COMPARING ASSET PRICING MODELS USING QUANTILE REGRESSIONS FOR DISTANCE-BASED METRICS

Ziwen Wang

Submitted in partial fulfillment
of the requirements for the degree of

Master of Science in Management (Finance)

Goodman School of Business, Brock University
St Catharines, Ontario

Abstract

This thesis compares the performance of ten well-known asset-pricing models for cross-sectional returns of various portfolios from January 1967 to December 2016. We rely on the distance-based metrics as the primary performance measure and use quantile regressions to compare models at a wide range of quantiles of the asset return distribution. The model performance is examined from both statistical and economic perspectives. We find that the Fama and French (2018) six-factor model reliably outperforms other competing models in pricing the selected portfolios. In particular, both the momentum factor and the value factor are necessary in asset-pricing models to explain the return variations in different quantiles. We also find that the performance of Barilla and Shanken (2018) six-factor model exhibits strong explanatory power in medium to high quantiles, despite some existing findings that their model performs poorly in OLS regressions. Overall, we show that the distance-based metrics coupled with quantile regressions provide a consistent and robust model-comparison methodology that largely enhances the existing OLS-based statistical measures.

Keywords: Multi-factor Asset Pricing Models, Quantile Regressions, Distance-Based Metrics

JEL classifications: G11, Z23

Table of Content

1. Introduction

As the asset pricing literature continues to develop, numerous multi-factor models are proposed to explain the average returns for a wide range of cross-sectional assets. The Capital Asset Pricing Model (CAPM) introduced by Sharpe (1964) and Lintner (1965) describes the relationship between a single market factor of systematic risks and expected returns. Fama and French (1993) develop a three-factor model (FF3) that includes the market factor, the size factor, and the value factor. Carhart (1997) introduces an additional momentum factor to the FF3, and the four-factor model dubbed Carhart4 has become the benchmark model in empirical asset pricing and performance evaluation during the first decade of the new millennium. In the current decade, Fama and French (2015) augment two additional factors, namely, the profitability factor and the investment factor to the FF3 to introduce their five-factor model (FF5). Fama and French (2018) further add the momentum factor so that their six-factor model (FF6) has now become the state-of-the-art benchmark model. In the meantime, inspired by the neoclassical q-theory, Hou, Xue and Lu (2015) construct the q-factor model (QF) that consists of alternative investment and profitability factors to the Fama-French models; Recently, they propose their new version of the five-factor model (Q5) by taking an expected growth factor into consideration. As for other factor models, Stambaugh and Yuan (SY, 2016) suggest that mispricing factors can be used in explaining the expected returns. Barillas and Shanken (2018) believe that the timely updated value factor can provide additional values to the descriptive power of multi-factor models.

All the multifactor models cited above are designed to explain expected returns of a cross-section of Left-Hand-Side (RHS) assets using a set of Right-Hand-Side (RHS) factors. Given the variety of portfolio-based factors that have been examined by researchers, it is important to understand how to best combine them in a parsimonious asset-pricing model for expected returns, one that excludes redundant factors. Fama and French (2018) caution that the multiple comparisons problem may arise in undisciplined searching for the best combination in a long list of potential factors, and the set of model factors should be limited to ensure the robustness of results. The main objective of this thesis is to compare the relative performance of the above-mentioned asset-pricing models, using quantile regressions based on the distance-based metrics recently developed by Goyal, He and Huh (2019).

In empirical asset pricing, there are standard econometric techniques (e.g., the GRS test) to evaluate the adequacy of a single model. However, a satisfactory statistical methodology to identify the best factor-pricing model(s) among several competing models has drawn research attention only in recent years. Some studies mention that traditional methods may not be suitable in comparing asset-pricing models. For example, the GRS-statistic, generally regarded as the authoritative statistical method, may induce the power problem, i.e., models that produce economically insignificant pricing errors tend to be over-rejected while models that produce large pricing errors with inflated residual covariance matrices tend to be under-rejected (Fama and French, 1993; Harvey, 2017). In addition, the alpha-based statistics is another widely used method that ignores the power problem but causes the extreme-error problem. The undisciplined use of the GRS and alpha-based statistics often leads to contradicting and counter-intuitive model choices.

To address the conflicting inferences caused by the power and extreme-error problems, in this thesis, we use distance-based metrics (Goyal, He and Huh, 2019) as the comparing metrics to measure model performance for a wide range of portfolios that consist of all NYSE, AMEX, and NASDAQ average stock returns from January 1967 to December 2016. The distance-based metrics proposed by Goyal, He and Huh (2019) measure model performance as the Euclidean distance between two distributions using the OLS regressions to estimate the mean and standard error of the mispricing parameter (alpha). However, the OLS approach only estimates model performance at the mean of returns. This is a valid method if alphas follow a normal distribution. Moreover, the assumption of normal distribution is often severely rejected for asset returns and alphas, which are widely known to exhibit skewness and fat tails. How do we compare models under the more general distributional assumption of alphas? Do we still obtain consistent model rankings across a wide range of LHS assets as documented in Goyal, He and Huh (2019) for different quantiles of the alpha distribution? In light of these questions, the motivation of this thesis is to test the distance-based metrics under a more general distributional assumption, checking whether the model ranking results are still consistent across different quantiles of the distribution of alphas. To this end, a more appropriate method is the quantile regression (QR) introduced by Koenker and Bassett (1978). The QR method is more robust to non-normal errors and outliers as it provides a richer characterization of data, allowing us to consider the impact of a covariate on the entire distribution of returns, not merely its conditional mean.

The contributions of this study are threefold. First, we employ the recently developed distance-based metrics (Goyal, He and Huh 2019) as the main performance measures and use quantile regressions instead of OLS regressions. In this regard, this thesis is the first to combine the two empirical methodologies in a systematic way. Specifically, as shown in Goyal, He and Huh (2019), the distance metrics effectively address the power problem and the extreme-error problems in empirical asset pricing. And the quantile-regression approach allows us to not only evaluate model performance by the mean returns, but also consider how the models perform at various quantiles of the entire distribution of asset returns. Second, the combined methodology of distance metrics and quantile regressions generates a comprehensive set of empirical results that largely expand the model-comparison findings in the existing literature. Overall, we find that the Fama and French (2018) six-factor model reliably outperforms other alternative models, including its close competitors such as the Q5 model (Hou, Mo, Xue and Zhang, 2018), the q-factor model (Hou, Xue and Zhang, 2015), and the SY model (Stambaugh and Yuan, 2017). The third contribution is with regard to the performance of some controversial factors in the extant asset-pricing literature. In particular, the FF5 model (Fama and French, 2015) does not include the momentum factor, which is reluctantly added by Fama and French (2018) to their FF6 model. Furthermore, Fama and French (2015) and Hou, Xue and Zhang (2015) find that the value factor is redundant. However, these findings are based on the statistical measures about the average returns. To what extent are the momentum factor and the value factor important if other distributional quantiles are considered by the distance metrics? The combined methodology allows us to answer this question. We find that both the momentum and the value factors (HML) are important in asset-pricing models. Specifically, the momentum factor shows its most descriptive power around the medium quantiles; and the value factor exhibits some significant pricing ability in some high quantiles. These findings are new to the existing asset-pricing literature and carry important implications to the performance evaluation of actively managed portfolios and risk management.

The remainder of the paper is structured as follows: Section 2 conducts the literature review. Section 3 presents the distance-based metrics and quantile regressions. Section 4 describes the data for the factors and test portfolios. Section 5 presents the empirical results of the comparison of asset-pricing models at various percentiles and analyzes them from both statistical and economic perspectives. Section 6 concludes and summarizes the main findings on model comparison.

2. Literature Reivew

2.1. Asset-Pricing Models

The development of asset-pricing models aims to explain cross-sectional expected returns using a small number of factors with high precision. The pioneer work of Sharpe (1964) interprets the relationship between an individual asset expected return and systematic risks and propose the original concept of a capital asset pricing model (CAPM). Lintner (1965) converts this conception into the corresponding formula and refines some important properties of the formula. Later researches realize that CAPM fails to describe the cross-sectional returns while some unmentioned variables in asset-pricing theory are able to provide additional explanatory power for average returns. Indeed, Ball (1978), Banz (1981), Basu (1983) and Lakonishok, Shleifer and Vishny (1994) confirm the existence of a relationship between average stock returns and firm size, book-to-market equity, earnings-to-price ratio, cash flow-to-price ratio and past sales growth. The follow-up research of Fama and French (1993) constructs a three-factor model (FF3) based on CAPM with the addition of the size and value factors as its second and third factors.

Numerous empirical researches assert that FF3 outperforms CAPM and better describes the cross-section of asset returns (Griffin & Lemmon, 2002; Liew & Vassalou, 2000). This makes the FF3 the benchmark model to price the variation in cross-sectional asset returns; many later proposed models tend to expand the FF3 with other additional factors.

Carhart (1997) finds that the common factors in stock returns and persistent differences in mutual fund expenses and transaction costs can explain almost all of the predictability in mutual fund returns, and takes the factors affecting mutual fund into consideration. Hendricks, Patel and Zeckhauser (1993), Goetzmann and Ibbotson (1994), Brown and Goetzmann (1995), and Wermers (1996) find evidence of persistence in mutual fund performance over short-term horizons of one to three years, and they attribute the persistence to the "hot hand" phenomenon or common investment strategies. Following these studies, Carhart (1997) indicates that Jegadeesh and Titman (1993) one-year momentum in stock returns accounts for the “hot hand” effect in mutual fund performance, and thus constructs a four-factor model (CAR4) that includes the FF3 factors and the momentum factor.

With regard to stock returns, Novy-Marx (2013) identifies a proxy for expected profitability that is strongly related to average returns. Aharoni, Grundy, and Zeng (2013) document a weaker but statistically reliable relation between investment and average return. Fama and French (2015) augment the previous FF3 with a new five-factor model (FF5) which further captures profitability and investment patterns in describing average stock returns. Although FF5 does not include any factors capturing momentum or anomalies, the necessity of these factors is mentioned in later studies. Meanwhile, the argument first proposed in Fama and French (2015) that whether the value factor (HML) is a redundant factor raises further concerns in empirical studies.

Hou et al. (2015) show that HML's high average return is fully captured by its exposures to the investment and profitability factors. Inspired by the neoclassical q-theory of investment, they construct their four-factor model (QF), which consists of the market, size, investment and profitability factors. The market factor is the same as the CAPM, while remaining factors are estimated by their methods. The authors believe that these additional factors not only capture the effect of HML, but also outperform FF3 and CAR4 in describing portfolio returns sorted on various variables. In contrast to Fama and French (1993, 2015), Hou et al. (2015) are primarily concerned with explaining the returns associated with anomaly variables not used to construct their factors, and they focus on value-weight portfolios from univariate sorts on each variable. Moreover, motivated by the theoretical model of Cochrane (1991), Hou et al. (2018) propose the Q5 model as their new version of the asset-pricing model. The Q5 model improves the QF model with a new expected growth factor, which adds additional explanatory power in the cross-section. Importantly, the authors claim that Q5 outperforms all versions of the Fama-French models.

Stambaugh and Yuan (2017) argue that given the proliferation of anomalies, an alternative factor model that can accommodate more anomalies is required. Similar to Hou et al. (2015), they construct a four-factor model (SY) with the market factor of CAPM and their construction of the size factor. Unlike FF5 and QF, the SY model does not include any profitability factors; instead, their new model attempts to capture both the value and momentum factors with two mispricing factors.

Fama and French (2018) compare all existing factors in the previous literature and emphasize the importance of using theory to limit the set of competing models. They argue that in the ideal case, the theory provides fully specified models that lead to precise statements about the relation between an asset's measurable characteristics and its expected returns. By ranking the maximum squared Sharpe ratio for various combinations of model factors, they find that the winner is the six-factor model (FF6): the FF5 plus the momentum factor. They show that the FF6 model outperforms other factor combinations; however, they argue that the momentum factor is somewhat suspicious due to the absence of theoretical justification.

Following Fama and French (2018), Barillas and Shanken (2018) examine the best combination of portfolio-based factors in a parsimonious asset-pricing model. The winner in their Bayesian asset pricing test is the six-factor model (BS) consisting of the market, the size and the momentum factors of CAR4 model, plus the investment and the profitability factors of QF model and a newly constructed value factor. The new value factor HMLm introduced in Asness and Frazzini (2013) is based on book-to-market rankings using the most recent monthly stock prices in the denominator. Barillas and Shanken (2018) emphasize that by substituting HMLm to HML, the value and the momentum factors are not redundant in their model.

2.2. Frequentist vs. Bayesian Views

To estimate how competent a set of selected factors explain cross-sectional returns, we compare the intercepts and the error terms from regressions of an asset's excess returns on the factor returns. An asset-pricing model is said to explain expected returns when the intercepts are indistinguishable from zero (Fama & French, 2015). While a flawless prediction on the basis of regression equations seems impractical, the asset-pricing theory suggests that models that generate smaller alphas have smaller pricing errors hence induce better pricing competency. The error terms indicate the precision of the estimated alphas. The smaller the value of the error terms, the less dispersed of the pricing errors in the time-series regressions thus the higher power of the model.

t-statistics for a single asset and the F-statistics for multiple assets (GRS henceforth) in Gibbons, Ross and Shanken (1989) are the most widely used statistical measures in testing asset-pricing models. They test whether the single intercept (t-statistic) or the joint intercepts (GRS) from regressions is (are) equal to zero. However, GRS-statistics tend to over-reject prominent models such as the Fama and French (2015) five-factor model (FF5). The terminology “too much power” describes the case where the GRS test over-rejects models that produce economically insignificant pricing errors. Another type of the "power problem” happens when the GRS-statistic fails to reject models that produce large pricing errors with inflated residual covariance matrices (dubbed "lack of power"). For instance, Fama and French (2012, 2018) report that GRS cannot reject global models in pricing Japanese stock returns. Furthermore, GRS is not suitable for comparing performances of non-nested models (Fama & French, 2018).

The mean absolute alpha (MAE), one of the alpha-based statistics, is another popular method frequently used jointly with the GRS statistic to compare model performance (Fama & French, 2018; Hou et al., 2015, Hou et al., 2018); Stambaugh & Yuan, 2017). It mainly focuses on model mispricing error, suggesting models with the lowest MAE value best describe returns. However, this method underestimates the impact of extreme alpha values on model performance. More importantly, MAE does not consider the power problem as it does not take the residual covariance matrix into account. For the same reason, the power problem also persists in other model performance measures such as the number of significant alpha's t-statistics, the number of GRS rejections (Hou et al, 2015, 2016), the number of alphas and the mean absolute t-statistics (Stambaugh & Yuan, 2017).

The power and extreme error problems of GRS and MAE statistics may induce contradicting model comparison results (Goyal, He and Huh, 2019), making model performance ranking ambiguous and challenging to interpret. Goyal, He and Huh (2019) point out that the root cause of these problems is that the two statistics are based on the Frequentist view of model tests, and they believe that their methods based on Bayesian view is more appropriate in comparing model performance. Their Bayesian view is consistent with the existing model comparison literature. For example, Pastor and Stambaugh (2000) propose utility-based metrics to examine the impact of varying degree of prior beliefs on portfolio choices from a Bayesian perspective. Recently, Barillas and Shanken (2018) use the Bayes factor to compute the posterior model probabilities and then choose the best set of factors.

Goyal, He and Huh (2019) compare the difference between Frequentist and Bayesian views and introduce the distance-based metrics that have intuitive Bayesian interpretations. They argue that the distance-based metrics effectively address the power problems and the extreme-error problems described above as the distance measures treat the size of pricing errors and the size of mispricing uncertainty not as a ratio but as the square root of the sum of pricing errors and standard errors.

