Investor Learning and Mutual Fund Family

Zhichao Zhang[a], Li Ding[a], Si Zhou[a*]

[a]Durham University Business School, Durham University

Abstract

In this paper we revisit the cross-fund learning method suggested by Jones and Shanken (2005) and construct a linear hierarchical model to consider the learning across funds within the fund family during the performance evaluation. We provide a full Bayesian treatment on all the factors of the pricing model and allow both the fund family and the individual manager to have dependent prior information regarding funds' alphas. The simulation results suggest that returns from peer funds within the family significantly affect investors' updating on fund alphas since the posterior distribution on fund alphas experience a faster shrinkage than those reported in the previous literature. The model can also be simulated with specific prior belief on different factors of the pricing model, i.e. fund alphas, betas and factor loadings of each pricing benchmark, to better address the learning issue.

JEL Classification: G12; C11

1. Introduction

---

[*] Corresponding author. Postal address: Durham University Business School, Mill Hill Lane, Durham, UK, DH1 3LB.
E-mail: si.zhou@durham.ac.uk

Information from the fund family can provide additional insight when evaluating the performance of its underlying funds. It is often the case that funds affiliated to the same fund family share the same investment advisor. Fund family and the fund manager combined contribute to the returns gained by a certain fund. The family can influence the performance of the individual funds not only from the administration perspective, but also in terms of the quality of analysis and information flows (Baks, 2003). In addition, the fund family conducts various investment strategies to affect the performance of its underlying funds (see for example Nanda, Wang and Zheng, 2004; Kempf and Ruenzi, 2008; Gaspar, Massa and Matos, 2006). However, standard performance evaluation literature usually examines the fund performance independently, neglecting the return information provided by the other parties. This research aims to conduct the performance evaluation procedure taking into consideration information provided by other funds as well as the fund family.

Jensen alpha, widely used as the risk adjusted performance of a mutual fund (Jensen, 1968), is conventionally calculated by applying the OLS estimation on the intercept of the capital asset pricing model (CAPM) by Sharpe (1964) and Lintner (1965). This performance evaluation has evolved to incorporate additional benchmark portfolios (see Fama and French, 1993; Jegadeesh and Titman, 1993; Carhart, 1997; Elton, Gruber and Blake, 1996). Other researches adopt alternative techniques to understand funds' abnormal performance. For example, studies by Kosowski, Timmermann, Wermers and White (2006) and Cuthbertson, Nitzsche and O'Sullivan (2008) apply a bootstrapping method to distinguish alphas that can be attributed to managers' genuine stock selecting skills from those resulting from sample variation.

More recently, a growing number of studies have shifted their interest to the additional information provided by benchmark (non-benchmark) pricing factors, investors' opinion and

returns from other funds. The Bayesian framework provides the opportunity to include information other than funds' historical data in the performance evaluation. Baks, Metrick and Wachter (2001) find that certain prior beliefs on managers' skill might justify the investment decision. Pástor and Stambaugh (2002) (PS hereafter) consider a seemingly unrelated model (SURE hereafter) with the Bayesian estimation to include the correlation between the pricing factors and the other non-benchmark portfolios, which incorporates the idea that information given by non-benchmark portfolios with longer return history provides more precise estimates (Stambaugh, 1997). Such Bayesian settings not only overcome the limited datasets in the estimation, but also improve understanding of how the so-called seemingly unrelated assets affect the performance evaluation of a certain fund. Busse and Irvine (2006) further conduct the performance persistence test based on the SURE in a similar Bayesian model suggested by PS. Their results indicate that higher predication power of the SURE model is more likely to be associated with the diffuse skill prior.

However, the independent prior based Bayesian settings considered in the previous research raise the issue of ignorance about cross-sectional influences from the peer funds. Such dependent nature of the variability of funds' alphas can be modelled in a hierarchical setting in which a dependent prior is designated on the cross-sectional mean. Jones and Shanken (2005) (JS hereafter) first consider a multilevel structure in the performance evaluation, with a dependent prior which can then be assigned to represent the investors' opinion on the mean of the cross-sectional fund returns distribution. They suggest that the alpha of a fund can be drawn from a common population distribution which is defined to describe the general belief on the cross-sectional performance. They find that the investors are more likely to believe that the manager of a certain fund is unskillful if more funds in the industry give them the same

impression. If the investors tend to have more homogeneous belief on the absence of fund managers' skill, i.e. the variance of the prior decreases, the shrinkage is also enhanced. In this case, a learning prior provides a compromise between the fund's own returns and the cross-sectional performance in the entire industry. However, since their research only considers the dependent nature of the prior belief on alpha, the evaluation model can be regarded as a special case of the hierarchical varying intercept and varying slope model. In reality, investors may also have heterogeneous belief on the pricing power of the certain factor model used in the performance evaluation, or on the risk exposure to a particular market benchmark. These concerns make it necessary to conduct a general case multilevel model.

We construct a linear hierarchical model to consider the learning across funds within the fund family to estimate fund alphas. Although a general solution to the linear hierarchical model is derived by Lindley and Smith (1972) and Smith (1973), the major problem encountered lies in adding appropriate prior information onto the covariance matrix when giving full Bayesian treatment on all the pricing factors in the model. In CAPM, only a general prior distribution, i.e. an inverse Wishart prior, is applied, to represent all the additional information regarding both the alpha and the market beta. However, given that the degree of freedom is the only variable used to define the distribution, use of the inverse Wishart as the prior belief in the estimation of Bayesian alphas is far from the situation in reality.

A separation strategy which decomposes the covariance matrix to produce the diagonal matrix with variance of each pricing factor and the correlation matrix, as suggested by several statistical studies, is able to overcome the restriction noted above and to include the return information from the other pricing factors (see for example Barnard, McCulloch and Meng, 2000). An important feature of the separation technique is its consideration of specific prior

beliefs on certain parameters of interest, i.e. the ability to strengthen the informative level on particular parameters and weaken it on others. In this research we apply a modified separation technique which not only maintains the original key feature, but also improves its efficiency (see Gelman and Hill, 2007; O'Malley and Zaslavsky, 2008). Specifically, a scaled inverse Wishart distribution is denoted as the prior on the covariance matrix through over-parameterization, the use of which enhances the convergence substantially. We can therefore define the prior information on each of the pricing factors as well as the between-factor correlation.

Our results from the simulation suggest that the separation strategy powered performance evaluation better addresses the learning issue. Firstly, we find that given a less dispersed prior belief on managers' inferior ability, the posterior mean on α of each of the underlying funds converges faster than when using the method suggested by JS. Our findings suggest that returns from peer funds within the same fund family can significantly affect investors' updating on fund alphas. Secondly, the covariance separation technique enables our method to provide a full Bayesian treatment on each of the pricing factors to generalize the learning process. Specifically, our model can grasp the specific prior information on the magnitude of funds' pricing factors deviating from the family mean. The results suggest that the posterior belief can provide a compromise between the observed data and the prior belief. Strong cross-sectional dispersion in the data can still mitigate the posterior shrinkage when a high skepticism prior is in place. Thirdly, after decomposing the individual fund α into the combination of the family mean and the idiosyncratic contribution from the manager, we find that the fund manager contributes positively to the overall fund performance whenever prior belief is applied. Finally, we place no restriction on the correlation matrix of different pricing factors in the pricing model. That is to say, we also include prior information to allow cross-factor learning, which is often impossible in the

conventional OLS estimated alphas.

The rest of the research is organized as follows: The learning model is derived in the following section. We also show the model given by JS, which can be regarded as a special case of our model. Section 3 discusses the model simulation results by using the hypothetical data as well as the real fund data. Conclusions and the implications of this research are summarized in the final section.

2. The Performance Evaluation Model

2.1 The general learning model

The learning process considered in this section is similar to the settings of JS, but in a more general framework. We adapt the Bayesian treatment for each of the pricing benchmarks in the factor model. In our evaluation model no restriction is applied on the correlation of different pricing factors, thus the co-movements can be viewed as an unknown variable which is decided by the information mixture of the prior belief and the true data, whereas the conventional OLS estimations might suffer substantial imprecision due to multicollinearity between different regressors. Another important feature of the general learning model is that the dependent prior on the pricing factors enables the model to explain the heterogeneous opinion on the pricing power of a particular factor model. Different prior beliefs on the pricing benchmarks can then be included to address the sensitivity issue of how funds' alphas respond to divergent views on the pricing power of benchmark portfolios. Thus the evaluation model of JS can be regarded as a special case of the general learning model. Meanwhile, instead of gaining information from the entirety of cross-sectional funds in the evaluation, we sort funds into different fund families,

since funds within the same family often share the same investment adviser, and they are more likely to set a similar market benchmark to compete with. In addition, fund companies often adopt various family strategies, such as reallocating capital or increasing cross-sectional variance, to achieve better performance for their underlying funds or entire fund families. Therefore, we construct the learning model from a family perspective to incorporate additional return information offered by funds within the same fund family.[1]

A hierarchical linear structure is applied to assess the manager's ability when performance is assumed to vary across the funds managed by the same fund company. To facilitate the estimation of the variables in the multilevel structure a Bayesian system is constructed to conduct the distribution of each variable as a weighted average of both prior belief and real data. We assume in our model that the risk adjusted performance, can be attributed to the fund family and the fund's idiosyncratic risk exposure, which are all assumed to be unknown to both the fund company and the investors. The posterior distribution is generated through Markov Chain Monte Carlo algorithm (MCMC hereafter). We derive the Gibbs sampler and the Metropolis-Hastings algorithm for each of the unknown variables in the Bayesian hierarchical linear model, since the posterior distribution of all the variables can be written in a closed form except that for the within variability in fund family. Gibbs sampler can update each variable directly at a time when its posterior distribution can be derived in a closed form, while a proposed distribution is needed for the Metropolis-Hastings algorithm to act as a reference for drawing.

2.1.1 Likelihood function

---

[1] Our model can be easily adapted to the research context of JS, where a diffuse prior is designated to each of the pricing factors. Meanwhile, the prior belief can be set to represent the opinion on the performance of the entirety of cross-sectional funds

Consider a fund family with M fund, for each of the fund j we assume that the excess returns $\mathbf{R}_j$ follows:

$$\mathbf{R}_j = f_j b_j + u_j, \quad (j = 1,\ldots,M) \tag{1}$$

where $\mathbf{R}_j$ is a n_j dimensional vector of fund's excess returns where n_j is the number of observation for fund j, and f_j is a $n_j \times K$ matrix of the excess returns from $K-1$ market benchmark portfolio(s), of which the first column is all 1. b_j is K dimensional factor loadings which include the risk adjusted return and pricing factors for each of the $K-1$ benchmark portfolio(s), i.e. $b_j = (\alpha_j, \beta_{j,2},\ldots,\beta_{j,K})'$. We assume that $u_j \sim N(0,\sigma_j^2)$, in which u_j is assumed to be homoscedastic and independent of each other. The prior belief on σ_j^2 are given by a scaled inverse χ^2 distribution, i.e. $\sigma_j^2 \sim Scale-inv-\chi^2(v_j,s_j^2)$.[2]

The family level likelihood function for fund j can be shown as:

$$b_j = \mathbf{X}_j \Theta + e, \quad (j = 1,\ldots,M) \tag{2}$$

where $\mathbf{X}_j = \mathbf{I}_k \otimes x_j$ is a $K \times K$ matrix of family level predictors, x_j. As suggested in the following simulation study, we assume that x_j equals to 1 for all j. Additional factors can also be incorporated as family level predictors, i.e. the non-benchmark assets in the SURE model. Θ is a K dimensional vector which describes the family level mean for each of the K pricing factors, specifically, $\Theta = (\theta_{j,1},\ldots,\theta_{j,K})'$. The risk adjusted return for fund j can therefore be

---

[2] Assuming Y follows an inversed Gamma distribution, i.e. $Y \sim Inv-gamma(a,b)$, where a and b are shape and scale parameter, respectively. Thus the probability density function for Y is $P(Y) = \dfrac{b^a}{\Gamma(\alpha)} Y^{-\alpha+1} exp(-\dfrac{b}{Y})$.

The $Scale-inv-\chi^2(v_j,s_j^2)$ distribution has a density function by letting $\alpha = \dfrac{v_j}{2}$ and $b = \dfrac{v_j s_j^2}{2}$.

given by $\alpha_j = \Theta_{j,1} + e_{j,1}$. The between-fund dispersion of the kth pricing factor is denoted as $\grave{o}_{j,k}$, i.e. $Var(\beta_{j,k}) = \grave{o}_{j,k}$ and $k = 1,\ldots,K$. Thus the covariance matrix of b_j can be expressed as a $K \times K$ matrix, λ_j, i.e. $\lambda_j = diag(\grave{o}_{j,1},\ldots,\grave{o}_{j,K})$.

Given Eq (1), let $\mathbf{R} = (\mathbf{R}'_1,\ldots,\mathbf{R}'_M)'$, $\mathbf{F} = diag(f_1,\ldots,f_M)$, $\mathbf{B} = (b_1',\ldots,b_M')'$ and $N = \sum_{j=1}^{M} n_j$, then we can rewrite Eq (1) for M funds as

$$\mathbf{R} = \mathbf{FB} + \mathbf{U}, \quad \mathbf{U} \sim N(0,\Sigma) \tag{3}$$

where $\Sigma = diag(\Sigma_1,\ldots,\Sigma_M)$ and $\Sigma_j = \sigma_j^2 \mathbf{I}_{n_j}$.

