



May/2018

Working Paper 18-23

rcea.org/RePEc/pdf/wp18-23.pdf

A BAYESIAN DYNAMIC MODEL TO TEST PERSISTENCE IN FUNDS' PERFORMANCE

Emmanuel Mamatzakis

University of Sussex, UK

RCEA

Mike Tsionas

Lancaster University Management School, UK

Athens University of Economics and Business, Greece

Copyright belongs to the author. Short sections of the text, not exceeding three paragraphs, can be used provided proper acknowledgement is given.

The *Rimini Centre for Economic Analysis* (RCEA) was established in March 2007. RCEA is a private, nonprofit organization dedicated to independent research in Applied and Theoretical Economics and related fields. RCEA organizes seminars and workshops, sponsors a general interest journal, the *Review of Economic Analysis* (REA), and organizes a biennial conference, the *Rimini Conference in Economics and Finance* (RCEF). Scientific work contributed by the RCEA Scholars is published in the RCEA Working Paper series.

The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Rimini Centre for Economic Analysis.

A Bayesian dynamic model to test persistence in funds' performance.

E. Mamatzakis^a and E. Tsionas^b

May 2018

We provide a Bayesian panel model to take into account persistence in US funds' performance while we tackle the important problem of errors in variables. Our modelling departs from prior strong assumptions such as error terms across funds being independent. In fact, we provide a novel, general Bayesian model for (dynamic) panel data that is stable across different priors as reported from the mapping of the prior to the posterior of the Bayesian baseline model with the adoption of different priors. We demonstrate that our model detects previously undocumented striking variability in terms of performance and persistence across funds categories and over time, and in particular through the financial crisis. The reported stochastic volatility exhibits a rising trend as early as 2003-2004 and could act as an early warning of future crisis.

Keywords: Bayesian panel model, time-varying stochastic heteroskedasticity, time-varying covariance, general autocorrelation, US mutual fund performance.

^a Department of Business and Management, University of Sussex, Falmer BN1 9SL, UK. E-mail: e.mamatzakis@sussex.ac.uk. ^bLancaster University Management School, Bailrigg, Lancaster, LA1 4YX, UK & Athens University of Economics and Business, Greece. E-mail: m.tsionas@lancaster.ac.uk.

1. INTRODUCTION

Assessment over fund managers' performance has received considerable attention since the seminal paper of Jensen (1968) with mixed findings, as it is openly challenged whether funds would outperform their passive benchmark (Gruber 1996, Carhart 1997; Lunde, et al. 1999; Fama and French 2010; Basak and Makarov, 2014; Cullen et al., 2012; Cabello et al. 2014; Utz et al. 2015). At the core of the dispute is accurately measuring the performance of funds. Traditional performance measures compare the returns of the examined portfolio to the returns of an unmanaged portfolio of comparable risk. Several drawbacks of these metrics such as their inability to incorporate funds' transaction costs or the issue of selecting the proper benchmark have fuelled the introduction of performance measures that rely on frontier analysis in the spirit of Koopmans (1951) and Farrell (1957). Performance is characterised by a number of measures, such as net return ratio (Khorana and Servaes, 2012; Blake, et al. 2014 & 2016), abnormal return is the difference between fund return and the return of a portfolio which share the same risk characteristics as the fund in consideration, a dummy equals to 1 if the family has at least one fund operating in the top 5% best performing funds of a given category in a given year (Khorana and Servaes, 2012), and the Sharpe ratio (Daraio and Simar, 2006), whereas risk is computed as the standard deviation of monthly returns (Huang et al., 2007). Ferson and Lin (2014) focus on alphas and argue that there should be some bounds that depend on cross sectional investor heterogeneity with the flow response to past fund alphas. This strand of research picks earlier findings (see Clode, 2011; and Busse 2001) arguing that alphas might not be without issues when it comes to select a fund. The underlying