2.3. Quantile Regression

All asset-pricing models mentioned in the previous section are based on the assumption that error terms are normally distributed. Hence, these studies use the ordinary least square (OLS) regressions to estimate the alphas and residual covariance metrics, which are then used to construct t-test and GRS test in asset pricing.

However, with the development of the asset pricing literature, normally distributed asset returns are severely rejected. For example, early studies such as Officer (1972) document that the distribution of returns has fat tails as compared to normal distribution. Levhari and Levy (1977) indicate that the stock returns carry fat tails and the beta estimates using monthly data are not the same as the beta estimated using yearly data. Some asset-pricing models' unsustainable pricing ability also implies the asymmetry of stock returns. Horowitz, Loughran and Savin (2000) argue that the results of the size effect are not robust across different sample periods and it disappears since 1982. These studies, among many others such as Chan and Lakonishok (1992), suggest the usage of more robust methods instead of OLS regressions.

The quantile regression method is proposed to be one of the favorable alternatives. Introduced in Koenker and Bassett (1978), quantile regression not only provides a complete coverage for the whole distribution of factor returns, but it also places no limitation on the distribution of asset returns (Mosteller & Tukey, 1977). Quantile regressions have been widely used in financial research. For example, Bassett and Chen (2001) use quantile regressions for portfolio analysis. Barnes and Hughes (2002) test the cross-sectional pricing ability of CAPM using quantile regressions. Ma and Pohlman (2008) analyze a similar relationship for different asset pricing factors. Chiang and Li (2012) employ quantile regressions to examine the risk-return relation by applying high-frequency data from four major stock indexes in the US market and find that the relationship between the mean of the excess returns and expected risk moves from negative to positive as percentile increases. Autchariyapanitkul, Chanaim and Sriboonchitta. (2015) use quantile regressions under asymmetric Laplace distribution to predict stock returns. Chen, So and Chiang (2016) propose a nonlinear threshold quantile GARCH model to estimate the relationship between return and lagged abnormal volume. Yamaka, Autchariyapanitkul, Mennejuk, and Sriboonchitta (2017) introduce the generalized maximum entropy (GME) approach proposed by Golan, Judge, and Miller (1997) to estimate the quantile regression model for capital asset pricing. Sharama, Gupta and Singh (2016) test the roles of size, value and market factors in explaining the returns of 30 Dow Jones Industrial Average Stocks using quantile regressions for the period of global financial crisis starting from January 2005 to December 2008.

Furthermore, the distributional assumption of OLS regressions leads some studies to cast doubt on the success of existing pricing models (Black, 1993; Kothari & Shanken, 1995; Levhari & Levy, 1977; Officer, 1972; Knez & Ready, 1997; Horowitz et al., 2000). Along this line of empirical research, Allen, Singh, and Powell (2011) test the pricing ability of the Fama and French three-factor model using quantile regressions. Their study not only shows that the factor models do not necessarily follow a linear relationship but also shows that the traditional method of OLS is less effective in analyzing the extremes within a distribution, which is often of key interest to investors and risk managers. More recently, Sharama, Gupta and Singh (2016) test the pricing ability of Carhart (1997) four-factor model using quantile regressions. The results of the study reveal that the quantile regression model has superior fitting across all percentile levels to the OLS that fails to fit these four factors across all percentile levels.

3. Methodology

3.1. Distance-based metrics

The distance metrics are derived from the classic research of optimal transportation theory rooted in mathematics and economics, with a wide range of applications in economics and econometrics. The optimal transportation problem was originally constructed in Monge (1781), where the author seeks to estimate the shortest distance or the minimum cost to move the mass of one probability distribution to another one by defining a quadratic Wasserstein distance between two probability distributions. Economically, the distance-based metric is the minimal cost of moving the mass of model-implied distribution to data-based distribution of cross-sectional asset returns. In other words, it is the minimum cost of holding a dogmatic belief in the model from a Bayesian perspective.

Let P_I and P_{II} be Gaussian measures on $\mathbb{R}^n$ with finite second moments such that $P_I \sim N(\alpha_I, V_I)$ and $P_{II} \sim N(\alpha_{II}, V_{II})$, where α_I and α_{II} are two $n \times 1$ vectors of mean, and V_I and V_{II} are two $n \times n$ symmetric, positive-definite covariance matrices. Then, the quadratic Wasserstein distance (WD2) between P_I and P_{II} is given by Equation 1.

$$WD_2 = \sqrt{\|\alpha_{II} - \alpha_I\|^2 + \|V_{II} - V_I\|} \quad (1)$$

$$\|V_{II} - V_I\| = \mathrm{Tr}\left(V_I + V_{II} - 2(V_I^{\frac{1}{2}} V_{II} V_I^{\frac{1}{2}})^{\frac{1}{2}}\right) \quad (2)$$

Where $\|\alpha_{II} - \alpha_I\|$ is the Euclidean 2-norm of the mean difference vector, $\|V_{II} - V_I\|$ is the distance between the two covariance matrices, and $V^{\frac{1}{2}}$ is the square root of the covariance matrix such that $V = V^{\frac{1}{2}} V^{\frac{1}{2}}$.

To use this distance measure in Bayesian setting, the first two moments, (α_I, V_I) of P_I and (α_{II}, V_{II}) of P_{II} are replaced with their model-generated posterior estimates of the alpha and its variance, $(\tilde{\alpha}_I, \tilde{V}_{\alpha_I})$ and $(\tilde{\alpha}_{II}, \tilde{V}_{\alpha_{II}})$, respectively, where I and II represent two distinct distribution specifications about prior mispricing uncertainty (σ_α) of for a given asset-pricing model.

In particular, let prior specification I be set as $\sigma_\alpha = 0$ (complete confidence in the model's pricing ability); under such dogmatic beliefs, there is no mispricing uncertainty and hence the posterior estimate of the alpha shrinks to its theoretical value of zero: i.e., both ($\tilde{\alpha}_I$ and $\tilde{V}_{\alpha_I}$) are zero. On the other hand, let prior specification II be set as $\sigma_\alpha = \infty$ (complete skepticism about the model's pricing ability), in which case the posterior estimates $(\tilde{\alpha}_{II}, \tilde{V}_{\alpha_{II}})$ shrink to their sample estimates based entirely on the sample of data. Given such prior specifications, the quadratic distance metric reduces to $WD_2 = \sqrt{\|\tilde{\alpha}_{II}\|^2 + Tr(\tilde{V}_{\alpha_{II}})}$, which also defined as the total distance (TD).

Given the non-informativeness in prior specification II, the posterior estimates $\tilde{\alpha}_{II}$ and $\tilde{V}_{\alpha_{II}}$ are identical to the maximum-likelihood estimates of the alpha, $\hat{\alpha}$, and its covariance matrix, $\hat{V}_\alpha$, respectively. Hence, the previous equation has its frequentist-equivalent form as

$$TD = \sqrt{\|\hat{\alpha}\|^2 + Tr(\hat{V}_\alpha)} \quad (3)$$

Where $\|\hat{\alpha}\|^2 = \sum_{i=1}^{n} \tilde{\alpha}_i^2$ is the sum of squared alphas of the LHS returns in asset-pricing test, and $Tr\left(\hat{V}_\alpha\right) = \sum_{i=1}^{n} \tilde{\sigma}_{\alpha_i}^{\ 2}$, where $\tilde{\sigma}_{\alpha i} = \hat{V}_\alpha^{\frac{1}{2}}(i,i)$ is the posterior estimate of the standard error of the alpha for asset i.

The total distance (TD) is measured as the shortest distance between the theoretical and reality model results, representing the divergence of absolute confidence and skepticism of a certain model. Similar to the GRS statistics, TD estimates model overall performance in a single measure, yet contrast to its counterpart, distance-based metrics consider neither significant alpha dispersion nor mispricing uncertainty good news.

We also include the average distance metric (AD) to compare model performance on average when the numbers of test assets are different:

$$AD = \sqrt{MSE(\tilde{\alpha}) + MSE(\tilde{\sigma}_\alpha)} = \sqrt{\sum_{i=1}^{n}(\tilde{\alpha}_i^2 + \tilde{\sigma}_{\alpha_i}^{\ 2})/n} \quad (4)$$

Where $MSE(\tilde{\alpha}) = \sum_{i=1}^{n} \tilde{\alpha}_i^2/n$ and $MSE(\tilde{\sigma}_\alpha) = \tilde{\sigma}_{\alpha_i}^{\ 2}/n$ are the mean squared errors of the pricing errors and their standard errors, respectively.

Both AD and mean absolute alpha (MAE) measure average model performance. While the latter disregards model power, hence equally weighs different alpha magnitudes, the former penalizes extreme pricing errors severely and favors models with insignificant alphas and higher power with low alpha dispersion.

3.2. Quantile Regression

Koenker and Bassett (1978) extend the conventional conditional mean least squares estimation to a range of models for various conditional quantile functions and derive the quantile regression that disregards OLS paramedic distributional assumption of the error terms. Quantile regression estimates the function of conditional median using the median estimation that minimizes the symmetrically weighted sum of absolute errors, in contrast to other conditional quantile functions where the weights are allocated accordingly to focused quantiles rather than 0.5, making it extremely potent when dealing with outliers. Overall, the quantile regression approaches excel in monitoring models for conditional functions of median as well as every other quantile. Such a mechanism allows quantile regression techniques to estimate the entire selection of conditional quantile functions, hence more comprehensively deriving statistical analyses of the inherent random relationships across distinct allocated weights or quantiles. Common quantile regression models minimize the weighted sum of absolute deviations, as expressed in the following Equation.

$$\min_{\beta \epsilon R_p} \sum_{i=1}^{n} \rho_\tau (\gamma_i - \xi(x_i, \beta)) \tag{5}$$

Where $\rho_\tau = \mu(\tau - I(\mu) < 0)$ is the check function as defined in Koenker and Bassett (1978).

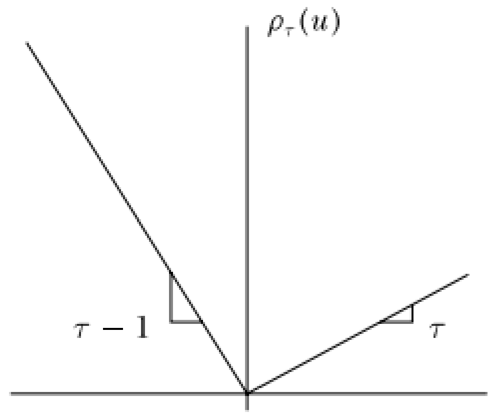

Figure 1: Check function

We set τ to $\frac{1}{2}$ for median regression, or any other percentiles for their respective weighted asymmetric regression.

There are two major differences between QR and OLS. First, quantile regression focuses on different percentiles, whereas OLS regression only focuses on the mean estimation. Second, quantile regression seeks to minimize absolute deviations from the weighted values, whereas OLS regression aims to minimize squared deviations from the mean.

4. Data

Our empirical tests compare the performance of ten asset-pricing models in describing the distribution of returns on a wide range of sorted portfolios. The sample period of our data covers 600-month from January 1967 to December 2016. The data are provided by Kenneth French's data library1 (for all test portfolios and most factor returns), Lu Zhang (for QF and Q5 factors) and AQR Capital Management (for HMLm).

[1] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

4.1. LHS portfolios

The left-hand-side (LHS) portfolios consist of four sets of bivariate-sorted portfolios, three sets of triple-sorted portfolios and four sets of univariate-sorted portfolios.

A bivariate- or triple-sorted portfolio is based on a combination of two or three sort factors from a selection of the five factors: Size (market capitalization – price times shares outstanding), B/M (book-to-market ratio), OP (profitability or performance ratio from operations), INV (Annual asset growth rate), MOM (return spread of momentum).

Consistent with the sample portfolios in Fama and French (2015, 2016b, 2018), the four sets of bivariate-sorted portfolios are from five 5*5 quantile sorts on Size and, independently, on B/M, OP or INV. For example, 5*5 Size-B/M portfolios include average monthly U.S one-month T-bill excess returns for 25 value-weight portfolios from independent sorts of stocks into five Size groups and five B/M groups.

The three sets of triple-sorted (2*4*4) portfolios are 32 Size-B/M-OP, 32 Size-B/M-Inv, and 32 Size-OP-Inv portfolios, respectively. They are formed by two Size groups, big and small, defined as the top half and bottom half of the market capitalization, and four quartiles groups for each of the other two sort variables. The breakpoints of all sort variables are based on NYSE stocks, while the portfolio sample also includes AMEX and NASDAQ stocks. Each set of LHS assets is individually and jointly examined over a cross-section of 196 (25*4+32*3) pooled portfolios.

In addition, following Fama and French (2016b) and Hou et al. (2015), we also examine 15 sets of univariate-sorted decile portfolios that cover numerous anomalies, most of which are not targeted by the respective factor model. The portfolios are also categorized into four groups: 1) the FF factors related, which contains 40 decile portfolios sorted on market capitalization (Size), book-to-market ratio (B/M), operating profitability (OP), and investment ratio (INV); 2) the valuation related, which contains 30 decile portfolios formed on earnings-to-price (E/P), cash flow-to-price (CF/P), and the dividend yield (D/P); 3) the past return related, which contains 30 decile portfolios formed on momentum (MOM), short-term reversal (STR) and long-term reversal (LTR); and 4) other anomalies related, which contains 50 decile portfolios formed on accruals (AC), net share issues (NI), market beta (Beta), return variance (VAR) and groups individually as well as jointly as another broad cross-section of 150 (40+30*2+50) to augmented portfolios. The two broad cross-sections of 196 and 150 pooled assets are used to cross-verify model ranking consistency generated by different performance metrics for the asset-pricing tests. The data on all LHS portfolios and definitions of the sorting variables are obtained from Kenneth French's website.

4.2. RHS models

[Please insert Table 1 about here]

We limit the comparative analyses to the ten asset-pricing models shown in Table 1, with two groups of nested models and two non-nested models.

The first group of nested models is FF6 versus FF6-hml, FF5, CAR4, FF3, and CAPM, including six factors at maximum. MKT (market factor) is the return on the value-weight market portfolio minus the risk free return; SMB (size factor) is the return on a diversified portfolio of small stocks minus the return on a diversified portfolio of big stocks; HML (value factor) is the difference between the returns on diversified portfolios of high and low B/M stocks; RMW (profitability factor) is the difference between the returns on diversified portfolios of stocks with robust and weak profitability; CMA (investment factor) is the difference between the returns on diversified portfolios of the stocks of low and high investment firms, which we call conservative and aggressive; and MOM (momentum factor) is the difference between the returns on diversified portfolios of up and down stocks.

The capital asset pricing model (CAPM) has MKT as the only explanatory variable. FF3 adds SMB and HML based on the CAPM. On top of FF3, the CAR4 model adds MOM, and FF5 extends RMW and CMA separately. The FF6 model has the maximum number of factors among this group, with the MOM added to the FF5. FF6-hml is a new five-factor model that excludes HML from the FF6 model.

The second group of nested models is Q5 versus QF, both of which are proposed by Hou et al. (2015, 2018). The QF model consists of MKT as the market factor, ME as the size factor, IA as the investment factor and ROE as the profitability factor. After that, EG as the expected growth factor added to the QF model to form a new five-factor model (Q5). The computation of MKT is the same as in the group of FF models, while other factors replaced by the authors' own construction.

In addition to comparing the factors in both FF models and q-factor models, Barillas and Shanken (2018) propose their version model by combining MKT of the CAPM model, MOM of the CAR4 model, IA and ROE of the QF model, SMBsy replacing SMB in FF model and HMLm as the monthly updated factor motivated by Asness and Frazzini (2013). As the new value factor, HMLm is based on book-to-market rankings that use the most recent monthly stock price in the denominator. This is in contrast to Fama and French (1993), who use annually updated logged prices in constructing HML.

