No. 1504 [EN]

IMI Working Paper

Risk-Adjusted Performance of Mutual Funds: Evidence from China

Gang Jianhua and Qian Zongxin

INTERNATIONAL MONETARY INSTITUTE

IMI Working Paper No. 1504 [EN]

Risk-Adjusted Performance of Mutual Funds: Evidence from China

By* GANG JIANHUA* and QIAN ZONGXIN*

April 2015

Abstract

In this paper, we evaluate the performance of individual mutual fund listed in China between 2006 and 2014. We build an indicator more consistent with investors' rationality to track funds' performance. More specifically, we firstly estimate the time-varying abnormal returns of each China's mutual fund by introducing an additional factor of active peer benchmark. An index of riskiness is then estimated and used to calculate the augmented performance measure (APM). The APM therefore addresses investors' preference towards managerial premium of a certain fund and their aversion to the tail risk. Empirical evidence shows that the APM incorporates information beyond the first and second moments of the distribution of fund returns, therefore it encompasses better fund-choosing decisions as compared with Sharpe ratio and the economic performance measure.

* Research Fellow of IMI, School of Finance, Renmin University of China.
**Research Fellow of IMI, School of Finance, Renmin University of China. Corresponding author, tel: 00861082500617; email: qianzx@ruc.edu.cn.

1. Introduction

Performance of mutual funds can be evaluated by comparing their historical abnormal return series, which are defined as excess fund returns excluding the compensations for various risk factors. Conditional on their systematic risk exposure, indicators as such are well-documented and designed to measure funds' ability to generate profits over safe assets. Therefore, investors may be attracted by funds with “superior" performance than other ones. Literature on the factor-model family to estimate the abnormal returns is rich, both theoretically and empirically. In his seminal study, Jensen (1968) uses a single factor model to estimate the abnormal return series. Jensen's single factor model is then quickly extended to include more factors in order to capture more explanatory variables to accurately explain the excess return. Fama and French (1993) introduces a three-factor model and Carhart (1997) adds a measurement of momentum as a fourth factor to pin-point the effect of persistence. Recent contribution (Ferson and Schadt, 1996; Christopherson et al., 1998; Avramov and Wermers, 2006; Mamaysky et al., 2008) allows time variations in the estimated abnormal returns. Hunter et al. (2014) proposes an active peer benchmark (APB)-augmented factor model, which accounts for commonalities in mutual fund strategies. Empirical evidence shows that this augmentation substantially reduce the correlation of residuals between categorized individual funds, which makes the alpha keen to reflect out-performance of management skill within fund categories.

Compared with the standard Sharpe ratio, the advantage of the abnormal return-based indicator acknowledges that mutual funds are not only affected by management skills of the funds but also affected by changes in the systematic risk. Funds' excess returns used to calculate standard Sharpe ratios do not exclude systematic risk compensations, and therefore, may bias the evaluation results. The idea of Sharpe ratio to formulate a risk-adjusted return while implying normality is reasonably consistent with individual rationality. The reason is simple: rational and risk-averse investors would tend to appreciate one unit of excess profit only if it would not incur too much excess volatility. Investors therefore will be indifferent towards different investments with identical return-risk ratio. The Sharpe ratio is then a typical implementation of the tradeoff between fund's returns and the corresponding volatility. It uses the standard deviation of excess returns of a given fund's as an indicator of risk.

However, the standard deviation being a proxy of riskiness has been widely criticized, because one of the prerequisites for it to be reasonable is the assumption of normality. As is well known, non-Gaussian properties such as non-zero skewness or excess kurtosis are the norms of the financial market yet completely discarded by lower moments. Therefore, investment decision based only on the lower moments would be misleading and expose investors to tail risk. Realizing this, Homm and Pigorsch (2012) formulates an economic performance measure (EPM) which includes the information contained in higher moments as well as in lower ones. Specifically, the denominator of the Sharpe ratio is replaced by the Aumann and Serrano (2008) economic index (the AS index) of riskiness. Despite its analytical elegance, the

Homm and Pigorsch (2012) EPM, as a single indicator of the refined risk-adjusted excess return by construction, fails to tell investors which fund may be chosen because of its superior management skills. Rather, a higher EPM can objectively tell a better balance of all four moments in the return series of a given fund, and vice versa. In other words, the EPM calculation is based on the excess fund returns as its nominator and the AS index as its denominator. Hence, although the EPM addresses the importance of higher moments of the fund's returns across time, it does not discriminate management premium from systematic risk compensations.