autocorrelation could explain results whilst the hypothesis of funds being cross-sectionally independent might not be valid (Goriaev et al. 2005).¹ Most studies, though, show that risk is important for performance (Utz et al. 2015; Basak and Makarov, 2014; Brown et al., 1996; Cullen et al., 2012; Goriaev et al., 2005; Koski and Pontiff, 1999).² Brown et al. (2001) examine competition and risk in the hedge fund, reporting similar results as in Brown et al., (1996). Busse (2001) show that poorly performing fund managers alter their risk to be able to catch up with interim winners at the end of the year. Basak and Makarov (2014) focus on the manager's portfolio choice with respect to the strategic interactions among managers competing for fund flows. Their model builds on the strategic behaviours of two risk-averse managers, revealing that a manager either wins or loses, and never opts for a draw.³ Other studies (Prather et al., 2004) report the link between fund performance and various operational characteristics such as expenses, size, past performance. This type of information could be rather beneficial to investors who decide among offered funds should a reliable relation exists between a fund's performance and some of its observable characteristics. Ferson and Mo (2016) argue that a well-specified performance measure should be based on the sum of covariance between the portfolio holdings and the subsequent abnormal, or risk-adjusted returns, with an underlying stochastic discount factor (see also Cabello et al. 2014). Their modelling has certain appeal, but it still does not address issues related to

¹ Kempf and Ruenzi (2008) study the competition between fund managers across funds' family. They argue that an optimal policy of fund managers is to alter their risk-taking. Studying US equity mutual funds between 1993 and 2001, they report the presence of the family tournament, which is more pronounced in large families.

² Brown et al. (1996) identify that interim losers who underperform the benchmark in the first half of the year are likely to increase their risk relative to mid-year winners. Funds are ranked according to their cumulative return, while risk is measured by the ratio of fund's standard deviation after the interim performance assessment to its standard deviation before that date. Another proxy for risk is the tracking error variance, which is the variance of the difference between fund's return and the value-weighted market index (Chevalier and Ellison, 1997).

³ Basak and Makarov (2014) show that, even when a manager is significantly ahead in the tournament, her investment behaviour and thus portfolio volatility is still influenced by the tournament incentives. In addition, Sato (2015) show the importance of flow-performance relationship and asset bubbles.

time-varying covariance where the evidence shows that indeed this is the case (Cabello et al. 2014; Utz et al. 2015; Ferson and Mo 2016; Basak and Makarov 2014; Blake et al. 2014 & 2016).

From the above literature becomes apparent that to date there is no silver bullet regarding an appropriate modelling of mutual fund performance and its underlying determinants across funds and over time. This paper bridges a gap in the literature by providing a novel way modelling mutual funds' performance, relaxing some of the strong assumptions in the literature. Moreover, we argue that time-varying heteroskedasticity and time-varying covariances (in line with Ferson and Mo 2016; Blake, D., et al. 2014 & 2016) are of importance for measuring mutual fund performance without resorting to strong assumptions regarding the unobservable underlying idiosyncratic characteristics of the fund managers. To this end, the purpose of our study is fourfold. First, we propose a new Bayesian panel model that captures time-varying heteroskedasticity and time-varying covariances in funds' performance as well as general autocorrelation and the underlying stochastic volatility. Second, this model allows measuring persistence and it takes also into account errors in the variables. This is commonly acknowledged (Ferson and Mo 2016; Basak and Makarov 2014; Blake, D., et al. 2014 & 2016), but we are not aware of any previous studies that have dealt with the issue. Our modelling departs from prior strong assumptions, for example that error terms across funds are independent (Ferson and Mo 2016; Basak and Makarov 2014; Blake, D., et al. 2014 & 2016). Third, as the estimation of this new model is cumbersome, we apply Bayesian techniques that facilitate the robustness of the estimation. Bayesian analysis is implemented using state-of-the-art Sequential Monte Carlo / Particle-Filtering (SMC/PF) techniques. Fourth, we broaden the findings

of the relatively few studies measuring fund performance for an up to date set concerning US mutual funds for which we demonstrate that results remain stable across different priors as reported from the mapping of the prior to the posterior.

