = \theta\frac{\Gamma\left(\frac{T-N+2}{2}\right)\Gamma\left(\frac{T}{2}\right)}{\Gamma\left(\frac{T-N+1}{2}\right)\Gamma\left(\frac{T+1}{2}\right)}\frac{\sqrt{T}\theta\Gamma\left(\frac{N+1}{2}\right)}{\sqrt{2}\Gamma\left(\frac{N+2}{2}\right)}{}_1F_1\left(\frac{1}{2};\frac{N+2}{2};-\frac{T\theta^2}{2}\right).$$

For $\mathbb{E}[\tilde{\theta}^2]$, we use (A22) and apply (A26) to obtan

$$\begin{aligned}\mathbb{E}[\tilde{\theta}^2] &= \theta^2\mathbb{E}\left[\frac{bz_1^2+(1-b)z_2^2+2\sqrt{b(1-b)}z_1z_2}{z_1^2+z_2^2+u_0}\right]\\ &= \theta^2\mathbb{E}\left[\frac{bz_1^2+(1-b)z_2^2}{z_1^2+z_2^2+u_0}\right]\\ &= \frac{\theta^2}{T}\mathbb{E}\left[\frac{(T-N+1)z_1^2+(N-1)z_2^2}{z_1^2+z_2^2+u_0}\right].\end{aligned} \tag{A31}$$

Note that the term $2\sqrt{b(1-b)}z_1z_2/(z_1^2+z_2^2+u_0)$ vanishes in the above expectation because of symmetry. The last expectation term can be written as $\mathbb{E}[(z'Az)/(z'z)]$, where $A = \mathrm{Diag}(T-N+1,\ N-1,\ 0'_{N-2})$. Using Theorem 4 of Hillier, Kan, and Wang (2014), we obtain the expectation of the ratio of quadratic form in z as

$$\begin{aligned}\mathbb{E}\left[\frac{z'Az}{z'z}\right] &= \frac{T}{N}{}_1F_1\left(1;\frac{N+2}{2};-\frac{T\theta^2}{2}\right) + \frac{T\theta^2(T-N+1)}{N+2}{}_1F_1\left(1;\frac{N+4}{2};-\frac{T\theta^2}{2}\right)\\ &= T-N+1-\frac{(N-1)(T-N)}{N}{}_1F_1\left(1;\frac{N+2}{2};-\frac{T\theta^2}{2}\right),\end{aligned} \tag{A32}$$

where the last equality follows from a recurrence relation of confluent hypergeometric function.[12] It follows that

$$\mathbb{E}[\tilde{\theta}^2] = \theta^2\left[\frac{T-N+1}{T} - \frac{(N-1)(T-N)}{NT}{}_1F_1\left(1;\frac{N+2}{2};-\frac{T\theta^2}{2}\right)\right] \quad \text{for } T \geq N+1. \tag{A34}$$

[12]The recurrence relation is

$$b{}_1F_1(a;b;z) - b{}_1F_1(a-1;b;z) = z{}_1F_1(a;b+1;z). \tag{A33}$$

The equality is obtained by setting $a = 1$, $b = (N+2)/2$ and $z = -T\theta^2/2$.

Finally, using (9) and (10), $\mathbb{E}[\hat{\theta}\tilde{\theta}]$ is given by

$$\begin{aligned}
\mathbb{E}[\hat{\theta}\tilde{\theta}] &= \theta\mathbb{E}\left[\frac{\tilde{z}}{\sqrt{u_1}}\right] \\
&= \theta\mathbb{E}[b^{\frac{1}{2}}]\sqrt{T}\theta\mathbb{E}[u_1^{-\frac{1}{2}}] \\
&= \theta\frac{\mathrm{B}\left(\frac{T-N+2}{2},\frac{N-1}{2}\right)}{\mathrm{B}\left(\frac{T-N+1}{2},\frac{N-1}{2}\right)}\sqrt{T}\theta\frac{\Gamma\left(\frac{T-N-1}{2}\right)}{\sqrt{2}\Gamma\left(\frac{T-N}{2}\right)} \\
&= \frac{\theta^2\sqrt{T}(T-N)\Gamma\left(\frac{T}{2}\right)}{\sqrt{2}(T-N-1)\Gamma\left(\frac{T+1}{2}\right)} \quad \text{for } T \geq N+2. \qquad \text{(A35)}
\end{aligned}$$

This completes the proof.

Proof of Lemma 2: We first prove $\mathbb{E}[\tilde{\theta}] < \theta$. Note that

$$\tilde{\theta} = \frac{\hat{w}'\mu}{(\hat{w}'\Sigma\hat{w})^{\frac{1}{2}}} \leq \frac{w^{*\prime}\mu}{(w^{*\prime}\Sigma w^*)^{\frac{1}{2}}} = \theta, \qquad \text{(A36)}$$

with the equality holds if and hold if $\hat{w}$ is proportional to w^*. Since this event has probability zero, we can write the above as $\tilde{\theta} < \theta$. Taking expectation, we obtain $\mathbb{E}[\tilde{\theta}] < \theta$.

Next we prove $\mathbb{E}[\hat{\theta}] > \theta$. Note that

$$\hat{\theta} = \frac{\hat{w}'\hat{\mu}}{(\hat{w}'\hat{\Sigma}\hat{w})^{\frac{1}{2}}} \geq \frac{w^{*\prime}\hat{\mu}}{(w^{*\prime}\hat{\Sigma}w^*)^{\frac{1}{2}}}, \qquad \text{(A37)}$$

with the equality holds if and only if $\hat{w}$ is proportional to w^*. Since this event has probability zero, we write the above as a strict inequality. Taking expectation on both sides and using the fact that $\hat{\mu}$ is independent of $\hat{\Sigma}$, we have

$$\mathbb{E}[\hat{\theta}] > \mathbb{E}\left[\frac{w^{*\prime}\hat{\mu}}{(w^{*\prime}\hat{\Sigma}w^*)^{\frac{1}{2}}}\right] = w^{*\prime}\mu\mathbb{E}\left[\frac{1}{(w^{*\prime}\hat{\Sigma}w^*)^{\frac{1}{2}}}\right] = \theta\mathbb{E}\left[\left(\frac{w^{*\prime}\Sigma w^*}{w^{*\prime}\hat{\Sigma}w^*}\right)^{\frac{1}{2}}\right]. \qquad \text{(A38)}$$

Using 3.2.5 of Muirhead (1982), we know $y \equiv w^{*\prime}\hat{\Sigma}w^*/(Tw^{*\prime}\Sigma w^*) \sim \chi^2_{T-1}$, so we have

$$\mathbb{E}\left[\left(\frac{w^{*\prime}\Sigma w^*}{w^{*\prime}\hat{\Sigma}w^*}\right)^{\frac{1}{2}}\right] = T^{\frac{1}{2}}\mathbb{E}\left[y^{-\frac{1}{2}}\right] \geq \frac{T^{\frac{1}{2}}}{(E[y])^{\frac{1}{2}}} = \frac{T^{\frac{1}{2}}}{(T-1)^{\frac{1}{2}}} > 1, \qquad \text{(A39)}$$

where the first inequality follows because of Jensen's inequality. Therefore, we have $\mathbb{E}[\hat{\theta}] > \theta$.

Finally, we use Lemma 1 to obtain the explicit expression of $\mathrm{Cov}[\hat{\theta},\tilde{\theta}]$ as

$$\mathrm{Cov}[\hat{\theta},\tilde{\theta}] = \frac{\sqrt{T}\theta^2(T-N)\Gamma\left(\frac{T}{2}\right)}{\sqrt{2}(T-N-1)\Gamma\left(\frac{T+1}{2}\right)}\left[1 - \frac{\Gamma\left(\frac{N+1}{2}\right)^2}{\Gamma\left(\frac{N+2}{2}\right)\Gamma\left(\frac{N}{2}\right)}\right.$$

$$\times\ {}_1F_1\left(-\frac{1}{2};\frac{N}{2};-\frac{T\theta^2}{2}\right){}_1F_1\left(\frac{1}{2};\frac{N+2}{2};-\frac{T\theta^2}{2}\right)\Bigg]. \tag{A40}$$

In order to prove $\mathrm{Cov}[\hat{\theta},\tilde{\theta}]>0$, we cite Theorem 3.1 from Kalmykov and Karp (2013), which suggests:

Suppose $\{f_k\}_{k=0}^{\infty}$ is a non-trivial nonnegative log-concave sequence without internal zeros. Then the function

$$f(y)=\sum_{k=0}^{\infty}\frac{f_k x^k}{\Gamma(y+k)k!} \tag{A41}$$

is strictly log-concave on $(0,\infty)$ for each fixed $x>0$. Moreover,

$$f(\alpha+y)f(\beta+y)-f(\alpha+\beta+y)f(y)>0 \tag{A42}$$

for $y\geq -1$, $\alpha>0$, $\beta>0$ and $\alpha+y\geq 0$, $\beta+y\geq 0$.

For our problem, we set $f_k=(a)_k$, $y=b$ and $\alpha=\beta=1/2$, which gives

$$\frac{{}_1F_1\left(a;b+\frac{1}{2};x\right)^2}{\Gamma\left(b+\frac{1}{2}\right)^2}=f\left(b+\frac{1}{2}\right)^2>f(b+1)f(b)=\frac{{}_1F_1(a;b+1;x){}_1F_1(a;b;x)}{\Gamma(b+1)\Gamma(b)}. \tag{A43}$$

It remains to show that $(a)_k$ is log-concave, i.e.,

$$\frac{(a)_{k+1}}{(a)_k}>\frac{(a)_k}{(a)_{k-1}}, \tag{A44}$$

or equivalently

$$\frac{\Gamma(a+k+1)}{\Gamma(a+k)}>\frac{\Gamma(a+k)}{\Gamma(a+k-1)}. \tag{A45}$$

Using the Cauchy-Schwarz inequality, it is easy to establish for x, $y>0$, we have $\Gamma((x+y)/2)^2\leq\Gamma(x)\Gamma(y)$ with equality holds if and only if $x=y$. Putting $x=a+k+1$ and $y=a+k-1$, we obtain (A45).

Putting $a=(N+1)/2$, $b=N/2$ and $x=T\theta^2/2$ in (A43), we obtain

$$\begin{aligned}
&\frac{{}_1F_1\left(\frac{N+1}{2};\frac{N}{2};\frac{T\theta^2}{2}\right)}{\Gamma\left(\frac{N}{2}\right)}\frac{{}_1F_1\left(\frac{N+1}{2};\frac{N+2}{2};\frac{T\theta^2}{2}\right)}{\Gamma\left(\frac{N+2}{2}\right)}<\left[\frac{{}_1F_1\left(\frac{N+1}{2};\frac{N+1}{2};\frac{T\theta^2}{2}\right)}{\Gamma\left(\frac{N+1}{2}\right)}\right]^2\\
\Rightarrow\quad&\frac{e^{\frac{T\theta^2}{2}}{}_1F_1\left(-\frac{1}{2};\frac{N}{2};-\frac{T\theta^2}{2}\right)}{\Gamma\left(\frac{N}{2}\right)}\frac{e^{\frac{T\theta^2}{2}}{}_1F_1\left(\frac{1}{2};\frac{N+2}{2};-\frac{T\theta^2}{2}\right)}{\Gamma\left(\frac{N+2}{2}\right)}<\left[\frac{e^{\frac{T\theta^2}{2}}}{\Gamma\left(\frac{N+1}{2}\right)}\right]^2
\end{aligned}$$

$$\Rightarrow \quad \frac{\Gamma\left(\frac{N+1}{2}\right)^2}{\Gamma\left(\frac{N+2}{2}\right)\Gamma\left(\frac{N}{2}\right)} {}_1F_1\left(-\frac{1}{2};\frac{N}{2};-\frac{T\theta^2}{2}\right) {}_1F_1\left(\frac{1}{2};\frac{N+2}{2};-\frac{T\theta^2}{2}\right) < 1, \tag{A46}$$

and we use the Kummer transformation of the confluent hypergeometric function in the second line. This completes the proof.