In this paper, we combine the APB-augmented four factor model of Hunter et al. (2014) with the Homm and Pigorsch (2012) EPM. The essential idea is to distill the superior management premium from the abnormal returns of funds whilst addressing the non-normality of the raw return distribution. More specifically, our model can be demonstrated as a two-step procedure: First, we estimate the abnormal fund returns using the APB-augmented four factor model; Second, we estimate the AS index of riskiness and then calculate the Homm and Pigorsch (2012) EPM using the estimated abnormal returns from the first step. In this way, we manage to separate the management skill premium of funds from systematic risk compensations and, in the meantime, normalize the premium by the corresponding risk in terms of weighted average of all four moments. The resulting per risk return is more likely to capture the contributions of active fund managers.

We apply our model to evaluate the performance of all funds that are publicly listed in China from 2006 to 2014. Existing literature on China's financial markets has focused on stock market reforms and their impacts (Firth et al., 2010; Li et al., 2011; Liao et al., 2011, 2014; Liu et al., 2014); asset pricing (Yao and Luo, 2009; Xiong and Yu, 2011); and spillover effects across sectors or markets in/between domestic China or/and outside China (see Chen et al., 2010; Moon and Yu, 2010; Qian and Luo, 2014).[1] However, very few studies focus on the performance of China-listed funds. One most relevant paper by Xu (2005) studies the investment funds' daily performance from 2000 to 2004 using a single-factor model. the contribution of our paper is then four-fold: Firstly, we focus on most China-listed funds' abnormal returns rather than the excess returns in order to pin-point the best ones because of their superior management skills; secondly, compared with the single-factor model, the four-factor model controls potential bias caused by omitted variables; thirdly, active peer benchmarks are then introduced in the baseline four-factor model, which reduces residual commonalities arising from similar trading strategies among fund managers. Therefore, management skill premium would accurately emerge as the APB-adjusted alpha; finally, we implement the EPM measurement which takes into account skew and fat-tail behavior in the fund's abnormal return series which is even more consistent with individual rationality.

The paper proceeds as follows: Section 2 introduces the background information on the development of China's mutual funds industry; Section 3 introduces our methodology; Section 4 describes the data-set; Section 5 presents and discusses the empirical results; Section 6 concludes the paper.

[1] See Xu (2005) for a survey of relevant literature.

2. The development of the mutual fund industry in China

China had the very first investment fund in 1991. Throughout the whole 1990's, China's investment fund industry was featured by the closed-end, trust-like investment funds and the lack of necessary regulation. In these early days, the number of China's closed-end funds (the "old funds") was small and the assets under management were insignificant. In late 1997, the State Council of China issued the Interim Measures for the Administration of Securities Investment Funds, which came into effect in 1998. Newly-established funds (the "new funds") that were of investment purpose must yield to regulations of this interim policy, and that all regulations with regard to investment funds must be carried out by the China Securities Regulatory Commission (the CSRC). But again, only the closed-end funds were allowed by then and operations of these funds are subject to many conservative regulatory constraints[2]. Not until 2001 was the first open-end fund introduced and publicly traded, which is also known as the first mutual fund in China. After 2002, China's mutual fund industry started to boom[3], especially after the launch of the Securities Investment Funds Law in June, 2004, and mutual funds quickly dominated China's fund market. Recent statistics from the CSRC show that there are 1,897 funds valued at 731.5 billion US dollars that are publicly listed by the end of 2014, among which 1,763 are mutual funds valued at 709.5 billion US dollars, accounting for 97 per cent of the aggregate market value of all public funds.[4]

3. Methodology

3.1 A two-stage augmented performance measure

The augmented performance measure (APM) that we derive generalizes the Hunter et al. (2014) APB-augmented four-factor model in two distinct dimensions: First, our model sets all parameters to be time-varying in order to satisfy optimal dynamic strategies. Second, the APM is capable to track low- and high-order risk. Specifically, our model not only eliminates common component (due to common strategies among fund managers) from the residuals regarding cross correlation but also evaluates risk element within abnormal returns (proxied by the APB-adjusted alphas). The APM model then has two stages as follows,

Stage 1: We generalize the Carhart (1997) four-factor model into a setting which allows parameters to be time-varying. This four-factor model is then used as the baseline model. We then follow the paper by Hunter et al. (2014) to include an APB factor in the baseline model. In this stage, we focus on the time-varying Jensen's alphas in the augmented-APB model.

Stage 2: We then introduce a risk gauge, the augmented performance mea-sure (APM), to further adjust Jensen's alphas derived from stage 1. We implement the calculation procedure of the economic performance measure (EPM) by Homm and

[2] See Xu (2005) for details.

[3] In 2002, China's first fund specialized in the bond market was established and China's first money market fund was introduced in 2003.

[4] Before 2006, institutional investors occupied over 50 per cent of mutual fund holdings, while individuals held more than 80 percent by the end of 2010.