The family level likelihood function for M funds can also been given by letting $\mathbf{X} = (\mathbf{X}_1,\ldots,\mathbf{X}_M)'$ and $\Lambda = \mathbf{I}_M \otimes \lambda_j$ in which $\mathbf{I}_M$ is a $M \times M$ identity matrix. Eq(2) for M funds then can be written as

$$\mathbf{B} = \mathbf{X\Theta} + \mathbf{E}, \quad \mathbf{E} \sim N(0,\Lambda) \tag{4}$$

where $\mathbf{\Theta}$ represents the mean value which remains the same across M funds, while Λ is the in-family dispersion level among the M funds. The prior on Λ can then be regarded as the magnitude of how factor loadings of an individual fund deviate from its group mean. Thus, a prior on fund's alpha with a higher variance suggests a higher cross-sectional variability on alpha within the fund family.

To address the dependence of the prior we further assume that $\mathbf{\Theta}$ is a random draw from a common multivariate normal distribution, $N(\zeta,\Delta)$, which represents the beliefs on the family's mean. In order to denote specific prior belief on each of the K pricing factors, we further consider a separation strategy to define the prior on the in-family variation, Λ, in which the family level covariance matrix is decomposed into a combination of diagonal scaled matrix and

an unscaled matrix that can describe the correlation of factor loadings among different funds within the same fund family, i.e. $\mathbf{\Lambda} = \mathbf{\Xi\Phi\Xi}$, where $\mathbf{\Xi}$ is a diagonal scaled matrix and $\mathbf{\Phi}$ is the unscaled matrix.[3]

2.1.2 Posterior distribution of $\mathbf{B}$

In this section we derive the posterior distribution of the factor loadings for M funds conditional on $\mathbf{R}$, $\mathbf{F}$ and $\mathbf{X}$. Assuming that $\mathbf{\Sigma}$, $\mathbf{\Theta}$ and $\mathbf{\Lambda}$ are all updated, ζ and $\mathbf{\Delta}$ are the prior belief on $\mathbf{B}$. The posterior belief of $\mathbf{B}$ can be derived as,

$$
\begin{aligned}
P(\mathbf{B} \mid \mathbf{R}, \mathbf{F}, \mathbf{\Sigma}, \mathbf{X}, \mathbf{\Theta}, \mathbf{\Lambda}) &\propto P(\mathbf{R} \mid \mathbf{B}, \mathbf{F}, \mathbf{\Sigma}) P(\mathbf{B} \mid \mathbf{X}, \mathbf{\Lambda}, \zeta, \mathbf{\Lambda}) \\
&= \mathrm{N}_M(\mathbf{R} \mid \mathbf{FB}, \mathbf{\Sigma}) \mathrm{N}_{MK}(\mathbf{B} \mid \mathbf{X}\zeta, \mathbf{\Lambda} + \mathbf{X}\mathbf{\Lambda}\mathbf{X}') \\
&\propto exp\left\{ -\frac{1}{2}\left[(\mathbf{R} - \mathbf{FB})'\mathbf{\Sigma}^{-1}(\mathbf{R} - \mathbf{FB}) + (\mathbf{B} - \mathbf{X}\zeta)'(\mathbf{\Lambda} + \mathbf{X}\mathbf{\Lambda}\mathbf{X}')^{-1}(\mathbf{B} - \mathbf{X}\zeta) \right] \right\} \\
&\propto exp\left\{ -\frac{1}{2}\left[\mathbf{B}'(\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F} + (\mathbf{\Lambda} + \mathbf{X}\mathbf{\Lambda}\mathbf{X}')^{-1})\mathbf{B} - 2\mathbf{B}'(\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{R} + (\mathbf{\Lambda} + \mathbf{X}\mathbf{\Lambda}\mathbf{X}')^{-1}\mathbf{X}\zeta) \right] \right\} \\
&\propto exp\left[-\frac{1}{2}(\mathbf{B} - \mathbf{D_1}\mathbf{V_1})'\mathbf{D_1}^{-1}(\mathbf{B} - \mathbf{D_1}\mathbf{V_1}) \right]
\end{aligned}
$$

So the posterior belief on the fund's factor loadings follows a $M K$ dimensional multivariate normal distribution,

$$
\mathbf{B} \mid \mathbf{R}, \mathbf{\Sigma} \sim \mathrm{N}_{MK}(\mathbf{D_1}\mathbf{V_1}, \mathbf{D_1}) \tag{5}
$$

where

$$
\mathbf{D_1} = [\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F} + (\mathbf{\Lambda} + \mathbf{X}\mathbf{\Lambda}\mathbf{X}')^{-1}]^{-1} \quad \text{and} \quad \mathbf{V_1} = \mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{R} + (\mathbf{\Lambda} + \mathbf{X}\mathbf{\Lambda}\mathbf{X}')^{-1}\mathbf{X}\zeta.
$$

---

[3] Gelman and Hill (2007) argue that such over parameterization not only enables the control of the dispersion level for the factor loadings within the same group, since $\mathbf{\Phi}$ is close to uniform, it also increases the convergence of the chain. See for example Barnard et al. (2000) and O'Malley and Zaslavsky (2008) for further discussion on the separation strategy and the scaled inverse Wishart distribution.

The posterior mean, $\mathbf{D_1 V_1}$, of $\mathbf{B}$ is a weighted average of the true return data and the prior belief on $\mathbf{B}$. We can further extend $(\mathbf{\Lambda} + \mathbf{X\Delta X'})^{-1}$ as,

$$(\mathbf{\Lambda} + \mathbf{X\Delta X'})^{-1} = \mathbf{\Lambda}^{-1} - \mathbf{\Lambda}^{-1}\mathbf{X}(\mathbf{X'}\mathbf{\Lambda}\mathbf{X} + \mathbf{\Delta'})^{-1}\mathbf{X'}\mathbf{\Lambda}^{-1}$$

Thus, when $\mathbf{\Lambda}^{-1} \to 0$ and $\mathbf{\Delta}^{-1} \to 0$, the posterior mean of $\mathbf{B}$ becomes $\mathbf{D_1 V_1} = (\mathbf{F'\Sigma^{-1}F})^{-1}(\mathbf{F'\Sigma^{-1}R})$, that is, the posterior mean of $\mathbf{B}$ reduces to its OLS estimates given a diffuse prior on both the cross-sectional variability and the variance of the family level mean.

2.1.3 Posterior distribution of $\mathbf{\Sigma}$

Given $\mathbf{B}$, $\mathbf{\Theta}$ and $\mathbf{\Lambda}$, we have $P(\mathbf{\Sigma} | \mathbf{R,F,B}) \propto P(\mathbf{R} | \mathbf{B,F,\Sigma})P(\mathbf{\Sigma})$. By assumption we have a homoscedastic error term for each fund j, which we can write as $\mathbf{\Sigma} = \mathbf{I_M} \otimes \mathbf{\Sigma_j}$. The posterior belief can then be shown as

$$P(\sigma_j^2 | R_{.j}) \propto \prod_{i=1}^{n_j} P(R_j | f_{i\,j}, \beta_{i\,j}; \sigma_j^2) P(\sigma_j^2)$$

$$\propto \prod_{i=1}^{n_j} N(R_{i,j} | f_{i\,j}\beta_{i\,j}; \sigma_j^2) Scale - inv - \chi^2(\sigma_j^2 | v_j, s_j)$$

$$\propto (\sigma_j^2)^{-(\frac{n+v_j}{2}+1)} exp\left[-\frac{1}{2\sigma_j^2}(nS_j + v_j s_j^2)\right]$$

Therefore, we have the posterior distribution for σ_j^2 as

$$\sigma_j^2 | R_{.j}, \beta_j \sim Scale - inv - \chi^2(n + v_j, \frac{nS_j + v_j s_j^2}{n + v_j}) \tag{6}$$

where $i = 1, \ldots, n_j$, $j = 1, \ldots, M$ and $S_j = \dfrac{\sum_{i=1}^{n_j}(R_{i,j} - f_j\beta_j)^2}{n_j}$.

2.1.4 Posterior distribution of Θ

When $\mathbf{B}$ and Λ are both updated by the distribution described in the previous two sections, the posterior belief on Θ can then be derived in a similar fashion given the information of prior distribution $\Theta \sim N(\zeta, \Delta)$,

$$
\begin{aligned}
P(\Theta \mid \mathbf{R}, \mathbf{F}, \Sigma, \mathbf{X}, \Lambda, \zeta, \Delta) &\propto P(\mathbf{R} \mid \mathbf{F}, \Sigma, \mathbf{X}, \Theta, \Lambda) P(\Theta \mid \zeta, \Delta) \\
&= \mathrm{N}_{MK}(\mathbf{R} \mid \mathbf{FX\Theta}, \Sigma + \mathbf{F\Lambda F}') \mathrm{N}_K(\Theta \mid \zeta, \Delta) \\
&\propto exp\left\{-\frac{1}{2}\left[(\mathbf{R} - \mathbf{FX\Theta})'(\Sigma + \mathbf{F\Lambda F}')^{-1}(\mathbf{R} - \mathbf{FX\Theta}) + (\Theta - \zeta)'\Delta^{-1}(\Theta - \zeta)\right]\right\} \\
&\propto exp\left\{-\frac{1}{2}\left[\Theta'(\Delta^{-1} + \mathbf{X}'\mathbf{F}'(\Sigma + \mathbf{F\Lambda F}')^{-1}\mathbf{FX})\Theta - \Theta'(\mathbf{X}'\mathbf{F}'(\Sigma + \mathbf{F\Lambda F}')^{-1}\mathbf{R} + \Delta^{-1}\zeta)\right]\right\}
\end{aligned}
$$

Thus Θ can be shown to follow a K dimensional multivariate normal distribution:

$$
\Theta \mid \mathbf{X}, \mathbf{F}, \Lambda \sim \mathrm{N}_K(\mathbf{D}_2\mathbf{V}_2, \mathbf{D}_2) \tag{7}
$$

where

$$
\mathbf{D}_2 = [\Delta^{-1} + \mathbf{X}'\mathbf{F}'(\Sigma + \mathbf{F\Lambda F}')^{-1}\mathbf{FX}] \text{ and } \mathbf{V}_2 = \mathbf{X}'\mathbf{F}'(\Sigma + \mathbf{F\Lambda F}')^{-1}\mathbf{R} + \Delta^{-1}\zeta
$$

The posterior mean of Θ has a similar form as those defined for $\mathbf{B}$ in the previous section. However, if we have $\Delta^{-1} \to 0$, the posterior mean of Θ equals to $[\mathbf{X}'\mathbf{F}'(\Sigma + \mathbf{F\Lambda F}')^{-1}\mathbf{FX}]^{-1}[\mathbf{X}'\mathbf{F}'(\Sigma + \mathbf{F\Lambda F}')^{-1}\mathbf{R}]$, where the prior mean ζ has little effect and the true data dominate the estimation of Θ.

2.1.5 Posterior distribution of Λ

The family level covariance matrix for M funds can be written as $\Lambda = \mathbf{I}_M \otimes \lambda$. Only $\mathbf{B}$ and Θ are related to the variation of λ, and the covariance λ can be written as a combination of the diagonal matrix of standard deviations and a matrix of correlation, i.e. $\Xi\Phi\Xi$. Thus the joint distribution of Ξ and Φ can be stated as,

$$P(\Xi, \Phi \mid \mathbf{B}, \mathbf{X}, \Theta) \propto P(\mathbf{B} \mid \Xi, \Phi, \Theta) P(\Xi) P(\Phi)$$
$$= \mathrm{N}_{MK}(\mathbf{B} \mid \mathbf{X}\Theta, \Xi\Phi\Xi) \mathrm{W}^{-1}(\Phi \mid K_0, \mathbf{I}) log - \mathrm{N}(\Xi \mid v, s^2)$$

We firstly derive the posterior distribution of the unscaled matrix that determines the correlation,

given the prior is $\Phi \sim \mathrm{W}^{-1}(K_0, \mathbf{I})$,

$$P(\Phi \mid \mathbf{B}, \Xi) \propto \prod_{j=1}^{M} N(\beta_j \mid \mathbf{X_j}\Theta, \Xi\Phi\Xi) \mathrm{W}^{-1}(K_0, \mathbf{I})$$

$$\propto |\Xi\Phi\Xi|^{-\frac{M}{2}} exp\left[-\frac{1}{2} tr \mathbf{S_0}(\Xi\Phi\Xi)^{-1}\right] |\Phi|^{\frac{K_0+K+1}{2}} exp\left[-\frac{1}{2} tr(\mathbf{I}\Phi^{-1})\right]$$

$$\propto |\Phi|^{\frac{(K_0+M)+K+1}{2}} exp\left[-\frac{1}{2} tr(\Xi^{-1}\mathbf{S_0}\Xi^{-1} + \mathbf{I})\Phi^{-1}\right]$$

Therefore, we can show that

$$\Phi \mid \mathbf{B}, \Theta, \Xi, K_0, M \sim Scaled - \mathrm{W}^{-1}(K_0 + M, \Xi^{-1}\mathbf{S_0}\Xi^{-1} + \mathbf{I}) \tag{8}$$

where $\mathbf{S_0} = \sum_{j=1}^{M} (\beta_j - x_j\Theta)(\beta_j - x_j\Theta)'$. For k^{th} factor in the learning model, its variance is