Proof of Proposition 2: We need to find out the range of $\tilde{z}$ for $\mathbb{P}[\hat{\theta} < c_1, \tilde{\theta} < c_2]$ from Proposition 1. When $c_2 > 0$, there are two cases to consider: (1) $0 < \tilde{u} < c_1^2(\theta^2 - c_2^2)u_1/\theta^2$. For this case, we need $c_2\sqrt{\tilde{u}/(\theta^2 - c_2^2)} < \tilde{z} < -\sqrt{u_1 c_1^2 - \tilde{u}}$. (2) $c_1^2(\theta^2 - c_2^2)u_1/\theta^2 < \tilde{u} < c_1^2 u_1$. For this case, we need $\sqrt{u_1 c_1^2 - \tilde{u}} < \tilde{z} < -\sqrt{u_1 c_1^2 - \tilde{u}}$. Together, we have when $c_2 > 0$,

$$\begin{aligned}
&\mathbb{P}[\hat{\theta} < c_1, \tilde{\theta} < c_2] \\
&= \int_0^\infty \int_0^1 \int_0^{\frac{c_1^2(\theta^2 - c_2^2)v}{\theta^2}} \left[\Phi\left(\frac{c_2\sqrt{u}}{\sqrt{\theta^2 - c_2^2}} - \sqrt{T}\theta\sqrt{b}\right) - \Phi\left(-\sqrt{c_1^2 v - u} - \sqrt{T}\theta\sqrt{b}\right)\right] \\
&\times f_{\tilde{u}}(u) f_b(b) f_{u_1}(v)\mathrm{d}u\mathrm{d}b\mathrm{d}v \\
&+ \int_0^\infty \int_0^1 \int_{\frac{c_1^2(\theta^2 - c_2^2)v}{\theta^2}}^{c_1^2 v} \left[\Phi\left(\sqrt{c_1^2 v - u} - \sqrt{T}\theta\sqrt{b}\right) - \Phi\left(-\sqrt{c_1^2 v - u} - \sqrt{T}\theta\sqrt{b}\right)\right] \\
&\times f_{\tilde{u}}(u) f_b(b) f_{u_1}(v)\mathrm{d}u\mathrm{d}b\mathrm{d}v \\
&= \int_0^\infty \int_0^1 \int_0^{c_1^2 v} \left[\Phi\left(\min\left[\frac{c_2\sqrt{u}}{\sqrt{\theta^2 - c_2^2}}, \sqrt{c_1^2 v - u}\right] - \sqrt{T}\theta\sqrt{b}\right) - \Phi\left(-\sqrt{c_1^2 v - u} - \sqrt{T}\theta\sqrt{b}\right)\right] \\
&\times f_{\tilde{u}}(u) f_b(b) f_{u_1}(v)\mathrm{d}u\mathrm{d}b\mathrm{d}v.
\end{aligned} \tag{A47}$$

When $c_2 \leq 0$, only the first case is possible and we have

$$\begin{aligned}
&\mathbb{P}[\hat{\theta} < c_1, \tilde{\theta} < c_2] \\
&= \int_0^\infty \int_0^1 \int_0^{\frac{c_1^2(\theta^2 - c_2^2)v}{\theta^2}} \left[\Phi\left(\frac{c_2\sqrt{u}}{\sqrt{\theta^2 - c_2^2}} - \sqrt{T}\theta\sqrt{b}\right) - \Phi\left(-\sqrt{c_1^2 v - u} - \sqrt{T}\theta\sqrt{b}\right)\right] \\
&\times f_{\tilde{u}}(u) f_b(b) f_{u_1}(v)\mathrm{d}u\mathrm{d}b\mathrm{d}v.
\end{aligned} \tag{A48}$$

Taking derivative of $\mathbb{P}[\hat{\theta} < c_1, \tilde{\theta} < c_2]$ with respect to c_1 and c_2 and using the Leibniz integral rule, we obtain the joint density of $(\hat{\theta}, \tilde{\theta})$ for both cases as

$$f_{\hat{\theta},\tilde{\theta}}(c_1, c_2) = \int_0^\infty \int_0^1 f_{\tilde{u}}\left(\frac{c_1^2(\theta^2 - c_2^2)v}{\theta^2}\right) \phi\left(\frac{c_1 c_2\sqrt{v}}{\theta} - \sqrt{T}\theta\sqrt{b}\right) \frac{2c_1^2 v^{\frac{3}{2}}}{\theta} f_b(b) f_{u_1}(v)\mathrm{d}b\mathrm{d}v. \tag{A49}$$

This completes the proof.

Proof of Lemma 3: We denote

$$U(a;b;x) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-xt} t^{a-1} (1+t)^{b-a-1} \mathrm{d}t \tag{A50}$$

as the confluent hypergeometric function of second kind. For $b-1 < \alpha$, we have the following integral identity (see, for example, Gradshtyen and Ryzhik (2007), 7.621.6)

$$\int_0^\infty x^{\alpha-1} e^{-x} U(a;b;x) \mathrm{d}x = \frac{\Gamma(\alpha)\Gamma(1-b+\alpha)}{\Gamma(a-b+\alpha+1)}. \tag{A51}$$

In order to prove (27) and (28), we first establish the following two identities. Suppose $a > 0$, $b \geq 1$, and $c > 0$. We have

$$\int_{-\infty}^\infty e^{-cz^2+hz} (z^2)^{b-1} z U(a;b;cz^2) \mathrm{d}z = \frac{h\Gamma\left(b+\frac{1}{2}\right)\Gamma\left(\frac{3}{2}\right)}{c^{b+\frac{1}{2}}\Gamma\left(a+\frac{3}{2}\right)} {}_1F_1\left(b+\frac{1}{2}; a+\frac{3}{2}; \frac{h^2}{4c}\right), \tag{A52}$$

$$\int_0^\infty e^{-w} w^{b+c-2} {}_0F_1(c;hw) U(a;b;w) \mathrm{d}w = \frac{\Gamma(c)\Gamma(b+c-1)}{\Gamma(a+c)} {}_1F_1(b+c-1; a+c; h). \tag{A53}$$

For (A52), we write

$$\begin{aligned}
\int_{-\infty}^\infty e^{-cz^2+hz} (z^2)^{b-1} z U(a;b;cz^2) \mathrm{d}z &= \int_0^\infty e^{-cz^2} (e^{hz} - e^{-hz}) z^{2b-1} U(a;b;cz^2) \mathrm{d}z \\
&= 2\int_0^\infty e^{-cz^2} \sum_{k=0}^\infty \frac{(hz)^{2k+1}}{(2k+1)!} z^{2b-1} U(a;b;cz^2) \mathrm{d}z. \qquad \text{(A54)}
\end{aligned}$$

Let $x = cz^2$, we have $\mathrm{d}x = 2cz\mathrm{d}z$ and

$$\begin{aligned}
\int_{-\infty}^\infty e^{-cz^2+hz} (z^2)^{b-1} z U(a;b;cz^2) \mathrm{d}z &= \frac{1}{c^b} \sum_{k=0}^\infty \int_0^\infty e^{-x} \frac{\left(\frac{h^2 x}{c}\right)^{k+\frac{1}{2}}}{(2k+1)!} x^{b-1} U(a;b;x) \mathrm{d}x \\
&= \frac{h}{c^{b+\frac{1}{2}}} \sum_{k=0}^\infty \frac{\left(\frac{h^2}{c}\right)^k}{4^k k! \left(\frac{3}{2}\right)_k} \int_0^\infty e^{-x} x^{b+k-\frac{1}{2}} U(a;b;x) \mathrm{d}x \\
&= \frac{h}{c^{b+\frac{1}{2}}} \sum_{k=0}^\infty \frac{\left(\frac{h^2}{4c}\right)^k}{k! \left(\frac{3}{2}\right)_k} \frac{\Gamma\left(b+\frac{1}{2}+k\right)\Gamma\left(\frac{3}{2}+k\right)}{\Gamma\left(a+\frac{3}{2}+k\right)} \\
&= \frac{h\Gamma\left(b+\frac{1}{2}\right)\Gamma\left(\frac{3}{2}\right)}{c^{b+\frac{1}{2}}\Gamma\left(a+\frac{3}{2}\right)} {}_1F_1\left(b+\frac{1}{2}; a+\frac{3}{2}; \frac{h^2}{4c}\right), \qquad \text{(A55)}
\end{aligned}$$

where the second last equality follows from (A51).

For (A53), we write

$$\begin{aligned}\int_0^\infty e^{-w} w^{b+c-2} {}_0F_1(c; hw) U(a; b; w) \mathrm{d}w &= \sum_{k=0}^\infty \int_0^\infty \frac{(hw)^k}{(c)_k k!} e^{-w} w^{b+c-2} U(a; b; w) \mathrm{d}w \\ &= \sum_{k=0}^\infty \frac{h^k}{(c)_k k!} \int_0^\infty e^{-w} w^{b+c-2+k} U(a; b; w) \mathrm{d}w \\ &= \sum_{k=0}^\infty \frac{h^k}{(c)_k k!} \frac{\Gamma(c+k)\Gamma(b+c-1+k)}{\Gamma(a+c+k)} \\ &= \frac{\Gamma(c)\Gamma(b+c-1)}{\Gamma(a+c)} {}_1F_1(b+c-1; a+c; h), \end{aligned} \tag{A56}$$

where the second last equality follows from (A51).

For $\mathbb{E}[\tilde{\theta}|\hat{\theta}]$, we use (A22) to write

$$\begin{aligned}\mathbb{E}[\tilde{\theta}|\hat{\theta}] &= \theta \mathbb{E}\left[\sqrt{b}\right] \mathbb{E}\left[\left. \frac{z_1}{\sqrt{z_1^2 + z_2^2 + u_0}} \right| \hat{\theta} \right] + \theta \mathbb{E}\left[\sqrt{1-b}\right] \mathbb{E}\left[\left. \frac{z_2}{\sqrt{z_1^2 + z_2^2 + u_0}} \right| \hat{\theta} \right] \\ &= \theta \mathbb{E}\left[\sqrt{b}\right] \mathbb{E}\left[\left. \frac{z_1}{\sqrt{z_1^2 + \tilde{u}_1}} \right| \hat{\theta} \right] + \theta \mathbb{E}\left[\sqrt{1-b}\right] \mathbb{E}\left[\left. \frac{z_2}{\sqrt{z_2^2 + \tilde{u}_2}} \right| \hat{\theta} \right] \\ &= \theta \frac{\mathrm{B}\left(\frac{T-N+2}{2}, \frac{N-1}{2}\right)}{\mathrm{B}\left(\frac{T-N+1}{2}, \frac{N-1}{2}\right)} \mathbb{E}\left[\left. \frac{z_1}{\sqrt{z_1^2 + \tilde{u}_1}} \right| \hat{\theta} \right] + \theta \frac{\mathrm{B}\left(\frac{N}{2}, \frac{T-N+1}{2}\right)}{\mathrm{B}\left(\frac{N-1}{2}, \frac{T-N+1}{2}\right)} \mathbb{E}\left[\left. \frac{z_2}{\sqrt{z_2^2 + \tilde{u}_2}} \right| \hat{\theta} \right], \end{aligned} \tag{A57}$$

where $\tilde{u}_1 = z_2^2 + u_0 \sim \chi^2_{N-1}$ and $\tilde{u}_2 = z_1^2 + u_0 \sim \chi^2_{N-1}(T\theta^2)$. The joint density of z_1, $\tilde{u}_1$, and u_1, and that of z_2, $\tilde{u}_2$, and u_1, where $u_1 \sim \chi^2_{T-N}$, are given by

$$f(z_1, \tilde{u}_1, u_1) = \phi\left(z_1 - \sqrt{T}\theta\right) f_{\tilde{u}_1}(\tilde{u}_1) f_{u_1}(u_1), \tag{A58}$$

$$f(z_2, \tilde{u}_2, u_1) = \phi(z_2) f_{\tilde{u}_2}(\tilde{u}_2) f_{u_1}(u_1). \tag{A59}$$

Using the representation of $\hat{\theta}$ in (A21), we make the change of variable $\hat{\theta} = \sqrt{z_1^2 + \tilde{u}_1}/\sqrt{u_1}$ and $\hat{\theta} = \sqrt{z_2^2 + \tilde{u}_2}/\sqrt{u_1}$ in (A58) and (A59), respectively, we obtain

$$f(z_1, \tilde{u}_1, \hat{\theta}) = \phi\left(z_1 - \sqrt{T}\theta\right) f_{\tilde{u}_1}(\tilde{u}_1) f_{u_1}\left(\frac{z_1^2 + \tilde{u}_1}{\hat{\theta}^2}\right) \frac{2(z_1^2 + \tilde{u}_1)}{\hat{\theta}^3}, \tag{A60}$$

$$f(z_2, \tilde{u}_2, \hat{\theta}) = \phi(z_2) f_{\tilde{u}_2}(\tilde{u}_2) f_{u_1}\left(\frac{z_2^2 + \tilde{u}_2}{\hat{\theta}^2}\right) \frac{2(z_2^2 + \tilde{u}_2)}{\hat{\theta}^3}. \tag{A61}$$