Pigorsch (2012), but replace the excess return series by the time-varying alphas. In contrast to the Sharpe ratio, the APM divides the "mean excess alphas" by its AS index instead of dividing by its standard deviation.

In the following three subsections, we break down the two-stage procedure and statistical features of the APM.

3.2 The APB-augmented factor model

Methodology to evaluate fund's performance varies from a simple Sharpe ratio comparison to a family of mutiple factor models, among which Carhart (1997) revises the Fama and French (1993) model by introducing a fourth factor to track momentum. Recent literature generalizes the settings of alpha and beta to be time-varying either with exogenous economic variables or Kalman filters. Another thread of literature addresses the problem of similar strategies used by fund managers, so that it is difficult to tell the star mangers from the crowd due to correlated residuals. Recent empirical evidence reiterates common practices of liquidity preference and momentum may cloud the real causes of funds' superior performance. Our method follows Hunter et al. (2014) to include an active peer benchmark (APB) in the conventional four-factor model. Ac-cording to Hunter et al. (2014), the APB then corrects the commonality bias by decomposing Jensen's alpha into some systematic component and the skill premium. Therefore, an econometric meaning of the APB is to stand as an additional independent variable to control common, unpriced idiosyncratic risks taken by mutual funds, while the practical idea to include an APB resides in the endogenous selection of a certain fund within its category, in addition to the exogenously determined factors in the standard regressions estimating the fund loadings and Jensen's alpha.

In this study, we firstly generalize the static Carhart (1997) four-factor model into a setting which allows parameters to be time-varying. This TVP four-factor model is then used as our baseline model. The model is as follows,

$$\begin{aligned} r_{i,t} &= \alpha_{i,t} + \beta_{i,t,rmrf} r_{rmrf,t} + \beta_{i,t,smb} r_{smb,t} + \beta_{i,t,hml} r_{hml,t} \\ &\quad + \beta_{i,t,umd} r_{umd,t} + e_{i,t} \end{aligned} \tag{1}$$

where, ri,t is the fund i0s monthly NAV return minus the three-month treasury bill rate, and rrmrf,t, rsmb,t, rhml,t, and rumd,t are the excess return on the market value-weighted portfolio [5], returns of size, book-to-market, and momentum [6]. Independent variables in terms of returns in model (1) are based on monthly time series. It is important to note that all coefficients in equation (1), αi,t, βi,t,rmrf , βi,t,smb, βi,t,hml, and βi,t,umd, are time-varying parameters (TVP) in certain predetermined rolling windows. This TVP setting is desirable and meaningful because

[5] China (excluding Hong Kong) has two stock exchanges: the Shanghai and Shengzhen stock markets and no stock in China can be cross-listed simultaneously in both markets. Because of the capital control, China also has two categories of stocks targeting different types of investors: the A-share for Chinese citizens and B-share for non-citizens to trade (insignificant in terms of trading volume). Here, the market value-weighted portfolio we use consists of listed stocks both in Shanghai and Shenzhen exchanges but only includes the A-shares.

[6] Data for momentum of China's stocks comes from the Resset Database, http://www.resset.cn/en/.

dynamic strategies would certainly prefer funds with competitive abnormal returns (positive Jensen's alpha or αi,t in model (1) and statistically significant) and the stability of it.

We then follow the paper by Hunter et al. (2014) to include an APB factor in our baseline model,

$$\begin{aligned} r_{i,t} &= \alpha_{i,t} + \beta_{i,t,rmrf} r_{rmrf,t} + \beta_{i,t,smb} r_{smb,t} + \beta_{i,t,hml} r_{hml,t} \\ &+ \beta_{i,t,umd} r_{umd,t} + \lambda_{i,t} \varepsilon_{APB_i,t} + e_{i,t}. \end{aligned} \quad (2)$$

And again, TVPs are assumed in the augmented model (2). According to Hunter et al. (2014), model (2) suggests if a fund manager's superior performance is truly unique and uncorrelated with his active peer group's average, then the alpha is identical as in the four-factor model (1). Otherwise the i,t should yield to adjustment by the term i,tAPBi,t. In many cases, the existence of i,tAPBi,t would even render Jensen's alpha insignificant, because fund managers may follow similar profitable strategies. Therefore, model (2) eliminates commonalities in idiosyncratic risk-taking by funds in the same active peer benchmark group. As a result, it improves the estimation efficiency of the standard four-factor model.

However, the APB-adjusted alpha is a measure of absolute abnormal return. Investors' income fluctuation risk associated with time variations in this setting is neglected. It could be the case that fund managers take on excess risk in order to offer competitive abnormal return. Hence, focusing on the value of APB-adjusted alpha may still incomplete even though it is reliable to reflect skill premium. Rational investors require reasonable risk levels in terms of low and high orders. A risk-adjusted performance indicator is then necessary to account for the additional risk associated with potential aggressiveness of fund managers.