Furthermore, Stambaugh and Yuan (2017) combine the two mispricing factors (MGMT and PERF) with market and size factors to produce another non-nested four-factor model (SY). They sort and rank each stock of NYSE, AMEX, and NASDAQ with respect to the available anomaly measures within each of the two clusters. Thus, each month a stock has two composite mispricing measures, P1 and P2. The computation of MGMT is the average returns on the two underpriced stocks (sorted by P1) minus the average returns on the two overpriced stocks (P1), and the computation of PERF is the same as MGMT but with stocks sorted by the P2 measurement.

[Please insert Table 2 about here]

Table 2 shows summary statistics for the total 14 factors mentioned above. Panel A of Table 2 describes the most common features in the distribution across the model factors. The average benchmark market excess return (MKT) factor is slightly above 0.5% with a small standard deviation of 4.53%. Among all factors, EG has the highest return spread at 0.82% and the lowest standard deviation of 1.88%, while SMB and RMW have lowest returns at 0.25% and 0.27% respectively. None of the remaining factors' standard deviation surpasses that of market excess returns. About the quantiles, MOM factor registers a minimum return of -34.39%, making it the only factor that has a lower minimum return than the -23.24% of market excess returns. While CMA has the highest minimum return of -6.88%, it also records the lowest maximum return at 9.58%. The three factors of SMB, MOM, and RERF have the highest maximum returns of over 18% among the factors. Regarding the distributional center tendency and skewness, market excess return, MOM and ROE are the only factors that have higher medians than mean returns while at the same time reporting a skewness of less than -0.5, indicating that they have more extreme returns on the left tail and are left-skewed. Meanwhile, ME and HMLm factors are significantly right-skewed with skewness scores of 0.61 and 0.88, and the remaining 9 factors are fairly symmetrical. Five factors, i.e., RMW, MOM, ME, ROE and HMLm, have kurtoses significantly greater than +3, with the first two exceeding +10, suggesting that their return distributions have less distinct central peaks and rather thick tails. SMB and RERF have kurtoses close to +3, indicating that their return distributions resemble a normal distribution, while other 7 low kurtoses imply sharper and higher central peaks.

Panel B of Table 2 reports the correlations between each two pair of factors. The market factor shows the highest positive correlation of 0.28 with SMB and the most negative of -0.54 with MGMT. The four pairs/groups of SMB-ME-SMBsy (size factor), HML-HMLm (value factor), RMW-ROE (profitability factor) and CMA-IA (investment factor) all record high positive correlation coefficients, with 0.97 as the highest. Besides, value factors and investment factors have relatively high correlation coefficients: 0.70 of HML-CMA; 0.67 of HML-IA; 0.52 of HMLm-CMA and 0.5 of HMLm-IA. Some factors within the same model have relatively high correlation as well. The pair of EG-ROE in the Q5 model has a positive correlation coefficient of 0.52. In the BS model, the pair of MOM-ROE shows a positive correlation (0.5), while the pair of MOM-HMLm shows a negative correlation (-0.65). The mispricing factor MGMT in the SY model shows a negative correlation to the market factor (MKT). Among all tested factors not in SY model, other three kinds of factors demonstrate a relatively high correlation to MGMT as well: value factors (HML-MGMT with 0.7 and HMLm-MGMT with 0.49), investment factors (CMA-MGMT with 0.79 and IA-MGMT with 0.78) and the expected growth factor (EG-MGMT with 0.54). Another mispricing factor of the SY model, RERF, shows a relatively high correlation with the other three factors, which are momentum factor (MOM-RERF with

0.72), profitability factor (ROW-RERF with 0.64) and timely updated value factor (HMLm-RERF with -0.63) respectively.

For each LHS portfolio, we explain excess returns (R_i-R_f) of each model by running time-series regressions on the corresponding sets of factors of the ten models, with the quantile points increasing by 5% at each distributional level:

$$R_{it} - R_{ft} = \alpha_i + \beta_i F_t + \varepsilon_{it} \tag{6}$$

Using the results from the time-series quantile regressions, we compute metrics for assessing the performance of asset-pricing models, such as TD, AD, RMSE($\tilde{\alpha}$), RMSE($\tilde{\sigma}_\alpha$). We also use the GRS-statistic and the MAE-statistic as the basis for comparison.

5. Empirical Results

In this section, we present the performance comparison of asset-pricing models. First, we analyze the results generated by three performance metrics. The distance-based metrics are compared to the GRS-statistics and MAE-statistics in measuring model performance. In doing so, we examine whether the distance-based metrics effectively address the power and extreme-error problems. Afterwards, we rank the model pricing ability from both the quantile-based statistical perspective and the economic significance perspective.

5.1. Performance Metrics

A large number of test-portfolios can eliminate potential biases possibly induced by appointed sorting variables (Goyal, He and Huh, 2019), and may aggravate the consequence caused by power or extreme-error problems (Harvey & Liu, 2017). We first conduct asset-pricing tests using two broad cross-sections of 196 and 150 pooled portfolios.

According to Goyal, He and Huh (2019), 196 pooled portfolios are composed of four sets of double-sorted (5*5) portfolios (25 Size-B/M portfolios, 25 Size-OP portfolios, 25 Size-Inv portfolios and 25 Size-Mom portfolios respectively) and three set of triple-sorted (2*4*4) portfolios (32 Size-B/M-Inv portfolios, 32 Size-B/M-OP portfolios and 32 Size-OP-Inv portfolios respectively). Following Fama and French (2016b) and Hou et al. (2015), we form 150 pooled portfolios by combining four sets of univariate-sorted decile portfolios (40 FF-related portfolios, 30 valuation-related portfolios, 30 return-related portfolios, and 50 anomalies-related portfolios).

We report the performance results of the ten models for five quantile estimates (0.1, 0.2, 0.5, 0.8 and 0.9) in Tables 3 and 4, where distance-based metrics and related components are contained in the second through sixth columns, together with the commonly used metrics (GRS and MAE) in the last two columns.

As explained before, TD and AD are, respectively, the total and average cost of holding dogmatic beliefs in an asset-pricing model. $RMSE(\tilde{\alpha})$, and $RMSE(\tilde{\sigma}_{\alpha})$ are the two components of AD: the square root of the mean squared posterior estimates of pricing errors $(\tilde{\alpha})$, and their standard errors $(\tilde{\sigma}_{\alpha})$ generated by the data-based model. $A|\tilde{\sigma}_{\alpha}^2|/$ $\mathrm{A}|\tilde{\alpha}^2|$ measures the contribution of mispricing uncertainty to the cost relative to that of pricing errors (alphas). This ratio is identical to $As^2(a_i)/Aa_i^2$ proposed by Fama and French (2016b). However, we use it differently from them. That is, the ratio is not used to rank the models but to compare the explanatory power of the models: the higher the ratio, the more imprecisely a model estimates alphas. The last two columns show the GRS-statistic and the mean absolute alpha (MAE) respectively.

[Please insert Table 3 about here]

As for the model ranking in both tables, all performance metrics generate diverse results across different quantiles. Table 3 shows that the distance-based metric identifies FF6 as the top model for all quantiles, with the AD value being the lowest at the median, which means that the minimal cost of moving the mass of FF6 model-implied distribution to its data-based distribution is 0.153% per month. The CAR4 model is ranked second at lower quantiles, replaced by FF6-hml at the median, while the BS model relatively outperforms FF6-hml at higher quantiles.

The conventional MAE statistic generates rather similar results to the distance-based metrics with some slight differences. MAE ranks FF5 higher than CAR4 at the 10th percentile and considers the performance of the BS model better than FF6 at the 80th percentile. SY is ranked third by MAE, while AD favors the Q-Factor model. On one hand, AD and MAE share nearly identical ranking in the best model because FF6 always produces the minimum level of pricing and sampling errors across all quantiles. On the other hand, the discrepancy is mainly attributed to the extreme-error problem: FF5 generates more extreme alphas than CAR4 at the 10th percentile (2.177% of $RMSE(\tilde{\alpha})$ for FF5 vs. 2.162% of $RMSE(\tilde{\alpha})$ for CAR4); BS generates more extreme alphas than FF6 at the 80th percentile (1.302% of $RMSE(\tilde{\alpha})$ for BS vs. 1.297% of $RMSE(\tilde{\alpha})$, for FF6) and SY generates more extreme alphas than QF at the median (0.155% of $RMSE(\tilde{\alpha})$ for SY vs. 0.148% of $RMSE(\tilde{\alpha})$ for QF).

GRS statistic, however, produces radically different ranking results compared to the other two measures. Besides its result at the median, the Q5 model always produces the smallest GRS values, followed by SY and QF models. This result implies strong evidence that GRS fails to consider the estimation precision: these three models generate the largest three values of $RMSE(\tilde{\alpha})$ and $RMSE(\tilde{\sigma}_{\alpha})$ among all. The highest Ao/Aa value produced by Q5 model under these quantiles (0.008 at the 10th percentile, 0.010 at the 20th percentile, 0.009 at the 80th percentile and 0.007 at the 90th percentile) show that the sampling error (mispricing uncertainty) contributes more to the cost (TD or AD) than the pricing error, confirming the lack of power problem of the GRS statistic from the Bayesian view. At the median, GRS ranks SY at the first place (2.725) and FF6 at the second (2.791). For the same reason, the SY model underperforms compared to the FF6 model with no comparative advantages of either sampling or pricing errors yet creates a smaller GRS value.

[Please insert Table 4 about here]

In contrast to the dominant performance of FF6 model in pricing 196 portfolios, Table 4 reports various ranking orders at different quantiles in pricing 150 portfolios. The AD value of FF6 is still the smallest at lower quantiles, whereas FF6-hml is ranked the top model at the median with the lowest AD value (i.e., 0.153% is the mimimum cost of holding a dogmatic belief in FF6-hml model from a Bayesian perspective), and BS model outperforms the rest at higher quantiles. Again, the ranking results produced by AD and MAE are similar to their selections of the best model are consistent, while the difference of the second and the third is still ambiguous. MAE statistic considers CAR4 better than FF6-hml at the 20th percentile, while the ranking of AD contradicts. The fact that CAR4 model produces more extreme pricing error (1.265%) than FF6-hml model (1.259%) strengthens the testimony of the extreme-error problem. The similar situation occurs with SY and FF6 at the median, yet it is due to another reason. Indeed, the values of $RMSE(\tilde{\alpha})$ indicate that SY model produces less extreme error than FF6 model, the values of

$RMSE(\tilde{\sigma}_\alpha)$ on the contrary, points out that the SY model has more mispricing uncertainty than the FF6 model. MAE statistic does not consider sampling errors, so it ranks the SY model higher than the FF6 model.

Similarly, according to the GRS statistic ranking with 196 portfolios, SY and Q5 models outperform others across most quantiles. The SY model is ranked the best model among four out of five quantiles in the table, and the second best at the remaining quantiles. Nonetheless, both $RMSE(\tilde{\alpha})$ and $RMSE(\tilde{\sigma}_\alpha)$ for SY model are relatively larger than the best model ranked by the distance-based metrics, implying the superior competence of distance-based metric over GRS when considering the impacts of both pricing and sampling errors.

To conclude, the FF6 model incurs the minimum cost of holding a dogmatic belief in the model; BS and FF6-hml have similar performance across all quantiles, showing increasingly competitive advantages at higher quantiles compared to FF6. Although the ranking of the GRS statistic is largely different from the other two methods, there is no convictive evidence that any models it favors can produce significantly small sampling or mispricing errors at any quantiles. Overall, in evaluating the performance of models, the distance-based metric is a more comprehensive method that can simultaneously address the power and extreme-error problems.

5.2. Mean Test

Next, we integrate the results generated by the distance-based metric method across all quantiles, comparing the model overall performance and analyze the difference of significance level among models.

We compare 19 AD values generated from the 5th to 95th percentiles for each model in pricing two large pooled portfolios and each small set of portfolios separately by using the paired t-test.

All tested values follow a normal distribution (meet the condition of t-test): i.e. w-value of FF5 model with Size-B/M portfolios by using the Shapiro-Wilk normality test is 0.92099 and the p-value is 0.1181. We report the testing results of two large pooled portfolios in Table 5.

[Please insert Table 5 about here]

As per the 196 portfolios in Panel A of Table 5, the paired t-value of the first group (CAPM vs FF3) is 5.634, with the p-value less than 0.001 and the mean of the difference of 0.628. This result indicates that the FF3 model significant outperforms CAPM on the portfolios, and the average cost of moving the mass of the model-implied distribution to the data-based distribution for CAPM is 0.628%, greater than for FF3 model on the monthly average. As the model with lower means outperforms the other, the top three models ranked by paired t-test are FF6, BS, and FF5 model, respectively. None of the p-values between the FF6 model and other models is greater than 0.01, suggesting that the difference between the FF6 model and others is significant. Unlike FF6, the results of paired t-tests indicate that three models have insignificant difference from the BS model, which are FF5 (p-value = 0.615), FF6-hml (p-value = 0.184) and CAR4 (p-value = 0.090) respectively. Moreover, a total ofnine groups of the pairwise relationship among models (p-value >0.05) fail to reject the null hypothesis of no significant difference, implying that we cannot differentiate these models solely by the smaller AD values.

Mean-test results from Panel B of Table 5 yield even greater p-values of two pair models of over 90% (0.994 for FF5 and CAR4, 0.933 for BS and FF6-hml). Furthermore, FF6-hml model at the third place in the model ranking by mean of difference replaces the FF5 model. Since both the significant differences between models and the ranking order are inconsistent in pricing two large pooled portfolios, we cannot draw a consistent conclusion based on these large-portfolio results. For this reason, we conduct the paired t-test for each small set of portfolios separately to check whether there exist more consistent results of individual portfolios.

[Please insert Table 6 about here]

Table 6 lists the ranking orders of the t-test for each portfolio. It is unsurprising that the FF6 model is ranked at the first place for most portfolios. Noticeably, consistent with OLS regression in Goyal et al. (2019), both FF5 and FF3 models fail to price the momentum-related portfolios: they are only ranked at the 8th and 9th places respectively in double-sorts on Size-MOM portfolios and in univariate-sorted on past return related portfolios. In contrast, all models considering the momentum factor (FF6, FF6-hml, BS, and CAR4) generate similar and relatively low AD values, consistent with the results in Barillas and Shanken (2018). The outperformance of the SY model in pricing the portfolios sorted on other anomaly variables is consistent with the original literature in Stambaugh and Yuan (2017). This may explain why the SY model is only ranked at the 10th place in 196 portfolios, but the 6th place in 150 portfolios.

To identify the overall ranking of model performance, we assign 10 points for the best model in each group, 9 points for the second, and likewise until the last model (by 1-point step). The points for each model under all portfolios are totaled (in parentheses) and ranked in the following order: FF6(105), FF5(90), BS(79), FF6-hml(75), CAR4(64), Q5(52), FF3(50), SY(44), QF(35), CAPM(11).

FF6 model is the best model with 105 points in total (6 first places and 5 second places: 6*10+5*9 = 105) and CAPM is the worst with 11 points in total (11 last places: 11*1 = 11).

[Please insert Table 7 about here]

While the score difference between FF5 and FF6 remains the same as FF5 and FF6-hml (15 points), we find that their difference in significance levels varies based on further comparison of the p-values. Tables 7 reports different levels of statistical significance between models. Panel A of Table 7 presents the numbers of portfolios where the p-value of tested models is greater than 0.05, indicating statistically insignificant differences. Panel B of Table 7 presents the numbers of portfolios whose the p-values for the tested models are less than or equal 0.05, but greater than 0.01, which indicates the difference is moderately insignificant in the statistical sense. We consider all situations where p-value of less than 0.01 may suggest either the difference between the models is significant or lack of evidence that the difference is insignificant. Noticeably, nine out of eleven p-values between FF5 and FF6 are greater than 0.05, indicating little evidence to support any distinct difference between the two. On the contrary, ten out of eleven p-values are less than 0.01 in the comparison between FF5 and FF6-hml, evidence that the two models are significantly different statistically.