A preliminary review of our results reveals that risk asserts a positive and significant impact on US mutual funds' performance across different specifications, whilst all Fama and French five factors also show strong positive and significant effect on funds' performance. There has been striking variability in terms of performance and persistence across funds categories and over time, and in particular through the financial crisis. The reported stochastic volatility exhibits a rising trend as early as 2003-2004 and could act as an early warning of future crisis. We show that our results are stable across different priors as reported from the mapping of the prior to the posterior of the Bayesian baseline model with the adoption of different priors. Using likelihood-based techniques, especially Bayesian methods organized around Sequential Monte Carlo, we avoid the need for asymptotically-based inferences which can be misleading in finite samples and in general models which take into account almost all features of commonly employed panel data sets.

The rest of the paper is organized as follows: Section 2 presents the new performance model for funds, whilst section 3 reports the data set. Section 4 discusses empirical results. Finally, Section 5 provides some concluding remarks and policy implications.

2. METHODOLOGY

2.1 Fund's performance and persistence model

So far in the literature performance is characterised by a number of measures, such as net return ratio, abnormal return, Sharpe ratio (Daraio and Simar, 2006; Khorana and

Servaes, 2012; Blake, D., et al. 2014 & 2016; Ferson and Mo 2016; Ferson and Lin 2014). Moreover, we build on the earlier research by Blake, et al. (2014) & (2016) suggest that an efficient way of measuring mutual funds' performance is to apply bootstrap methods. The authors effectively pool observations over time, whereas some cross-correlation of fund returns is allowed. We argue that such modelling is too restrictive in the underlying assumptions as heteroskedasticity, errors in variables, covariance across funds, and volatility are not effectively captured. In addition, we augment Ferson and Mo (2016) by allowing time varying heteroskedasticity and covariance of fund performance. Moreover, we propose the following model:

$$y_{it} = \gamma_{it} + x'_{it}\beta + v_{it}, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \quad (1)$$

where the dependent variable, y_{it} , is the return of mutual fund i at date t , x_{it} is a $k \times 1$ vector of covariates, γ_{it} captures fund- and time-specific effects and v_{it} is an error term. We intend to propose a general model for panel data to allow for structure in returns.

The first novelty that we introduce is that error terms across funds cannot be independent: $v_t = [v_{1t}, \dots, v_{nt}] \stackrel{iid}{\sim} N_n(0, S)$, $t = 1, \dots, T$. Later on we intend to modify the iid assumption. Secondly, we intend to model the γ_{it} in a non-parametric way. This is of importance and complements previous research by Blake, et al. (2014) & (2016) that apply bootstrap methods to measure γ_{it} with some strong underlying assumptions.

Herein we build on Koop and Poirier (2004), arguing that $\gamma_{it} - \gamma_{i,t-1} \sim N(0, \omega^2)$ which is equivalent to a spline model⁴. It is important to emphasize that this is a Bayesian interpretation of standard non-parametric procedures because it imposes a prior notion

⁴ Initial conditions are treated as unknown parameters.

about smoothness in the sense that, as a function of time, the γ_{it} s are likely to behave smoothly.

We extend such model as follows:

$$g_{it} - g_{i,t-1} \sim N(a_i, W_i^2), \quad (2)$$

where a_i is a ‘persistence effect’ across funds and we allow W_i s to be also fund-specific. Moreover, we extend below this to a model where first derivatives are likely to be smooth as well.

This modelling allows revealing volatility across funds whilst we also account for persistence in funds’ performance at fund level. Moreover, if we take first differences,

we have:
$$\Delta y_{it} = \alpha_i + \Delta x_{it}' \beta + \Delta v_{it} + \xi_{it}, \quad i = 1, \dots, n, \quad t = 2, \dots, T, \quad (3)$$

where $\xi_{it} \sim N(0, \omega_i^2)$ and Δv_{it} follows an MA(1) process with a unit coefficient. This, in fact, shows that persistence can be measured using the model in first differences while performance can be measured using directly the g_{it} s.