3.3 APM as a measurement of riskiness

There are several ways to interpret the word “risk” in finance. Rothschild and Stiglitz (1970, 1971) considers increases in risks to be defined in terms of dynamic innovations in the probability density functions. Objective riskiness as such is absolute and risk-neutral. However, agents in the financial market exhibit various subjective preferences towards risk-return relationship. Therefore, Diamond and Stiglitz (1974) generalizes the concept of riskiness by revealing an element of asymmetry that jitters an agent who is risk averse significantly more than if he is otherwise less risk averse. Diamond and Stiglitz (1974) defines the riskiness of a gamble based on two distinct considerations: (i) riskiness of the gamble in absolute terms; and (ii) how risk averse an agent truly is. Hence, a single index capable of addressing both considerations is valuable. Aumann and Serrano (2008) quantifies the riskiness (the AS index) based on the risk aversion. The index is positively homogeneous, continuous, and sub-additive and respects first- and second-order stochastic dominance. The riskiness suggests duality to risk aversion to which a risk-averter is averse. So on the whole, the index reflects the following natural notion of less risky: given that an investment is accepted by some agent, less risk-averse individuals accept riskier investments (Aumann and Serrano, 2008; Homm and Pigorsch, 2012).

The APM is a single index based on the dynamics of Jensen's alpha and the AS index of the alphas, such that it correctly represents skill premium as per risk. While Homm and Pigorsch (2012) directly use the excess fund returns to build the EPM of funds, our approach uses the abnormal fund returns based on the APB-augmented four-factor model which excludes the systematic risk compensations from the excess returns. The APM has the following form,

$$APM(\alpha^*_{i,t}) = \frac{E\left(\alpha^*_{i,t}\right)}{AS(\alpha^*_{i,t})} \tag{3}$$

where AS(:) stands for the AS index (see Aumann and Serrano (2008)). Thus, in contrast to the Sharpe ratio, the APM divides the \mean excess alphas" by its AS index. According to Homm and Pigorsch (2012), the APM can be constructed with various settings: normal distribution assumption (equivalent to the Sharpe ratio), inverse normal distribution (NIG), and the nonparametric procedures. In order to make our results practically applicable (minimizing sample selection bias) and also distinct from the Sharpe ratio, we focus on the NIG calculation. Therefore, the representation of the AS index and of the APM in terms of the moments is given by:

$$\begin{aligned} \widetilde{AS}^{(NIG)}\left(\mu,\sigma^2,\chi,\kappa\right) &= \frac{\left(3\kappa\mu - 4\mu\chi^2 - 6\chi\sigma + \frac{9\sigma^2}{\mu}\right)}{18} \\ \widetilde{APM}^{(NIG)}\left(\mu,\sigma^2,\chi,\kappa\right) &= \frac{18\mu}{\left(3\kappa\mu - 4\mu\chi^2 - 6\chi\sigma + \frac{9\sigma^2}{\mu}\right)} \end{aligned} \tag{4}$$

where μ, σ^2, χ, and κ stands for the first four moments. They all obey conditions that are assumed in Aumann and Serrano (2008).

3.4 Properties of the augmented performance measure

Most properties of the APM can be easily inferred from the properties of the AS index. First of all, because both the numerator and the denominator of the APM as in (3) are homogeneous, so the APM is scale invariant. And because the APM adopts the strictly monotone index by Aumann and Serrano (2008) as its denominator, therefore the APM can be proved to be also strictly first- and second-order monotonic. It is similar to the proof of EPM by Homm and Pigorsch (2012). Provided that the AS index is monotonic with respect to the first- and second-order, and, without generality, assume that first- and second-order dominates, then we have the AS index to be first-(second-) order monotonic AS() < AS().[7] We also know that the mean is monotonic with respect to stochastic dominance, therefore E()E(). It then immediately follows that the APM()APM(), which can be interpreted as the APM itself has the first- (second-) order monotonicity with respect to stochastic dominance. We then follow the generalized continuity properties proved by Homm and Pigorsch (2012) to see if the APM can do well in terms of continuity. The APM clearly satisfies the following assumptions:

Assumption 1. The economic index of riskiness $AS(\alpha_t)$ exists for all $n \geq 0$.

[7] See Aumann and Serrano (2008) for the proof of AS index to be of monotonicity with respect to stochastic dominance.

Assumption 2. There exists a real number $b > \frac{1}{AS(\alpha_t)}$, such that $\sup_n M_n(-b) < \infty$:

Homm and Pigorsch (2012) proves that given the above assumptions, the generalized continuity holds as follows[8],

Proposition 1. (Generalized Continuity). If Assumption 1 and 2 hold, then $\alpha_t \xrightarrow{d} \alpha_0$ implies $AS(\alpha_t) \rightarrow AS(\alpha_0)$.