$\xi_k^2 \Phi_{kk}$, where Φ_{kk} is the kth value on the diagonal of Φ. The posterior distribution on Ξ

can be estimated by using the Metropolis-Hastings algorithm since the distribution function is not

in a convenient form. Given $\Xi = diag(\xi_1, \ldots, \xi_K)$, its conditional posterior distribution function

can be written as

$$P(\xi_k \mid \mathbf{B}, \Phi) \propto \prod_{j=1}^{M} (\beta_{k,j} \mid \mu_{\beta_k}, \sigma_{\beta_k}^2) log - \mathrm{N}(\xi_k \mid v_k, s_k^2)$$

where[4]

$$\mu_{\beta_k} = E(\beta_k) + Cov(\beta_k, \beta_{-k}) Var([\beta_{-k}])^{-1}(\beta_{-k} - E(\beta_{-k}))$$
$$\sigma_{\beta_k}^2 = Var(\beta_k) - Cov(\beta_k, \beta_{-k}) Var([\beta_{-k}])^{-1} Cov(\beta_k, \beta_{-k})$$

---

[4] Since λ is a $K \times K$ matrix, thus $[\beta_{-k}]$ indicates a matrix without the kth element.

and the prior on ξ_k is given by $\xi_k \sim log - N\ (v_k, s_k^2)$. We then use a log-normal distribution as the proposed distribution to simulate the target distribution with the acceptance rate over 44%.[5]

2.2 The non-learning and partial learning models

After deriving the general learning model, we then look at the difference between the non-learning model, the partial learning model and the general learning model. The non-learning model can be regarded as an evaluation model with independent prior belief, while the partial learning model considers dependent prior only on the fund alphas. Given the likelihood function Eq (1), we can derive the non-learning model for fund j as,

$$R_j = \alpha_j + f_j \beta_j + u_j, \ \ u_j \sim N(0, \sigma^2)$$

Then we can draw α_j given R_j, f_j and β_j, assuming σ^2 is drawn from another procedure. The posterior belief then follows,

$$P(\alpha_j \mid R_{.j}, f, \beta_j, \sigma^2) \propto \prod_{n=1}^{N_j} N(R_{.j} \mid \alpha_j + f\beta_j, \sigma^2) N(\alpha_j \mid \mu_j, \sigma_j^2)$$

$$\propto exp\left\{ -\frac{1}{2\sigma_\alpha^2}(\alpha_j - \mu_j)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{N_j}[\alpha_j - (R_{.j} - f\beta_j)]^2 \right\}$$

$$\propto exp\left[-\frac{1}{2\tilde{\sigma}_\alpha^2}(\alpha_j - \tilde{\alpha}_j)^2 \right]$$

where

$$\tilde{\alpha}_j = (\frac{\mu_j}{\sigma_j^2} + \frac{S_0}{\sigma^2})(\frac{1}{\sigma_j^2} + \frac{N}{\sigma^2})^{-1}$$

$$\tilde{\sigma}_\alpha^2 = (\frac{1}{\sigma_j^2} + \frac{N_j}{\sigma^2})^{-1}$$

$$S_0 = \sum_{i=1}^{N_j}(R_{.j} - f\beta_j)$$

---

[5] Similar argument can be found in, for example, Gelman and Hill (2007) and O'Malley and Zaslavsky (2008).

and assuming that α_j follows a prior belief, $\alpha_j \sim N(\mu_j, \sigma_j^2)$, for fund j. Thus, each of the M funds in the fund family is denoted with independent prior beliefs. In the simulation of the non-learning model we denote a non-informative prior on the variance parameter of β_j, thus its posterior distribution is the OLS estimation. For the prior distribution on α, we denote prior beliefs independent of each other. Therefore, the non-learning model can be regarded as the no pooling model, with specific prior on each of the funds.

JS applies a hierarchical model with dependent prior on individual funds' α_j. Their model can therefore be regarded as a varying intercept model while the factor loadings of other market benchmarks are left without Bayesian treatment. Given the same likelihood function Eq (1), the prior belief of α_j states

$$\alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2), \quad (j = 1, \ldots, M)$$

The posterior mean of α_j can be derived in the same fashion as the non-learning model:

$$P(\alpha_j \mid R_{.j}, f, \beta_j, \sigma^2) \propto \prod_{n=1}^{N_j} N(R_{.j} \mid \alpha_j + f\beta_j, \sigma^2) N(\alpha_j \mid \mu_\alpha, \sigma_\alpha^2)$$

$$\propto \exp\left\{-\frac{1}{2\sigma_\alpha^2}(\alpha_j - \mu_\alpha)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{N_j}[\alpha_j - (R_{.j} - f\beta_j)]^2\right\}$$

$$\propto \exp\left[-\frac{1}{2\tilde{\sigma}_\alpha^2}(\alpha_j - \tilde{\alpha}_j)^2\right]$$

where

$$\tilde{\alpha}_j = \left(\frac{\mu_\alpha}{\sigma_\alpha^2} + \frac{S_0}{\sigma^2}\right)\left(\frac{1}{\sigma_\alpha^2} + \frac{N}{\sigma^2}\right)^{-1}$$

$$\tilde{\sigma}_\alpha^2 = \left(\frac{1}{\sigma_\alpha^2} + \frac{N_j}{\sigma^2}\right)^{-1}$$

$$S_0 = \sum_{i=1}^{N_j}(R_{.j} - f\beta_j)$$

For the mean performance μ_α, we denote prior belief as $\mu_\alpha \sim N(m_\alpha, V_\alpha)$, thus the posterior belief

of μ_α is given by

$$P(\mu_\alpha, \alpha, \sigma_\alpha^2, m_\alpha, V_\alpha) \propto \prod_{j=1}^{M} N(\alpha_j, \sigma_\alpha^2) N(\mu_\alpha \mid m_\alpha, V_\alpha)$$

$$\propto exp\left\{-\frac{1}{2}\left[(\frac{M}{\sigma_\alpha^2} + \frac{1}{V_\alpha})\mu_\alpha^2 - 2\mu_\alpha(\frac{M\bar{\alpha}_j}{\sigma_\alpha^2} + \frac{m_\alpha}{V_\alpha})\right]\right\}$$

$$= exp\left[-\frac{1}{2\tilde{\sigma}_\alpha^2}(\mu_\alpha - \tilde{\mu}_\alpha)^2\right]$$

and we have $\mu_\alpha \sim N(\tilde{\mu}_\alpha, \tilde{\sigma}_\alpha^2)$, where

$$\tilde{\mu}_\alpha = (\frac{M\bar{\alpha}_j}{\sigma_\alpha^2} + \frac{m_\alpha}{V_\alpha})(\frac{M}{\sigma_\alpha^2} + \frac{1}{V_\alpha})^{-1}$$

$$\tilde{\sigma}_\alpha^2 = (\frac{M}{\sigma_\alpha^2} + \frac{1}{V_\alpha})^{-1}$$

$$\bar{\alpha}_j = \frac{1}{M}\sum_{j=1}^{M}\alpha_j$$

In the model derived above, for each fund i, a common prior belief is applied to all the M

funds in the same fund family. Thus, the common prior can be viewed as the additional information

on the mean performance of the entire fund family. In the simulation, we include a diffuse prior on

μ_α, i.e. apply a large value on V_α to eliminate the influence from m_α. The precision parameter

σ_α is given a prior belief following inverse χ^2 distribution. Since JS include no prior

information on other pricing factors in addition to α_j, a diffuse prior is then denoted to each of the

factor loadings to let them converge to the OLS estimations, which is similar with the settings

considered by JS.[6]

2.3 Prior beliefs

---

[6] JS rearrange Eq (1) to have $R_j - f\beta_j = \alpha_j + u_j$, and β_j is obtained directly from the OLS regression.

In this section we discuss the prior distribution we use for drawing from the posterior distribution of the parameters in the learning model. Although there are several possible choices of prior beliefs on all unknown variables, we concentrate on the family level variance, $\mathbf{\Lambda}$, since it is closely related to the cross-sectional variability within the fund family. In particular, a diffuse prior would allow the data to dominate the posterior distribution, while contracted prior leads to a high degree of shrinkage.

We consider three log normal distributions as the prior belief on ξ_k. Figure 1 shows our first choice, $log(\xi_k) \sim N(-1,1)$. The prior mean is then centered at around 0.25 suggesting ξ_k has a variance over 0.0625, which is far beyond the actual value observed in the data. This highly informative prior maintains the degree of shrinkage at a low level for the reason shown in section 2.1.2, where extremely diffuse variance drives the posterior mean to approach the OLS estimation. Thus, prior belief provides no information in this case.

The second and third choices of prior belief are illustrated in Figure 2. The dashed line given by $log(\xi_k) \sim N(-5,1)$ is centered at 0.003, that is 9×10^{-6} for variance. This prior is chosen to represent a plausible actual prior of alphas across funds within the fund family, since it is close to the highest cross-sectional variance among alphas given by the data. But its long right tail also enables the prior to provide sufficient deviation from the mean. The third choice of prior is given by the solid line in Figure 2. It is centered at 2×10^{-5}, which can match the lowest between variability suggested by the data. Such a prior is expected to substantially increase the shrinkage toward the prior mean in order to address the situation where information is heavily shared within the fund family.

We also designate prior distribution to other parameters in the learning model. The prior on

$\theta_{j,1}$ is centered at zero, as we assume that no manager is found to have superior stock selection ability. This is consistent with the settings given by PS and JS. The prior on the rest of the kth parameters in Θ are centered at one. The prior on the correlation matrix Φ is given by an inverse Wishart distribution with a degree of freedom that is higher than the dimension of its scale matrix. This setting allows Φ to have a uniform prior distribution on the correlation parameters, since information regarding the correlation among family-level predictors is normally unknown by assumption. Moreover, the prior settings for the correlation matrix and the standard deviation for the kth pricing factor considered in our research are different from those discussed in Barnard et al. (2000), where the group-level covariance matrix is decomposed into product of the correlation matrix and the diagonal matrix of standard deviations. Then a certain prior can be allocated on the particular predictor with a marginal uniform prior on the correlation parameters. But our technique can achieve the same objective with simpler computation and faster convergence.

<Please insert Figure 1 here>

<Please insert Figure 2 here>

3 Simulation analysis

3.1 Simulation with returns of hypothetical fund family

In this section we report the simulation results of the learning models under the various prior beliefs chosen in section 2.3. We conduct the simulations based on hypothetical returns generated by different compositions of the benchmark. For the CAPM model, the abnormal performance α is assumed to have a normal distribution with mean of 7.1×10^{-4} and standard

deviation of 2.23×10^{3}, and β_{market} has a normal distribution with a mean value of 0.979 and standard deviation of 0.087. The standard deviation of the error term follows $log(\sigma) \sim N(-4.044, 0.398^{2})$. The 3-factor model α follows a normal distribution with a mean of 0.0767% and standard deviation of 0.226, and β_{market} follows $N(1.0293, 0.122^{2})$, while β_{hml} and β_{smb} follow $N(-0.036, 0.093^{2})$ and $N(0.053, 0.086^{2})$, respectively. The residual standard deviation has $log(\sigma) \sim N(-4.079, 0.378^{2})$. For the 4-factor model, α is drawn from a normal distribution with mean of 0.0806% and standard deviation of 0.231%, and β_{market} is assumed to follow $N(1.029, 0.121^{2})$. The other pricing factors are drawn independently from the following distribution: $\beta_{hml} \sim N(-0.037, 0.093^{2})$, $\beta_{smb} \sim N(0.0523, 0.0844^{2})$, $\beta_{mom} \sim N(0.006, 0.017^{2})$. The standard deviation of the error term $log(\sigma)$ is drawn from the log-normal distributed with $N(-4.086, 0.376^{2})$. The returns and pricing factors are then drawn independently of each other to form returns for a particular fund.

Table 1 reports the posterior simulations from three types of learning model under the chosen prior beliefs using hypothetical fund returns. All the three types of learning model exhibit a substantial degree of shrinkage with a rapid decrease in dispersion of the cross-sectional variability in funds' α, that is, λ_{α} declines sharply under the chosen prior beliefs. The results from the general learning model seem to have the highest degree of shrinkage, since the λ_{α} in Panel C is lower than those in Panel A by over 60 basis points and lower than the partial learning model by almost 90 basis points under the high skepticism prior. It is also worth noting that the degree of shrinkage decreases considerably across the prior beliefs when using both the partial learning model and the non-learning model. Specifically, λ_{α} from the CAPM model drops by

almost 1000 basis points in Panel B, from 0.1234 to 0.0101, while it behaves more stably in Panel C with only an 80-basis point change. The general learning model incorporates the prior belief on the variance from both the pricing factors and it also works as the scale factor in the denominator of the posterior mean, thus the prior belief is more likely to have significant influence on the cross-sectional variability of fund alphas.

Figures 3, 4 and 5 provide further evidence to confirm the shrinkage. Figure 3 illustrates the boxplot of the posterior mean of the α for the 5 hypothetical funds. The value is quite dispersed when $\log(\xi_k)$ has a diffuse prior, and the median of each α's posterior distribution is close to the OLS estimates since a highly close-diffuse prior would mitigate the influence from the prior mean. The dispersion on α reduces significantly when turning to a less diffuse prior. In the extreme case where $\log(\xi_k)$ has the high skepticism prior, the funds' α converges to a common mean which is close to zero. There is also some evidence supporting the notion that the shrinkage is sensitive to the evaluation model chosen. The boxplot shown in Figure 4 suggests that the α estimated by the 3-factor model has a low degree of shrinkage under all three prior beliefs compared to those from the CAPM. Similar results can also be found in Figure 5 where the 4-factor model is considered.