Finally, we address another important problem that has not received enough attention in the mutual fund evaluation literature. This is the problem of errors in the variables. In addition, we also allow the β 's to be time varying. Therefore, we modify the model

as⁵:
$$y_{it} = \gamma_{it} + x_{it}^* \beta_t + v_{it}, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \quad (4)$$

where x_{it}^* denotes the actual data, and $x_{it} = x_{it}^* + \varepsilon_{it}$, (5)

where ε_{it} denotes measurement error.⁶

⁵ It is easy to show that the interpretation of γ and α goes through even when β s are time-varying.

⁶ Given the model of equation (4) the following applies: i) we can allow for different temporal coefficients, β_t , and ii) we can allow for arbitrary patterns of autocorrelation, since we can assume: $E(v_s v_t') = \Sigma_{st}$, $s, t \in \{1, \dots, T\}$. In addition, we allow for arbitrary autocorrelation and arbitrary forms of

We assume: $\varepsilon_t = [\varepsilon_{1t}, \dots, \varepsilon_{nt}]' \sim N_n(0, \Omega)$. (6)

To determine a prior for Ω we use the decomposition $\Omega = C'C$ where C is an upper triangular matrix. Let $c = \text{vec}(C)$, where vec vectorizes the elements in the upper diagonal. Our prior is: $c \sim N(\bar{c}, \bar{V}_c)$. (7)

For simplifying the notation for a given time period the model in equation (4) can be written as: $y_t = X_t \beta_t + \gamma_t + v_t, t = 1, \dots, T$, (8)

where $y_t = [y_{1t}, \dots, y_{nt}]'$, $X_t = [x'_{1t}, \dots, x'_{nt}]$, $\gamma_t = [\gamma_{1t}, \dots, \gamma_{nt}]'$, and $v_t = [v_{1t}, \dots, v_{nt}]'$.

If we modify X_t to include an identity matrix with dimension $n \times n$ to X_t and expand β_t with a vector of g_{it} with dimension $n \times 1$, the model can be written without loss of generality as: $y_t = X_t \beta_t + v_t, t = 1, \dots, T$. (9)

If we define $X = \text{diag}[X_1, \dots, X_T]$ we can write the model as: $y = X\beta + v$, (10)

where $v = [v'_1, \dots, v'_T]'$, and $E(vv') = \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1T} \\ \Sigma_{12} & \Sigma_{22} & \dots & \Sigma_{2T} \\ \dots & \dots & \dots & \dots \\ \Sigma_{1T} & \Sigma_{2T} & \dots & \Sigma_{nT} \end{bmatrix}$. (11)

As there are $\frac{nT(nT+1)}{2}$ free elements in Σ , there is a huge number of parameters to estimate. We intend to place priors on Σ which tend to favour lack of autocorrelation but allow for heteroskedasticity. This means that Σ_{st} should be 'close' to a zero $n \times n$

heteroskedasticity as well. This comes at the cost of allowing for $\frac{T(T+1)}{2}$ matrices of the form Σ_{st} each of which is $n \times n$.

matrix ($s \neq t \in \{1, \dots, T\}$), say $O_{(n \times n)}$.⁷ If we leave Σ_{tt} ($t \in \{1, \dots, T\}$) unrestricted, this means that we have $\frac{n(n+1)}{2}$ parameters for each ones of the T matrices, which is still excessive.⁸

From our previous discussion on g_{it} s it is clear that a spline or smoothness prior will essentially result in a non-parametric model. Thus, for the elements of b_t , the smoothness prior is of the form: $b_t - b_{t-1} \sim N(d, \chi)$, $\chi = \text{diag}[\chi_1^2, \dots, \chi_n^2]$, (12)

$$\text{or } \beta_t - 2\beta_{t-1} + \beta_{t-2} \sim N(\delta, \Xi), \Xi = \text{diag}[\xi_1^2, \dots, \xi_n^2]. \quad (13)$$