Therefore, if we further assume that the time series of Jensen's alphas, $\{\alpha_t\}_{t \geq 1}$, to be uniformly integrable, then we have the following results with regards to generalized continuity for our APM,

Corollary 1. (Generalized Continuity for APM). If Assumption 1, 2 hold, then $\alpha_t \xrightarrow{d} \alpha_0$ implies $APM(\alpha_t) \rightarrow APM(\alpha_0)$.

Just as the EPM, the APM approximates the APM of normally distributed returns as the sampling frequency decreases. While the Sharpe ratio is appropriate for low frequency returns, the APM is appropriate for both low and high frequency returns, with no disadvantages compared with the Sharpe ratio in the former case.

There is one more argument about why the APM differs from the Sharpe ratio and the EPM. As criticized by many works, the mean-variance decision framework (the Sharpe ratio) typically ignore either higher moments which matters in the asymmetry of return distribution (at least in finite sample and low frequency data) or much higher probability of extreme events in practice. Therefore, the similarity of APM and EPM resides in the fact that both index consider skewness and kurtosis in addition to location and scale. However, APM and EPM are also very different in terms of reflecting managers' skill premium. The EPM does include this premium implicitly, but as an unobservable component within the excess return, the EPM also includes systematic compensation which shed difficulties to separate the premium. Therefore, investment decisions based on the EPM may exhibit superior features compared with the mean-variance measures but are still incomplete. The reason is simple: investors would typically prefer better managed fund with reasonable risk profile. Hence, the APM that we calculate in this section satisfies both objectives by considering the APB-augmented Jensen's alpha as per risk[9] in general term.

4. Data

According to Mamaysky et al. (2008), we use monthly data rather than the daily data for fund performance to eliminate micro structural problems[10]. The sample starts from 2006 until September of 2014. Our data-set includes official data of China's three-month treasury bill rate, which is only available from China Central Depository and Clearing Co., Ltd. (CCDC) ever since 2006[11]. Time series of dividend-adjusted

---

[8] See Homm and Pigorsch (2012) for very detailed proof.

[9] The risk here refers to as the economic index of riskiness (the AS index, Aumann and Serrano (2008)).

[10] More detailed discussion about this can be found in Mamaysky et al. (2008).

[11] There are several unofficial sources that have quite different quotes of the China's T-bill rates, but this paper prefers the official series that reflects the whole market in aggregate terms.

NAV (net asset values) returns of China's mutual funds is from the Wind[12] database, which classifies China's mutual funds into nine categories (or groups) in terms of their strategic characteristics: the large-cap value (LV), large-cap balance (LB), large-cap growth (LG), mid-cap value (MV), mid-cap balance (MB), mid-cap growth (MG), small-cap value (SV), small-cap balance (SB), and small-cap growth (SG). The classification of these types is similar to the one in Hunter et al. (2014) for the US funds. Because the APB-augmented model requires cross-sectional data-sets to be large enough to have statistical meanings, therefore, each category in our sample contains at least 30 funds. In addition, only the active funds that have been listed for at least 36 months are chosen. According to our criteria as stated, only the cluster that includes LV, LB, LG, and MG forms a satisfactory data-set. Hence, yielding to some data limitations as above, our analysis focuses on the funds that fall into these four categories. There are 41 funds in the LV, 152 funds in the LB; 224 funds in the LG, and 87 funds in the MG group.

5. Empirical Results

5.1 Group distribution of estimated abnormal returns

Empirical Results are demonstrated through Table 1 to Table 7. Table 1 summarizes distributional properties of the estimated average abnormal returns of individual funds within each fund group. It suggests that, on average, funds in the LV group produce higher abnormal returns than the funds in other groups. Differences in abnormal returns across individual funds are also smaller in the LV group. The group of MG funds has the largest variation of 3.32 percent compared with the other three groups. Normality tests (the last column in Table 1) of the returns are all rejected and highly significant in all groups, which suggests the decision-making based on the mean-variance approach will be misleading at least for finite samples. Test results in Table 1 also imply the non-normality is partially caused by dramatic skewness (positive or negative). Therefore, the violation of normality due to higher-order moments leads to the necessity of ranking individual funds according to a much general risk-return trade-off.