The first expectation term has the following explicit expression:

$$\mathbb{E}\left[\left. \frac{z_1}{\sqrt{z_1^2 + \tilde{u}_1}} \right| \hat{\theta} \right]$$

$$
\begin{aligned}
&= \int_0^{\infty}\int_{-\infty}^{\infty} \frac{z_1}{\sqrt{z_1^2+\tilde{u}_1}} \frac{f(z_1,\tilde{u}_1,\hat{\theta})}{f_{\hat{\theta}}(\hat{\theta})} \mathrm{d}z_1 \mathrm{d}\tilde{u}_1 \\
&= \frac{1}{f_{\hat{\theta}}(\hat{\theta})} \int_{-\infty}^{\infty} \frac{\sqrt{2} z_1 e^{-\frac{z_1^2}{2\hat{\theta}^2}} \left(\frac{z_1^2}{2\hat{\theta}^2}\right)^{\frac{T-2}{2}} \hat{\theta}^{N-3}}{\Gamma\left(\frac{T-N}{2}\right)} U\left(\frac{N-1}{2};\frac{T}{2};\frac{z_1^2(1+\hat{\theta}^2)}{2\hat{\theta}^2}\right) \phi(z_1-\sqrt{T}\theta)\mathrm{d}z_1 \\
&= \frac{e^{-\frac{T\theta^2}{2}}\hat{\theta}^{N-3}}{f_{\hat{\theta}}(\hat{\theta})\sqrt{\pi}\Gamma\left(\frac{T-N}{2}\right)} \int_{-\infty}^{\infty} z_1 e^{-\frac{z_1^2(1+\hat{\theta}^2)}{2\hat{\theta}^2}+\sqrt{T}\theta z_1} \left(\frac{z_1^2}{2\hat{\theta}^2}\right)^{\frac{T-2}{2}} U\left(\frac{N-1}{2};\frac{T}{2};\frac{z_1^2(1+\hat{\theta}^2)}{2\hat{\theta}^2}\right)\mathrm{d}z_1 \\
&= \frac{e^{-\frac{T\theta^2}{2}}\sqrt{2}\sqrt{T}\theta\Gamma\left(\frac{T+1}{2}\right)\hat{\theta}^N}{f_{\hat{\theta}}(\hat{\theta})\Gamma\left(\frac{T-N}{2}\right)\Gamma\left(\frac{N+2}{2}\right)(1+\hat{\theta}^2)^{\frac{T+1}{2}}} {}_1F_1\left(\frac{T+1}{2};\frac{N+2}{2};\frac{T\theta^2\hat{\theta}^2}{2(1+\hat{\theta}^2)}\right), \qquad \text{(A62)}
\end{aligned}
$$

where the last equality follows from (A52). For the second expectation term, we have

$$
\mathbb{E}\left[\frac{z_2}{\sqrt{z_2^2+\tilde{u}_2}}\middle|\hat{\theta}\right] = \int_0^{\infty}\int_{-\infty}^{\infty} \frac{z_2}{\sqrt{z_2^2+\tilde{u}_2}} \frac{f(z_2,\tilde{u}_2,\hat{\theta})}{f_{\hat{\theta}}(\hat{\theta})}\mathrm{d}z_2\mathrm{d}\tilde{u}_2 = 0 \qquad \text{(A63)}
$$

because of symmetry. It follows that

$$
\mathbb{E}[\tilde{\theta}|\hat{\theta}] = \frac{\theta^2 e^{-\frac{T\theta^2}{2}}\hat{\theta}^N}{f_{\hat{\theta}}(\hat{\theta})(1+\hat{\theta}^2)^{\frac{T+1}{2}}} \frac{\sqrt{T}(T-N)\Gamma\left(\frac{T}{2}\right)}{\sqrt{2}\Gamma\left(\frac{N+2}{2}\right)\Gamma\left(\frac{T-N+1}{2}\right)} {}_1F_1\left(\frac{T+1}{2};\frac{N+2}{2};\frac{T\theta^2\hat{\theta}^2}{2(1+\hat{\theta}^2)}\right). \qquad \text{(A64)}
$$

Using the expression of $f_{\hat{\theta}}(\hat{\theta})$ from (19) and denoting $y = T\theta^2\hat{\theta}^2/[2(1+\hat{\theta}^2)]$, we obtain

$$
\mathbb{E}[\tilde{\theta}|\hat{\theta}] = \frac{2\theta\sqrt{y}}{N} \frac{\Gamma\left(\frac{T-N+2}{2}\right){}_1F_1\left(\frac{T+1}{2};\frac{N+2}{2};y\right)}{\Gamma\left(\frac{T-N+1}{2}\right){}_1F_1\left(\frac{T}{2};\frac{N}{2};y\right)}. \qquad \text{(A65)}
$$

For $\mathbb{E}[\tilde{\theta}^2|\hat{\theta}]$, we use (A22) to write

$$
\begin{aligned}
\mathbb{E}[\tilde{\theta}^2|\hat{\theta}] &= \theta^2\mathbb{E}[b]\mathbb{E}\left[\frac{z_1^2}{z_1^2+z_2^2+u_0}\middle|\hat{\theta}\right] + 2\theta^2\mathbb{E}[\sqrt{b(1-b)}]\mathbb{E}\left[\frac{z_1z_2}{z_1^2+z_2^2+u_0}\middle|\hat{\theta}\right] \\
&\quad + \theta^2\mathbb{E}[1-b]\mathbb{E}\left[\frac{z_2^2}{z_1^2+z_2^2+u_0}\middle|\hat{\theta}\right] \\
&= \theta^2\left(\frac{T-N+1}{T}\mathbb{E}\left[\frac{z_1^2}{z_1^2+\tilde{u}_1}\middle|\hat{\theta}\right] + \frac{N-1}{T}\mathbb{E}\left[\frac{z_2^2}{z_2^2+\tilde{u}_2}\middle|\hat{\theta}\right]\right), \qquad \text{(A66)}
\end{aligned}
$$

where the second term drops out because of symmetry. Again, by symmetry, we have

$$
\mathbb{E}\left[\frac{z_1^2}{z_1^2+z_2^2+u_0}\middle|\hat{\theta}\right] + (N-1)\mathbb{E}\left[\frac{z_2^2}{z_1^2+z_2^2+u_0}\middle|\hat{\theta}\right] = 1. \qquad \text{(A67)}
$$

It follows that

$$
\mathbb{E}\left[\frac{z_1^2}{z_1^2+\tilde{u}_1}\middle|\hat{\theta}\right] = 1-(N-1)\mathbb{E}\left[\frac{z_2^2}{z_2^2+\tilde{u}_2}\middle|\hat{\theta}\right] \qquad \text{(A68)}
$$

and we have

$$\mathbb{E}[\tilde{\theta}^2|\hat{\theta}] = \theta^2 \left(\frac{T-N+1}{T} - \frac{(T-N)(N-1)}{T} \mathbb{E}\left[\frac{z_2^2}{z_2^2+\tilde{u}_2} \middle| \hat{\theta} \right] \right). \tag{A69}$$

Following the same transformation as before, we have

$$\begin{aligned}
&\mathbb{E}\left[\frac{z_2^2}{z_2^2+\tilde{u}_2} \middle| \hat{\theta} \right] \\
&= \int_0^\infty \int_{-\infty}^\infty \frac{z_2^2}{z_2^2+\tilde{u}_2} \frac{f(z_2,\tilde{u}_2,\hat{\theta})}{f_{\hat{\theta}}(\hat{\theta})} \mathrm{d}z_2 \mathrm{d}\tilde{u}_2 \\
&= 2\int_0^\infty \int_0^\infty \frac{z_2^2}{z_2^2+\tilde{u}_2} \frac{f(z_2,\tilde{u}_2,\hat{\theta})}{f_{\hat{\theta}}(\hat{\theta})} \mathrm{d}z_2 \mathrm{d}\tilde{u}_2 \\
&= \frac{4}{f_{\hat{\theta}}(\hat{\theta})\hat{\theta}^3\sqrt{2\pi}2^{\frac{T-N}{2}}\Gamma\left(\frac{T-N}{2}\right)} \int_0^\infty f_{\tilde{u}_2}(\tilde{u}_2) \left[\int_0^\infty z_2^2 e^{-\frac{z_2^2}{2}} e^{-\frac{\tilde{u}_2+z_2^2}{2\hat{\theta}^2}} \left(\frac{\tilde{u}_2+z_2^2}{\hat{\theta}^2} \right)^{\frac{T-N-2}{2}} \mathrm{d}z_2 \right] \mathrm{d}\tilde{u}_2 \\
&= \frac{1}{f_{\hat{\theta}}(\hat{\theta})\hat{\theta}^{T-N+1}\sqrt{\pi}2^{\frac{T-N-3}{2}}\Gamma\left(\frac{T-N}{2}\right)} \int_0^\infty e^{-\frac{\tilde{u}_2}{2\hat{\theta}^2}} f_{\tilde{u}_2}(\tilde{u}_2) \left[\int_0^\infty z_2^2 e^{-\frac{z_2^2(1+\hat{\theta}^2)}{2\hat{\theta}^2}} (\tilde{u}_2+z_2^2)^{\frac{T-N-2}{2}} \mathrm{d}z_2 \right] \mathrm{d}\tilde{u}_2.
\end{aligned} \tag{A70}$$

Making a transformation of $z = z_2^2/\tilde{u}_2$, we have $\tilde{u}_2 \mathrm{d}z = 2z_2 \mathrm{d}z_2$ and the term in the squared brackets can be simplified to

$$\begin{aligned}
\int_0^\infty z_2^2 e^{-\frac{z_2^2(1+\hat{\theta}^2)}{2\hat{\theta}^2}} (\tilde{u}_2+z_2^2)^{\frac{T-N-2}{2}} \mathrm{d}z_2 &= \frac{\tilde{u}_2^{\frac{T-N+1}{2}}}{2} \int_0^\infty e^{-\frac{(1+\hat{\theta}^2)\tilde{u}_2 z}{2\hat{\theta}^2}} z^{\frac{1}{2}} (1+z)^{\frac{T-N-2}{2}} \mathrm{d}z \\
&= \frac{\tilde{u}_2^{\frac{T-N+1}{2}}\sqrt{\pi}}{4} U\left(\frac{3}{2}; \frac{T-N+3}{2}; \frac{(1+\hat{\theta}^2)\tilde{u}_2}{2\hat{\theta}^2} \right),
\end{aligned} \tag{A71}$$

where in the last equality we use the definition of $U(a;b;x)$ from (A50). Using this expression and the expression of $f_{\tilde{u}_2}(\tilde{u}_2)$, we have

$$\begin{aligned}
&\mathbb{E}\left[\frac{z_2^2}{z_2^2+\tilde{u}_2} \middle| \hat{\theta} \right] \\
&= \frac{1}{f_{\hat{\theta}}(\hat{\theta})\hat{\theta}^{T-N+1}\sqrt{\pi}2^{\frac{T-N-3}{2}}\Gamma\left(\frac{T-N}{2}\right)} \\
&\quad \times \int_0^\infty e^{-\frac{\tilde{u}_2}{2\hat{\theta}^2}} f_{\tilde{u}_2}(\tilde{u}_2) \frac{\tilde{u}_2^{\frac{T-N+1}{2}}\sqrt{\pi}}{4} U\left(\frac{3}{2}; \frac{T-N+3}{2}; \frac{(1+\hat{\theta}^2)\tilde{u}_2}{2\hat{\theta}^2} \right) \mathrm{d}\tilde{u}_2 \\
&= \frac{e^{-\frac{T\theta^2}{2}}}{f_{\hat{\theta}}(\hat{\theta})\hat{\theta}^{T-N+1}2^{\frac{T}{2}}\Gamma\left(\frac{T-N}{2}\right)\Gamma\left(\frac{N-1}{2}\right)}
\end{aligned}$$

$$\times \int_0^\infty e^{-\frac{(1+\hat{\theta}^2)\tilde{u}_2}{2\hat{\theta}^2}} \tilde{u}_2^{\frac{T-2}{2}} {}_0F_1\left(\frac{N-1}{2}; \frac{T\theta^2\tilde{u}_2}{4}\right) U\left(\frac{3}{2}; \frac{T-N+3}{2}; \frac{(1+\hat{\theta}^2)\tilde{u}_2}{2\hat{\theta}^2}\right) \mathrm{d}\tilde{u}_2$$

$$= \frac{e^{-\frac{T\theta^2}{2}}}{f_{\hat{\theta}}(\hat{\theta})\hat{\theta}^{T-N+1}2^{\frac{T}{2}}\Gamma\left(\frac{T-N}{2}\right)\Gamma\left(\frac{N-1}{2}\right)}$$

$$\times \int_0^\infty e^{-w}\left(\frac{2\hat{\theta}^2}{1+\hat{\theta}^2}\right)^{\frac{T}{2}} w^{\frac{T-2}{2}} {}_0F_1\left(\frac{N-1}{2}; \frac{T\theta^2\hat{\theta}^2 w}{2(1+\hat{\theta}^2)}\right) U\left(\frac{3}{2}; \frac{T-N+3}{2}; w\right) \mathrm{d}w$$

$$= \frac{\Gamma\left(\frac{T}{2}\right)\hat{\theta}^{N-1}e^{-\frac{T\theta^2}{2}}}{f_{\hat{\theta}}(\hat{\theta})\Gamma\left(\frac{N+2}{2}\right)\Gamma\left(\frac{T-N}{2}\right)(1+\hat{\theta}^2)^{\frac{T}{2}}} {}_1F_1\left(\frac{T}{2}; \frac{N+2}{2}; \frac{T\theta^2\hat{\theta}^2}{2(1+\hat{\theta}^2)}\right), \tag{A72}$$

where we use the transformation of $w = (1+\hat{\theta}^2)\tilde{u}_2/(2\hat{\theta}^2)$ in the second last equality, and the last equality follows from (A53). Using (19) and denoting $y = T\theta^2\hat{\theta}^2/[2(1+\hat{\theta}^2)]$ in (A69), we obtain

$$\mathbb{E}[\tilde{\theta}^2|\hat{\theta}] = \theta^2\left[\frac{T-N+1}{T} - \frac{(N-1)(T-N)}{NT}\frac{{}_1F_1\left(\frac{T}{2}; \frac{N}{2}+1; y\right)}{{}_1F_1\left(\frac{T}{2}; \frac{N}{2}; y\right)}\right]. \tag{A73}$$

This completes the proof.