We call these models spline-I and spline-II, respectively and our underlying priors

$$\text{are: } \delta \sim N(\bar{\delta}, \bar{V}_\delta), \quad (14)$$

$$\log \xi_i^2 \sim N(\bar{\varphi}, \bar{V}_\xi), \quad i = 1, \dots, n. \quad (15)$$

2.2 Specification of the baseline prior and its variations

⁷ Regarding the elements of Σ_{ts} , $t \neq s$, since the prior belief is that these are all zero matrices, we can adopt a ‘model selection prior’ (see Koop 2013) of the form: $\frac{\sigma_{ij,t}}{\sqrt{\sigma_{ii,t}\sigma_{jj,t}}} = \begin{cases} 0, & \text{with probability } p, \\ \rho, & \text{with probability } 1-p. \end{cases}$

In this prior we set $p = \frac{1}{2}$ and we treat ρ as unknown parameter.

⁸ Regarding Σ_{tt} the diagonal elements, say $\sigma_{ii,t}$, allow for arbitrary time-varying heteroskedasticity while $S_{ij,t}$ allows for contemporaneous correlation of returns. The matrix has received a great deal of attention in DCC and similar models. Suppose that: $\Sigma_{tt} = H_t H_t'$, where H_t is an $n \times n$ upper diagonal elements and its non-zero elements can be vectorized as: $h_t = \text{vec}[h_{ij,t}, i, j = 1, \dots, \frac{n(n+1)}{2}]$. We proceed with a prior assuming that: $h_t = a + A h_{t-1} + u_t$, where a and A have dimension $\frac{n(n+1)}{2} \times 1$ and $\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}$, respectively, and the error term $u_t \sim N(0, \Upsilon)$. To determine a prior for Ψ we use the decomposition $\Upsilon = C_\Psi' C_\Psi$ where C_Ψ is an upper triangular matrix. Let $c_\Psi = \text{vec}(C_\Psi)$. Our prior is: $c_\Psi \sim N(\bar{c}_\Psi, \bar{V}_\Psi)$. In effect, we place a multivariate stochastic volatility prior. Although we have high dimensional objects a and A , we can proceed with priors to resolve the curse of dimensionality: $a \sim N(\bar{a}, \bar{V}_a), \text{vec}(A) \sim N(\bar{A}, \bar{V}_A)$.

In (7) we have $c \sim N(\bar{c}, \bar{V}_c)$ and we set $\bar{c} = 0, V_c = h_c I$ for $h_c = 1$. Similarly we have $c_\psi \sim N(\bar{c}_\psi, \bar{V}_\psi)$ (see footnote 6) and we set $\bar{c}_\psi = 0, \bar{V}_\psi = h_\psi I$ for $h_\psi = 1$ and set $\bar{a} = 0, \bar{V}_a = h_a I$ with $h_a = 1$ and $\bar{V}_A = h_A I$ with $h_A = 1$. Thus, for \bar{A} we adopt a Minnesota-like prior where the elements corresponding to the diagonal are $d_{\bar{A}} \sim (0,1)$ and the others are zero. We start by setting $p = \frac{1}{2}$ (see footnote 6r) and we treat r as unknown parameter. The prior of r is taken to be $\rho \sim N_{(-1,1)}(0, h_\rho)$, $h_\rho = 1$. In (14) we have $\delta \sim N(\bar{\delta}, \bar{V}_\delta)$, where $\bar{\delta} = 0$ and $\bar{V}_\delta = h_\delta I$ for $h_\delta = 1$.

Finally, in (15) we have $\log \xi_i^2 \sim N(\bar{\varphi}, \bar{V}_\xi)$, $i = 1, \dots, n$. We set $\bar{f} = 0$ and $\bar{V}_\xi = h_\xi I$ where $h_\xi = 1$. In the baseline prior, there are many parameters that we can vary to perform sensitivity analysis. To facilitate the analysis, we present the variations of baseline priors in Table 1.