The results of Panel A suggest compelling insignificant differences between the models in the combination of FF5-FF6, FF3-CAR4 and CAR4-BS, with 9, 8 and 7 p-values over 0.05 respectively, while both FF6-hml-BS and Q5-BS register 5 p-values over 0.05.

In Panel B, the two pairs of FF6-hml-BS and Q5-BS show moderately significant differences with 4 and 3 p-values over 0.01. Combining the two tables, the two pairs of FF6-hml-BS and Q5-BS aggregate for 9 and 8 p-values greater than 0.01 correspondingly.

Consistent with the results generated from three performance metrics, BS model outperforms FF5 under the mean-test. This conclusion conflicts with our previous results under mean-test, as we find that the FF5 model has better performance on average (90 points for FF5 but 79 points for BS). This situation is not caused by the larger number of test assets that can eliminate potential biases. First, although both ranking results of pooled portfolios suggest that BS outperforms FF5 and FF6-hml, the BS model shares significant similarities with the other two. Second, nine out of eleven times the mean of FF5 model is less than that of BS model, whereas the opposite only occurs twice. This means that most of the time FF5 performs better. Finally and importantly, we pinpoint the causes of the conflict to be precisely the two opposite circumstances (mentioned in the analysis of Table 7): the 25 double-sorted Size-MOM portfolios in 196-pooled portfolios and the univariate-sorted 30 past return portfolios in 150-pooled portfolios. With other portfolios, the mean difference between FF5 and BS is not greater than 0.1. However, the mean difference of the two models for Size-MOM portfolios is 0.411, considerably greater than others and dominates the results of 196-pooled portfolios. Similarly, the mean difference of the two models for 30 past return portfolios is 0.201, dominating the results of 150-pooled portfolios.

5.3. Percentage Difference

We now proceed to discuss the question of whether Q-factor models outperform Fama-French models.

From the previous ranking order and t-test results, we conclude that the FF6 model outperforms Q-factor models (both QF and Q5) in pricing all tested portfolios and the difference is statistically significant at below 1%. Since the current literature argues that economic significance has the same importance as the statistical significance in evaluating factors, we therefore compare the performance of Q-factor models with Fama-French models from economic significance.

It can be seen that AD's lowest value is at the median and its highest is at extreme percentiles (5th percentile and 95th percentile) where the most outliers reside. Therefore, if we only judge the economic significance by looking at the difference between AD values under each quantile, the economic significance may only be visible adjacent to the extreme quantiles. To eliminate the impact of AD values generated from different quantiles, we divide the difference of AD values between two models by the smaller AD to observe the percentage difference.

Tables 8 and 9 show the percentage differences of the distances between Q-factor and two Fama-French models FF5 and FF6 in pricing differently sorted portfolios. Panel A consists of four sets of double-sorted portfolios, three sets of triple sorted portfolios and a combination of 196 pooled portfolios. Panel B consists of four sets of univariate-sorted portfolios and the combined 150 pooled portfolios.

QF & FF5

[Please insert Table 8 about here]

As for 196 portfolios in Table 8, the AD value of QF is 10.6% higher than FF5 at the 5th percentile. In order to distinguish the percentage differences, we say that the two models show a significant economic difference when the absolute value of the percentage difference between them is greater than 10%. The 10% threshold value is chosen intuitively based on the distribution of percentage differences on all tested percentiles. Thus, the percentage differences greater than 0.10 suggest that the latter (FF5 in Table 8) significantly outperforms the former (QF in Table 8) economically and are marked in pale pink, and those lower than -10% suggest the other way around and are marked in pale yellow, respectively.

Consistent with what we observe above, the FF5 model underperforms the QF model in pricing double-sorted Size-MOM portfolios and univariate-sorted past return related portfolios due to its omission of prior-return related factors such as momentum. Nonetheless, FF5 exhibits its superiority and significantly outperforms QF in most portfolios at higher quantiles above the median. This means that the FF5 model is better at pricing the portfolios with higher returns. Additionally, the performance of the FF5 model in the portfolios sorted on B/M or OP significantly surpasses the QF model at low extreme quantiles. This means that the FF5 model can be more useful than the QF model in risk management since risk managers mostly pay attention to the extreme lower quantile cases. Conversely, the QF model demonstrates its obvious advantage on the median in the portfolios sorted on momentum factors, with the highest percentage distance difference being 83.4%. Moreover, the QF model outperforms the FF5 model in two large-pooled portfolios with significant economic differences. In particular, in the 150 large-pooled portfolios, the differences between the 45th percentile and the 70th percentile exemplify the superior pricing competency of the QF model.

QF & FF6

[Please insert Table 9 about here]

Compared to the QF and FF5 models, the differences between QF and FF6 models suggest the FF6 model's superiority and the importance of the momentum factor. First, when compared to the QF model, the performance of the FF6 model shows significant economic differences in momentum-sorted portfolios across all quantiles. In Table 8, the largest advantage of QF versus FF5 is 83.4% at the median, while in Table 9, the AD value of the FF6 model is 31.7% lower than the QF model at the same quantile. Second, judged by relative performance, FF6 outperforms QF in nearly all portfolios across all the quantiles. This dominance suggests FF6's superior pricing competency to QF when analyzed from the economic perspective.

Though QF considerably outperforms FF5 in momentum-sorted portfolios and shows similar competitiveness with Fama-French models in anomalies-related portfolios, it underperforms compared to the FF6 model at most quantiles. Recently, Hou et al. (2018) propose that the QF model overlooks the dimension of the expected return and they expand the QF model with the expected growth factor to form the Q5 model. The authors believe that the Q5 model shows stronger explanatory power in the cross section and outperforms the FF6 model. We therefore apply a similar approach to compare Q5 and the Fama-French models and present the results in Table 10 (for FF5) and Table 11 (for FF6).

Q5 & FF5

[Please insert Table 10 about here]

The differences between QF and FF5 models and between Q5 and FF5 models are quite similar. FF5 continues to outperform Q5 at percentiles above the median and low percentiles for certain divisions of portfolios, and Q5 still shows significant economic differences compared to FF5 in pricing two momentum-sorted portfolios. Remarkably, the addition of the expected growth factor increases the Q5's advantages. Indeed, the number of significant differences that occur between Q5 and FF5 at low extreme quantiles decreases, particularly in the OP-sorted portfolios. Moreover, the predominance of FF5 at higher quantiles is weakened (though still significant), and the significant advantages of Q5 in two momentum-sorted portfolios are further magnified. Surprisingly, compared to QF and FF5, the percentage difference between Q5 and FF5 increases at and near the median in pricing Size-B/M, Size-MOM, Size-OP, Size-B/M-OP, and FF-related portfolios. This indicates Q5's underperformance compared to QF in measuring the middle-level returns of these portfolios. Thus, this finding conflicts with Hou et al. (2018) who show that Q5 is significantly better than QF when using OLS regressions. One possible reason is that they ignore the possibility that the distribution of the value-weighted monthly returns of these portfolios is asymmetric, and the outliers that greatly influence OLS regression results could be another reason leading to the evalution bias.

Q5 & FF6

[Please insert Table 11 about here]

Like in Table 11, the addition of the expected growth factor improves the pricing ability of the QF model at lower quantiles, narrowing the performance gap relative to the FF6 model. Hou et al. (2018) indicate that R&D expenses depress current earnings, but induce future growth, implying that firms with lower returns currently have a closer relationship with the expected growth factor. This is consistent with the insignificant difference between Q5 and FF6 in investment-sorted portfolios. Firms with above-average returns, however, are less affected by the expected growth since they tend to be more stable and mature. This explains why the AD difference between Q5 and FF6 increases at the quantiles at and above the median in some portfolios. Therefore, Q5 shows similar or better performance than QF in pricing anomalies-related portfolios. Furthermore, compared to the Q5 and FF5 models, the FF6 model with the addition of the momentum factor mitigates its difference from the Q5 model in pricing anomalies-related portfolios.

In summary, though the addition of expected growth factor makes Q5 capture more average returns than QF, FF6 outperforms both Q-factor models (QF and Q5) by a wide margin, more so at the higher quantiles. Consequently, we confirm that the differences between the two Q models (QF and Q5) and the two Fama-French models of FF5 and FF6 have both statistical and economic significance. Furthermore, the existing literature (e.g., Goyal et al., 2019) show that both the Q-factor models and the Fama-French models outperform the BS model. It is remarkable that our previous mean-test results show that BS ranks the third, higher than the ranking of the Q-factor models. To explain this contradiction and examine the performance of BS from a comparative perspective, we compare the difference between BS and FF6, and present the results in Table 12.

BS & FF6

[Please insert Table 12 about here]

Compared to the standard FF6 model, the performance of the BS model improves as quantile increases. BS underperforms FF6 with significant economic differences at lower quantiles of B/M-sorted and OP-sorted portfolios. That indicates that the joint effect of HMLm, IA and ROE is rather negligible in the presence of HML, RMW, and CMA in capturing lower returns of these portfolios. Surprisingly, the performance of BS in capturing higher-than-median returns demonstrates its superior pricing ability, even better than the performance of FF6 in most portfolios. The following five percentage differences show BS's significant outperformance over FF6: 12.2% and 11.4% in pricing Size-MOM portfolios at the 55th and the 60th percentile respectively; 17.9% in pricing Size-OP-INV portfolios at the 55th percentile, 14.3% in pricing past return-related portfolios at 55th percentile, and 12.3% in pricing value-related portfolios at the 60th percentile. The former two in the Size-MOM portfolios are the only significant economic differences between BS and FF6, confirming BS's superior pricing ability to FF6 when considering the momentum factor. Unlike other models, the outperformance of the BS model exhibits its regularity in pricing slightly above median returns for all the portfolios. This is in contrast to its underperformance for under median returns. In particular, the pricing ability of BS is low when only the central tendency of returns is concerned; this possibly justifies the model's low ranking in OLS regressions.

Although we find the outperformance of the BS model in pricing higher than median returns, the FF6 model remains its superiority in the overall comparison. All of our previous results show that the FF6 model has the best performance among all models. However, we have not demonstrated the effect of individual factors. In Section 2, we present the argument on the value and momentum factors of the FF6 model following the recent literature. Next, we compare the difference between FF5-FF6 and FF6-hml-FF6 to document the impact of the value factor and the momentum factor in each portfolio at different quantiles and examine to what extent these two factors are important in asset pricing. Tables 14 and 15 present the results of percentage differences between FF5-FF6 and FF6-hml-FF6, respectively.

FF5 & FF6

[Please insert Table 13 about here]

In the previous sections, we have analyzed the importance of the momentum factor in pricing past-return-sorted portfolios. Table 13 shows that the differences between FF5 and FF6 across all quantiles are significant for the Size-MOM portfolios. The largest difference between the two models is 141.6% at the median, which means that a 1.416 times average cost is saved from choosing FF6 instead of FF5. While portfolios sorted on other past returns produce similar results, not all the generated differences are economically significant, indicating that the momentum factor cannot completely explain both short-term and long-term reversal. The lack of significant difference between 65th and 85th percentile levels reveals that a proportion of the larger value-weighted past return portfolios are rather unaffected by the momentum factor compared to lower returns, with the exclusion of the two highest extreme percentiles. The influences of the momentum factor between these two and other portfolios differ sharply. Consistent with Fama and French (2017), the inclusion of the momentum factor marginally affects Fama-French models' portfolio pricing competence.

Except for momentum portfolios, there is barely any significant economic difference between FF5 and FF6 in all other portfolios, suggesting that the addition of momentum factor does not significantly improve the model's pricing ability of the remaining portfolios from the economic perspective.

Overall, because the momentum factor plays a pivotal role in pricing momentum-related portfolios while in the meantime it does not harm the performance of other portfolios, we conclude that the momentum factor is a necessary addition to the Fama-French factor models.

FF6-hml & FF6

[Please insert Table 14 about here]

Fama and French (2015) argue that the value effect is completely subsumed by the newly added profitability and investment factors; hence the value factor, HML, seems to become redundant in describing average returns.

Nevertheless, we show that the value factor is necessary in pricing the B/M-sorted and dividend-related portfolios. Many differences between FF6 and FF6-hml in these portfolios are significant in the 196 large pooled portfolios between the 55th and 65th percentiles. This is evidence that the HML factor supplements additional explanatory power to the model and the value effefct is not fully captured by its exposures to the other factors of the five-factor model. We find that HML is essential in explaining average returns in the portfolios related to the company's valuation as both book-to-market ratio and dividend yield are highly associated with company's valuation. Indeed, the impact of HML is not that significant when we only focus on the central tendency at both mean and median. However, HML is indispensable at some other percentiles when we consider the whole distribution of returns. The significant difference between FF6 and FF6-hml at extreme percentiles shows that the inclusion of HML can be a favored choice in such applications as risk management, in which the primary focus of risk managers is on the tails rather than the central tendency of the return distribution.

6. Conclusion

This thesis compares the performance of ten selected asset pricing models using a combined methodology of the distance-based metrics and quantile regressions. Our main findings are summarized below.

Overall, FF6 is the best model to describe the returns of each set of portfolios, followed by FF5. Specifically, FF6 has a significant comparative advantage in pricing low-quantile returns and FF5 is more reliable in the case of high-quantile returns. Both models are superior to other competing models in pricing the distribution of returns for most portfolios. The comparison results of the percentage difference between FF6 and other models demonstrate the consistency and robustness of FF6's pricing ability. Besides, we find that the BS model exhibits a superior pricing ability at some higher-than-median quantiles. These findings are generally consistent but differ from the OLS results of Goyal et al. (2019). They find that FF6 and FF6-hml are the two best models with indistinguishable performance, and that the BS model underperforms other competing models. In comparison with other performance measures, we find that the alpha-based statistics generates similar results to the distance-based metrics; however, the alpha-based statistics are subject to the extreme-error problems and do not consider the influence of estimation precision. The GRS-statistic, however, generates inconsistent model ranking results mainly due to its lack-of-power problem. The GRS test suggests that Q5 and SY can better explain the quantile returns of test portfolios than other models.

To check the robustness of the results, we conduct pairwise mean tests, which further confirm the overall rankings of model performance. Among all test portfolios, FF6 performs the best, followed by FF5 and BS. However, it should be cautioned that a pure reliance on statistical interpretations of the model comparison results can be ambiguous. For example, the p-values of mean tests indicate that, although FF6 outperforms FF5 by economic significance, there is a lack of statistical evidence that the difference between the two models is significant in some quantiles. This result adds further support to Harvey's (2017) argument against p-hacking.

While showing some regularities in the comparison, we find that the BS model also bears certain weaknesses in explaining median or lower returns, yet is equally competent in explaining higher returns compared to FF6. The OLS regressions fail to capture the pricing ability of BS in the higher-than-median returns; that is why the BS model is ranked low by the OLS. It is interesting to find that the BS pricing ability in higher-than-median returns does not depend on the type of test portfolios. This provides a further research avenue in the future.

Finally, in the evaluation of the momentum and value factors in FF6, we provide evidence that both factors significantly improve the pricing ability of a model. The momentum factor exhibits its absolute necessity in pricing test portfolios sorted by past returns. Something similar can be said about the value factor (HML) in pricing B/M-sorted and dividend-related portfolios. Although Fama and French (2015) indicate that HML is redundant and the value effect is mostly subsumed by the investment factor (CMA) and the profitability factor (RMW), we find that the pricing ability of the FF models in the absence of the value factor significantly underperforms in these specific portfolios.