Table 1 also reports the results of the mean performance for a particular fund family, θ_α. As shown in Section 2.1.4, the posterior mean of θ_α is weighted by both the OLS estimates and prior information. Since we apply no predictors at the family level likelihood, Eq (2), X_j is assumed to be an identity matrix for all M funds. Thus Eq (2) is by design a sum of family's mean performance and the fund's idiosyncratic performance. In the case where Λ has diffuse prior beliefs on its diagonal, each of the elements in **B** should converge to its OLS estimates.

On the other hand, when a least dispersed prior is considered for $\mathbf{\Lambda}$ and $\mathbf{B}$ should reduce to its mean, $\mathbf{\Theta}$. We therefore expect the factor loadings and the α within the same family to converge to a common mean which can be attributed as the mean performance of the fund family. One may argue that the fund manager can also contribute to the mean performance; thus a feasible extension to the general learning model is to further decompose the mean value $\mathbf{\Theta}$ and to designate particular predictors representing the difference between the contribution from the manager and that from the fund family.

<Please insert Table 1 here>

The posterior mean of θ_α reported in Table 1 seems to be very close to zero across all the learning models considered. The results document some weak support for a decreasing pattern of the value of θ_α with a diminishing dispersion on the prior variance, i.e. it reduces from 0.17% to 0.08% when the general learning model based on the 4-factor model is considered. The value is even lower when using a CAPM based partial learning model. This may be explained with the aid of Figure 3, in which the posterior α of each fund is more concentrated around zero and the extreme values at both ends offset each other. The distribution of each fund's α in Figure 4 has a more extreme value at the positive end, implying a more positive θ_α in a 3-factor based learning model. A similar situation can be found in Figure 5, where the 4-factor based learning model is considered, i.e. the median of the posterior α is further from zero compared to the others. Consequently, θ_α provides additional information on the common performance across funds within the same fund family. Our simulation results suggest that the abnormal performance

can be attributed mainly to funds' idiosyncratic behavior, since the common mean reduces to zero under the least dispersed prior.

Figure 6 illustrates the posterior mean of market beta in the general learning model. Since we put non-informative prior on **B**, its posterior mean is expected to converge to the OLS estimates. The boxplot in Figure 6 shows a steady shrinkage across the chosen prior beliefs. We further extend the simulation to incorporate the influence of informative prior beliefs on other pricing factors. The results are discussed in Table 3.

<Please insert Figure 3 here>

<Please insert Figure 4 here>

<Please insert Figure 5 here>

<Please insert Figure 6 here>

3.2 Simulation with returns of hypothetical funds universe

We consider a more extreme case where, instead of considering a hypothetical family with 5 funds, we enlarge the sample size to incorporate 200 funds to analyze the degree of shrinkage of funds' α. Results are reported in Table 2. The posterior mean of δ_α suggests that of the three models, the general learning model exhibits the highest degree of shrinkage of funds' α, which is consistent with the results reported in Table 1. The posterior means of λ_α given by the non-learning and the partial learning models have a similar value under the same prior. Moreover, compared to the results in Table 1, simulations with a larger group of funds produces smaller value of λ_α for a given prior, indicating that it becomes easier to converge to the common mean

when they are able to gain information from more funds. However, there is only weak evidence to support the declining pattern of λ_α with less dispersed prior beliefs. This is because the growing sample size may lead to more heterogeneous beliefs among individual funds' alpha, which therefore slows down the efficiency of the convergence process.

The posterior means of θ_α reported in Panels A and B are higher than those in Table 1. Meanwhile, we find that funds' idiosyncratic performance seems to have limited impact on the overall mean performance, since θ_α remains almost unchanged across different prior beliefs from both the partial learning and the general learning models. This implies that managers' superior (inferior) performances offset each other in a large funds population, and such mean performance is also independent of the prior information. The steady nature of θ_α documented in Panel B is apparently different from that discovered by JS. This is because we only incorporate prior beliefs in the cross-sectional variability, and leave the prior on **B** non-informative, whereas JS put decreasingly dispersed prior on the group mean, and find that the posterior mean is driven toward zero.

Figure 7 plots the density of the posterior mean of the 200 funds' alpha with respect to the chosen prior beliefs. Not surprisingly, the solid line, which indicates the density of α given a close-diffuse prior, has the lowest degree of kurtosis among the three densities, while the dashed line has more values around its mean. However, Figure 7 shows there is a limited margin on the shrinkage level between different prior beliefs, which is consistent with the results found in λ_α in Table 2. Furthermore, it seems that more funds are found to have a positive α given a left skewed density no matter which prior belief is chosen. Our results in relation to the simulation of returns of hypothetical fund universe suggest a low degree of shrinkage of the cross-sectional α

compared to that found in fund families. But a steady mean performance, θ_α, implies a feasible estimate of the mean performance for the fund universe through the general learning model.

<Please insert Table 2 here>

<Please insert Figure 7 here>

The general learning model enables us to consider specific prior beliefs on factor loadings. Using this property, we further extend our simulation to allocate information on all of the elements in **B** in addition to α. In other words, the factor loadings of each market benchmark in a particular pricing model are assumed with informative prior beliefs before undertaking the estimation. Table 3 reports the results.

We find that the chosen prior belief on other pricing factors can influence the posterior dispersion of α. For example, the posterior mean of λ_α increases from 0.035% to 0.118% when the prior belief on k^{th} pricing factors turns to a dispersed one. This finding is also consistent when different pricing model is considered. From the fund family's perspective, if investors are only certain that funds in the same family have similar risk exposure to the market benchmarks, they may choose more concentrated prior beliefs on the corresponding factor loadings. On the other hand, investors might have limited knowledge on the overall skill of the fund family, and hence they choose a more dispersed prior belief on α. Such a situation can be represented by column 3 of Table 3 under the settings of CAPM. The results imply a slightly lower degree of shrinkage of λ_α relative to those reported in Panel C of Table 1, indicating that adding prior information from other pricing benchmarks can improve the shrinkage of the α

value. In other words, if the prior information considered happens to be correct, the general learning model can provide a more precise estimation of the cross-sectional mean performance of the family. Intuitively, the situation discussed above could be a more common case in reality. Since fund companies publish their top holdings frequently, investors are more likely to form their own opinions on the co-movements between the fund and the market portfolio. Hence, a less dispersed prior can be used to represent the investors' belief before seeing the data. However, it is often the case that the manager's stock selection skill is unknown to the investors. Therefore a diffuse prior on α could be a reasonable setting.

Moreover, we consider another extreme situation, in which investors are more convinced that the fund family contains no skilled managers, but they are also unsure that the market benchmark can completely price the fund return. Therefore, a highly concentrated prior is defined on both the α and the market beta. Such a scenario is considered in the second column of the CAPM settings. The result shows a significant degree of shrinkage of the cross-sectional market beta as λ_{β} decreases with less dispersed priors than that in the fourth column. However, compared with the results on λ_{α} given by Panel C of Table 1, where market beta has a diffuse prior, we find that λ_{α} increases by over 10 bps. This is because the general learning model provides a compromise estimation of λ_{α} between the real data and the prior belief, since the hypothetical returns still contain evidence to support the existence of skilful managers. Thus, the posterior cross-sectional variability on α increases to signal such concerns.[7]

In Table 4, we look further into the non-equal prior problem by computing the posterior

---

[7] Unlike the hypothetical returns generated by JS, in which the abnormal return has been centered to have a zero mean, we draw the abnormal returns from $N(0.071\%, 0.223\%^{2})$, which matches the general empirical findings in the real fund industry.

correlation coefficients of the parameters considered in Table 3. In general the correlation between the different pricing factors and the abnormal returns remains at a low level. However, this does not contradict the results found in Table 3, since we place only a diffuse prior on the correlation matrix of all the pricing factors. It is not only the correlation coefficient but also the cross-sectional variability of a certain pricing factor that can decide the learning outcome. Therefore, such low correlation coefficient can suggest a low level of cross-fund learning only in the correlation itself. Our method provides a way to define an informative prior on the correlation matrix. However, the construction of an efficient prior remains an open question in the statistics literature.

<center><Please insert Table 3 here></center>

<center><Please insert Table 4 here></center>

3.3 Simulation with the universe of real funds

In this section, we consider the simulation using the returns from the actual mutual funds. We select monthly returns of 220 unit trusts from 47 fund families in the UK fund industry from 2001 to 2010. All the sampled funds are UK equity unit trusts. We screen out the non-equity funds and the mixed funds since our performance evaluation focuses only on fund managers' stock selection skill. Within each of the fund families, we also screen out the new funds due to splitting and keep funds with longest return history for each share class. To focus solely on the domestic funds, the funds in our sample are all UK domicile equity funds, indicating that most of their capital should be

invested in UK companies. Meanwhile, the UK domicile funds share the similar market benchmarks, which facilitate estimation of the funds' alphas by the factor models.

We employ three sets of benchmark returns to form the baseline performance evaluation model. We choose the FTSE All Shares as the excess market return factor motivated by CAPM. The returns of the additional size and book to market factors in the Fama French 3-factor model are computed by two pairs of market portfolios: the size factor is generated by the difference between the FTSE 100 index and the FTSE small capital index; the book to market factor is calculated by taking the difference between the MSCI UK Growth index and the MSCI UK Value index. The returns of the additional momentum factor in the 4-factor model are generated by using the 1-year high return portfolio minus the low return portfolio.

Firstly, in Table 5 we consider the situation in which informative prior beliefs are only given to the within variability of the cross-sectional alphas, δ_α. In other words, the investors are presumed to have prior information on how individual alphas deviate from each other within the same fund family, which is similar with the settings considered in the fake data simulation. We include the simulation results given by the non-learning model with independent prior beliefs, the partial learning model with dependent prior beliefs only on funds' alphas, and the general learning model, which we design to provide a full Bayesian treatment on each of the pricing factors. For each type of learning model, we conduct the simulation through three different baseline evaluation models, i.e. CAPM, 3-factor model and 4-factor model. The prior beliefs selected for λ_α are the same as those implemented in the fake data simulation. In addition, since no information is given on the mean performance of the fund family or on the family mean value of other pricing factors, we apply a diffuse prior distribution on the prior variance of θ_α, θ_β,

θ_{HML}, θ_{SMB} and θ_{MOM}, and the prior means are centered at 0. The scale parameters on each of the pricing factors, except those on θ_α, all have diffuse priors.

The posterior mean of λ_α, which indicates how individual funds deviate from the family mean, decrease rapidly with the increase in skepticism on both the skill level and the within variation. Compared with the value of λ_α from the non-learning and partial learning models, the general learning model seems to be more sensitive to the priors chosen. λ_α in the partial learning model is about 20 basis points higher than that from the general learning model under the low skepticism prior. The difference is even larger under the diffuse prior, but they all turn to zero when a high skepticism prior is given. The results in Table 5 also suggest that such a decreasing pattern is not sensitive to the model specification, since the difference in magnitude is robust in all the types of baseline evaluation models.

The posterior mean of the family mean performance θ_α, and individual funds' alphas reported by Panels B and C of Table 5, also experience a decrease in value with the increasing skepticism in the prior beliefs, indicating that the prior information on family's mean performance would alter investors' view of individual funds' performance. Particularly, the funds' alphas reduce to θ_α when the high skepticism prior belief is applied, because λ_α approaches 0 under the least dispersed variance and the cross-sectional variation among alphas is almost eliminated within the same fund family. Although the prior mean of θ_α is centered at 0, its diffuse prior variance mitigates the influence from the prior mean and allows the real data to dominate the posterior distribution of θ_α. Meanwhile, Table 5 also provides some evidence to support the presence of managers' skill. By the assumption of the general learning model, the

difference between θ_α and alphas under the diffuse and low skepticism priors can be regarded as the gain from funds' cross-sectional variation. Panels B and C both document that the funds' average alphas exceed θ_α for more than 20 basis points under the diffuse and low skepticism priors for all types of baseline evaluation model. However, since we provide no further decomposition on the family mean performance, we presume that apart from the fund families, θ_α might still incorporate a contribution by individual fund managers. But such a portion in θ_α should be limited, since for each fund family we keep only one fund for each share class, in order to maintain the variety of funds with distinct investment objectives in a fund family. One may argue that θ_α should maintain a stable value instead of decreasing with the skepticism prior. Since Eq (7) suggests that the posterior belief of Θ is a weighted average of prior information and the real data, the posterior mean of Θ is conditional only on the posterior distribution of the in-family covariance matrix when $\Delta^{-1} \to 0$ and $\Lambda^{-1} \to 0$. Thus, the posterior mean of Θ may also shift with the changing value of the prior belief. However, given a high skepticism prior the prior variance approaches zero, and can hardly affect Θ, which drives the average alpha to θ_α.

Baks (2003) provides an alternative way to extract the family contribution out of funds' individual alphas through a Cobb-Douglas production function by denoting arbitrary weights on the performance of managers and fund organizations, respectively. The performance attribution is therefore sensitive to the weights chosen. Moreover, we are not surprised to see that the posterior means of β and θ_β documented in Table 5 remain almost unchanged, since no informative prior beliefs are applied to both θ_β and λ_β throughout the simulation.

<Please insert Table 5 here>

We further investigate how prior information from other pricing factors affects the posterior distribution of the cross-sectional alphas and θ_α, by placing prior beliefs simultaneously on θ_α and the family mean and the scale parameters of all the other pricing factors. This is also an important feature of the general learning model, given that it enables us to denote specific prior beliefs on each of the pricing factors in the baseline evaluation model. The priors on the scale parameters are similar to those discussed in Table 5. The family mean value of each of the pricing factors (including θ_α) are also assigned with prior beliefs to address the learning issue, i.e. the kth element in vector $\boldsymbol{\Theta}$ is set to have $\theta_k \sim (0,100)$ as the diffuse prior; for the low skepticism prior we set $\theta_k \sim (0,1)$; for the high skepticism prior we have $\theta_k \sim (0,0.001)$. Table 6 reports the posterior results of the parameters of interest.