5.2 Fund ranking

Table 2 to 5 summarize the rankings of individual funds (the top 20) within each group according to their augmented performance measures (the APM) that we have built in section 3. To facilitate comparisons of the results, we assign exclusive ticker to each fund for identification. Our ticker system has a general format of "the group ticker" plus "individual fund ID".[13] Table 2 to 5 report rankings based on three different criteria: the APM, Sharpe ratio, and average abnormal return (the APB-augmented alpha or the APB-). In general, risk-averse investors may not want to pick a certain fund only according to its dominating APB-, because after adjusting its

[12] http://www.wind.com.cn/En/Default.aspx

[13] Due to the length limit, we are unable to provide all fund names as well as their tickers anywhere in this paper. An appendix which lists the Wind codes, names of individual funds and their corresponding tickers can be obtained upon request from the author.

accompanying risk (low-and high-order moments) this particular fund does not seem attractive in the aspect of superior management. For example, the one with the highest APB-in the large-cap value (LV) group (see Table 2, fund ID: 37) only ranks the thirteenth after risk-adjustment based on Sharpe ratio and APM. Likewise, the top fund under APB in the large-cap balance (LB) group (see Table 3, fund ID: 181) ranks the twenty-eighth under Sharpe ratio, but the third under the APM. Therefore, results show that the skill premium itself is volatile, and some funds may trump their peers only accidentally but in no sense sustainably. It is also true that some funds exhibit heavy volatility in low order but rarely in higher order. Similarly, the large-cap growth (LG) group as in Table 4 puts the best APB-fund (ID: 204) into a place outside top-ten under the Sharpe ratio, but its performance is still preferable, though not the best, if we put weights on superior higher order moments of its return distribution. The ranking of the top-one fund by APB-in the mid-cap growth funds (see Table 5, ID: 25) plummets after risk adjustments.

Table 6 reports the correlations between rankings by different performance indicators. As is clear from this table, the three performance indicators generate systematically different fund rankings. Results in Table 6 suggest that a rational and risk averse investor most likely would choose a set of completely different funds than an investor who is risk neutral. Table 6 therefore implies high-order moments of the abnormal return distribution, such as skewness and kurtosis, are non-negligible in calculating investment risk to investors who are risk averse. Fund selection under the Sharpe ratio may expose investors to excessive risk as suggested by exotic skewness and excess kurtosis, which generate welfare loss.

5.3 Performance comparison across groups

Table 7 summarizes the distributional properties of the APMs in different groups. On average, funds belonging to the LV category perform much better than other groups in terms of the APM (3.65 in LV versus 0.61 in LB, 0.47 in LG and 1.00 in MG). The LV fund group also has the highest median performance of 3.28 (medians of 0.16, 0.29 and 1.07 in the LB, LG and MG, respectively). However, dispersion of the performances in the LV group is also substantial. Results as such suggest that categorizing mutual funds solely based on their strategic characteristics (value, balance or growth) does not apply in China. Table 7 also gives the normality tests (the fifth column), which are all highly significant and therefore reject the null of Gaussian distribution. The pairwise Wilcoxon signed-rank tests are also provided in the last column of Table 7, which intend to test the null hypothesis of zero difference between the observed signed-rank medians and the zero signed-rank median. And results returned show the group medians are significantly different from zero. Hence, the median funds in all four groups generate positive utility gains for risk-averse fund investors.

6. Conclusion

We study the economic performance of mutual funds in China over the sample period from 2006 to 2014 using monthly data. In order to separate fund managers' skill premium from systematic risk compensations, we implement the active peer benchmark (APB) augmented four factor model. The APB-augmented model is then used as the baseline model. Our approach is a two-stage procedure which is designed to firstly (stage one) estimate the APB-augmented abnormal returns of individual funds. A new feature is then introduced (stage two): we consider the general distributional properties in terms of high-order moments that the abnormal returns may exhibit and calculate the augmented performance measure (the APM). The APM is a measure that takes into account the risk (revealed by low- and high-order moments) associated with time variations in abnormal returns of individual funds. The APM therefore differs from the Sharpe ratio and the EPM. It satisfies risk averse investors' needs of preferring some fund which delivers better returns due to superior managerial skills and has reasonable risk profile at the same time. In essence, the APM can achieve both objectives by considering the APB-augmented Jensen's alpha as per risk in a general term. Using a sample from China's mutual fund market, we are able to show that fund selections based on the rankings of average abnormal returns may not be sufficient for risk-averse investors. In principal, choosing a fund that has some modest return but less risk may improve investors' welfare. Empirical evidence in this paper further confirms that skewness and kurtosis contain valuable information for the true riskiness of funds and therefore should not be neglected. Hence, investment decisions based on mean-variance measures are sometimes misleading. Finally, although mutual funds' performance in China is diverse, funds with median performance of each group are still producing positive values for their investors, which shows the benefit that institutional investors can bring.

References:

Aumann, R. J. and Serrano, R. (2008). An Economic Index of Riskiness. Journal of Political Economy, 116(5):810-836.