To conclude, FF6 is the model that best describes portfolio returns in our quantile-based performance comparison. Compared to the conventional GRS and alpha-based statistics, the distance-based metrics can be used as a robust performance measure that provides highly consistent model ranking results.

References

Aharoni, G., Grundy, B., and Zeng, Q., (2013). Stock Returns and the Miller Modigliani Valuation Formula: Revisiting the Fama French Analysis. *Journal of Financial Economics*, 110(2), 347-357.

Allen, D. E., Singh, A. K., and Powell, R. (2011), Asset pricing, the Fama French factor model and the implications of quantile regression analysis. *Financial Econometrics Modeling: Market Microstructure, Factor Models and Financial Risk Measures*. UK: Palgrave Macmillan. p176-193.

Asness, C., and Frazzini, A., (2013). The Devil in HML's Details. *Journal of Portfolio Management*, 39, 49-68.

Autchariyapanitkul K., Chanaim S., and Sriboonchitta S. (2015). Quantile Regression Under Asymmetric Laplace Distribution in Capital Asset Pricing Model. Econometrics of Risk. Studies in Computational Intelligence, vol 583. Springer, Cham.

Ball, R., (1978). Anomalies in Relationships Between Securities' Yields and Yield-Surrogates. *Journal of Financial Economics*, 6(2-3), 103-126.

Banz, R., (1981). The Relationship between Return and Market Value of Common Stocks. *Journal of Financial Economics*, 9(1), 3-18.

Barnes, M., and Hughes, A. W., (2002). A Quantile Regression Analysis of the Cross Section of Stock Market Returns. *Federal Reserve Bank of Boston Working Paper*.

Barillas, F., and Shanken, J., (2018). Comparing Asset Pricing Models. *Journal of Finance*, 73(2), 715-753.

Bassett, G., and Chen, H., (2001), Quantile style: Return-based attribution using regression quantiles. *Empirical Economics*, 26, 293-305.

Basu, S., (1983). The Relationship between Earnings' Yield, Market Value and Return for NYSE Common Stocks: Further Evidence. *Journal of Financial Economics*, 12(1), 129-156.

Black, F. (1993), Beta and return. *Journal of Portfolio Management*, 20, 8-18

Brown, S., and Goetzmann, W. N., (1995). Performance Persistence. *The Journal of Finance*, 50(2), 679-698.

Carhart, M., (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, 52(1), 57-82.

Chan, L., and Lakonishok, J., (1992). Robust Measurement of Beta Risk. *The Journal of Financial and Quantitative Analysis*, 27(2), 265-282.

Chen, W. S., So, K. P., and Chiang, C., (2016). Evidence of Stock Returns and Abnormal Trading Volume: A Threshold Quantile Regression Approach. *The Japanese Economic Review*. 67(1), 96-124.

Chiang T. C., and Li J., (2012). Stock Returns and Risk: Evidence from Quantile. *Journal of Risk and Financial Management*, 5(1), 1-39.

Cochrane, J., (1991). Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations. *The Journal of Finance*, 46(1), 209-237.

Fama, E., and French, R., (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427-465.

Fama, E., and French, R., (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics,* 33(1), 3-56.

Fama, E., and French, R., (2012). Size, Value, and Momentum in International Stock Returns. *Journal of Financial Economics*, 105(3), 457-72.

Fama, E., and French, R., (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics,* 116(1), 1-22.

Fama, E., and French, R., (2016b). Dissecting Anomalies with A Five-Factor Model, *Review of Financial Studies*, 29(1), 69-103.

Fama, E., and French, R., (2017). International Tests of a Five-Factor Asset Pricing Model. *Journal of Financial Economics*, 123(3), 441-463.

Fama, E., and French, R., (2018). Choosing Factors. *Journal of Financial Economics*, 128(2), 234-252.

Gibbons, M., Ross, S., and Shanken, J., (1989). A Test of the Efficiency of a Given Portfolio. *Econometrica*, 57(5), 1121-1152.

Goetzmann, N., and Ibbotson, R., (1994). Do Winners Repeat? *The Journal of Portfolio Management*, 20(2), 9-18.

Golan, A., Judge, G., and Miller, D., (1997), The Maximum Entropy Approach to Estimation And Inference. *Applying Maximum Entropy to Econometric Problems* (Advances in Econometrics, Volume 12). Emerald Group Publishing Limited, pp.3 - 24

Goyal, A., He Z., and Huh S., (2019). Distance-Based Metrics: A Bayesian Solution to the Power and Extreme-Error Problems in Asset-Pricing Tests. Working Paper.

Griffin, J., and Lemmon, M., (2002). Book-to-Market Equity, Distress Risk, and Stock Returns. *The Journal of Finance,* 57(5), 2317-2336.

He, Z. (2018), Comparing Asset Pricing models: Distance-based Metrics and Bayesian Interpretations. Working Paper.

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

Harvey, C. R., and Liu, Y., (2017), Lucky Factors, *SSRN Working Paper*.

Hendricks, D., Patel, J., and Zeckhauser, R., (1993). Hot Hands in Mutual Funds: Short-Run Persistence Of Relative Performance, 1974-1988. *The Journal of Finance*, 48(1), 93-130.

Hou, K., Chen, X., and Lu, Z., (2015). Digesting Anomalies: An Investment Approach. *Review of Financial Studies,* (28), 650-705.

Hou, K., Chen, X., and Lu, Z., (2016), A Comparison of New Factor Models. *Working paper*.

Hou, K., Mo, H., Xue, C., and Zhang, L., (2018). Q5. *NBER Working Paper*.

Horowitz, J. L., Loughran, T., and Savin, N. E. (2000), Three analyses of the firm size premium. *Journal of Empirical Finance*, 7, 143-153.

Jegadeesh, N., and Titman, S., (1993). Returns To Buying Winners And Selling Losers: Implications For Stock Market Efficiency. *The Journal of Finance*, 48(1), 65-91.

Knez, P. J., and Ready, M., (1997). On The Robustness of Size and Book-to-Market in Cross-Sectional Regressions. *The Journal of Finance*, 52(4), 1355-1382.

Koenker, R., and Bassett, G., (1978). Regression Quantiles. *Econometrica*, 46(1), 33-50.

Kothari, S., and Shanken, J., and Sloan, R., (1995). Another Look at The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 50(1), 185-224.

Lakonishok, J., Shleifer, A., and Vishny, R., (1994). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*, 49(5), 1541-1578.

Levhari, D., and Levy, H., (1977). The Capital Asset Pricing Model and the Investment Horizon. *The Review of Economics and Statistics*, 59(1), 92-104.

Liew, J., and Vassalou, M., (1994). Can Book-to-Market, Size and Momentum be Risk Factors that Predict Economic Growth? *Journal of Financial Economics*, 57(2), 221-245.

Lintner, J., (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), 13-37.

Ma, L., and Pohlman, L., (2008). Return Forecasts and Optimal Portfolio Construction: A Quantile Regression Approach. *The European Journal of Finance*, 14(5), 409-425.

Monge, G., (1781). Mémoire Sur la Théorie des Déblais et des Remblais. *In Histoire de l'Académie Royale des Sciences de Paris*, 666-704.

Mosteller, F., and Tukey, J. W., (1977). Data Analysis and Regression. A second course in statistics. Pearson.

Novy-Marx, R., (2013). The Other Side of Value: The Gross Profitability Premium. *Journal of Financial Economics*, 108(1), 1-28.

Officer, R. R. (1972). The Distribution of Stock Returns. *Journal of the American Statistical Association*, 67, 340.

Pastor, L., and Stambaugh, R. F., (2000), Comparing Asset Pricing Models: An Investment Perspective, *Journal of Financial Economics*, 56(2000), 335-381.

Sharma, P., Gupta, P., and Singh A., (2016). Pricing Ability of Four Factor Model using Quantile Regression: Evidences from India. *International Journal of Economics and Financial Issues*, 6(4), 2146-4138.

Sharpe, W., (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions Of Risk. *The Journal of Finance*, 19(3), 425-442.

Stambaugh, R., and Yuan, Y., (2017). Mispricing Factors. *Review of Financial Studies*, 30, 1270-1315.

Wermers, R., (1996), Momentum Investment Strategies of Mutual Funds, Performance Persistence, and Survivorship Bias, *Working paper*.

Yamaka, W., Autchariyapanitkul, K., Mennejuk, P., and Sriboonchitta, S., (2017). Capital Asset Pricing Model Through Quantile Regression: An Entropy Approach. *Thai Journal of Mathematics*. 53-65.

Tables

Table 1: Ten Tested Asset-Pricing Models and Their Factors

Model	**Factors**					
CAPM	MKT					
FF3	MKT	SMB	HML			
CAR4	MKT	SMB	HML	MOM		
FF5	MKT	SMB	HML	RMW	CMA	
FF6	MKT	SMB	HML	RMW	CMA	MOM
FF6-hml	MKT	SMB	RMW	CMA	MOM	
BS	MKT	SMB	HML^m	IA	ROE	MOM
QF	MKT	ME	IA	ROE		
Q5	MKT	ME	IA	ROE	EG	
SY	MKT	SMB_{sy}	RERF	MGMT		

Note: Table 1 reports the tested models and their factors. CAPM is the Capital Asset Pricing Model. FF3 is the Fama and French three-factor model. CAR4 is the Carhart four-factor model. FF5 is the Fama and French five-factor model. FF6 is the Fama and French (2018) six-factor model. FF6-hml is the Fama and French six-factor model except for the value factor HML. BS is a six-factor model proposed by Barilla and Shanken (2018). QF is a four-factor model proposed by Hou, Xue and Zhang (2015). Q5 is a five-factor model proposed by Hou, Mo, Xue and Zhang (2018). SY is a four-factor model proposed by Stambaugh and Yuan (2017). MKT is the market factor. SMB, ME and SMB_{sy} are size factors. HML is the value factor. MOM is the momentum factor. RMW and ROE are profitability factors. CMA and IA are investment factors. HML^m is the timely updated value factor in BS. EG is the expected growth factor. RERF and MGMT are two mispricing factors in SY.

Table 2: Descriptive Statistics of Factors

Variable	Mkt_rf	SMB	HML	RMW	CMA	MOM	ME	IA	ROE	EG	HML^m	SMB_{sy}	MGMT	RERF
Panel A: Descriptive Statistics														
Mean	0.005	0.002	0.004	0.003	0.003	0.006	0.003	0.004	0.005	0.008	0.003	0.004	0.006	0.007
Std	0.045	0.031	0.029	0.022	0.020	0.043	0.031	0.019	0.025	0.019	0.035	0.029	0.029	0.039
Min	-0.232	-0.149	-0.111	-0.180	-0.069	-0.344	-0.144	-0.072	-0.138	-0.062	-0.180	-0.111	-0.089	-0.215
Q1	-0.022	-0.015	-0.012	-0.008	-0.010	-0.009	-0.015	-0.008	-0.007	-0.003	-0.016	-0.014	-0.011	-0.013
Median	0.008	0.001	0.003	0.002	0.002	0.008	0.002	0.003	0.007	0.007	0.001	0.003	0.006	0.007
Q3	0.036	0.021	0.018	0.013	0.016	0.029	0.021	0.016	0.019	0.019	0.019	0.022	0.023	0.028
Max	0.161	0.184	0.129	0.128	0.096	0.184	0.221	0.092	0.104	0.109	0.269	0.160	0.146	0.185
Skew.	-0.517	0.403	0.070	-0.349	0.329	-1.334	0.607	0.116	-0.699	0.230	0.876	0.383	0.178	-0.082
Kurt.	1.820	3.310	1.992	11.312	1.529	10.295	5.231	1.477	4.677	2.090	8.308	2.149	1.644	3.619
Panel B: Correlation Matrix														
Mkt_rf	1.000													
SMB	0.277	1.000												
HML	-0.270	-0.078	1.000											
RMW	-0.233	-0.361	0.080	1.000										
CMA	-0.397	-0.097	0.699	-0.014	1.000									
MoM	-0.143	-0.054	-0.187	0.115	-0.001	1.000								
ME	0.267	0.973	-0.037	-0.370	-0.059	-0.022	1.000							
IA	-0.385	-0.185	0.672	0.088	0.913	0.026	-0.146	1.000						
ROE	-0.202	-0.373	-0.138	0.668	-0.084	0.502	-0.313	0.038	1.000					
EG	-0.466	-0.419	0.199	0.431	0.335	0.343	-0.371	0.387	0.515	1.000				
HML^m	-0.119	-0.011	0.775	-0.064	0.516	-0.648	-0.005	0.495	-0.451	-0.041	1.000			
SMB_{sy}	0.260	0.942	-0.047	-0.285	-0.069	0.003	0.927	-0.145	-0.283	-0.334	-0.035	1.000		
MGMT	-0.540	-0.341	0.705	0.218	0.786	0.041	-0.307	0.775	0.078	0.539	0.491	-0.289	1.000	
RERF	-0.260	-0.151	-0.309	0.447	-0.063	0.721	-0.146	-0.055	0.642	0.464	-0.635	-0.088	0.008	1.000

Note: Panel A of Table 2 reports descriptive statistics of all tested factors. MKT is the market factor. SMB, ME and SMB_{sy} are size factors. HML is the value factor. MOM is the momentum factor. RMW and ROE are profitability factors. CMA and IA are investment factors. HML^m is the timely updated value factor. EG is the expected growth factor. RERF and MGMT is mispricing factors. Mean is the mean of factors. Sd is the log standard deviation of factors. Min s the minimum value of factors. Q1 is the first quantile of factors. Median is the median of factors. Q3 is the third quantile of factors. Max is the maximum value of factors. Skew the skewness of factors. Kurt is the kurtosis of factors. Panel B of Table 2 reports correlation coefficients.