We find a similar decreasing pattern in λ_α with the increasing level of skepticism in the prior beliefs. Such a pattern is also robust throughout different baseline evaluation models. However, the in-family variation on $\mathbf{B}$ seems to increase with the prior belief, e.g. λ_β of CAPM equals to 0.117 under the diffuse prior and it increases to 0.692 given the high skepticism prior. A possible reason is that the prior beliefs we apply are far below the real in-family variance of the market beta, which makes the MCMC simulation hard to converge. Because of the power in place of the extreme priors, most of the posterior $\mathbf{B}$ shrinks towards the prior mean, leaving several outliers which enlarge the posterior in-family variance. However, the posterior correlations between each pair of the pricing factors in Table 7 are too low to affect the convergence of other pricing factors.

The averaged alpha and beta both experience a decrease with the increasing level of

skepticism on a larger scale than those reported in Table 5, for the reason that prior information is included on both the family mean and the in-family variation. For instance, alpha equals to 0.495% in Panel C of Table 5 given a high skepticism prior, while it is 0.366% when prior belief on θ_α is included. This finding is robust in all the types of baseline evaluation models we consider. Our simulation in Table 6 validates that performance evaluation of individual funds can be affected by including information on the prior view of the mean performance from the family as a whole, as well as the variation of performance among funds within the fund family. Given the situation that the sets of prior beliefs on the pricing factors do provide additional information regarding the population of returns for a particular fund family, i.e. risk shifting in different market condition, adjustment in investment strategy when facing new information or engaging in tournament among fund managers within the family, the general learning model can incorporate this information so as to provide a more precise evaluation result.

<Please insert Table 6 here>

However, we find no strong evidence to support the presence of cross-factor learning in the general learning model during the simulation. The averaged posterior correlations between alphas and market betas under the three sets of prior beliefs are reported in Table 7. The posterior correlation, $\rho_{\theta_\alpha,\theta_\beta}$ remains at a very low level at all times, indicating that the prior information of other pricing factors has no substantial impact on the changes of the posterior family mean performance. But such a low correlation does not affect the outcome of learning, since as mentioned previously, the posterior mean of λ_k is conditional on the covariance matrix, which includes the prior information on the in-family variation and family mean value of all pricing

31

factors. Therefore, if correlation among the mean value of different pricing factors can be omitted, the prior information can be applied to family mean value directly, which can significantly speed up the convergence of the Markov chain, but bears the loss of the co-movements of the pricing factors. On the other hand, we place no informative prior on the correlation matrix in the simulation, i.e. an inverse Wishart distribution, $W^{-1}(K+1, \mathbf{I}_{K \times K})$, is applied on the correlation matrix to represent a uniform prior on the correlation. It would certainly be possible to include an informative prior on the correlation matrix to address the dependence issue of the pricing factors if necessary. However, such a setting might involve denoting specific correlation among different market portfolios, which is beyond the scope of this research.

<Please insert Table 7 here>

To provide further insight into the slow shrinkage on λ_β detected above, we further extend the research to analyze the posterior shrinkage from an empirical Bayes perspective. We denote an informative prior on the family-level mean of both α and each of the pricing factors in Θ. Such prior beliefs are initially given by the historical cross-section value and then updated by the previous generated posterior mean. Specifically, we use the fund returns in 2001 to compute the OLS estimation of α and market beta for each fund, then the cross-sectional in-family mean can be computed. These values are applied to the general learning model as the initial prior for Θ to simulate the posterior distribution in 2002, then the prior is updated by the newly generated posterior mean of α and market beta for the simulation of the following year. We also consider two groups of prior settings for the in-family shrinkage level on ξ_α and ξ_β

to utilize the full Bayesian treatment of the general learning model, i.e. (1) low skepticism prior on both cross-sectional α and market beta; (2) low (high) skepticism prior on α (market beta).

Figures 8a and 8b illustrate the averaged posterior mean of λ_α and λ_β for each year under the two groups of prior settings. Both the cross-sectional in-family α and market beta experience significant levels of shrinkage compared with the OLS estimation (solid line), since the estimation is driven by the skepticism prior. The posterior mean of λ_α under both prior settings indicates a similar pattern to that of the OLS value in Figure 8a, and the dotted line, which represents the posterior belief under Prior setting (1), closely matches the estimated value under Prior setting (2), the dashed line. Since a low skepticism prior is considered in the simulation of α, we are not surprised to see that investors are more likely to believe in a similar value of α within the same family. Meanwhile, given a moderate prior on λ_α, the actual data still have substantial power to lead the posterior value to follow a similar changing pattern.

Unlike the results given by Figure 8a, the posterior mean of λ_β in Figure 8b deviates from the conventional estimation when a high skepticism is considered, i.e. the dashed line moves toward the opposite direction in the years 2003, 2004, 2007 and 2009. This is due to the fact that the strong prior belief enables the posterior mean to mitigate the increasing volatility on the cross-sectional market beta. However, it is often the case that the observed data strongly disagree with the prior belief, particularly when the diffuse estimation experiences a significant deviation from the family-level mean over the period 2005 to 2007 (solid line). The posterior distribution of λ_β therefore contains more value in the right tail to promote the increase in the mean. In other words, the posterior simulation absorbs the information given by the observed data to

provide a compromise with the strong prior. If the investors are more convinced that the funds within the same family might have similar market risk exposure, but the observed return delivers highly dispersed market beta due to risk shifting or portfolio reconstruction. We then expect the model to present a decreasing level of shrinkage, in order not to overstate the influence of the learning process.

<Please insert Figure 8 here>

4. Conclusion

In this research, we devote our attention to the analysis of how returns from other parallel funds affect the alpha of particular funds within the same fund family. We consider a general learning model in a Bayesian framework to incorporate the additional information given by other funds in the prior beliefs. We decompose the Jenson alpha as well as the loadings of each market portfolio in the factor pricing model into the combination of a family mean value and the fund's idiosyncratic variation. The family mean value represents the investors' opinion on the cross-sectional mean of both alpha and factor loadings, while the in-family variation addresses how parameters from the individual fund deviate from the family mean.

To simulate this general learning model we construct the combined Gibbs samplers with the Metropolitan Hastings algorithm by using data given by the monthly NAV from the UK domicile equity fund. We incorporate three sets of prior belief to simulate the possible prior information on the family mean of each pricing factor and their in-family deviation. The simulation results suggest that the posterior mean of in-family variation decreases given a less dispersed prior belief, indicating that individual funds' alphas might concentrate around their

family means if prior information implies a serious lack of skilful managers. Moreover, we find that the general learning model is more sensitive to the chosen prior belief, compared to the non-learning and the partial learning model discussed by JS. Thus, the higher level of shrinkage from our model can better address cross-fund learning.

The general learning model can also provide a compromise of performance evaluation between the observed returns delivered by funds and additional information on how other funds behave in the same fund family. The proposed model utilizes the full Bayesian treatment by specifying the prior information on the certain pricing factors, which enables the incorporation of investors' view on different family strategies in the performance evaluation. Since most of the family strategies would involve allocating more capital to certain funds or encouraging fund managers to compete with each other, i.e. family tournament, star fund phenomenon, and family favoritism, which may lead to an increase of the cross-sectional variability among alphas and other factor loadings, the prior beliefs can be used to simulate these strategies or to capture the pattern of in-family risk shifting implied by the historical data.

While a separation strategy is considered to enable the full Bayesian treatment, a uniform prior is denoted on the correlation matrix of the family mean to simplify the algorithm. One extension of the research therefore, would be to further incorporate the prior belief on the strength of the correlation among the family-level means of fund alphas and different pricing factors. This would require more advanced settings for the correlation matrix, which could take the form of those discussed in Liechty, Liechty and Müller (2004).

Despite the efforts to consider a simple empirical Bayes setting in forming the prior beliefs, as shown in Figure 8, our research could also be extended by incorporating the process of prior elicitation discussed in several studies, i.e. PS, Baks et al. (2001) and Busse and Irvine (2006).

Certain prior beliefs, representing investors' specific views regarding the managers' risk taking, performance persistence, external market conditions or managers' style, might provide further insights into the interplay between investors' behavior and fund performance.

References

ALLEN, D. E. and TAN, M. L. (1999) A Test of the Persistence in the Performance of UK Managed Funds. *Journal of Business Finance & Accounting* 26(5-6)**:** 559-593.

BAKS, K. P. (2003) On the Performance of Mutual Fund Managers. US: Emory University.

BAKS, K. P., METRICK, A. and WACHTER, J. (2001) Should Investors Avoid All Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation. *The Journal of Finance* 56(1)**:** 45-85.

BARNARD, J., MCCULLOCH, R. and MENG, X.-L. (2000) Modeling Covariance Matrices in Terms of Standard Deviations and Correlations, with Application to Shrinkage. *Statistica Sinica* 4(10)**:** 1281-1311.

BOLLEN, N. P. B. and BUSSE, J. A. (2005) Short-Term Persistence in Mutual Fund Performance. *Review of Financial Studies* 18(2)**:** 569-597.

BUSSE, J. A. and IRVINE, P. J. (2006) Bayesian Alphas and Mutual Fund Persistence. *The Journal of Finance* 61(5)**:** 2251-2288.

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

CUTHBERTSON, K., NITZSCHE, D. and O'SULLIVAN, N. (2008) UK mutual fund performance: Skill or luck? *Journal of Empirical Finance* 15(4)**:** 613-634.

ELTON, E. J., GRUBER, M. J. and BLAKE, C. R. (1996) The Persistence of Risk-Adjusted Mutual Fund Performance. *Journal of Business* 69(2)**:** 133-157.

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.

GASPAR, J.-M., MASSA, M. and MATOS, P. (2006) Favoritism in Mutual Fund Families? Evidence on Strategic Cross-Fund Subsidization. *The Journal of Finance* 61(1)**:** 73-104.

GELMAN, A. and CARLIN, J. B. (2011) *Bayesian data analysis.* Boca Raton: CRC Press.

GELMAN, A. and HILL, J. (2007) *Data analysis using regression and multilevel/hierarchical models.* Cambridge: Cambridge University Press.

GRUBER, M. J. (1996) Another Puzzle: The Growth in Actively Managed Mutual Funds. *The Journal of Finance* 51(3)**:** 783-810.

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.

JENSEN, M. C. (1968) THE PERFORMANCE OF MUTUAL FUNDS IN THE PERIOD 1945–1964. *The Journal of Finance* 23(2)**:** 389-416.

JONES, C. S. and SHANKEN, J. (2005) Mutual fund performance with learning across funds. *Journal of Financial Economics* 78(3)**:** 507-552.

KEMPF, A. and RUENZI, S. (2008) Tournaments in Mutual-Fund Families. *Review of Financial Studies* 21(2)**:** 1013-1036.

KOSOWSKI, R., TIMMERMANN, A., WERMERS, R. and WHITE, H. A. L. (2006) Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis. *The Journal of Finance* 61(6)**:** 2551-2595.

LIECHTY, J. C., LIECHTY, M. W. and MÜLLER, P. (2004) Bayesian correlation estimation. *Biometrika* 91(1)**:** 1-14.

LINDLEY, D. V. and SMITH, A. F. M. (1972) Bayes Estimates for the Linear Model. *Journal of the Royal Statistical Society. Series B (Methodological)* 34(1)**:** 1-41.

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.

NANDA, V., WANG, Z. J. and ZHENG, L. (2004) Family Values and the Star Phenomenon: Strategies of Mutual Fund Families. *Review of Financial Studies* 17(3)**:** 667-698.

O'MALLEY, A. J. and ZASLAVSKY, A. M. (2008) Domain-Level Covariance Analysis for Multilevel Survey Data With Structured Nonresponse. *Journal of the American Statistical Association* 103(484)**:** 1405-1418.

PÁSTOR, Ľ. and STAMBAUGH, R. F. (2002) Mutual fund performance and seemingly unrelated assets. *Journal of Financial Economics* 63(3)**:** 315-349.

PÁSTOR, Ľ. and VERONESI, P. (2009) Learning in Financial Markets. *Annual Review of*

Financial Economics 1(1)**:** 361-381.

SHARPE, W. F. (1964) CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK. *The Journal of Finance* 19(3)**:** 425-442.

SMITH, A. F. M. (1973) A General Bayesian Linear Model. *Journal of the Royal Statistical Society. Series B (Methodological)* 35(1)**:** 67-75.