Avramov, D. and Wermers, R. (2006). Investing in mutual funds when returns are predictable. Journal of Financial Economics, 81(2):339-377.

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

Chen, Z., Jiang, H., Li, D., and Sim, A. B. (2010). Regulation Change and Volatility Spillovers: Evidence from China's Stock Markets. Emerging Markets Finance and Trade, 46(6):140-157.

Christopherson, J. A., Ferson, W. E., and Glassman, D. A. (1998). Conditioning Manager Alphas on Economic Information: Another Look at the Persistence of Performance. Review of Financial Studies, 11(1):111-42.

Diamond, P. A. and Stiglitz, J. E. (1974). Increases in risk and in risk aversion. Journal of Economic Theory, 8(3):337-360.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1):3-56.

Ferson, W. E. and Schadt, R. W. (1996). Measuring Fund Strategy and Performance in Changing Economic Conditions. Journal of Finance, 51(2):425-61.

Firth, M., Lin, C., and Zou, H. (2010). Friend or Foe? The Role of State and Mutual Fund Ownership in the Split Share Structure Reform in China. Journal of Financial and Quantitative Analysis, 45(03):685-706.

Homm, U. and Pigorsch, C. (2012). Beyond the Sharpe ratio: An application of the Aumann Serrano index to performance measurement. Journal of Banking & Finance, 36(8):2274-2284.

Hunter, D., Kandel, E., Kandel, S., and Wermers, R. (2014). Mutual fund performance evaluation with active peer benchmarks. Journal of Financial Economics, 112(1):1-29.

Jensen, M. C. (1968). The Performance of Mutual Funds In The Period 19451964. Journal of Finance, 23(2):389-416.

Li, K., Wang, T., Cheung, Y.-L., and Jiang, P. (2011). Privatization and Risk Sharing: Evidence from the Split Share Structure Reform in China. Review of Financial Studies, 24(7):2499-2525.

Liao, L., Liu, B., and Wang, H. (2011). Information Discovery in Share Lockups: Evidence from the Split Share Structure Reform in China. Financial Management, 40(4):1001-1027.

Liao, L., Liu, B., and Wang, H. (2014). Chinas secondary privatization: Perspectives from the Split-Share Structure Reform. Journal of Financial Economics, 113(3):500-518.

Liu, C., Uchida, K., and Yang, Y. (2014). Controlling shareholder, split-share structure reform and cash dividend payments in China. International Review of Economics and Finance, 29(C):339-357.

Mamaysky, H., Spiegel, M., and Zhang, H. (2008). Estimating the Dynamics of Mutual Fund Alphas and Betas. Review of Financial Studies, 21(1):233-264.

Moon, G.-H. and Yu, W.-C. (2010). Volatility Spillovers between the US and China Stock Markets: Structural Break Test with Symmetric and Asymmetric GARCH Approaches. Global Economic Review, 39(2):129-149.

Qian, Z. and Luo, Q. (2014). Regime dependent determinants of sovereign credit spread in China. forthcoming in the Emerging Markets Finance and Trade.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. A definition. Journal of Economic Theory, 2(3):225-243.

Rothschild, M. and Stiglitz, J. E. (1971). Increasing risk II: Its economic consequences. Journal of Economic Theory, 3(1):66-84.

Xiong, W. and Yu, J. (2011). The Chinese Warrants Bubble. American Economic Review, 101(6):2723-53.

Xu, X. E. (2005). Performance of Securities Investment Funds in China. Emerging Markets Finance and Trade, 41(5):28-42.

Yao, S. and Luo, D. (2009). The Economic Psychology of Stock Market Bubbles in China. The World Economy, 32(5):667-691.

Table 1 Summary statistics of individual funds' average abnormal returns

Group	Group Mean	Group Median	Group S.D.	Group Min	Group Max	Normality
Largecap Value	0.43%	0.49%	0.42%	-1.08%	1.07%	0.00
Largecap Balance	0.14%	0.20%	0.91%	-5.59%	2.30%	0.00
Largecap Growth	0.18%	0.17%	0.86%	-3.43%	3.86%	0.00
Midcap Growth	0.09%	-0.21%	3.32%	-17.77%	20.86%	0.00

Notes: Abnormal returns are non-annualized monthly returns. The normality reports *p* values of the Jarque-Bera test.