Table 3: Performance Metrics for the 196 Pooled Portfolios

Models	TD	AD	RMSE ($\widetilde{\alpha}$)	RMSE($\tilde{\sigma}_{\alpha}$)	$A\widetilde{\sigma}_{\alpha}^{2}$ /$A\widetilde{\sigma}^{2}$	GRS	A|$\widetilde{\alpha}$|
Panel A: Model performance at 10th percentile							
CAPM	45.379	3.241	3.234	0.212	0.004	4.720	3.087
FF3	32.178	2.298	2.292	0.167	0.005	4.527	2.154
CAR4	30.351	2.168	2.162	0.161	0.006	4.366	2.062
FF5	30.564	2.183	2.177	0.164	0.006	4.246	2.061
FF6	28.767	2.055	2.049	0.152	0.005	4.149	1.965
FF6-hml	30.513	2.179	2.173	0.165	0.006	4.159	2.077
BS	30.914	2.208	2.201	0.174	0.006	4.144	2.111
QF	32.973	2.355	2.347	0.197	0.007	3.923	2.246
Q5	31.768	2.269	2.261	0.197	0.008	3.607	2.162
SY	33.124	2.366	2.359	0.188	0.006	3.889	2.273
Panel B: Model performance at 20th percentile							
CAPM	28.855	2.061	2.055	0.159	0.006	4.664	1.939
FF3	20.488	1.463	1.459	0.112	0.006	4.437	1.360
CAR4	19.244	1.375	1.370	0.113	0.007	4.327	1.296
FF5	19.424	1.387	1.383	0.117	0.007	4.116	1.310
FF6	18.264	1.305	1.300	0.111	0.007	4.034	1.249
FF6-hml	19.250	1.375	1.370	0.120	0.008	4.086	1.305
BS	19.768	1.412	1.406	0.128	0.008	4.101	1.350
QF	20.436	1.460	1.453	0.138	0.009	3.931	1.388
Q5	19.733	1.409	1.403	0.139	0.010	3.636	1.339
SY	21.006	1.500	1.494	0.137	0.008	3.800	1.435
Panel C: Model performance at 50th percentile							
CAPM	4.686	0.335	0.309	0.128	0.172	4.099	0.227
FF3	3.669	0.262	0.247	0.089	0.129	3.368	0.165
CAR4	2.787	0.199	0.177	0.091	0.267	3.130	0.130
FF5	2.841	0.203	0.182	0.090	0.242	3.040	0.117
FF6	2.136	0.153	0.123	0.090	0.539	2.791	0.091
FF6-hml	2.298	0.164	0.134	0.095	0.509	2.935	0.096
BS	2.549	0.182	0.152	0.100	0.429	3.196	0.118
QF	2.534	0.181	0.148	0.104	0.489	3.310	0.112
Q5	2.626	0.188	0.150	0.112	0.557	2.927	0.120
SY	2.628	0.188	0.155	0.106	0.470	2.725	0.109
Panel D: Model performance at 80th percentile							
CAPM	30.946	2.210	2.203	0.178	0.007	4.843	2.121
FF3	19.605	1.400	1.395	0.117	0.007	4.591	1.343
CAR4	19.442	1.389	1.384	0.118	0.007	4.409	1.333
FF5	18.399	1.314	1.309	0.117	0.008	4.285	1.256
FF6	18.159	1.297	1.292	0.116	0.008	4.164	1.243
FF6-hml	19.483	1.392	1.387	0.119	0.007	4.109	1.328
BS	18.228	1.302	1.296	0.126	0.009	4.207	1.236
QF	21.101	1.507	1.501	0.135	0.008	4.290	1.438
Q5	21.051	1.504	1.497	0.143	0.009	3.780	1.431
SY	20.689	1.478	1.471	0.139	0.009	3.957	1.412
Panel E: Model performance at 90th percentile							
CAPM	50.172	3.584	3.574	0.258	0.005	4.796	3.438
FF3	31.870	2.276	2.270	0.173	0.006	4.609	2.185
CAR4	31.431	2.245	2.238	0.173	0.006	4.436	2.158
FF5	29.695	2.121	2.115	0.164	0.006	4.296	2.023
FF6	28.996	2.071	2.065	0.161	0.006	4.228	1.985
FF6-hml	31.082	2.220	2.214	0.167	0.006	4.272	2.119
BS	29.586	2.113	2.106	0.175	0.007	4.232	2.009
QF	33.673	2.405	2.397	0.192	0.006	4.129	2.292
Q5	33.454	2.390	2.381	0.203	0.007	3.762	2.270
SY	33.716	2.408	2.401	0.193	0.006	3.945	2.299

Table 4: Performance Metrics for the 150 Pooled Portfolios

Models	TD	AD	RMSE ($\widetilde{\alpha}$)	RMSE($\tilde{\sigma}_{\alpha}$)	$A\widetilde{\sigma}_{\alpha}^{2}$ /$A\widetilde{\sigma}^{2}$	GRS	A|$\widetilde{\alpha}$|
Panel A: Model performance at 10th percentile							
CAPM	29.587	2.416	2.411	0.159	0.004	5.476	2.248
FF3	26.114	2.132	2.126	0.160	0.006	5.350	2.011
CAR4	24.706	2.017	2.012	0.152	0.006	5.241	1.937
FF5	25.401	2.074	2.068	0.158	0.006	5.026	1.977
FF6	24.150	1.972	1.966	0.151	0.006	4.940	1.909
FF6-hml	24.537	2.003	1.997	0.154	0.006	4.962	1.938
BS	25.264	2.063	2.056	0.174	0.007	5.350	1.997
QF	25.963	2.120	2.112	0.179	0.007	5.101	2.040
Q5	25.240	2.061	2.053	0.184	0.008	4.450	1.987
SY	24.971	2.039	2.032	0.165	0.007	4.464	1.965
Panel B: Model performance at 20th percentile							
CAPM	18.817	1.536	1.532	0.117	0.006	5.405	1.409
FF3	16.568	1.353	1.349	0.106	0.006	5.288	1.262
CAR4	15.551	1.270	1.265	0.106	0.007	5.169	1.207
FF5	16.233	1.325	1.321	0.111	0.007	4.890	1.258
FF6	15.261	1.246	1.241	0.108	0.008	4.692	1.201
FF6-hml	15.475	1.264	1.259	0.109	0.007	4.823	1.217
BS	15.881	1.297	1.291	0.121	0.009	5.326	1.251
QF	16.254	1.327	1.322	0.121	0.008	5.006	1.273
Q5	15.752	1.286	1.280	0.127	0.010	4.389	1.236
SY	15.631	1.276	1.271	0.119	0.009	4.150	1.227
Panel C: Model performance at 50th percentile							
CAPM	3.095	0.253	0.236	0.091	0.148	3.348	0.147
FF3	2.963	0.242	0.228	0.081	0.125	2.790	0.123
CAR4	2.251	0.184	0.163	0.084	0.265	2.693	0.099
FF5	2.441	0.199	0.180	0.086	0.230	2.508	0.111
FF6	1.887	0.154	0.128	0.086	0.454	2.303	0.095
FF6-hml	1.878	0.153	0.126	0.087	0.476	2.287	0.092
BS	2.089	0.171	0.143	0.093	0.427	2.493	0.114
QF	2.012	0.164	0.135	0.094	0.485	2.865	0.096
Q5	1.996	0.163	0.129	0.100	0.603	2.535	0.100
SY	1.917	0.156	0.125	0.094	0.567	2.066	0.094
Panel D: Model performance at 80th percentile							
CAPM	18.288	1.493	1.488	0.124	0.007	5.548	1.431
FF3	15.197	1.241	1.236	0.109	0.008	5.422	1.206
CAR4	15.216	1.242	1.237	0.112	0.008	5.344	1.210
FF5	14.554	1.188	1.183	0.110	0.009	4.991	1.148
FF6	14.580	1.190	1.185	0.109	0.009	4.905	1.152
FF6-hml	15.096	1.233	1.227	0.112	0.008	4.984	1.189
BS	14.518	1.185	1.179	0.121	0.010	5.036	1.137
QF	15.757	1.287	1.281	0.121	0.009	5.065	1.232
Q5	15.703	1.282	1.275	0.132	0.011	4.581	1.223
SY	15.569	1.271	1.265	0.123	0.009	4.281	1.219
Panel E: Model performance at 90th percentile							
CAPM	30.393	2.482	2.474	0.193	0.006	5.467	2.368
FF3	25.295	2.065	2.058	0.168	0.007	5.336	2.002
CAR4	25.073	2.047	2.040	0.166	0.007	5.229	1.987
FF5	24.223	1.978	1.971	0.167	0.007	5.058	1.900
FF6	23.843	1.947	1.940	0.162	0.007	4.987	1.882
FF6-hml	24.616	2.010	2.003	0.164	0.007	5.003	1.938
BS	23.971	1.957	1.950	0.170	0.008	5.152	1.885
QF	25.984	2.122	2.114	0.182	0.007	4.975	2.026
Q5	25.851	2.111	2.102	0.191	0.008	4.592	2.006
SY	25.551	2.086	2.079	0.176	0.007	4.340	2.001

Note: This table reports the distance-based metrics and alpha-based statistic generated by various models for 196 and 150 large pooled portfolios. TD is the total distance. AD is the average distance. RMSE ($\tilde{\alpha}$) is the square root of the mean square pricing error. RMSE ($\tilde{\sigma}_{\alpha}$) is the square root of the mean square error. A$\tilde{\sigma}_{\alpha}^2$ /A$\tilde{\sigma}^2$ is the ratio o to the mean square standard error to the mean square pricing error. A|$\tilde{\alpha}$| is the mean absolute pricing error (MAE). CAPM is the Capital Asset Pricing Model. FF3 is the Fama and French three-factor model. CAR4 is the Carhart four-factor model. FF5 is the Fama and French (2017) five-factor model. FF6 is the Fama and French (2018) six-factor model. FF6-hml is the Fama and French six-factor model except for the value factor. BS is a six-factor model proposed by Barilla and Shanken (2018). QF is a four-factor model proposed by Hou, Xue and Zhang (2015). Q5 is a five-factor model proposed by Hou, Mo, Xue and Zhang (2018). SY is a four-factor model proposed by Stambaugh and Yuan (2017).

Table 5: Mean-Test Results

	CAPM	FF3	CAR4	FF5	FF6	FF6-hml	BS	QF	Q5	SY
Panel A: Mean-test Result for 196 Pooled Portfolios										
CAPM	0.000									
T. Value	n/a									
FF3	-0.628***	0.000								
T. Value	-5.634	n/a								
CAR4	-0.688***	-0.06***	0.000							
T. Value	-5.970	-5.497	n/a							
FF5	-0.713***	-0.086***	-0.026**	0.000						
T. Value	-5.834	-7.457	-2.189	n/a						
FF6	-0.771***	-0.143***	-0.083***	-0.058***	0.000					
T. Value	-6.007	-7.063	-6.155	-5.267	n/a					
FF6-hml	-0.694***	-0.067***	-0.007	0.019	0.077***	0.000				
T. Value	-6.003	-5.469	-1.515	1.563	5.777	n/a				
BS	-0.716***	-0.088***	-0.028*	-0.003	0.055***	-0.021	0.000			
T. Value	-5.827	-6.849	-1.791	-0.512	3.777	-1.383	n/a			
QF	-0.597***	0.031	0.091***	0.116***	0.174***	0.097***	0.119***	0.000		
T. Value	-6.169	1.735	4.557	4.264	5.335	4.978	4.137	n/a		
Q5	-0.622***	0.006	0.066***	0.091***	0.149***	0.073***	0.094***	-0.025***	0.000	
T. Value	-6.171	0.318	4.129	3.518	5.182	4.506	3.269	-2.953	n/a	
SY	-0.595***	0.032**	0.092***	0.118***	0.176***	0.099***	0.120***	0.002	0.026*	0.000
T. Value	-6.156	2.121	4.549	4.594	5.369	4.762	4.551	0.199	1.963	n/a
Panel B: Mean-test Result for 150 Pooled Portfolios										
CAPM	0.000									
T. Value	n/a									
FF3	-0.189***	0.000								
T. Value	-5.540	n/a								
CAR4	-0.239***	-0.051***	0.000							
T. Value	-6.354	-4.601	n/a							
FF5	-0.239***	-0.051***	0.000	0.000						
T. Value	-5.993	-6.690	0.007	n/a						
FF6	-0.288***	-0.099***	-0.049***	-0.049***	0.000					
T. Value	-6.222	-6.126	-5.016	-4.305	n/a					
FF6-hml	-0.258***	-0.069***	-0.019***	-0.019*	0.03***	0.000				
T. Value	-6.398	-5.543	-4.887	-1.799	4.653	n/a				
BS	-0.259***	-0.07***	-0.019*	-0.02***	0.029***	-0.001	0.000			
T. Value	-6.135	-7.153	-1.778	-4.168	3.225	-0.085	n/a			
QF	-0.196***	-0.008	0.043***	0.043***	0.092***	0.062***	0.062***	0.000		
T. Value	-6.907	-0.847	4.211	3.070	4.964	4.936	4.163	n/a		
Q5	-0.217***	-0.029**	0.022***	0.022	0.07***	0.041***	0.041**	-0.021***	0.000	
T. Value	-6.840	-2.445	3.371	1.475	4.485	4.330	2.705	-3.599	n/a	
SY	-0.231***	-0.042***	0.009	0.009	0.057***	0.027***	0.028*	-0.034***	-0.013***	0.000
T. Value	-6.967	-3.511	1.636	0.609	3.993	3.406	1.930	-4.863	-6.463	n/a

Note: Panel A reports mean-test results for 196 pooled portfolios. Panel B reports mean-test results for 150 pooled portfolios. ***, **, and * indicates statistical significance at 1%, 5%, and 10%, respectively. T.V. stands for t-value. CAPM is the Capital Asset Pricing Model. FF3 is the Fama and French three-factor model. CAR4 is the Carhart four-factor model. FF5 is the Fama and French (2015) five-factor model. FF6 is the Fama and French (2018) six-factor model. FF6-hml is the Fama and French six-factor model except for the value factor. BS is a six-factor model proposed by Barilla and Shanken (2018). QF is a four-factor model proposed by Hou, Xue and Zhang (2015). Q5 is a five-factor model proposed by Hou, Mo, Xue and Zhang (2018). SY is a four-factor model proposed by Stambaugh and Yuan (2017).

Table 6: Mean-Test Ranking Order

Portfolios/Orders	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Size-B/M	FF6	FF5	CAR4	FF3	BS	FF6-hml	Q5	SY	QF	CAPM
Size-INV	FF6	FF5	FF6-hml	BS	Q5	CAR4	FF3	QF	SY	CAPM
Size-MOM	FF6	BS	FF6-hml	CAR4	SY	Q5	QF	FF5	FF3	CAPM
Size-OP	FF6	FF5	FF6-hml	BS	Q5	CAR4	FF3	QF	SY	CAPM
Size-B/M-INV	FF5	FF6	BS	FF3	CAR4	FF6-hml	Q5	QF	SY	CAPM
Size-B/M-OP	FF5	FF6	BS	CAR4	FF3	FF6-hml	SY	Q5	QF	CAPM
Size-OP-INV	FF5	FF6	FF6-hml	BS	Q5	QF	CAR4	FF3	SY	CAPM
FF-related	FF6	FF5	FF6-hml	BS	CAR4	FF3	Q5	QF	SY	CAPM
Anomalies	SY	FF6	FF5	FF6-hml	Q5	BS	QF	CAR4	FF3	CAPM
Past return	FF6	FF6-hml	BS	CAR4	SY	Q5	QF	FF5	FF3	CAPM
Value	FF5	FF6	FF3	BS	CAR4	SY	FF6-hml	Q5	QF	CAPM

Note: This table reports mean-test ranking order. CAPM is the Capital Asset Pricing Model. FF3 is the Fama and French three-factor model. CAR4 is the Carhart four-factor model. FF5 is the Fama and French (2015) five-factor model. FF6 is the Fama and French (2018) six-factor model. FF6-hml is the Fama and French six-factor model except for the value factor. BS is a six-factor model proposed by Barilla and Shanken (2018). QF is a four-factor model proposed by Hou, Xue and Zhang (2015). Q5 is a five-factor model proposed by Hou, Mo, Xue and Zhang (2018). SY is a four-factor model proposed by Stambaugh and Yuan (2017). The 1st to 4th rows are 5*5 double-sorted portfolios. Size-B/M is the portfolios sorted on size and book-to-market. Size-INV is the portfolios sorted on size and investment. Size-MOM is the portfolios sorted on size and momentum. Size-OP is the portfolios sorted on size and operation profitability. The 5th to 7th rows are 2*4*4 triple-sorted portfolios. Size-B/M-INV is the portfolios sorted on size, book-to-market and investment. Size-B/M-OP is the portfolios sorted on size, book-to-market and operation profitability. Size-OP-INV is the portfolios sorted on size, operation profitability and investment. The 8th to 11th rows are four sets of univariate-sorted decile portfolios. FF-related is the portfolios related to Fama and French factors. Anomalies is the portfolios related to other anomalies. Past return is the portfolios related on past returns. Value is the portfolios related to valuation.