STAMBAUGH, R. F. (1997) Analyzing investments whose histories differ in length. *Journal of Financial Economics* 45(3)**:** 285-331.
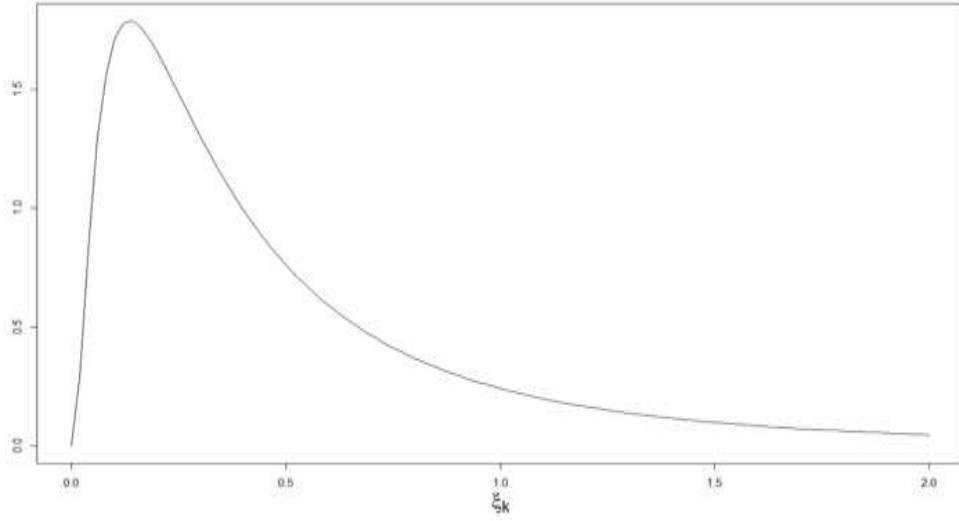
Figure 1: Prior distribution on ξ_k for $log(\xi_k) \sim N(-1,1)$

This figure illustrates the choice of the prior distribution considered for the cross-sectional variability parameter, ξ_k. Its logarithm value has a normal distribution with mean as -1 and variance as 1.
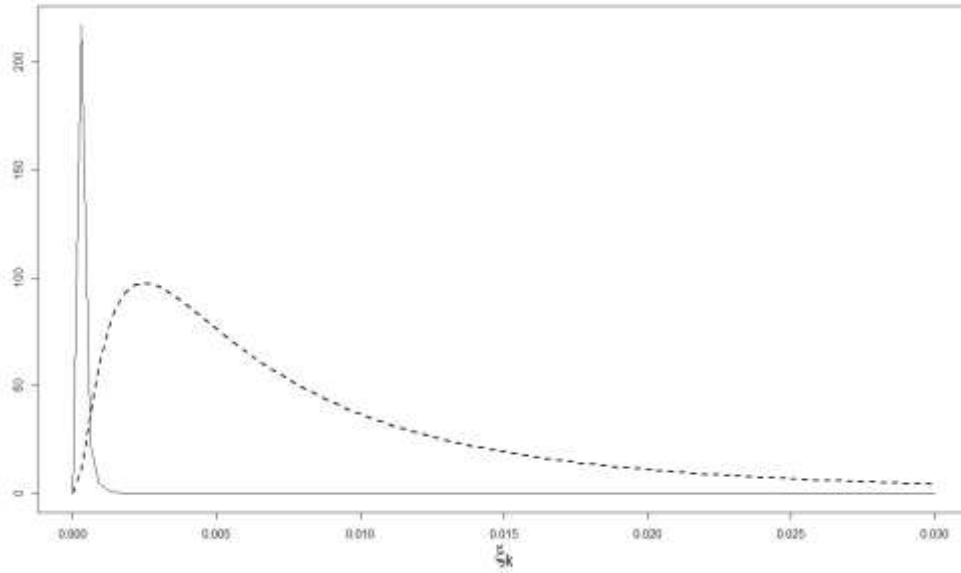
Figure 2: Prior distribution on ξ_k for $log(\xi_k) \sim N(-5,1)$ and
$log(\xi_k) \sim N(-10,1)$

This figure illustrates the choice of the prior distribution considered for the cross-sectional variability parameter, ξ_k. The dashed line represents the distribution of $log(\xi_k) \sim N(-5,1)$, while the solid line is for $log(\xi_k) \sim N(-10,1)$.

Table 1: Simulation of learning within fund family

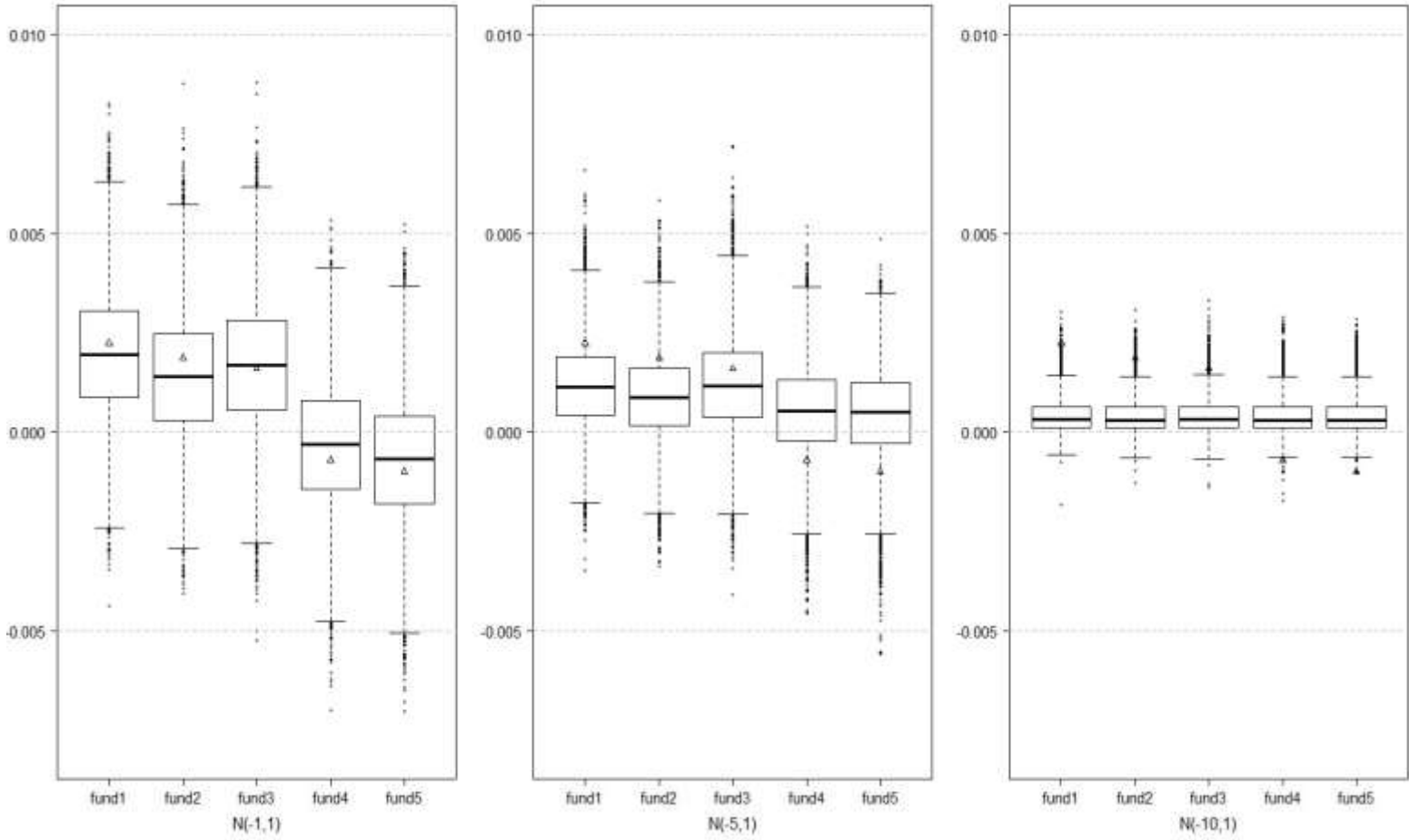| Prior Beliefs	CAPM			3-factor model			4-factor model		
	Diffuse	Low	High	Diffuse	Low	High	Diffuse	Low	High
Panel A				*Non-learning model*					
λ_α	0.0969	0.0574	0.0092	0.0858	0.0411	0.0077	0.0927	0.0529	0.0079
σ_R	0.0157	0.0157	0.0159	0.0194	0.0194	0.0194	0.0168	0.0168	0.0169
Panel B				*Partial learning model*					
λ_α	0.1234	0.0631	0.0101	0.1117	0.0434	0.0076	0.1169	0.0515	0.0079
θ_α	0.0003	0.0001	0.0000	0.0007	0.0008	0.0008	0.0017	0.0017	0.0015
σ_R	0.0157	0.0157	0.0159	0.0194	0.0194	0.0194	0.0167	0.0168	0.0168
Panel C				*General learning model*					
λ_α	0.0082	0.0034	0.0001	0.0058	0.0012	0.0001	0.0069	0.0017	0.0001
θ_α	0.0001	0.0001	0.0002	0.0007	0.0008	0.0004	0.0017	0.0017	0.0008
σ_R	0.0157	0.0157	0.0159	0.0194	0.0194	0.0194	0.0168	0.0168	0.0169

This table presents the simulation results from three evaluation models: the non-learning model, the partial learning model and the general learning model. The posterior mean of the variables, i.e. the in-family variability λ_α, the family level mean performance θ_α, and the fund's individual risk level σ_R, are reported. We control the prior belief on the scaled parameter of the cross-sectional variability in α, ξ_α to be three distinct distributions, i.e. diffuse prior, $log(\xi_\alpha) \sim N(-1,1)$; low skepticism, $log(\xi_\alpha) \sim N(-5,1)$ and high skepticism, $log(\xi_\alpha) \sim N(-10,1)$. The prior belief on the mean value of the kth pricing factor, θ_k, is centered at zero with a diffuse variance. The scaled parameter of θ_β is also assumed to have a diffuse distribution. Panels A, B and C report the simulation results from the CAPM, Fama French 3-factor model and the 4-factor model. The posterior distributions of the variables considered are simulated by the MCMC technique by using hypothetical returns from 5 funds. The fund returns are generated through Eq (1), in which the factor loadings and the market benchmarks are drawn independently across funds. The distribution parameters are chosen to match the empirical results.
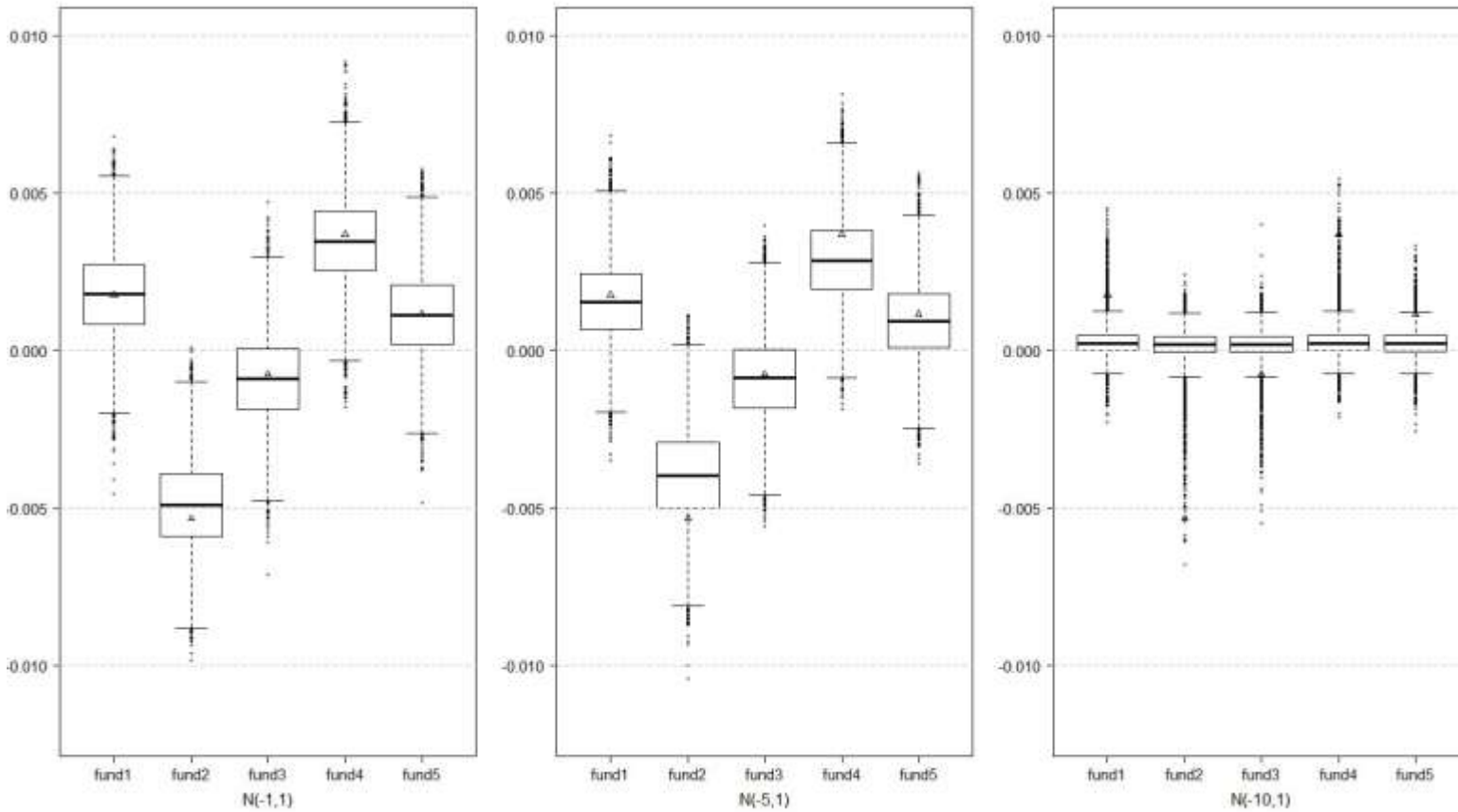
Figure 3: Boxplot of Posterior Draws of CAPM α

This figure illustrates 6000 posterior draws from 5 hypothetical funds' α given the decreasingly dispersed prior beliefs on Λ in the CAPM formed general learning model.