Table 2 Rankings in the Largecap Value (LV) group

Ranking	Criteria		
	APM	Sharpe	APB − α
1	7	36	37
2	53	43	39
3	36	25	29
4	38	42	36
5	52	9	32
6	54	8	53
7	16	53	35
8	15	39	49
9	9	35	26
10	19	51	22
11	40	29	51
12	49	32	15
13	37	37	25
14	23	47	19
15	42	22	52
16	32	26	3
17	39	21	8
18	3	49	18
19	27	34	43
20	34	3	47
21	29	7	12
22	35	19	48
23	20	16	38
24	43	18	16
25	8	38	54
26	21	54	40
27	51	48	21
28	47	15	23
29	48	52	27
30	18	23	7

Notes: Numbers that are under the three criteria (APM, Sharpe and APB − α) stand for the fund IDs within the LV category (group). The rankings are done independently based on the APM, Sharpe and APB − α, respectively. Table 2 to 5 only lists the top-30 funds as suggested in the column with the title "Ranking" (the first column).

Table 3 Rankings in the Largecap Balance (LB) group

Ranking	Criteria		
	APM	Sharpe	APB − α
1	167	36	181
2	41	157	98
3	181	208	162
4	178	211	152
5	29	183	67
6	229	188	207
7	163	84	85
8	82	65	228
9	23	40	105
10	183	78	46
11	65	27	33
12	49	37	210
13	36	160	231
14	117	24	227
15	188	231	41
16	43	158	29
17	233	216	117
18	84	118	189
19	40	70	212
20	152	49	167
21	46	167	65
22	97	23	166
23	102	217	154
24	208	46	125
25	157	29	163
26	78	119	120
27	211	97	224
28	118	181	180
29	216	152	178
30	27	221	36

Notes: Numbers that are under the three criteria (APM, Sharpe and APB − α) stand for the fund IDs within the LB category (group). The rankings are done independently based on the APM, Sharpe and APB − α, respectively.

Table 4 Rankings in the Largecap Growth (LG) group

Ranking	Criterion		
	APM	Sharpe	APB − α
1	52	57	204
2	40	202	210
3	154	231	31
4	148	182	242
5	204	240	37
6	179	50	142
7	57	29	263
8	50	1	195
9	45	40	282
10	151	179	120
11	82	150	57
12	240	48	249
13	182	178	150
14	1	128	145
15	263	47	274
16	265	165	52
17	48	167	95
18	29	161	1
19	47	239	134
20	150	130	175
21	130	172	130
22	65	58	183
23	128	95	240
24	274	66	62
25	185	72	185
26	239	96	110
27	235	204	254
28	202	263	66
29	178	52	212
30	231	156	179

Notes: Numbers that are under the three criteria (APM, Sharpe and APB − α) stand for the fund IDs within the LG category (group). The rankings are done independently based on the APM, Sharpe and APB − α, respectively.

Table 5 Rankings in the Midcap Growth (MG) group

Ranking	Criterion		
	APM	Sharpe	APB − α
1	65	61	25
2	78	19	63
3	68	116	125
4	122	98	52
5	66	41	75
6	69	93	113
7	57	113	100
8	110	63	98
9	46	85	90
10	30	115	124
11	109	22	115
12	88	52	61
13	32	75	93
14	106	36	116
15	61	125	41
16	82	86	19
17	71	25	22
18	113	95	36
19	79	45	95
20	94	37	49
21	81	49	45
22	59	100	24
23	51	96	37
24	116	90	111
25	105	111	96
26	108	42	34
27	56	24	85
28	72	124	58
29	50	34	17
30	28	58	42

Notes: Numbers that are under the three criteria (APM, Sharpe and APB − α) stand for the fund IDs within the MG category (group). The rankings are done independently based on the APM, Sharpe and APB − α, respectively.

Table 6 Correlations between rankings by different performance indicators

Fund Group	APM, Sharpe	Sharpe, APB − α	APB − α, APM
Largecap Value	-0.19	0.21	0.09
	(0.23)	(0.18)	(0.58)
Largecap Balance	0.06	0.06	0.04
	(0.45)	(0.47)	(0.63)
Largecap Growth	0.00	-0.07	0.01
	(0.95)	(0.30)	(0.88)
Midcap Growth	-0.04	0.08	-0.19
	(0.72)	(0.46)	(0.09)

Notes: *p* values are in the parenthesis.

Table 7 APM distributions of different groups

Fund Group	Mean	Median	S.D.	Normality	Signed-rank
Largecap Value	3.65	3.28	3.01	0.00	0.00
Largecap Balance	0.61	0.16	2.95	0.00	0.00
Largecap Growth	0.47	0.29	2.26	0.00	0.00
Midcap Growth	1.00	1.07	1.03	0.00	0.00

Notes: *p* values are reported for the normality test (fifth column) and the Wilcoxon signed-rank test (sixth column) of median for statistical significance. The pairwise Wilcoxon signed-rank test has a null hypothesis of zero expected median by comparing the signed-rank of the observed series with a synthesized series with zero median. A significant signed-rank test means the observed median differs from zero significantly.