Table 7: Different Levels of Statistical Significance between Models

	FF6	FF5	SY	BS	FF6-hml	CAR4	FF3	Q5	QF	CAPM
Panel A: Numbers of portfolios where p > 0.05										
FF6	0									
FF5	9	0								
SY	1	1	0							
BS	2	1	1	0						
FF6-hml	2	1	2	5	0					
CAR4	1	1	1	7	0	0				
FF3	1	1	1	3	0	8	0			
Q5	1	1	2	5	3	2	2	0		
QF	0	0	2	0	0	2	3	0	0	
CAPM	0	0	0	0	0	0	0	0	0	0
Panel B: Numbers of portfolios where 0.05 < p < 0.01										
FF6	0									
FF5	0	0								
SY	1	1	0							
BS	2	1	0	0						
FF6-hml	1	0	1	4	0					
CAR4	0	1	1	0	1	0				
FF3	0	1	1	3	0	0	0			
Q5	0	1	2	3	1	2	2	0		
QF	0	3	2	2	0	1	2	3	0	
CAPM	0	0	0	1	0	0	0	1	1	0

Note: Tables 7 reports different levels of statistical significance between models. Panel A of table 7 presents the numbers of portfolios where the p-value of tested models is greater than 0.05, indicating strong insignificant difference statistically. Panel B of table 7 presents the numbers of portfolio where the p-value for the tested models is less than or equal 0.05, but greater than 0.01, which indicates the difference is moderately insignificant in statistical. CAPM is the Capital Asset Pricing Model. FF3 is the Fama and French three-factor model. CAR4 is the Carhart four-factor model. FF5 is the Fama and French (2015) five-factor model. FF6 is the Fama and French (2018) six-factor model. FF6-hml is the Fama and French six-factor model except for the value factor. BS is a six-factor model proposed by Barilla and Shanken (2018). QF is a four-factor model proposed by Hou, Xue and Zhang (2015). Q5 is a five-factor model proposed by Hou, Mo, Xue and Zhang (2018). SY is a four-factor model proposed by Stambaugh and Yuan (2017).

Table 8: Percentage Differences between QF and FF5

	Panel A: 196 portfolios of QF and FF5								Panel B: 150 portfolios of QF and FF5				
Percentile	**196P**	**S-B/M**	**S-INV**	**S-MOM**	**S-OP**	**S-B/M-INV**	**S-B/M-OP**	**S-OP-INV**	**150P**	**FF**	**Anomaly**	**Past**	**Value**
5th	0.106	0.182	0.076	-0.034	0.136	0.147	0.166	0.102	0.043	0.115	0.030	-0.020	0.098
10th	0.079	0.157	0.070	-0.101	0.128	0.136	0.125	0.110	0.022	0.100	0.014	-0.047	0.068
15th	0.064	0.120	0.053	-0.105	0.093	0.113	0.120	0.078	0.018	0.082	-0.004	-0.024	0.064
20th	0.052	0.089	0.043	-0.143	0.106	0.102	0.124	0.079	0.001	0.073	-0.015	-0.050	0.041
25th	0.030	0.042	0.025	-0.183	0.089	0.089	0.082	0.074	-0.004	0.071	-0.012	-0.066	0.029
30th	0.014	0.024	0.017	-0.248	0.105	0.073	0.084	0.060	-0.013	0.059	-0.018	-0.078	0.024
35th	-0.014	-0.008	-0.003	-0.283	0.070	0.041	0.088	0.047	-0.038	0.057	-0.012	-0.147	-0.021
40th	-0.070	-0.040	-0.033	-0.480	0.013	-0.002	0.013	0.040	-0.064	0.030	-0.017	-0.212	-0.036
45th	-0.151	-0.044	-0.135	-0.693	-0.057	-0.009	-0.038	0.037	-0.170	-0.032	-0.074	-0.405	-0.126
50th	-0.124	0.267	0.064	-0.834	0.064	0.145	0.360	0.090	-0.216	0.140	-0.152	-0.639	0.134
55th	0.183	0.500	0.232	-0.298	0.359	0.378	0.543	0.164	0.080	0.280	0.102	-0.226	0.446
60th	0.223	0.396	0.173	-0.004	0.240	0.291	0.365	0.145	0.121	0.202	0.072	0.016	0.290
65th	0.192	0.317	0.129	0.079	0.235	0.223	0.308	0.134	0.118	0.168	0.071	0.069	0.217
70th	0.170	0.297	0.144	0.082	0.187	0.194	0.278	0.124	0.104	0.128	0.092	0.068	0.148
75th	0.154	0.250	0.112	0.076	0.179	0.183	0.220	0.122	0.085	0.134	0.076	0.043	0.109
80th	0.147	0.243	0.091	0.045	0.175	0.171	0.205	0.112	0.083	0.125	0.068	0.049	0.113
85th	0.138	0.243	0.112	0.045	0.168	0.172	0.175	0.104	0.071	0.110	0.047	0.042	0.114
90th	0.134	0.230	0.106	0.054	0.177	0.164	0.172	0.111	0.073	0.124	0.056	0.033	0.113
95th	0.120	0.194	0.101	0.012	0.172	0.167	0.175	0.114	0.050	0.134	0.047	-0.033	0.115

Note: Panel A of Table 8 reports the percentage differences between QF and FF5 of 196 portfolios across 19 percentiles. 196P is 196 large pooled portfolios combining all the subset portfolios within the panel. S-B/M is the portfolios double-sorted on size and book-to-market. S-INV is the portfolios double-sorted on size and investment. S-MOM is the portfolios double-sorted on size and momentum. S-OP is the portfolios double-sorted on size and operation profitability. S-B/M-INV is the portfolios triple-sorted on size, book-to-market and investment. S-B/M-OP is the portfolios triple-sorted on size, book-to-market and operation profitability. S-OP-INV is the portfolios triple-sorted on size, operation profitability and investment. Panel B of Table 8 reports the percentage differences between QF and FF5 of 150 portfolios across 19 percentiles. 150P is 150 large pooled portfolios combining all the subset portfolios within the panel. FF is the portfolios related to the Fama and French factors. Anomaly is the portfolios related to other anomalies. Past is the portfolios related on past returns. Value is the portfolios related to valuation. We mark the percentage differences greater than 0.10 in pale pink, indicating that FF5 significantly outperforms QF economically, and those lower than -0.10 in pale yellow, indicating that QF is significantly better than FF5 economically.

Table 9: Percentage Difference between QF and FF6

	Panel A: 196 portfolios of QF and FF6								Panel B: 150 portfolios of QF and FF6				
Percentile	**196P**	**S-B/M**	**S-INV**	**S-MOM**	**S-OP**	**S-B/M-INV**	**S-B/M-OP**	**S-OP-INV**	**150P**	**FF**	**Anomaly**	**Past**	**Value**
5th	0.176	0.213	0.076	0.389	0.152	0.144	0.180	0.101	0.105	0.132	0.040	0.183	0.107
10th	0.146	0.169	0.075	0.328	0.141	0.136	0.129	0.111	0.075	0.102	0.023	0.143	0.069
15th	0.130	0.142	0.065	0.301	0.113	0.121	0.126	0.079	0.076	0.097	0.012	0.174	0.066
20th	0.119	0.094	0.055	0.269	0.121	0.113	0.128	0.090	0.065	0.085	0.011	0.151	0.052
25th	0.103	0.071	0.031	0.224	0.118	0.098	0.112	0.081	0.065	0.076	0.023	0.138	0.052
30th	0.095	0.072	0.020	0.208	0.115	0.085	0.115	0.068	0.065	0.068	0.027	0.157	0.032
35th	0.083	0.022	0.021	0.260	0.087	0.049	0.141	0.055	0.054	0.078	0.028	0.138	-0.003
40th	0.059	0.017	-0.007	0.198	0.037	0.027	0.074	0.039	0.058	0.057	0.048	0.144	-0.010
45th	0.049	0.007	-0.097	0.250	-0.042	0.039	0.018	0.042	0.018	0.001	0.002	0.146	-0.072
50th	0.188	0.280	0.046	0.317	0.040	0.221	0.337	0.081	0.063	0.128	-0.016	0.102	0.158
55th	0.258	0.420	0.160	0.240	0.245	0.296	0.344	0.132	0.141	0.178	0.080	0.073	0.335
60th	0.228	0.339	0.140	0.240	0.195	0.246	0.259	0.137	0.114	0.164	0.025	0.124	0.227
65th	0.190	0.288	0.121	0.259	0.206	0.189	0.244	0.112	0.102	0.168	0.004	0.131	0.189
70th	0.177	0.269	0.136	0.262	0.179	0.168	0.245	0.111	0.091	0.132	0.035	0.126	0.116
75th	0.170	0.222	0.104	0.265	0.164	0.162	0.219	0.115	0.078	0.115	0.026	0.120	0.091
80th	0.162	0.228	0.084	0.254	0.150	0.156	0.177	0.114	0.081	0.115	0.024	0.130	0.096
85th	0.158	0.215	0.101	0.294	0.148	0.157	0.157	0.096	0.078	0.099	0.020	0.131	0.099
90th	0.161	0.216	0.105	0.316	0.173	0.146	0.163	0.095	0.090	0.113	0.032	0.148	0.102
95th	0.168	0.192	0.104	0.370	0.157	0.152	0.170	0.106	0.104	0.134	0.050	0.166	0.105

Note: Panel A of Table 8 reports the percentage differences between QF and FF6 of 196 pooled portfolios across 19 percentiles. 196P is 196 large pooled portfolios combining all subset portfolios within the panel. S-B/M is the portfolios double-sorted on size and book-to-market. S-INV is the portfolios double-sorted on size and investment. S-MOM is the portfolios double-sorted on size and momentum. S-OP is the portfolios double-sorted on size and operation profitability. S-B/M-INV is the portfolios triple-sorted on size, book-to-market and investment. S-B/M-OP is the portfolios triple-sorted on size, book-to-market and operation profitability. S-OP-INV is the portfolios triple-sorted on size, operation profitability and investment. Panel B of Table 8 reports the percentage differences between QF and FF6 of 150 pooled portfolios across 19 percentiles. 150P is 150 large pooled portfolios combining all subset portfolios within the panel. FF is the portfolios related to the Fama and French factors. Anomaly is the portfolios related to other anomalies. Past is the portfolios related on past returns. Value is the portfolios related to valuation. We mark the percentage differences greater than 0.10 in pale pink, indicating that FF6 significantly outperforms QF economically, and those lower than -0.10 in pale yellow, indicating that QF is significantly better than FF6 economically.

Table 10: Percentage Difference between Q5 and FF5

	Panel A: 196 portfolios of Q5 and FF5								Panel B: 150 portfolios of Q5 and FF5				
Percentile	**196P**	**S-B/M**	**S-INV**	**S-MOM**	**S-OP**	**S-B/M-INV**	**S-B/M-OP**	**S-OP-INV**	**150P**	**FF**	**Anomaly**	**Past**	**Value**
5th	0.063	0.129	0.046	-0.092	0.058	0.113	0.118	0.069	0.012	0.079	-0.007	-0.054	0.081
10th	0.039	0.113	0.022	-0.198	0.058	0.100	0.097	0.078	-0.006	0.078	-0.021	-0.085	0.053
15th	0.024	0.070	-0.008	-0.199	0.039	0.080	0.092	0.052	-0.015	0.052	-0.048	-0.061	0.043
20th	0.016	0.042	-0.005	-0.232	0.052	0.066	0.090	0.052	-0.031	0.052	-0.065	-0.082	0.023
25th	-0.006	-0.009	-0.023	-0.260	0.013	0.055	0.071	0.036	-0.034	0.048	-0.063	-0.087	0.011
30th	-0.023	-0.027	-0.037	-0.331	0.030	0.031	0.059	0.037	-0.050	0.020	-0.077	-0.114	0.018
35th	-0.060	-0.067	-0.069	-0.449	-0.028	0.031	0.037	0.015	-0.079	0.033	-0.097	-0.178	-0.019
40th	-0.131	-0.116	-0.124	-0.613	-0.087	-0.036	-0.023	0.009	-0.131	0.016	-0.155	-0.251	-0.057
45th	-0.190	-0.070	-0.177	-0.931	-0.119	-0.021	-0.009	0.017	-0.237	-0.012	-0.270	-0.436	-0.079
50th	-0.078	0.284	0.084	-0.788	0.351	0.099	0.320	0.107	-0.217	0.232	-0.211	-0.589	0.118
55th	0.254	0.524	0.338	-0.169	0.596	0.410	0.560	0.179	0.129	0.363	0.168	-0.135	0.409
60th	0.242	0.409	0.228	0.064	0.348	0.268	0.361	0.118	0.141	0.198	0.150	0.026	0.239
65th	0.204	0.287	0.164	0.107	0.257	0.187	0.300	0.121	0.133	0.180	0.131	0.064	0.183
70th	0.174	0.270	0.130	0.109	0.196	0.174	0.233	0.099	0.097	0.121	0.114	0.043	0.114
75th	0.157	0.256	0.123	0.079	0.169	0.159	0.211	0.093	0.083	0.121	0.089	0.045	0.087
80th	0.144	0.236	0.087	0.049	0.169	0.164	0.197	0.092	0.079	0.126	0.071	0.050	0.089
85th	0.130	0.227	0.089	0.046	0.166	0.138	0.171	0.085	0.064	0.097	0.063	0.035	0.077
90th	0.127	0.211	0.090	0.050	0.153	0.129	0.168	0.091	0.067	0.111	0.060	0.030	0.095
95th	0.112	0.157	0.082	0.024	0.139	0.142	0.163	0.084	0.040	0.116	0.044	-0.041	0.094

Note: Panel A of Table 8 reports the percentage differences between Q5 and FF5 of 196 pooled portfolios across 19 percentiles. 196P is 196 large pooled portfolios combining all subset portfolios within the panel. S-B/M is the portfolios double-sorted on size and book-to-market. S-INV is the portfolios double-sorted on size and investment. S-MOM is the portfolios double-sorted on size and momentum. SOP is the portfolios double-sorted on size and operation profitability. S-B/M-INV is the portfolios triple-sorted on size, book-to-market and investment. S-B/M-OP is the portfolios triple-sorted on size, book-to-market and operation profitability. S-OP-INV is the portfolios triple-sorted on size, operation profitability and investment. Panel B of Table 8 reports the percentage differences between Q5 and FF5 of 150 pooled portfolios across 19 percentiles. 150P is 150 large pooled portfolios combining all subset portfolios within the panel. FF is the portfolios related to the Fama and French factors. Anomaly is the portfolios related to other anomalies. Past is the portfolios related on past returns. Value is the portfolios related to valuation. We mark the percentage differences greater than 0.10 in pale pink, indicating that FF5 significantly outperforms Q5 economically, and those lower than -0.10 in pale yellow, indicating that Q5 is significantly better than FF5 economically.