Figure 4: Boxplot of Posterior Draws of 3-factor model α

This figure illustrates 6000 posterior draws from the 5 hypothetical funds' α given the decreasingly dispersed prior beliefs on Λ in the 3-factor based general learning model.
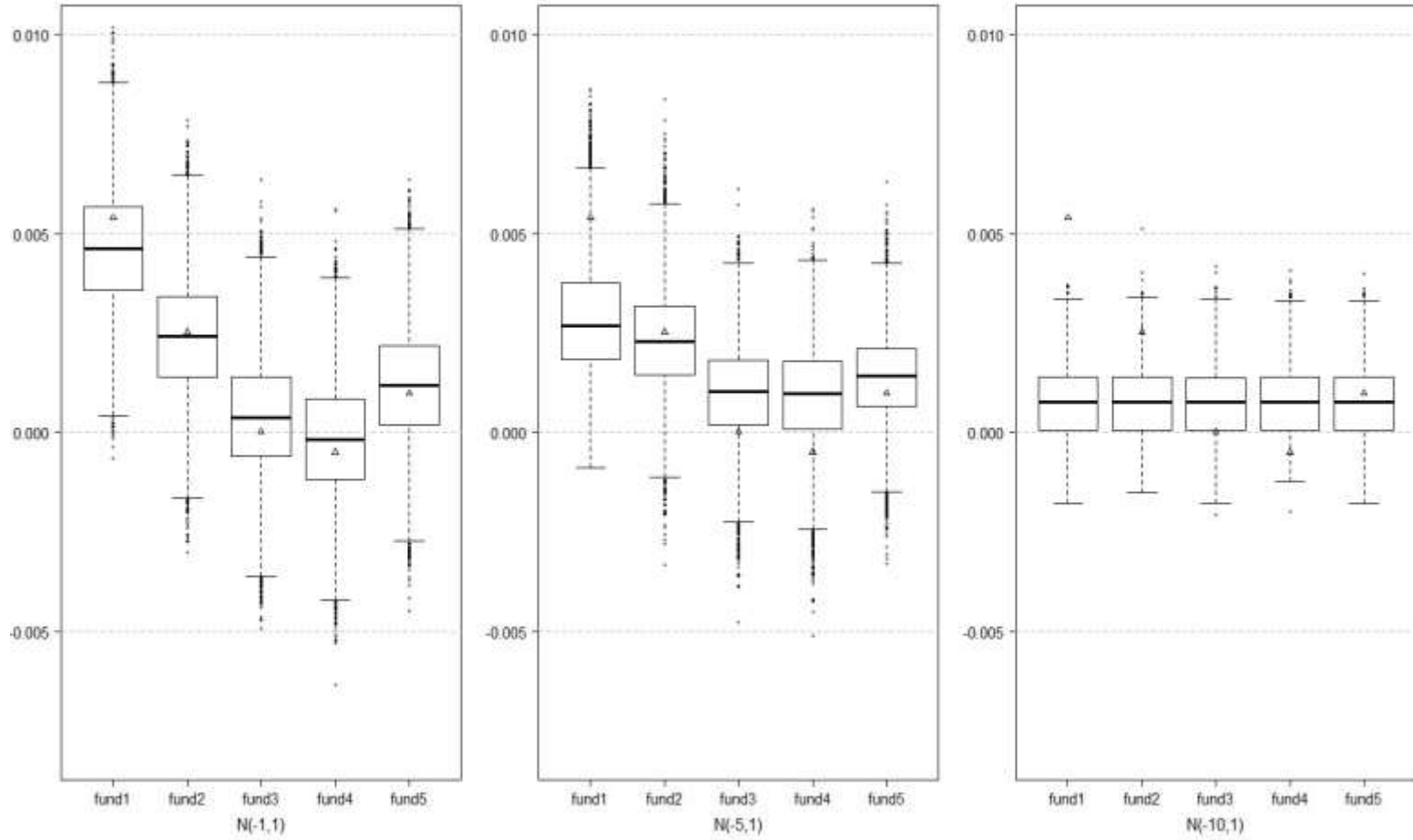
Figure 5: Boxplot of Posterior Draws of 4-factor model α

This figure illustrates 6000 posterior draws from 5 hypothetical funds' α given the decreasingly dispersed prior beliefs on Λ in the 4-factor based general learning model.

Figure 6: Boxplot of Posterior Draws of CAPM β_{market}

This figure illustrates 6000 posterior draws from 5 hypothetical fund's β_{market} given the decreasingly dispersed prior beliefs on Λ in the CAPM formed general learning model.
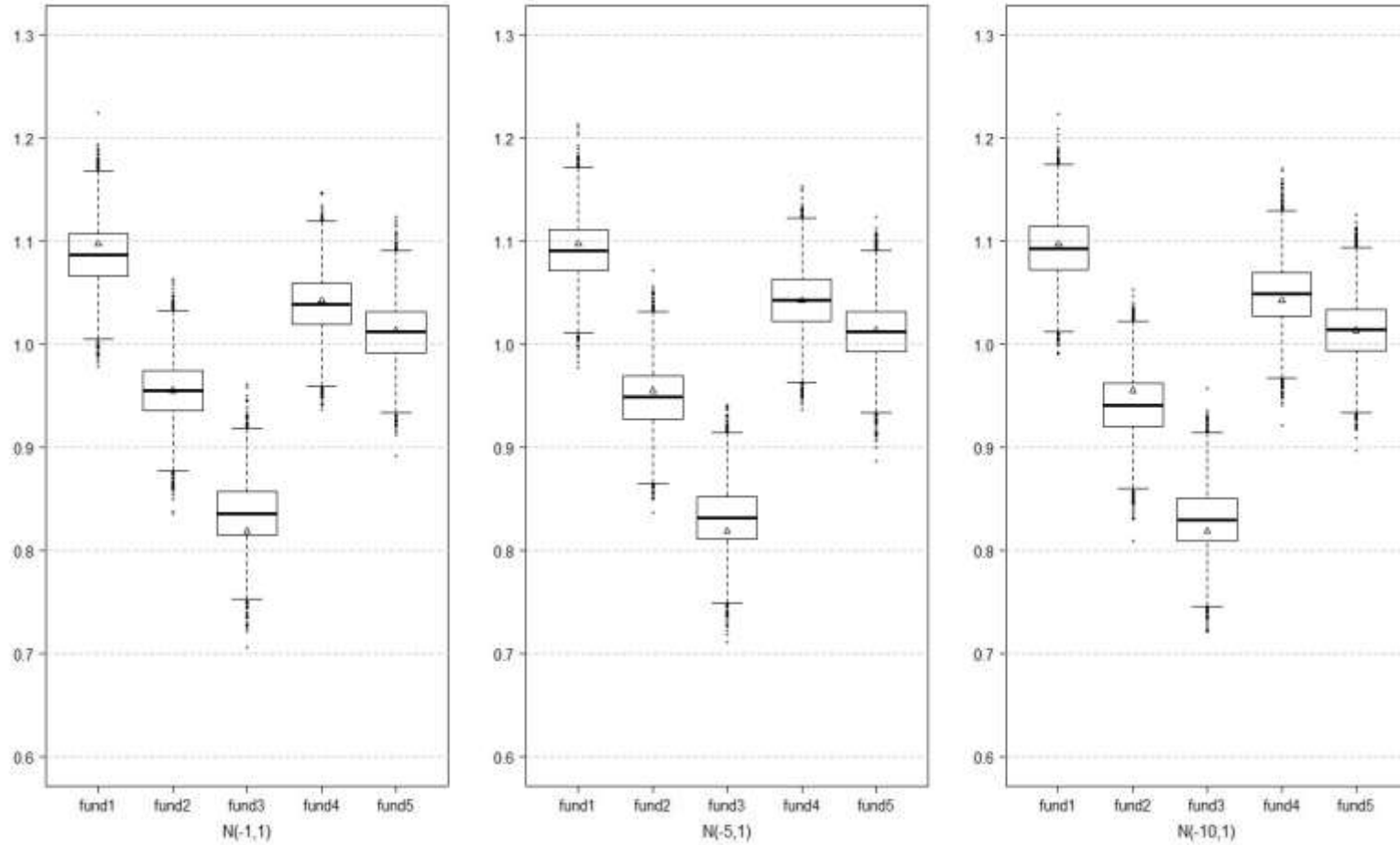
Table 2: Simulation of learning across funds universe

Prior Beliefs	CAPM			3-factor model			4-factor model		
	Diffuse	Low	High	Diffuse	Low	High	Diffuse	Low	High
Panel A					*Non- learning model*				
λ_α	0.05020	0.04956	0.04876	0.04950	0.04888	0.04819	0.05188	0.05135	0.05058
σ_R	0.02111	0.02111	0.02111	0.01887	0.01887	0.01887	0.01855	0.01856	0.01855
Panel B					*Partial learning model*				
λ_α	0.05002	0.04928	0.04844	0.04792	0.04734	0.04655	0.04979	0.04919	0.04844
θ_α	0.00039	0.00040	0.00040	0.00086	0.00086	0.00087	0.00106	0.00105	0.00106
σ_R	0.02111	0.02111	0.02111	0.01887	0.01887	0.01887	0.01856	0.01855	0.01856
Panel C					*General learning model*				
λ_α	0.00248	0.00242	0.00238	0.00228	0.00222	0.00217	0.00247	0.00242	0.00236
θ_α	0.00040	0.00040	0.00039	0.00086	0.00087	0.00086	0.00106	0.00105	0.00105
σ_R	0.02111	0.02111	0.02111	0.01887	0.01887	0.01887	0.01856	0.01856	0.01856

This table presents the simulation results from three evaluation models: the non-learning model, the partial learning model and the general learning model. The posterior mean of the variables, i.e. the in-family variability λ_α, the family level mean performance θ_α, and the fund's individual risk level σ_R, are reported. We control the prior belief on the scaled parameter of the cross-sectional variability in α, ξ_α to be three distinct distributions, i.e. diffuse prior, $log(\xi_\alpha) \sim N(-1,1)$; low skepticism, $log(\xi_\alpha) \sim N(-5,1)$ and high skepticism, $log(\xi_\alpha) \sim N(-10,1)$. The prior belief on the mean value of the kth pricing factor, θ_k, is centered at zero with a diffuse variance. The scaled parameter of θ_β is also assumed to have a diffuse distribution. Panels A, B and C report the simulation results from the CAPM, Fama French 3-factor model and the 4-factor model. The posterior distributions of the variables considered are simulated using the MCMC technique based on hypothetical returns of 200 funds. The fund returns are generated through Eq (1), in which the factor loadings and the market benchmarks are drawn independently across funds. The distribution parameters are chosen to match the empirical results.
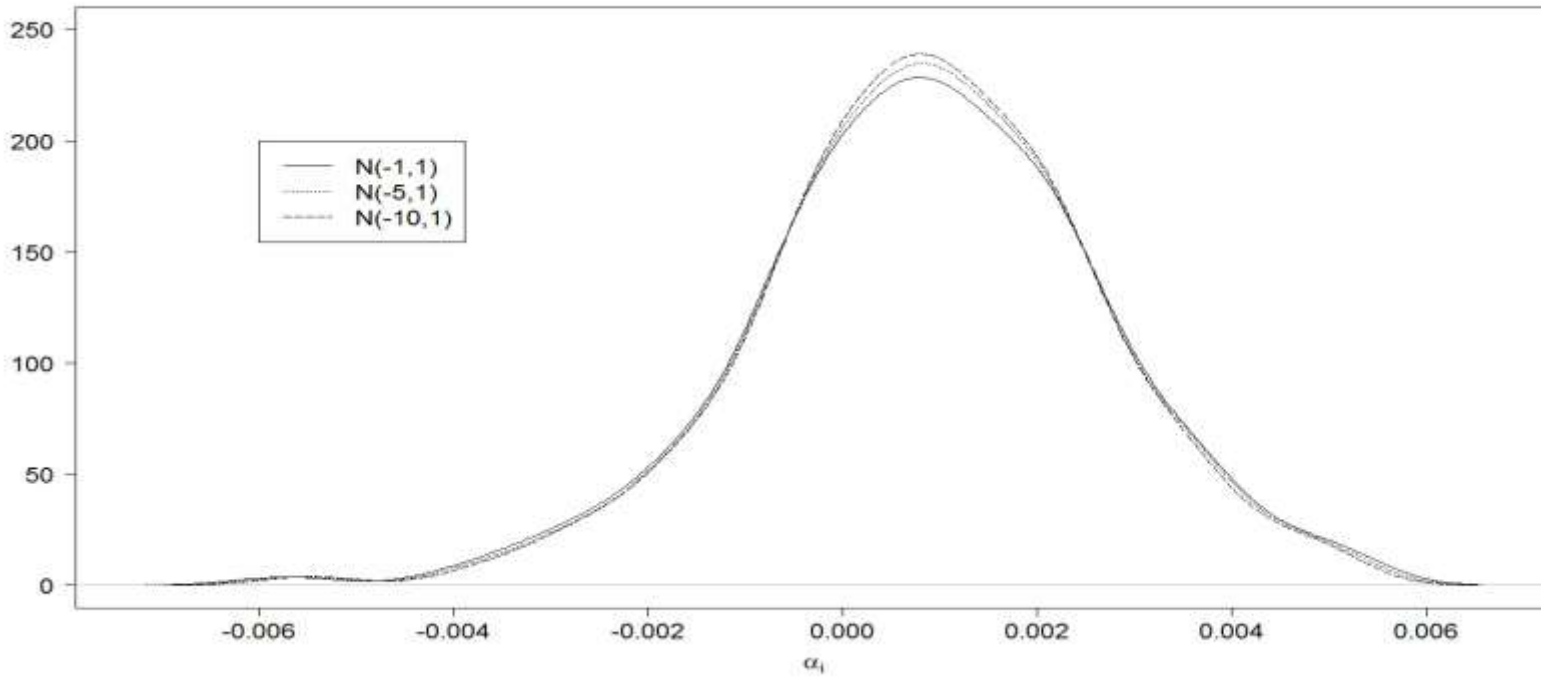
Figure 7: Density of the posterior draws of CAPM α

This figure illustrates 6000 posterior draws from the CAPM α by applying the general learning model to the hypothetical fund population with decreasingly dispersed prior beliefs on Λ.