Proof of Proposition 3: Using the representation of $\hat{\theta}$ in (A21) and defining w_1 and w_2 as

$$z_1 - \sqrt{T}\theta = w_1 \sim \mathcal{N}(0,1), \tag{A74}$$

$$\frac{T-u_1}{\sqrt{2T}} \xrightarrow{d} w_2 \sim \mathcal{N}(0,1), \tag{A75}$$

where the limiting distribution of $(T-u_1)/\sqrt{2T}$ is obtained by using the central limit theorem, we can write

$$\begin{aligned}
\sqrt{T}(\hat{\theta}-\theta) &= \sqrt{T}\left(\frac{\sqrt{z_1^2+z_2^2+u_0}}{\sqrt{u_1}} - \theta\right) \\
&= \sqrt{T}\left(\frac{1}{\sqrt{u_1}} - \frac{1}{\sqrt{T}}\right)\sqrt{z_1^2+z_2^2+u_0} + \left(\sqrt{z_1^2+z_2^2+u_0} - \sqrt{T}\theta\right) \\
&= \frac{\sqrt{T}(T-u_1)}{\sqrt{T}\sqrt{u_1}(\sqrt{T}+\sqrt{u_1})}\sqrt{z_1^2+z_2^2+u_0} + \frac{z_1^2+z_2^2+u_0-T\theta^2}{\sqrt{z_1^2+z_2^2+u_0}+\sqrt{T}\theta} \\
&= \frac{\sqrt{2}(T-u_1)}{\sqrt{2T}}\left[\frac{1}{\frac{\sqrt{u_1}}{\sqrt{T}}\left(1+\frac{\sqrt{u_1}}{\sqrt{T}}\right)}\right]\left(\frac{z_1^2+z_2^2+u_0}{T}\right)^{\frac{1}{2}} + \frac{2\theta w_1 + \frac{w_1^2+z_2^2+u_0}{\sqrt{T}}}{\left(\frac{z_1^2+z_2^2+u_0}{T}\right)^{\frac{1}{2}}+\theta} \\
&\xrightarrow{d} \frac{\theta w_2}{\sqrt{2}} + w_1 \equiv X \sim \mathcal{N}\left(0, 1+\frac{\theta^2}{2}\right).
\end{aligned} \tag{A76}$$

The last equality is obtained by using the fact that $u_1/T \xrightarrow{p} 1$, $(z_1^2 + z_2^2 + u_0)/T \xrightarrow{p} \theta^2$, and $(w_1^2 + z_2^2 + u_0)/\sqrt{T} \xrightarrow{p} 0$.

Using the representation of $\tilde{\theta}$ in (A22) and defining v as

$$T(1-b) \xrightarrow{d} v \sim \chi^2_{N-1}, \tag{A77}$$

we can write

$$\begin{aligned}
T(\tilde{\theta} - \theta) &= T\theta \left(\frac{\sqrt{b}z_1 + \sqrt{1-b}z_2 - \sqrt{z_1^2 + z_2^2 + u_0}}{\sqrt{z_1^2 + z_2^2 + u_0}} \right) \\
&= -T\theta \left[\frac{(1-b)z_1^2 + bz_2^2 + u_0 - 2\sqrt{b(1-b)}z_1 z_2}{\sqrt{z_1^2 + z_2^2 + u_0}(\sqrt{b}z_1 + \sqrt{1-b}z_2 + \sqrt{z_1^2 + z_2^2 + u_0})} \right] \\
&= -\theta \left[\frac{(1-b)(T\theta^2 + 2\sqrt{T}\theta w_1 + w_1^2) + bz_2^2 + u_0 - 2\sqrt{b(1-b)}(\sqrt{T}\theta + w_1)z_2}{\left(\frac{z_1^2+z_2^2+u_0}{T}\right)^{\frac{1}{2}} \left[\frac{\sqrt{b}z_1}{\sqrt{T}} + \frac{\sqrt{1-b}z_2}{\sqrt{T}} + \left(\frac{z_1^2+z_2^2+u_0}{T}\right)^{\frac{1}{2}} \right]} \right] \\
&\xrightarrow{d} -\theta \left[\frac{\theta^2 v + z_2^2 + u_0 - 2\theta\sqrt{v}z_2}{\theta(\theta+\theta)} \right] \\
&= -\frac{\theta^2 v + z_2^2 + u_0 - 2\theta\sqrt{v}z_2}{2\theta} \equiv Y.
\end{aligned} \tag{A78}$$

The second last equality follows because $b \xrightarrow{p} 1$, $(1-b)\sqrt{T}\theta w_1 \xrightarrow{p} 0$, $(1-b)w_1^2 \xrightarrow{p} 0$, $\sqrt{b(1-b)}w_1 z_2 \xrightarrow{p} 0$, $(z_1^2 + z_2^2 + u_0)/T \xrightarrow{p} \theta^2$, $\sqrt{b}z_1/\sqrt{T} \xrightarrow{p} \theta$, $\sqrt{1-b}z_2/\sqrt{T} \xrightarrow{p} 0$. It remains to show that $Y \sim -(1+\theta^2)/(2\theta)\chi^2_{N-1}$. In order to show that, we let

$$W = \begin{bmatrix} \sqrt{v} & 0 \\ z_2 & \sqrt{u_0} \end{bmatrix} \begin{bmatrix} \sqrt{v} & z_2 \\ 0 & \sqrt{u_0} \end{bmatrix}. \tag{A79}$$

From the Bartlett decomposition of Wishart distribution, we know $W \sim \mathcal{W}_2(N-1, I_2)$. Then using 3.2.8 of Muirhead (1982), we have

$$Y = -\frac{\theta^2 v + z_2^2 + u_0 - 2\theta\sqrt{v}z_2}{2\theta} = -\frac{[\theta,\ -1]W[\theta,\ -1]'}{2\theta} \sim -\frac{(1+\theta^2)\chi^2_{N-1}}{2\theta}. \tag{A80}$$

Finally, X is independent of Y because X is a function z_1 and u_1, and Y is a function of z_2, b, u_0, and (z_1, u_1) are independent of (z_2, b, u_0) from the proof of Proposition 1. This completes the proof.

Proof of Proposition 4: Based on the definition of random variables in Proposition 1, we let

$$\tilde{z}_1 = \frac{z_1}{\sqrt{T}}, \tag{A81}$$

$$\tilde{z}_2 = \frac{z_2}{\sqrt{T}}, \tag{A82}$$

$$w_1 = \frac{u_0}{T}, \tag{A83}$$

$$w_2 = \frac{u_1}{T}. \tag{A84}$$

Using the central limit theorem, we can easily show that when $N \to \infty$, $T \to \infty$, and $N/T \to \rho$, we have

$$\sqrt{T}(\tilde{z}_1 - \theta) \sim \mathcal{N}(0, 1), \tag{A85}$$

$$\sqrt{T}\tilde{z}_2 \sim \mathcal{N}(0, 1), \tag{A86}$$

$$\sqrt{T}(w_1 - \rho) \xrightarrow{d} \mathcal{N}(0, 2\rho), \tag{A87}$$

$$\sqrt{T}(w_2 - (1-\rho)) \xrightarrow{d} \mathcal{N}(0, 2(1-\rho)), \tag{A88}$$

$$\sqrt{T}(b - (1-\rho)) \xrightarrow{d} \mathcal{N}(0, 2\rho(1-\rho)), \tag{A89}$$

and these five random variables are independent of each other. From (A21) and (A22), we can write $\hat{\theta}$ and $\tilde{\theta}$ as

$$\hat{\theta} = \frac{(\tilde{z}_1^2 + \tilde{z}_2^2 + w_1)^{\frac{1}{2}}}{w_2^{\frac{1}{2}}}, \tag{A90}$$

$$\tilde{\theta} = \frac{\theta(\sqrt{b}\tilde{z}_1 + \sqrt{1-b}\tilde{z}_2)}{(\tilde{z}_1^2 + \tilde{z}_2^2 + w_1)^{\frac{1}{2}}}, \tag{A91}$$

and both of them are functions of $(\tilde{z}_1, \tilde{z}_2, w_1, w_2, b)$. Then using the delta method and upon simplification, we obtain

$$\sqrt{T}\left(\begin{bmatrix} \hat{\theta} \\ \tilde{\theta} \end{bmatrix} - \begin{bmatrix} \frac{\sqrt{\theta^2+\rho}}{\sqrt{1-\rho}} \\ \frac{\sqrt{1-\rho}\theta^2}{\sqrt{\theta^2+\rho}} \end{bmatrix}\right) \xrightarrow{d} N\left(0_2, \begin{bmatrix} \frac{\theta^4+2\theta^2+\rho}{2(1-\rho)^2(\theta^2+\rho)} & \frac{\rho\theta^2}{2(\theta^2+\rho)^2} \\ \frac{\rho\theta^2}{2(\theta^2+\rho)^2} & \frac{\rho\theta^2}{2(\theta^2+\rho)}\left[\frac{(1-\rho)(2\rho+\theta^2)}{(\theta^2+\rho)^2} + 2 + \theta^2\right] \end{bmatrix}\right). \tag{A92}$$

This completes the proof.

References

Ao, Mengmeng, Yingying Li, and Xinghua Zheng, 2019, Approaching mean-variance efficiency for large portfolios, *Review of Financial Studies* 32, 2890–2919.

Asness, Clifford, and Andrea Frazzini, 2013, The devil in HML's details, *Journal of Portfolio Management* 39, 49–68.

Asness, Clifford, Andrea Frazzini, and Lasse H. Pedersen, 2019, Quality minus junk, *Review of Accounting Studies* 24, 34–112.

Barillas, Francisco, Raymond Kan, Cesare Robotti, and Jay Shanken, 2019, Model comparison with Sharpe ratios, *Journal of Financial and Quantitative Analysis*, forthcoming.

Barillas, Francisco, and Jay Shanken, 2017, Which alpha? *Review of Financial Studies* 30, 1316–1338.

Barillas, Francisco, and Jay Shanken, 2018, Comparing asset pricing models, *Journal of Finance* 73, 715–754.

Carhart, Mark, 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Casella, George, and Roger L. Berger. 1990. *Statistical Inference*. Duxbury Press, CA, USA.

Cochrane, John H., and Jesús Saá-Requejo, 2000, Beyond arbitrage: Good-deal asset price bounds in incomplete markets, *Journal of Political Economy* 108, 79–119.

Detzel, Andrew, Robert Novy-Marx, and Mihail Velikov, 2019, Model selection with transaction costs, working paper.

Dickey, James M., 1967, Matricvariate generalizations of the multivariate t distribution and the inverted multivariate t distribution, *Annals of Mathematical Statistics* 38, 511–518.

El Karoui, Noureddine, 2010, High-dimensionality effects in the Markowitz problem and other quadratic programs with linear constraints: risk underestimation, *Annals of Statistics* 38, 3487–3566.

El Karoui, Noureddine, 2013, On the realized risk of high-dimensional Markowitz portfolios, *SIAM Journal on Financial Mathematics* 4, 737–783.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and Kenneth R. French, 2018, Choosing factors, *Journal of Financial Economics* 128, 234–252.

Frazzini, Andrea, David Kabiller, and Lasse H. Pedersen, 2018, Buffett's alpha, *Financial Analysts Journal* 74, 35–55.

Frazzini, Andrea, and Lasse H. Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Gibbons, Michael, Stephen Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Gradshteyn, I. S., and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th edition (Academic Press, London).

Harvey, Campbell, and Yan Liu, 2013, Multiple testing in economics, working paper, Duke University.

Hillier, Grant, Raymond Kan, and Xiaolu Wang, 2014, Generating functions and short recursions, with applications to the moments of quadratic Forms in noncentral normal vectors, *Econometric Theory* 30, 436–473.

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang, 2019, q^5, working paper, Ohio State University.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Kan, Raymond, and Xiaolu Wang, 2019, On the economic value of alpha, Working paper, University of Toronto.

Kan, Raymond, and Cesare Robotti, 2016, The exact distribution of the Hansen-Jagannathan bound, *Management Science* 62, 1915–1943.

Kalmykov, S. I., and D. B. Karp, 2013, Log-concavity for series in reciprocal gamma functions and applications, *Integral Transforms and Special Functions* 24, 859–872.

Krishnan, M., 1967, The moments of a doubly noncentral t-distribution. *Journal of the American Statistical Association* 62, 278–287.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13–37.

Lo, Andrew W., 2002, The statistics of Sharpe ratios, *Financial Analyst Journal* 58, 36–52.

Lo, Andrew W., and A. Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431–467.

MacKinlay, A. Craig, 1995, Multi-factor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3–28.

Muirhead, Robb J., 1982, *Aspects of Multivariate Statistical Theory* (Wiley, New York).