Table 11: Percentage Difference between Q5 and FF6

	Panel A: 196 portfolios of Q5 and FF6								Panel B: 150 portfolios of Q5 and FF6				
Percentile	196P	S-B/M	S-INV	S-MOM	S-OP	S-B/M-INV	S-B/M-OP	S-OP-INV	150P	FF	Anomaly	Past	Value
5th	0.130	0.158	0.046	0.315	0.073	0.110	0.131	0.068	0.073	0.095	0.003	0.145	0.089
10th	0.104	0.124	0.027	0.221	0.070	0.100	0.101	0.079	0.045	0.080	-0.012	0.102	0.053
15th	0.088	0.091	0.003	0.200	0.058	0.088	0.098	0.053	0.041	0.066	-0.031	0.134	0.046
20th	0.081	0.048	0.007	0.178	0.066	0.077	0.095	0.063	0.032	0.064	-0.038	0.117	0.035
25th	0.065	0.018	-0.017	0.149	0.040	0.063	0.101	0.043	0.034	0.053	-0.027	0.116	0.034
30th	0.056	0.020	-0.035	0.133	0.040	0.043	0.090	0.045	0.027	0.030	-0.030	0.120	0.026
35th	0.037	-0.036	-0.044	0.116	-0.011	0.040	0.089	0.022	0.013	0.053	-0.055	0.108	-0.001
40th	0.002	-0.055	-0.097	0.099	-0.063	-0.007	0.037	0.008	-0.005	0.042	-0.083	0.108	-0.030
45th	0.014	-0.018	-0.138	0.096	-0.103	0.027	0.047	0.022	-0.038	0.021	-0.180	0.121	-0.027
50th	0.238	0.297	0.065	0.351	0.321	0.172	0.298	0.098	0.061	0.219	-0.069	0.137	0.142
55th	0.333	0.443	0.260	0.378	0.462	0.327	0.358	0.147	0.192	0.254	0.145	0.160	0.300
60th	0.247	0.351	0.193	0.325	0.298	0.224	0.255	0.109	0.134	0.160	0.099	0.135	0.178
65th	0.203	0.260	0.155	0.293	0.228	0.154	0.237	0.099	0.117	0.180	0.061	0.127	0.155
70th	0.182	0.242	0.122	0.292	0.188	0.149	0.201	0.086	0.084	0.124	0.056	0.100	0.083
75th	0.173	0.229	0.115	0.269	0.153	0.138	0.210	0.086	0.076	0.103	0.039	0.123	0.069
80th	0.159	0.221	0.080	0.259	0.144	0.149	0.169	0.094	0.077	0.115	0.027	0.131	0.072
85th	0.150	0.200	0.079	0.295	0.145	0.124	0.153	0.077	0.072	0.087	0.036	0.123	0.062
90th	0.154	0.198	0.089	0.310	0.149	0.112	0.159	0.076	0.084	0.101	0.036	0.145	0.084
95th	0.160	0.155	0.084	0.387	0.125	0.128	0.158	0.077	0.093	0.116	0.048	0.157	0.084

Note: Panel A of Table 8 reports the percentage differences between Q5 and FF6 of 196 pooled portfolios across 19 percentiles. 196P is 196 large pooled portfolios combining all subset portfolios within the panel. S-B/M is the portfolios double-sorted on size and book-to-market. S-INV is the portfolios double-sorted on size and investment. S-MOM is the portfolios double-sorted on size and momentum. S-OP is the portfolios double-sorted on size and operation profitability. S-B/M-INV is the portfolios triple-sorted on size, book-to-market and investment. S-B/M-OP is the portfolios triple-sorted on size, book-to-market and operation profitability. S-OP-INV is the portfolios triple-sorted on size, operation profitability and investment. Panel B of Table 8 reports the percentage differences between Q5 and FF6 of 150 pooled portfolios across 19 percentiles. 150P is 150 large pooled portfolios combining all subset portfolios within the panel. FF is the portfolios related to the Fama and French factors. Anomaly is the portfolios related to other anomalies. Past is the portfolios related on past returns. Value is the portfolios related to valuation. We mark the percentage differences greater than 0.10 in pale pink, indicating that FF6 significantly outperforms Q5 economically, and those lower than -0.10 in pale yellow, indicating that Q5 is significantly better than FF6 economically.

Table 12: Percentage Difference between BS and FF6

	Panel A: 196 portfolios of BS and FF6								Panel B: 150 portfolios of BS and FF6				
Percentile	**196P**	**S-B/M**	**S-INV**	**S-MOM**	**S-OP**	**S-B/M-INV**	**S-B/M-OP**	**S-OP-INV**	**150P**	**FF**	**Anomaly**	**Past**	**Value**
5th	0.076	0.122	0.033	0.035	0.087	0.089	0.078	0.070	0.054	0.083	0.031	0.044	0.079
10th	0.075	0.104	0.038	0.039	0.100	0.081	0.072	0.079	0.046	0.073	0.019	0.029	0.088
15th	0.078	0.107	0.051	0.038	0.089	0.082	0.094	0.062	0.041	0.088	0.011	0.029	0.069
20th	0.082	0.092	0.058	0.031	0.118	0.073	0.094	0.092	0.041	0.091	0.000	0.027	0.084
25th	0.087	0.102	0.046	0.035	0.124	0.093	0.104	0.077	0.047	0.104	0.003	0.029	0.096
30th	0.099	0.121	0.055	0.065	0.109	0.102	0.122	0.083	0.055	0.128	0.002	0.043	0.095
35th	0.109	0.114	0.060	0.076	0.130	0.115	0.144	0.083	0.051	0.135	-0.027	0.055	0.108
40th	0.154	0.191	0.093	0.089	0.142	0.177	0.206	0.104	0.061	0.180	-0.031	0.033	0.136
45th	0.225	0.276	0.120	0.085	0.220	0.309	0.285	0.140	0.087	0.257	-0.037	0.036	0.209
50th	0.198	0.206	0.009	-0.043	0.264	0.396	0.370	0.045	0.108	0.367	0.010	-0.064	0.338
55th	-0.061	-0.010	-0.053	-0.122	0.060	-0.019	-0.067	-0.179	-0.003	-0.024	0.069	-0.143	-0.006
60th	-0.062	-0.053	-0.058	-0.114	0.007	-0.063	-0.065	-0.066	-0.023	-0.037	0.043	-0.069	-0.123
65th	-0.042	-0.054	-0.042	-0.086	0.010	-0.041	-0.032	-0.056	-0.013	-0.016	0.019	-0.026	-0.064
70th	-0.016	-0.030	-0.008	-0.079	0.032	-0.040	0.011	-0.018	-0.014	-0.011	0.030	-0.036	-0.082
75th	-0.003	-0.023	-0.015	-0.059	0.021	-0.022	0.022	0.019	-0.008	-0.005	0.025	-0.017	-0.069
80th	0.004	-0.001	-0.021	-0.045	0.026	-0.008	0.014	0.034	-0.004	0.002	0.018	-0.011	-0.045
85th	0.010	0.013	-0.019	-0.039	0.028	-0.004	0.034	0.022	0.002	0.011	0.020	-0.015	-0.021
90th	0.020	0.024	0.007	-0.024	0.044	0.005	0.042	0.024	0.005	0.029	0.013	-0.018	-0.002
95th	0.026	0.032	0.020	-0.004	0.047	0.006	0.036	0.039	0.017	0.035	0.030	-0.004	0.000

Note: Panel A of Table 8 reports the percentage differences between BS and FF6 of 196 pooled portfolios across 19 percentiles. 196P is 196 large pooled portfolios combining all subset portfolios within the panel. S-B/M is the portfolios double-sorted on size and book-to-market. S-INV is the portfolios double-sorted on size and investment. S-MOM is the portfolios double-sorted on size and momentum. S-OP is the portfolios double-sorted on size and operation profitability. S-B/M-INV is the portfolios triple-sorted on size, book-to-market and investment. S-B/M-OP is the portfolios triple-sorted on size, book-to-market and operation profitability. S-OP-INV is the portfolios triple-sorted on size, operation profitability and investment. Panel B of Table 8 reports the percentage differences between BS and FF6 of 150 pooled portfolios across 19 percentiles. 150P is 150 large pooled portfolios combining all subset portfolios within the panel. FF is the portfolios related to the Fama and French factors. Anomaly is the portfolios related to other anomalies. Past is the portfolios related on past returns. Value is the portfolios related to valuation. We mark the percentage differences greater than 0.10 in pale pink, indicating that FF6 significantly outperforms BS economically, and those lower than -0.10 in pale yellow, indicating that BS is significantly better than FF6 economically.

Table 13: Percentage Difference between FF5 and FF6

	Panel A: 196 portfolios of FF5 and FF6								Panel B: 150 portfolios of FF5 and FF6				
Percentile	**196P**	**S-B/M**	**S-INV**	**S-MOM**	**S-OP**	**S-B/M-INV**	**S-B/M-OP**	**S-OP-INV**	**150P**	**FF**	**Anomaly**	**Past**	**Value**
5th	0.063	0.026	0.000	0.437	0.014	-0.003	0.012	-0.001	0.060	0.015	0.010	0.207	0.008
10th	0.062	0.011	0.005	0.463	0.011	0.000	0.004	0.000	0.052	0.002	0.009	0.196	0.001
15th	0.062	0.020	0.011	0.438	0.018	0.007	0.005	0.001	0.057	0.014	0.017	0.202	0.003
20th	0.064	0.005	0.012	0.451	0.013	0.010	0.004	0.010	0.064	0.011	0.026	0.208	0.011
25th	0.071	0.027	0.006	0.448	0.027	0.008	0.028	0.006	0.069	0.004	0.035	0.213	0.023
30th	0.080	0.047	0.002	0.508	0.010	0.012	0.029	0.008	0.078	0.009	0.046	0.247	0.007
35th	0.099	0.031	0.024	0.617	0.017	0.008	0.049	0.007	0.093	0.020	0.040	0.305	0.018
40th	0.134	0.058	0.025	0.773	0.023	0.029	0.061	-0.002	0.125	0.026	0.066	0.386	0.026
45th	0.207	0.051	0.035	1.116	0.015	0.049	0.057	0.005	0.191	0.033	0.077	0.610	0.051
50th	0.335	0.010	-0.018	1.416	-0.023	0.066	-0.017	-0.008	0.292	-0.011	0.134	0.806	0.021
55th	0.063	-0.057	-0.062	0.610	-0.092	-0.063	-0.149	-0.028	0.057	-0.087	-0.020	0.316	-0.084
60th	0.004	-0.043	-0.029	0.246	-0.038	-0.036	-0.084	-0.007	-0.007	-0.033	-0.046	0.106	-0.051
65th	-0.001	-0.022	-0.007	0.167	-0.023	-0.028	-0.051	-0.020	-0.015	0.000	-0.067	0.058	-0.024
70th	0.006	-0.022	-0.008	0.166	-0.007	-0.022	-0.027	-0.012	-0.012	0.003	-0.055	0.055	-0.029
75th	0.014	-0.023	-0.007	0.176	-0.014	-0.018	-0.001	-0.006	-0.006	-0.017	-0.048	0.074	-0.017
80th	0.013	-0.013	-0.006	0.200	-0.022	-0.013	-0.024	0.002	-0.002	-0.009	-0.043	0.077	-0.016
85th	0.018	-0.023	-0.009	0.239	-0.018	-0.013	-0.016	-0.007	0.007	-0.010	-0.026	0.085	-0.014
90th	0.024	-0.011	-0.002	0.248	-0.003	-0.016	-0.008	-0.014	0.016	-0.009	-0.023	0.112	-0.010
95th	0.043	-0.001	0.002	0.354	-0.012	-0.013	-0.004	-0.007	0.051	0.001	0.003	0.204	-0.009

Note: Panel A of Table 8 reports the percentage differences between FF5 and FF6 of 196 pooled portfolios across 19 percentiles. 196P is 196 large pooled portfolios combining all subset portfolios within the panel. S-B/M is the portfolios double-sorted on size and book-to-market. S-INV is the portfolios double-sorted on size and investment. S-MOM is the portfolios double-sorted on size and momentum. S-OP is the portfolios double-sorted on size and operation profitability. S-B/M-INV is the portfolios triple-sorted on size, book-to-market and investment. S-B/M-OP is the portfolios triple-sorted on size, book-to-market and operation profitability. S-OP-INV is the portfolios triple-sorted on size, operation profitability and investment. Panel B of Table 8 reports the percentage differences between FF5 and FF6 of 150 pooled portfolios across 19 percentiles. 150P is 150 large pooled portfolios combining all subset portfolios within the panel. FF is the portfolios related to the Fama and French factors. Anomaly is the portfolios related to other anomalies. Past is the portfolios related on past returns. Value is the portfolios related to valuation. We mark the percentage differences greater than 0.10 in pale pink, indicating that FF6 significantly outperforms FF5 economically, and those lower than -0.10 in pale yellow, indicating that FF5 is significantly better than FF6 economically.

Table 14: Percentage Difference between FF6-hml and FF6

	Panel A: 196 portfolios of FF6-hml and FF6								Panel B: 150 portfolios of FF6-hml and FF6				
Percentile	**196P**	**S-B/M**	**S-INV**	**S-MOM**	**S-OP**	**S-B/M-INV**	**S-B/M-OP**	**S-OP-INV**	**150P**	**FF**	**Anomaly**	**Past**	**Value**
5th	0.071	0.159	0.013	0.021	0.027	0.113	0.087	0.030	0.031	0.068	0.008	0.011	0.063
10th	0.061	0.127	0.013	0.007	0.022	0.102	0.081	0.016	0.016	0.043	0.000	-0.004	0.042
15th	0.057	0.113	0.014	-0.006	0.011	0.087	0.094	0.016	0.021	0.051	0.007	0.002	0.038
20th	0.054	0.090	0.009	-0.003	0.010	0.082	0.100	0.005	0.014	0.030	0.003	-0.005	0.041
25th	0.044	0.071	-0.004	-0.005	0.005	0.054	0.096	0.005	0.015	0.024	0.007	-0.007	0.044
30th	0.035	0.056	0.002	-0.015	0.001	0.055	0.074	0.005	0.012	0.015	0.016	-0.005	0.020
35th	0.035	0.033	-0.018	-0.026	-0.009	0.052	0.106	-0.005	0.005	0.019	0.003	0.000	0.002
40th	0.015	0.035	-0.031	-0.049	-0.014	0.019	0.071	-0.014	0.003	0.015	0.018	-0.018	-0.017
45th	0.000	0.039	0.001	-0.035	-0.024	0.002	0.005	-0.011	-0.010	-0.006	0.012	-0.056	-0.025
50th	0.083	0.239	-0.044	0.000	-0.016	0.070	0.149	0.088	-0.006	0.050	-0.002	-0.048	-0.008
55th	0.156	0.290	0.014	0.067	0.093	0.171	0.275	0.049	0.072	0.123	0.022	0.046	0.165
60th	0.132	0.233	0.019	0.044	0.029	0.192	0.228	0.032	0.057	0.083	0.001	0.031	0.167
65th	0.103	0.183	0.040	0.026	0.041	0.152	0.167	0.022	0.049	0.088	-0.014	0.026	0.158
70th	0.097	0.162	0.033	0.029	0.046	0.132	0.159	0.028	0.042	0.087	-0.003	0.012	0.112
75th	0.083	0.166	0.039	0.019	0.031	0.116	0.125	0.024	0.036	0.073	0.002	0.006	0.096
80th	0.073	0.163	0.025	0.010	0.017	0.098	0.111	0.027	0.035	0.074	-0.002	0.015	0.087
85th	0.071	0.164	0.031	0.018	0.021	0.101	0.100	0.019	0.032	0.063	0.007	0.005	0.081
90th	0.072	0.168	0.028	0.002	0.014	0.108	0.107	0.019	0.032	0.076	0.004	0.002	0.080
95th	0.066	0.132	0.020	0.021	0.016	0.123	0.082	0.018	0.037	0.065	0.014	0.007	0.090

Note: Panel A of Table 8 reports the percentage differences between FF6-hml and FF6 of 196 pooled portfolios across 19 percentiles. 196P is 196 large pooled portfolios combining all subset portfolios within the panel. S-B/M is the portfolios double-sorted on size and book-to-market. S-INV is the portfolios double-sorted on size and investment. S-MOM is the portfolios double-sorted on size and momentum. SOP is the portfolios double-sorted on size and operation profitability. S-B/M-INV is the portfolios triple-sorted on size, book-to-market and investment. S-B/M-OP is the portfolios triple-sorted on size, book-to-market and operation profitability. S-OP-INV is the portfolios triple-sorted on size, operation profitability and investment. Panel B of Table 8 reports the percentage differences between FF6-hml and FF6 of 150 pooled portfolios across 19 percentiles. 150P is 150 large pooled portfolios combining all subset portfolios within the panel. FF is the portfolios related to the Fama and French factors. Anomaly is the portfolios related to other anomalies. Past is the portfolios related on past returns. Value is the portfolios related to valuation. We mark the percentage differences greater than 0.10 in pale pink, indicating that FF6 significantly outperforms FF6-hml economically, and those lower than -0.10 in pale yellow, indicating that FF6-hml is significantly better than FF6 economically.