Table 3: Simulation of learning across funds with non-equal prior

	CAPM			3-factor model			4-factor model		
α prior	N(-10,1)	N(-1,1)	N(-10,1)	N(-10,1)	N(-1,1)	N(-10,1)	N(-10,1)	N(-1,1)	N(-10,1)
k^{th} prior	N(-10,1)	N(-10,1)	N(-5,1)	N(-10,1)	N(-10,1)	N(-5,1)	N(-10,1)	N(-10,1)	N(-5,1)
λ_α	0.00118	0.00779	0.00035	0.00093	0.00553	0.00016	0.00057	0.00598	0.00018
λ_β	0.09256	0.08614	0.09813	0.03127	0.02678	0.05620	0.14649	0.14246	0.15307
θ_α	0.00020	0.00001	-0.00003	-0.00046	0.00058	0.00015	0.00096	0.00157	0.00100
θ_β	0.98026	0.98038	0.98394	1.01606	1.00374	1.00852	1.10090	1.09377	1.09721
λ_{HML}	-	-	-	0.02350	0.04729	0.05253	0.00209	0.00092	0.02231
λ_{SMB}	-	-	-	0.11161	0.10807	0.10445	0.05469	0.03193	0.06296
θ_{HML}	-	-	-	-0.04559	-0.06392	-0.13764	-0.00153	0.00021	-0.01518
θ_{SMB}	-	-	-	0.02872	0.08703	0.09251	0.04877	0.06429	0.06808
λ_{MOM}	-	-	-	-	-	-	0.00096	0.00071	0.00804
θ_{MOM}	-	-	-	-	-	-	0.00117	-0.00020	-0.00251

This table presents the simulation results from the general learning model. The posterior mean of the variables, i.e. the in-family variability on all the pricing factors, λ, and the mean performance of all the pricing factors, θ, are reported. We control the prior belief on the scaled parameters of the cross-sectional variability in factor loadings, ξ_k. The prior beliefs on the mean value of the kth pricing factor, θ_k, are centered at 1 with various prior beliefs. We report results based on three pricing models: CAPM, the 3-factor and the 4-factor models. The posterior distributions of the variables considered are simulated by the MCMC technique by using hypothetical returns of 5 funds. The fund returns are generated through equation 1, in which the factor loadings and the market benchmarks are drawn independently across funds. The distribution parameters are chosen to match the empirical results.

	CAPM			3-factor model			4-factor model		
α prior	N(-1,1)	N(-1,1)	N(-10,1)	N(-1,1)	N(-1,1)	N(-10,1)	N(-1,1)	N(-1,1)	N(-10,1)
k^{th} prior	N(-1,1)	N(-10,1)	N(-5,1)	N(-1,1)	N(-10,1)	N(-5,1)	N(-1,1)	N(-10,1)	N(-5,1)
$\rho_{\alpha,\beta}$	0.19	0.26	0.17	0.01	-0.07	0.02	-0.08	-0.09	-0.02
$\rho_{\alpha,\text{HML}}$	-	-	-	0.08	0.07	-0.08	-0.01	0.01	-0.07
$\rho_{\alpha,\text{SMB}}$	-	-	-	0.10	0.17	-0.09	0.09	0.07	0.04
$\rho_{\alpha,\text{MOM}}$	-	-	-	-	-	-	0.11	0.03	0.04

This table reports the posterior correlation coefficients from the general learning model. We control the prior belief on the scaled parameters of the cross-sectional variability in factor loadings, ξ_k. The prior beliefs on the mean value of the kth pricing factor, θ_k, are centered at 1 with various prior beliefs. We report results based on three pricing models: CAPM, the 3-factor and the 4-factor models. The posterior distributions of the variables considered are simulated by the MCMC technique by using hypothetical returns of 5 funds. The fund returns are generated through equation 1, in which the factor loadings and the market benchmarks are drawn independently across funds. The distribution parameters are chosen to match the empirical results.

<div align="center">Table 5: Simulation of learning within fund family</div>

Prior belief	CAPM			3-factor model			4-factor model		
	Diffuse	Low	High	Diffuse	Low	High	Diffuse	Low	High
Panel A					*Non-learning model*				
α (%)	0.782	0.512	-0.040	0.854	0.592	-0.038	0.904	0.637	-0.035
β	0.977	0.977	0.977	1.025	1.025	1.025	1.025	1.025	1.025
λ_α	0.142	0.050	0.008	0.142	0.050	0.008	0.142	0.050	0.008
Panel B					*Partial learning model*				
α (%)	0.838	0.835	0.156	0.913	0.909	-0.050	0.961	0.961	0.223
β	0.977	0.977	0.971	1.025	1.025	1.025	1.025	1.025	1.021
θ_α (%)	0.618	0.519	0.156	0.658	0.507	-0.050	0.982	0.963	0.223
λ_α	0.210	0.052	0.008	0.211	0.053	0.008	0.211	0.053	0.008
Panel C					*General learning model*				
α (%)	0.845	0.837	0.495	0.913	0.904	0.408	0.967	0.973	0.503
β	0.977	0.977	0.977	1.025	1.025	1.025	1.025	1.025	1.025
θ_α (%)	0.704	0.636	0.495	0.671	0.538	0.408	0.763	0.674	0.501
λ_α	0.023	0.002	0.000	0.024	0.002	0.000	0.024	0.002	0.000
θ_β	0.976	0.977	0.977	1.024	1.013	1.023	1.025	1.023	1.025
λ_β	0.117	0.116	0.133	0.134	0.148	0.132	0.138	0.138	0.136

This table presents the simulation results from three evaluation models: the non-learning model, the partial learning model and the general learning model. The posterior mean of the variables, i.e. the in-family variability, δ_α, the family level annualized mean performance, θ_α, and the cross-sectional averaged annualized alpha, are reported. We control the prior belief on the scaled parameter of the cross-sectional variability in α, ξ_α and priors on θ_α to be three distinct distributions, i.e. a diffuse prior has $log(\xi_\alpha) \sim N(-1,1)$; the low skepticism has $log(\xi_\alpha) \sim N(-5,1)$; the high skepticism has $log(\xi_\alpha) \sim N(-10,1)$. θ_α, θ_β and the scaled parameter of θ_β is assumed to have diffuse prior distribution. Panels A, B and C report the simulation results from the CAPM, 3-factor model and 4-factor model. The posterior distributions are generated by applying the MCMC method on monthly returns from 220 UK unit trusts (47 fund families).

Table 6: Simulation of learning within fund family

λ_k prior	CAPM			3-factor model			4-factor model		
	Diffuse	Low	High	Diffuse	Low	High	Diffuse	Low	High
	General learning model								
α (%)	0.845	0.726	0.366	0.913	0.815	0.492	0.967	0.859	0.627
β	0.977	0.974	0.736	1.025	1.016	0.933	1.025	1.016	0.963
θ_α (%)	0.704	0.511	0.000	0.671	0.412	0.000	0.763	0.374	0.000
λ_α	0.023	0.002	0.001	0.024	0.002	0.001	0.024	0.002	0.001
θ_β	0.976	0.757	0.000	1.024	0.706	0.000	1.025	0.745	0.000
λ_β	0.117	0.321	0.692	0.134	0.391	0.874	0.138	0.363	0.902

This table presents the simulation results from the general learning model. The posterior mean of the variables, i.e. the cross-sectional annualized averaged alpha, the cross-sectional averaged β, the annualized mean performance of alpha (θ_α), the mean performance of β (θ_β) and the in-family variability (λ_α) are reported. We control the prior belief on the scaled parameters of the cross-sectional variability in factor loadings, ξ_k. The prior beliefs on the mean value of the kth pricing factor, θ_k, are centered at zero with a diffuse variance. The scaled parameter of θ_β is also assumed to have a diffuse prior. Panels A, B and C report the simulation results from the CAPM, the 3-factor and the 4-factor models. The posterior distributions of the variables considered are simulated by the MCMC technique by using monthly returns from 220 UK unit trusts.

Table 7: Posterior correlation coefficients

	CAPM			3-factor model			4-factor model		
Panel A λ_α prior	Diffuse	Low	High	Diffuse	Low	High	Diffuse	Low	High
$\rho_{\theta_\alpha,\theta_\beta}$	-0.025	-0.046	-0.005	0.018	0.030	0.007	0.025	0.041	0.011
Panel B Prior beliefs	Diffuse	Low	High	Diffuse	Low	High	Diffuse	Low	High
$\rho_{\theta_\alpha,\theta_\beta}$	-0.025	0.011	0.065	0.018	0.078	0.038	0.025	0.089	0.201

This table reports the posterior correlation coefficients from the general learning model. We control the prior belief on the scaled parameters of the cross-sectional variability in factor loadings, ξ_k. The prior beliefs on the mean value of the kth pricing factor, θ_k, are centered at 1 with various prior beliefs. We report results based on three pricing models: CAPM, the 3-factor and the 4-factor models. The posterior distributions of the variables considered are simulated by the MCMC technique by using monthly returns from 220 UK unit trusts.
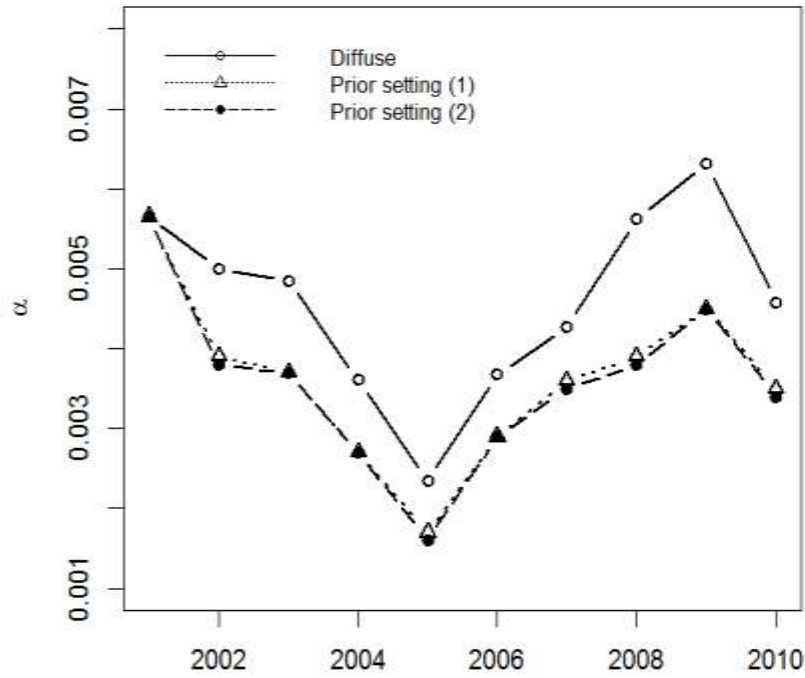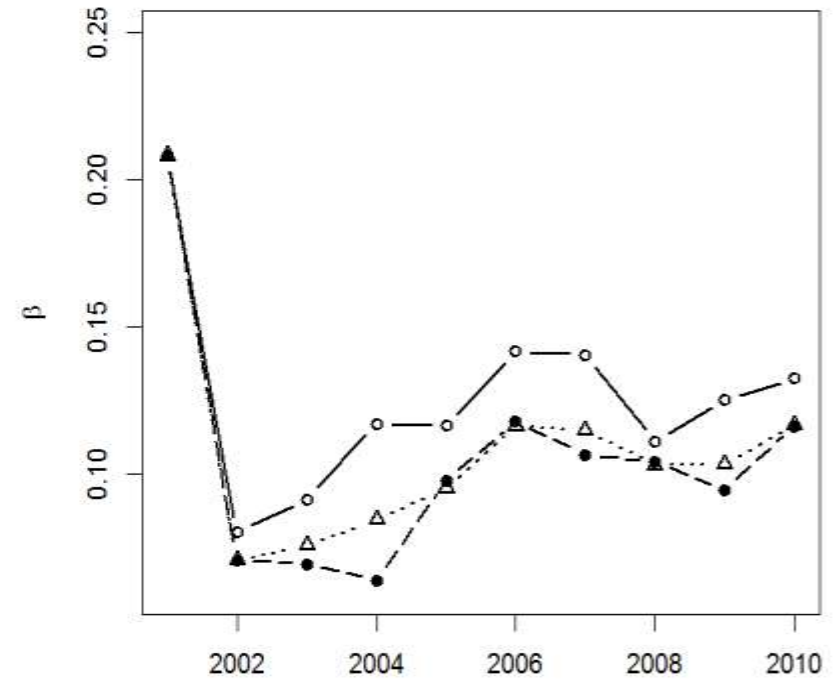
Figure 8: Posterior cross sectional in-family dispersion

Figure (a) and (b) present the posterior mean of λ_α and λ_β from a CAPM based general learning model, respectively. Both of the parameters are generated by the mean value of 6000 draws from their posterior distribution.