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Patton, Andrew J., and Brian M. Weller, 2019, What you see is not what you get: the costs of trading market anomalies, *Journal of Financial Economics*, forthcoming.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

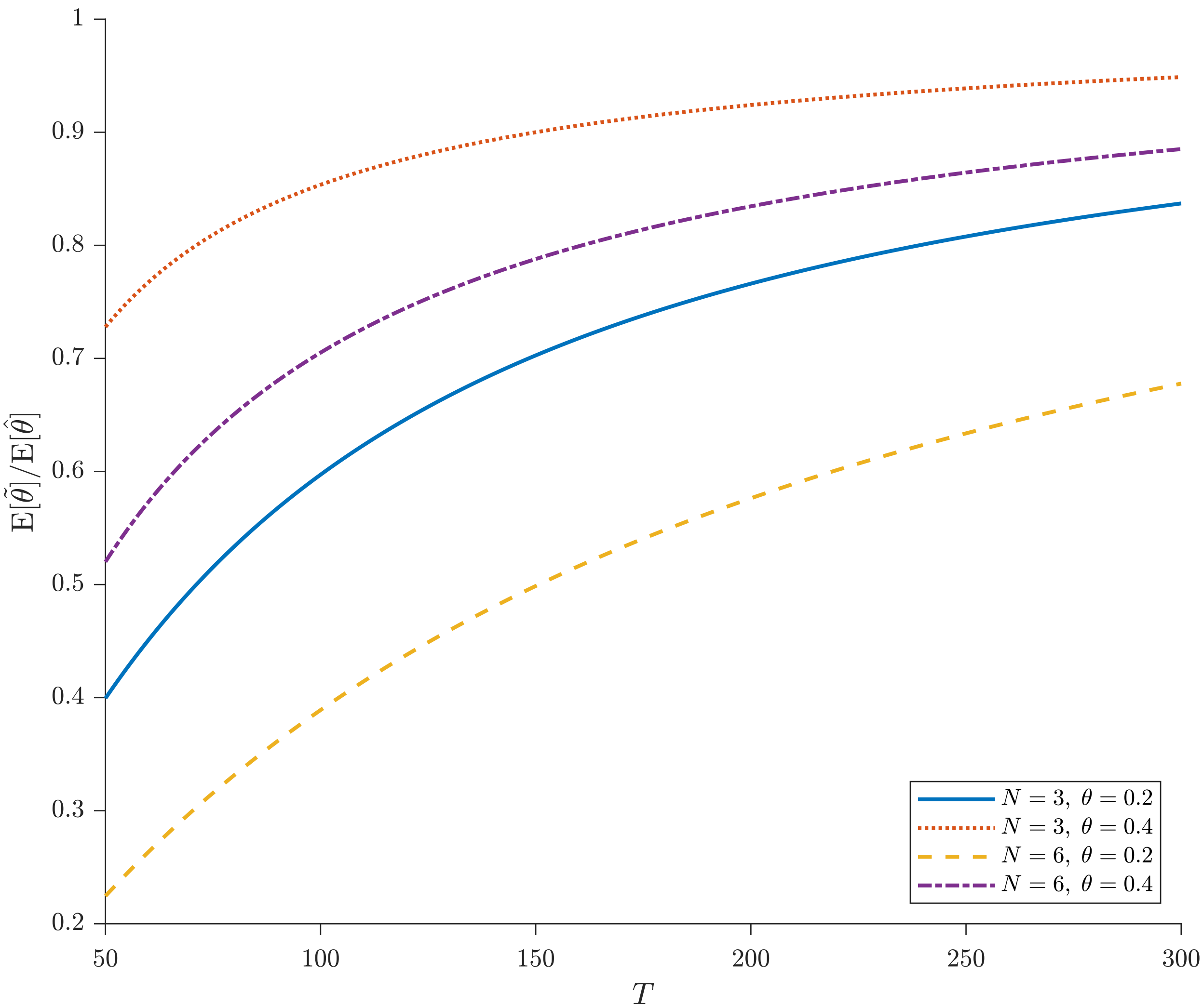

Figure 1: **Ratio of Expected Out-of-sample Sharpe Ratio to Expected In-sample Sharpe Ratio of an Asset Pricing Model**

The figure plots the ratio of expected out-of-sample Sharpe ratio to expected in-sample Sharpe ratio, $\mathbb{E}[\tilde{\theta}]/\mathbb{E}[\hat{\theta}]$, as a function of sample size, T, for an asset pricing model with N traded factors and a population Sharpe ratio of θ. Plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

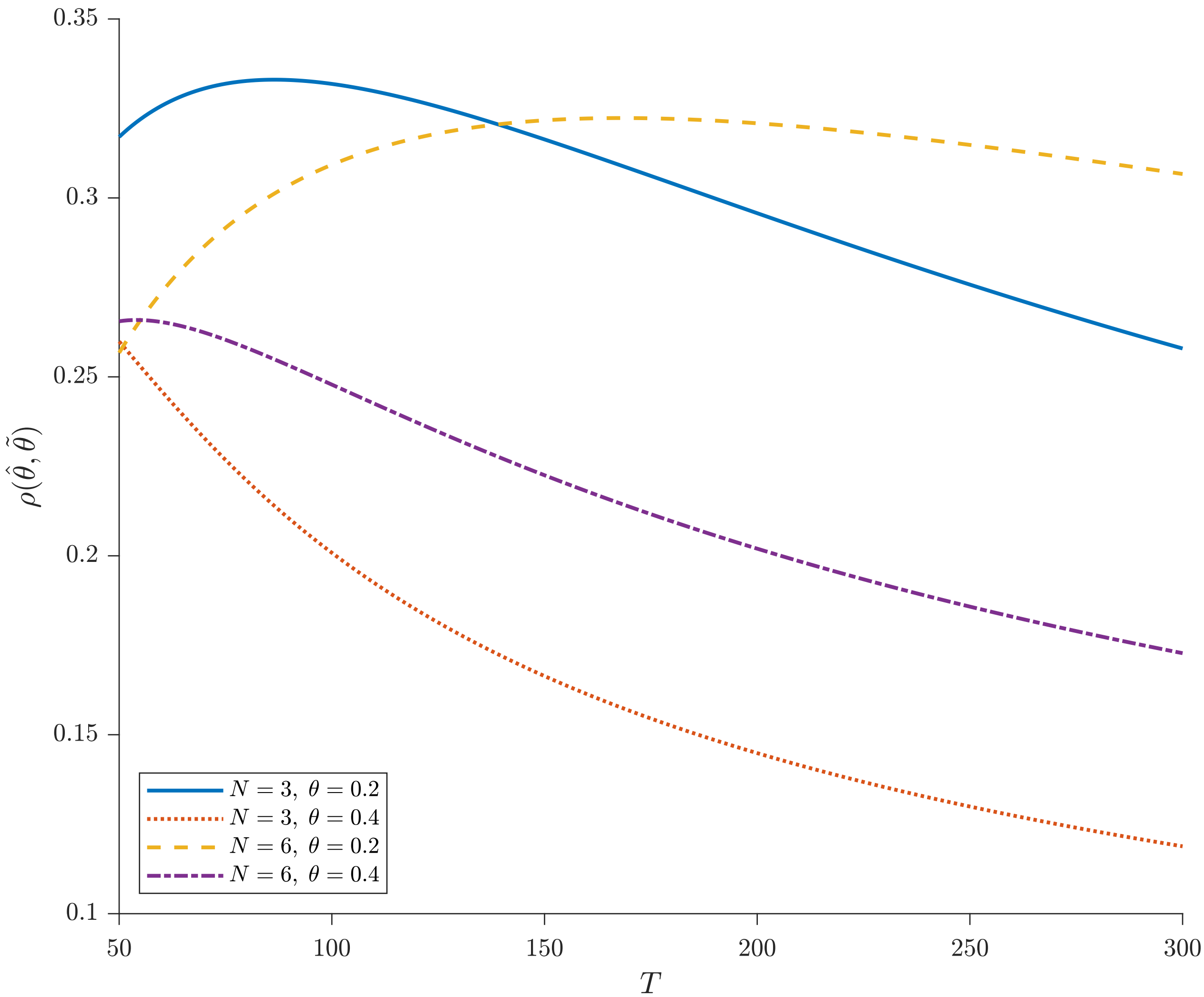

Figure 2: **Correlation between In-sample and Out-of-sample Sharpe Ratios of an Asset Pricing Model**

The figure plots the correlation coefficient between the in-sample and out-of-sample Sharpe ratios, $\rho(\hat{\theta}, \tilde{\theta})$, as a function of sample size, T, for an asset pricing model with N traded factors and a population Sharpe ratio of θ. Plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

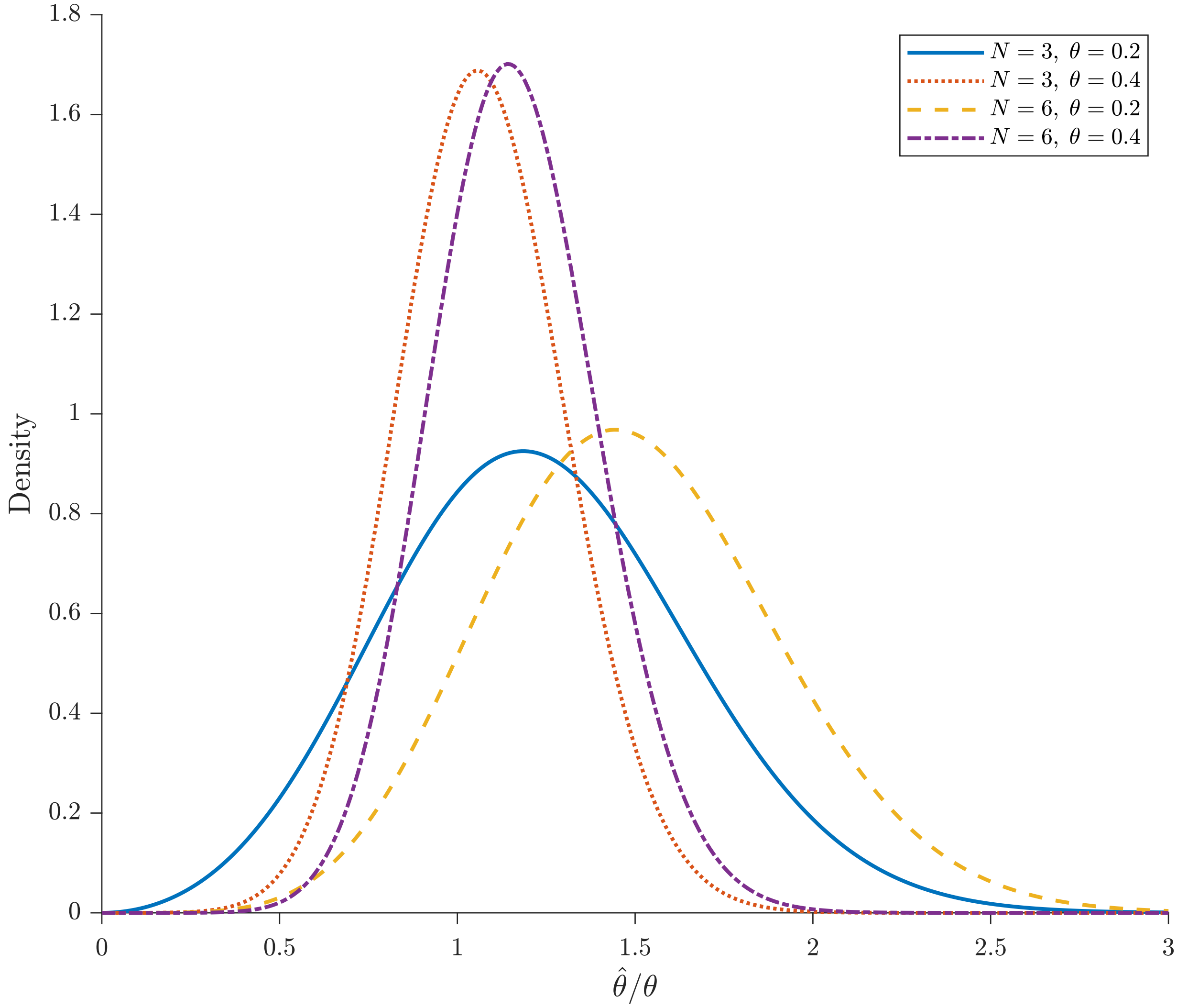

Figure 3: **Density of In-sample Sharpe Ratio of an Asset Pricing Model ($T = 120$)**

The figure plots the density of $\hat{\theta}/\theta$ of an asset pricing model with N traded factors when the length of time series is $T = 120$, where $\hat{\theta}$ is the in-sample Sharpe ratio and θ is the population Sharpe ratio. Plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

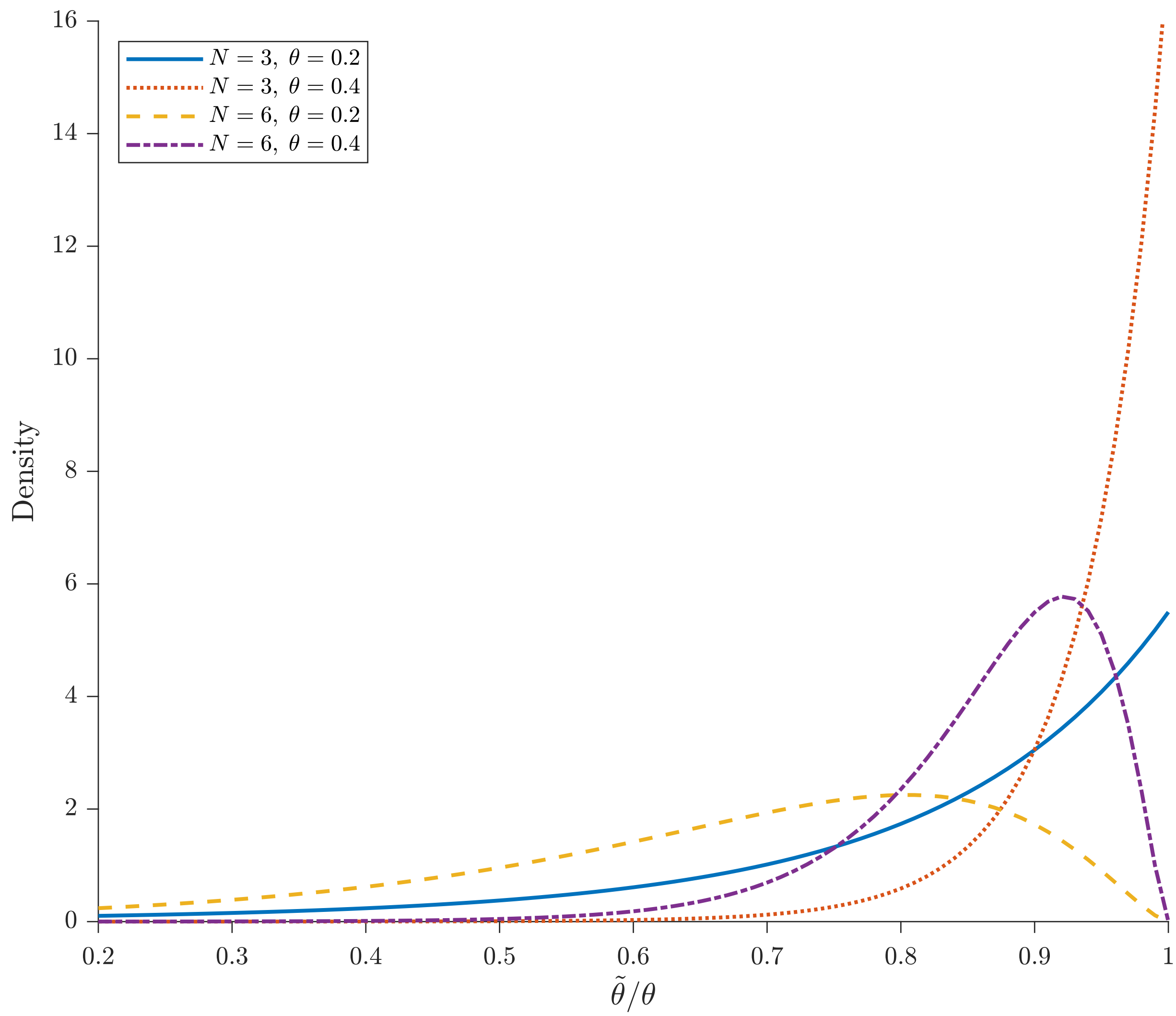

Figure 4: **Density of Out-of-sample Sharpe Ratio of an Asset Pricing Model ($T = 120$)**

The figure plots the density of $\tilde{\theta}/\theta$ of an asset pricing model with N traded factors when the length of time series is $T = 120$, where $\tilde{\theta}$ is the out-of-sample Sharpe ratio and θ is the population Sharpe ratio. Plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

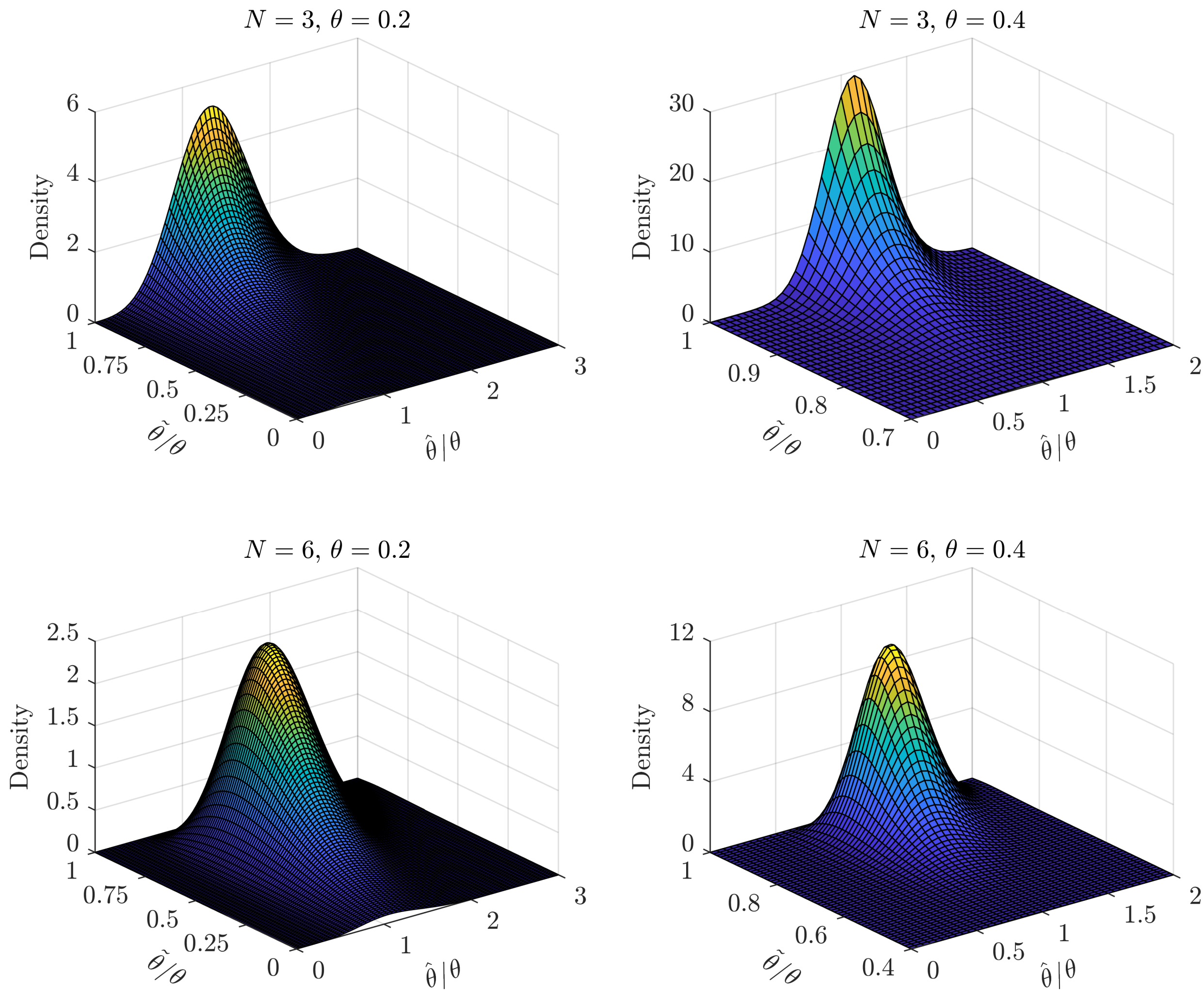

Figure 5: **Joint Density of In-sample and Out-of-sample Sharpe Ratios of an Asset Pricing Model ($T = 120$)**

The figure plots the joint density of $\hat{\theta}/\theta$ and $\tilde{\theta}/\theta$ of an asset pricing model with N traded factors when the length of time series is $T = 120$, where $\hat{\theta}$ is the in-sample Sharpe ratio, $\tilde{\theta}$ is the out-of-sample Sharpe ratio, and θ is the population Sharpe ratio. Plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

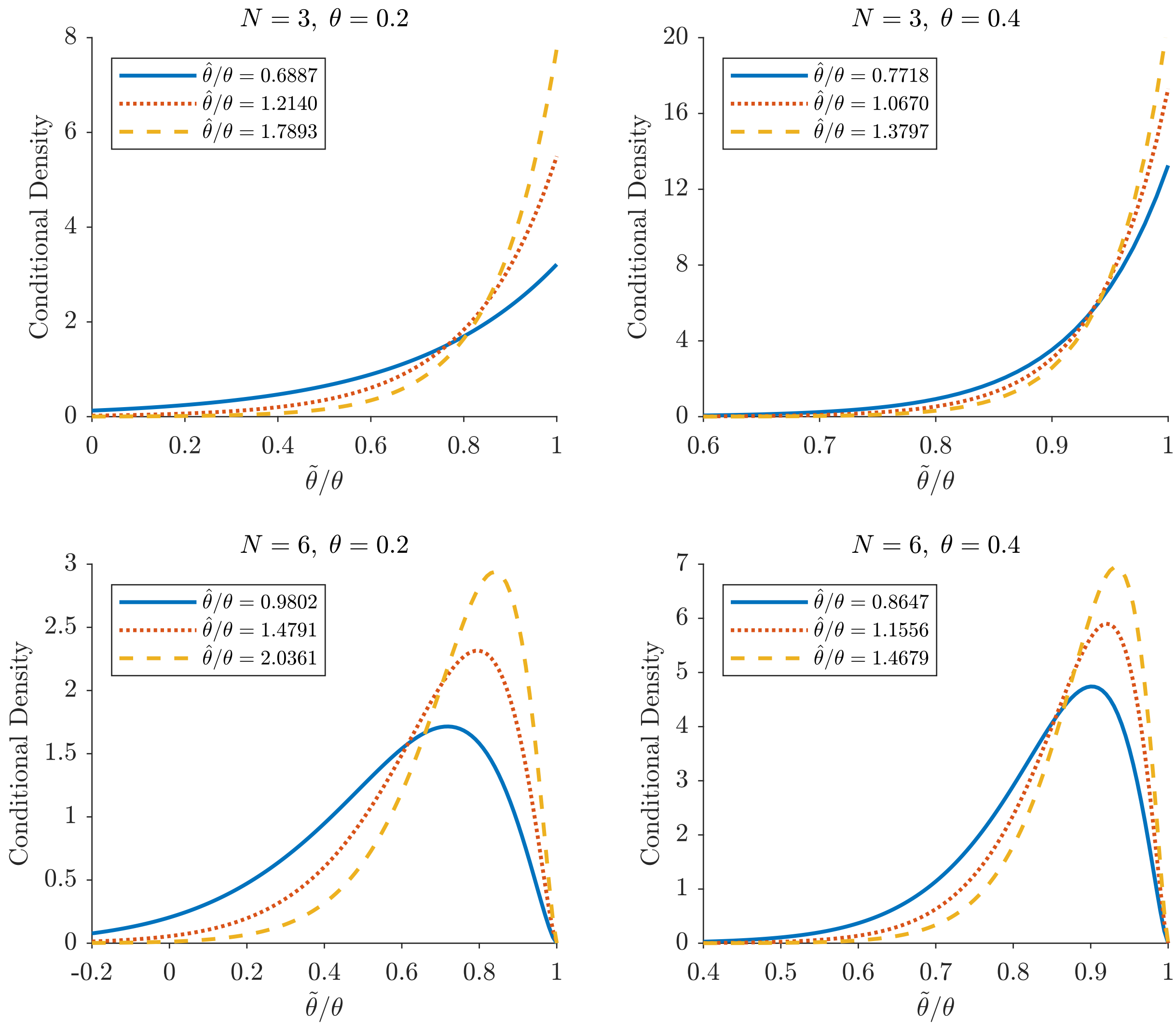

Figure 6: **Conditional Density of Out-of-sample Sharpe Ratio of an Asset Pricing Model ($T = 120$)**

The figure plots the conditional density of the normalized out-of-sample Sharpe ratio ($\tilde{\theta}/\theta$) of an asset pricing model when conditional on the in-sample Sharpe ratio ($\hat{\theta}$) is at its 10th (solid line), 50th (dotted line), and 90th (dashed line) percentiles. The length of time series is $T = 120$ and plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

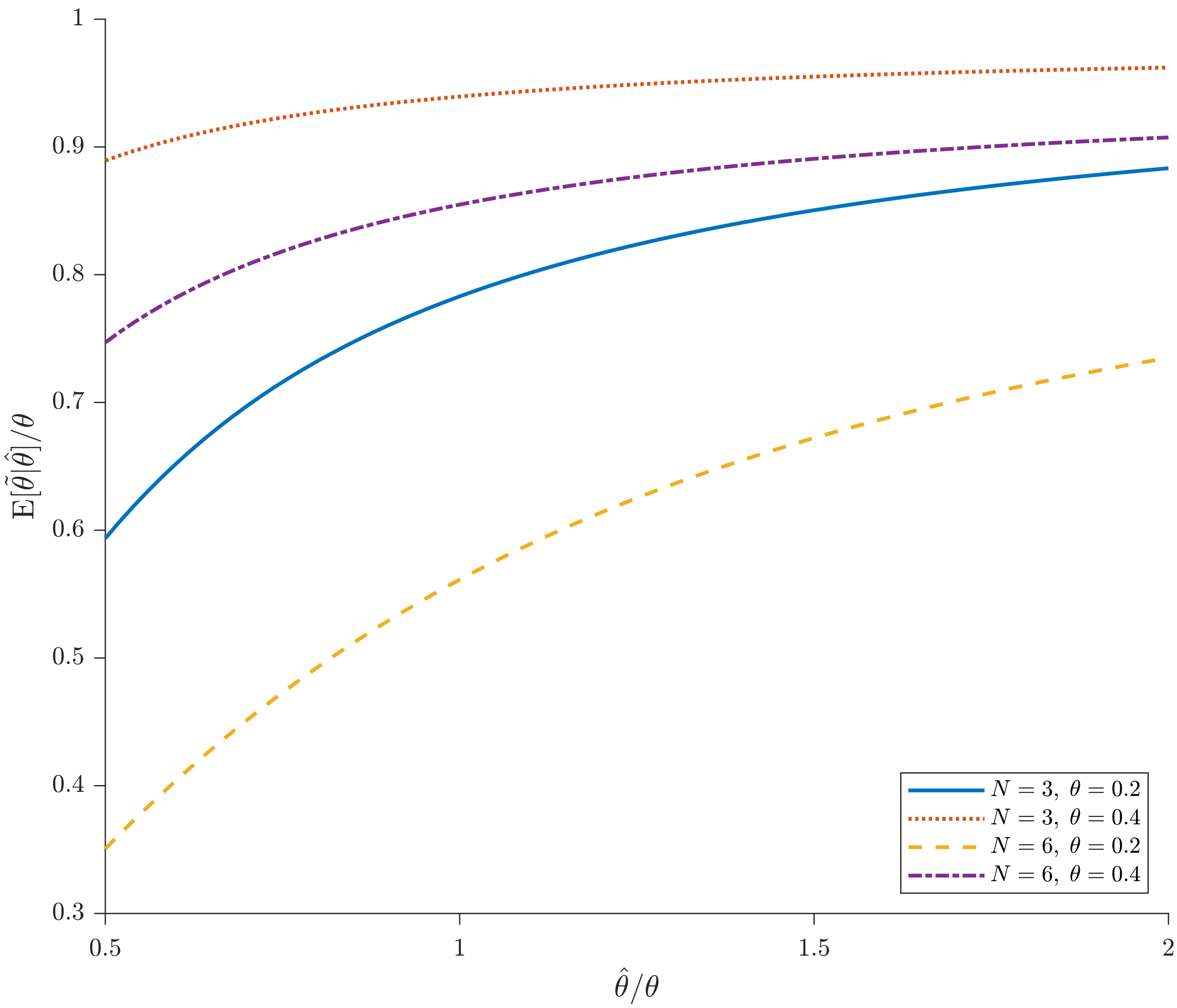

Figure 7: **Expected Conditional Out-of-sample Sharpe Ratio of an Asset Pricing Model ($T = 120$)**

The figure plots the expected conditional normalized out-of-sample Sharpe ratio ($\mathbb{E}[\tilde{\theta}|\hat{\theta}]/\theta$) of an asset pricing model as a function of normalized in-sample Sharpe ratio ($\hat{\theta}/\theta$). The length of time series is $T = 120$ and plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

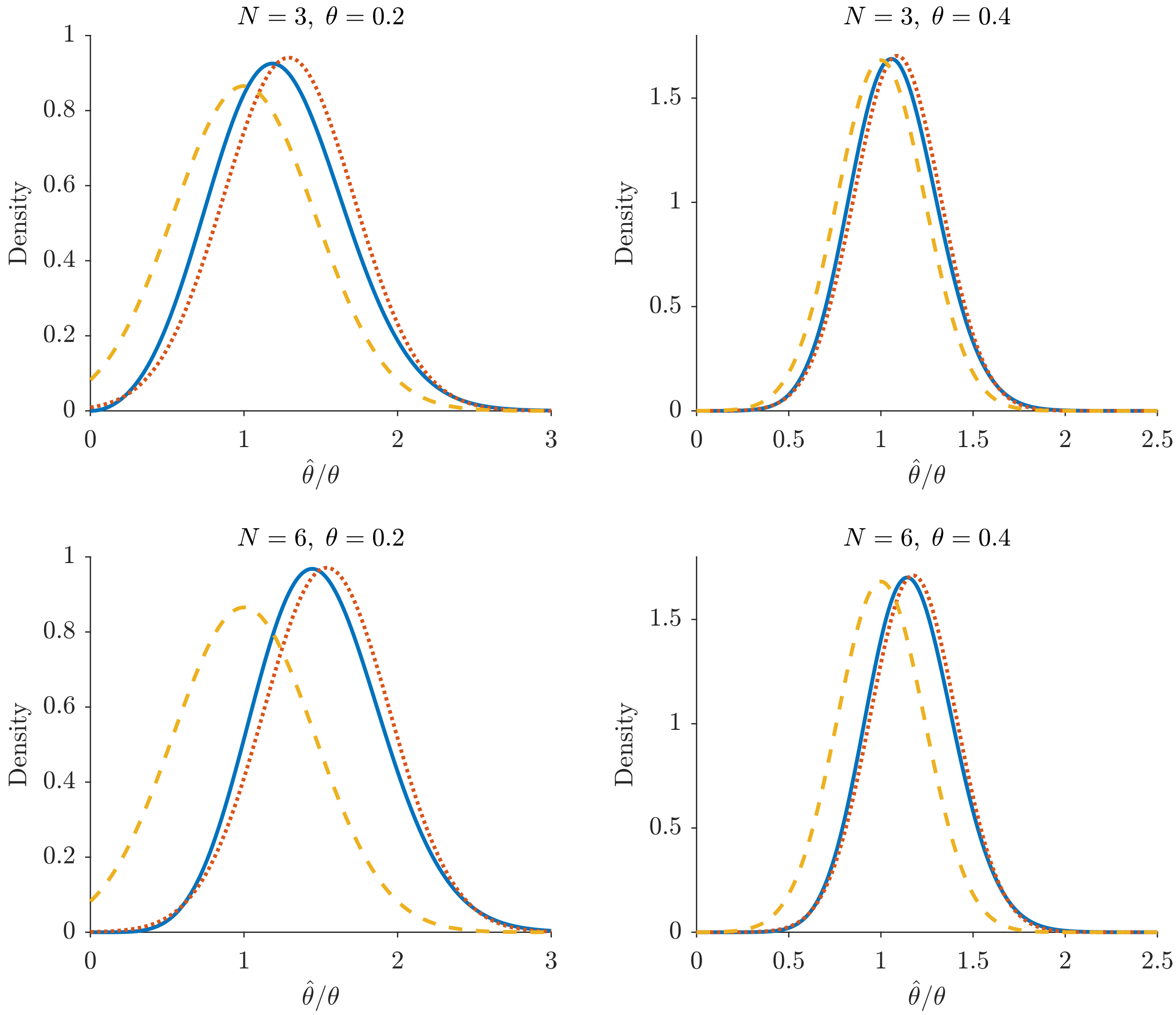

Figure 8: **Exact and Approximate Densities of In-sample Sharpe Ratio of an Asset Pricing Model ($T = 120$)**

The figure plots the exact density (solid line) of the normalized in-sample Sharpe ratio ($\hat{\theta}/\theta$) and two different approximated density, the first one assumes N and $T \to \infty$, but $N/T \to \rho$ (dotted line) and the second one assumes N is fixed is fixed but $T \to \infty$ (dashed line). The length of time series is $T = 120$ and plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

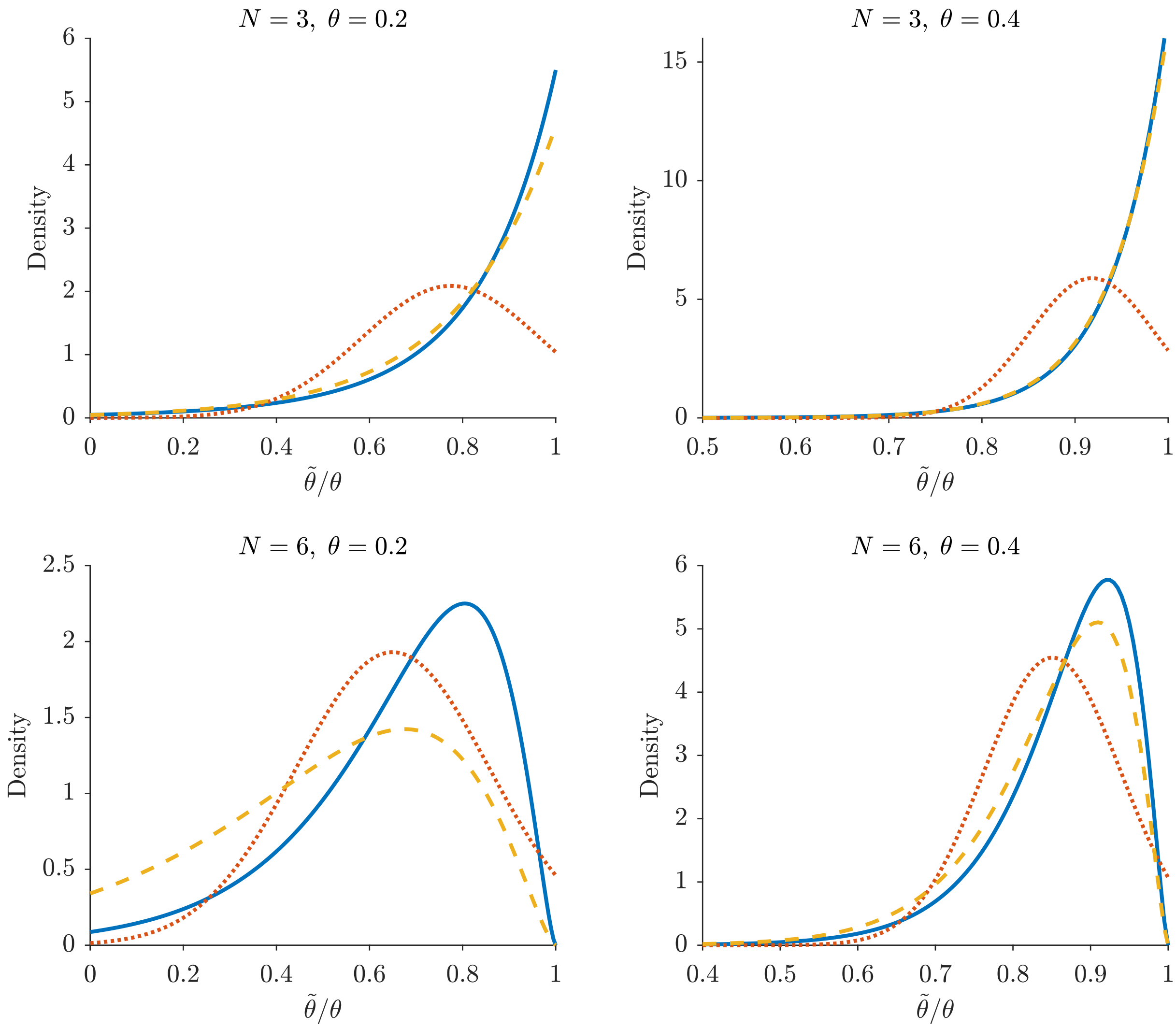

Figure 9: **Exact and Approximate Densities of Out-of-sample Sharpe Ratio of an Asset Pricing Model** ($T = 120$)

The figure plots the exact density (solid line) of the normalized out-of-sample Sharpe ratio ($\tilde{\theta}/\theta$) and two different approximated density, the first one assumes N and $T \to \infty$, but $N/T \to \rho$ (dotted line) and the second one assumes N is fixed is fixed but $T \to \infty$ (dashed line). The length of time series is $T = 120$ and plots for two different values of number of traded factors ($N = 3$ and 6) and population Sharpe ratio ($\theta = 0.2$ and 0.4) are presented.

Table I: **In-sample Sharpe Ratios of the Sample Optimal Portfolio of Various Asset Pricing Models**

This table reports the in-sample Sharpe ratios (IS-SR) of eight asset pricing models for the full sample period (1967–2018) as well as the two subperiods (1967–1992 and 1993–2018). The asset pricing models included are: CAPM, Fama-French 3-factor model, Carhart 4-factor model, Betting-against-beta (BAB) 2-factor model, Fama-French 5-factor model, HXZ's q-factor model, Barillas and Shanken (BS-6) 6-factor model, and HMXZ's q^5 model. The column "Year" presents the year that the model was first published. The p-values from the Gibbons-Ross-Shanken F-test that compares the Sharpe ratio of a given model with that of the CAPM are also reported in the table.

		1967–2018		1967–1992		1993–2018	
	Year	IS-SR	p-value	IS-SR	p-value	IS-SR	p-value
CAPM	1964	0.117		0.089		0.148	
FF-3	1993	0.195	0.001	0.239	0.001	0.175	0.270
Carhart-4	1997	0.288	0.000	0.379	0.000	0.242	0.013
BAB	2014	0.306	0.000	0.369	0.000	0.315	0.000
FF-5	2015	0.322	0.000	0.488	0.000	0.339	0.000
HXZ q	2015	0.416	0.000	0.611	0.000	0.346	0.000
BS-6	2018	0.485	0.000	0.786	0.000	0.373	0.000
HMXZ q^5	2019	0.634	0.000	0.767	0.000	0.562	0.000

Table II: **Mutual Fund Performance Relative to In-sample Sharpe Ratios of Asset Pricing Models**

This table presents mutual fund performance in terms of sample Sharpe ratio over the period of 1993–2018 as well as the two subperiods 1993–2005 and 2006–2018, and compares mutual fund performance with the in-sample Sharpe ratio of various asset pricing models. Sample Sharpe ratios computed based on after-fee returns (AF-SR) and before-fee returns (BF-SR) are both obtained. The upper panel reports the number of funds, the mean, the first quartile, the median, and the third quartile of the cross-sectional sample Sharpe ratios. The bottom panel reports the proportion (in percentage points) of the mutual funds outperforming the asset pricing models.

	1993–2018		1993–2005		2006–2018	
	AF-SR	BF-SR	AF-SR	BF-SR	AF-SR	BF-SR
N	3494		2107		2513	
Mean	0.107	0.126	0.099	0.119	0.121	0.138
Q1	0.072	0.091	0.051	0.069	0.092	0.109
Median	0.113	0.130	0.104	0.124	0.119	0.136
Q3	0.145	0.162	0.151	0.171	0.148	0.164
	AF-%	BF-%	AF-%	BF-%	AF-%	BF-%
CAPM	22.90	34.83	27.91	38.54	22.16	33.98
FF-3	11.91	17.60	0.66	1.00	11.94	16.99
Carhart-4	3.86	4.75	0.24	0.28	11.10	15.64
BAB	0.77	1.40	0.19	0.28	5.41	6.29
FF-5	0.40	0.80	0.14	0.28	0.40	0.80
HXZ q	0.34	0.72	0.09	0.14	3.10	4.14
BS-6	0.17	0.37	0.05	0.09	1.67	2.39
HMXZ q^5	0.03	0.03	0.00	0.05	0.04	0.04

Table III: **Out-of-sample Sharpe Ratios of the Sample Optimal Portfolio of Various Asset Pricing Models**

This table reports the out-of-sample Sharpe ratios (OS-SR) of various asset pricing models for three different out-of-sample periods, 1993–2018, 1993–2005, and 2006–2018. For a given out-of-sample period, we estimate the sample optimal portfolio using monthly data from 1967 to the beginning of the out-of-sample period, and hold the portfolio weights constant throughout the out-of-sample period. Out-of-sample Sharpe ratios are computed using the out-of-sample returns of the sample optimal portfolio. For comparison, the in-sample Sharpe ratios (IS-SR) over the out-of-sample periods are also reported in the table. Gibbons-Ross-Shanken F-test is conducted to compare IS-SR of a given model with that of the CAPM. We also conduct one-sided test based on asymptotic distribution to compare the OS-SR of a given model with that of the CAPM. ***, **, * denote that the Sharpe ratio of the given model is higher than that of the CAPM at 1%, 5%, and 10% significance levels.

	1993–2018		1993–2005		2006–2018	
	IS-SR	OS-SR	IS-SR	OS-SR	IS-SR	OS-SR
CAPM	0.148	0.148	0.145	0.145	0.152	0.152
FF-3	0.175	0.147	0.331***	0.294*	0.180	0.049
Carhart-4	0.242**	0.192	0.403***	0.319*	0.182	0.063
BAB	0.315***	0.221	0.406***	0.266	0.229**	0.221
FF-5	0.339***	0.193	0.425***	0.213	0.329**	0.197
HXZ q	0.346***	0.247	0.474***	0.253	0.259*	0.189
BS-6	0.373***	0.261	0.562***	0.335*	0.286	0.174
HMXZ q^5	0.562***	0.475***	0.731***	0.555***	0.437***	0.419***

Table IV: **Mutual Fund Performance Relative to Out-of-sample Sharpe Ratios of Asset Pricing Models**

This table reports the proportion (in percentage points) of mutual funds with sample Sharpe ratio higher than the out-of-sample Sharpe ratios of various asset pricing models for three out-of-sample periods: 1993–2018, 1993–2005, and 2006–2018. The sample optimal portfolios of various asset pricing models are estimated using monthly data from 1967 to the beginning of the out-of-sample period. Both after-fee (AF) and before-fee (BF) Sharpe ratios are obtained for mutual funds, and they are compared with the out-of-sample Sharpe ratios of the asset pricing models.

	1993–2018		1993–2005		2006–2018	
	AF-%	BF-%	AF-%	BF-%	AF-%	BF-%
CAPM	22.90	34.83	27.91	38.54	22.16	33.98
FF-3	24.16	36.46	1.23	1.90	90.05	93.24
Carhart-4	8.67	11.76	0.81	1.23	86.87	91.09
BAB	5.09	6.78	1.99	3.61	5.97	7.00
FF-5	8.56	11.31	6.69	11.01	8.71	10.94
HXZ q	3.55	4.49	0.85	1.33	9.83	13.09
BS-6	2.52	3.55	0.62	1.00	13.21	19.30
HMXZ q^5	0.06	0.06	0.05	0.09	0.04	0.08

Table V: **Confidence Intervals of Population and Conditional Distribution of Out-of-sample Sharpe Ratios of Various Asset Pricing Models**

Using h months of factor return data ending at 2018/12, we obtain the in-sample Sharpe ratios ($\hat{\theta}$) as well as the 95% confidence interval (θ_L, θ_U) of the population Sharpe ratios for various asset pricing models. With the population Sharpe ratio set to θ_L or θ_U, we further obtain the 10th, 50th, and 90th percentiles of the distribution of the out-of-sample Sharpe ratios based on the conditional distribution, $F(\tilde{\theta}|\hat{\theta})$. Panels A, B, and C present results for $h = 60$, 120, and 240 months, respectively.

A. $h = 60$

				$F^{-1}(\tilde{\theta}\|\hat{\theta})$				$F^{-1}(\tilde{\theta}\|\hat{\theta})$	
Model	$\hat{\theta}$	θ_L	10th-%	50th-%	90th-%	θ_U	10th-%	50th-%	90th-%
CAPM	0.201	0.000				0.454			
FF-3	0.249	0.000	0.000	0.000	0.000	0.448	0.280	0.397	0.441
Carhart-4	0.271	0.000	0.000	0.000	0.000	0.445	0.242	0.367	0.426
BAB	0.555	0.248	0.197	0.240	0.248	0.803	0.746	0.793	0.803
FF-5	0.271	0.000	0.000	0.000	0.000	0.413	0.175	0.308	0.379
HXZ q	0.336	0.000	0.000	0.000	0.000	0.523	0.350	0.457	0.507
BS-6	0.345	0.000	0.000	0.000	0.000	0.482	0.249	0.370	0.440
HMXZ q^5	0.567	0.153	0.036	0.101	0.136	0.762	0.607	0.694	0.741

Table V: **Confidence Intervals of Population and Conditional Distribution of Out-of-sample Sharpe Ratios of Various Asset Pricing Models (Cont'd)**

B. $h = 120$

Model	$\hat{\theta}$	θ_L	$F^{-1}(\tilde{\theta}\|\hat{\theta})$ 10th-%	50th-%	90th-%	θ_U	$F^{-1}(\tilde{\theta}\|\hat{\theta})$ 10th-%	50th-%	90th-%
CAPM	0.275	0.091				0.455			
FF-3	0.326	0.089	0.027	0.070	0.086	0.484	0.417	0.464	0.481
Carhart-4	0.327	0.048	−0.011	0.023	0.042	0.472	0.383	0.438	0.464
BAB	0.554	0.347	0.321	0.342	0.347	0.734	0.706	0.730	0.734
FF-5	0.374	0.095	0.017	0.061	0.084	0.513	0.414	0.470	0.499
HXZ q	0.302	0.000	0.000	0.000	0.000	0.445	0.351	0.409	0.436
BS-6	0.333	0.000	0.000	0.000	0.000	0.454	0.329	0.394	0.432
HMXZ q^5	0.493	0.243	0.169	0.211	0.233	0.642	0.558	0.606	0.630

C. $h = 240$

Model	$\hat{\theta}$	θ_L	$F^{-1}(\tilde{\theta}\|\hat{\theta})$ 10th-%	50th-%	90th-%	θ_U	$F^{-1}(\tilde{\theta}\|\hat{\theta})$ 10th-%	50th-%	90th-%
CAPM	0.103	0.000				0.229			
FF-3	0.141	0.000	0.000	0.000	0.000	0.244	0.174	0.223	0.241
Carhart-4	0.174	0.000	0.000	0.000	0.000	0.271	0.195	0.242	0.263
BAB	0.243	0.100	0.074	0.096	0.100	0.363	0.338	0.359	0.363
FF-5	0.327	0.157	0.106	0.135	0.150	0.431	0.377	0.408	0.424
HXZ q	0.274	0.106	0.059	0.088	0.102	0.381	0.330	0.362	0.376
BS-6	0.295	0.103	0.043	0.074	0.092	0.389	0.319	0.356	0.377
HMXZ q^5	0.469	0.308	0.268	0.291	0.302	0.581	0.538	0.562	0.575

Table VI: **Posterior Predictive Mean and Standard Deviation of Out-of-sample Sharpe Ratios of Various Asset Pricing Models**

Using h months of factor return data ending at 2018/12, we obtain the in-sample Sharpe ratios ($\hat{\theta}$) for various asset pricing models. Under the assumption that the prior distribution of θ is $\theta \sim 0.6 \times \text{Beta}(2, \beta)$ with $\beta = 6$, 4, and 2, we obtain the posterior predictive mean and standard deviation of $\tilde{\theta}$. Panels A, B, and C present results for $h = 60$, 120, and 240 months, respectively.

A. $h = 60$

		$\beta = 6$		$\beta = 4$		$\beta = 2$	
Model	$\hat{\theta}$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\text{Std}[\tilde{\theta}\|\hat{\theta}]$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\text{Std}[\tilde{\theta}\|\hat{\theta}]$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\text{Std}[\tilde{\theta}\|\hat{\theta}]$
CAPM	0.201	0.159	0.076	0.192	0.086	0.241	0.098
FF-3	0.249	0.080	0.084	0.109	0.098	0.155	0.116
Carhart-4	0.271	0.067	0.077	0.093	0.092	0.136	0.111
BAB	0.555	0.253	0.094	0.313	0.096	0.396	0.096
FF-5	0.271	0.053	0.070	0.074	0.083	0.110	0.103
HXZ q	0.336	0.090	0.085	0.125	0.099	0.183	0.118
BS-6	0.345	0.061	0.071	0.087	0.086	0.132	0.107
HMXZ q^5	0.567	0.160	0.099	0.220	0.108	0.308	0.113

Table VI: **Posterior Predictive Mean and Standard Deviation of Out-of-sample Sharpe Ratios of Various Asset Pricing Models (Cont'd)**

B. $h = 120$

		$\beta = 6$		$\beta = 4$		$\beta = 2$	
Model	$\hat{\theta}$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\mathrm{Std}[\tilde{\theta}\|\hat{\theta}]$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\mathrm{Std}[\tilde{\theta}\|\hat{\theta}]$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\mathrm{Std}[\tilde{\theta}\|\hat{\theta}]$
CAPM	0.275	0.209	0.072	0.240	0.077	0.280	0.082
FF-3	0.326	0.182	0.082	0.218	0.086	0.265	0.092
Carhart-4	0.327	0.156	0.083	0.192	0.089	0.238	0.094
BAB	0.554	0.344	0.070	0.389	0.071	0.450	0.070
FF-5	0.374	0.168	0.085	0.208	0.090	0.259	0.095
HXZ q	0.302	0.138	0.082	0.171	0.088	0.214	0.094
BS-6	0.333	0.117	0.080	0.150	0.088	0.194	0.096
HMXZ q^5	0.493	0.254	0.083	0.300	0.084	0.362	0.086

C. $h = 240$

		$\beta = 6$		$\beta = 4$		$\beta = 2$	
Model	$\hat{\theta}$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\mathrm{Std}[\tilde{\theta}\|\hat{\theta}]$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\mathrm{Std}[\tilde{\theta}\|\hat{\theta}]$	$\mathbb{E}[\tilde{\theta}\|\hat{\theta}]$	$\mathrm{Std}[\tilde{\theta}\|\hat{\theta}]$
CAPM	0.103	0.111	0.049	0.122	0.051	0.134	0.054
FF-3	0.141	0.081	0.058	0.093	0.061	0.107	0.064
Carhart-4	0.174	0.089	0.060	0.104	0.063	0.120	0.066
BAB	0.243	0.190	0.061	0.209	0.062	0.230	0.064
FF-5	0.327	0.217	0.065	0.241	0.066	0.268	0.068
HXZ q	0.274	0.183	0.065	0.204	0.066	0.228	0.068
BS-6	0.295	0.170	0.067	0.193	0.069	0.218	0.071
HMXZ q^5	0.469	0.333	0.059	0.363	0.060	0.403	0.063