Table 1: Variations of the baseline prior.

parameters and baseline prior	variations
$c \sim N(\bar{c}, \bar{V}_c)$, $\bar{c} = 0$, $V_c = h_c I$ for $h_c = 1$.	$c \sim U(-5, 5), h_c \sim U(0.1, 10)$
$a \sim N(\bar{a}, \bar{V}_a), \text{vec}(A) \sim N(\bar{A}, \bar{V}_A), \bar{a} = 0, \bar{V}_a = h_a I$ with $h_a = 1$. Also $\bar{V}_A = h_A I$ with $h_A = 1$.	$\bar{a} \sim U(-5, 5), h_a, h_A \sim U(0.1, 10)$
For \bar{A} we adopt a Minnesota-like prior where the elements corresponding to the diagonal are $d_{\bar{A}} \in (0, 1]$ and the others are zero.	$d_{\bar{A}} \sim U(0, 1)$
$p = \frac{1}{2}$, $\rho \sim N_{(-1,1)}(0, h_\rho), h_\rho = 1$	$p \sim U(0, 1), h_\rho \sim U(0.1, 10)$
$\delta \sim N(\bar{\delta}, \bar{V}_\delta)$, where $\bar{\delta} = 0$ and $\bar{V}_\delta = h_\delta I$ for $h_\delta = 1$.	$\bar{\delta} \sim U(-5, 5), h_\delta \sim U(0.1, 10)$
$\log \xi_i^2 \sim N(\bar{\varphi}, \bar{V}_\xi)$, $i = 1, \dots, n$. We set $\bar{f} = 0$ and $\bar{V}_\xi = h_\xi I$ where $h_\xi = 1$.	$\bar{\varphi} \sim U(-5, 5), h_\xi \sim U(0.1, 10)$

Notes: $U(a, b)$ denotes the uniform distribution in the interval (a, b) , $b > a$.

In the empirical application, our intention is to run the baseline model and then adopt different priors as in Table 1 to examine how the results change so that we have some sense of the mapping from the prior to the posterior. To explore the posterior, we use SMC/PF techniques (results are available under request).

3. DATA

3.1. Sample description

We obtain mutual fund data from Morningstar database for the 2000-2014 period. There are 10,391 funds (94,670 observations), 459 families (5,689 observations), and 25 Morningstar categories (366 observations). Table 2 provides the descriptive statistics at the fund family level for every five years in the sample. Our sample includes US funds in different categories available during the observed period. There are 1623 funds, which charge redemption fee, 1875 funds charging front load, and 260 funds charging both types of fees. Some studies of mutual funds exclude load funds to avoid the problem of addressing different sales fees in the fund's operating costs (Babalos et al., 2015; Gil-Bazo and Ruiz-Verdú, 2009). Our study, in contrast, accounts for both load and no-load funds, and use the information on front-end and back-end loads as fund's characteristic variables as in Daraio and Simar (2006). As indicated in Ferris and Chance (1987) funds' expenses do not comprise load charges. While front-end loads are sales charges paid to brokers or financial advisors for selling the fund, back-end loads are levied on customers for redeeming their shares (Daraio and Simar, 2006; Khorana and Servaes, 2012). The fact that redemption fee exists could inhibit families' competitiveness as investors could be hindered from leaving the funds, especially when funds appear to be underperforming (Khorana and Servaes, 2012). As a result, one may

presume that no-load funds attract more investors. However, for those who would need professional advice for their investment choices, front-end load could be a reasonable premium they are willing to compensate for financial advisors (Ferris and Chance, 1987). Additionally, there could be the probability of no-load funds imposing higher other fees on their investors (Tran-Dieu, 2015) or incurring higher expense ratio compared to their load peers (Ferris and Chance, 1987). Hence, analysing this comprehensive sample would produce inclusive results on the competitiveness of different types of funds in the US mutual fund industry.

Table 2: Summary statistics for funds.

	2000-2004	2005-2009	2010-2014
Total assets (mil USD)	4860	7220	9670
No of families	459	459	459
No of funds per family	23.876	26.792	21.663
Number of funds started	5.710	4.241	4.550
Number of categories	2.617	2.998	2.501
Gross expense ratio	1.308	1.712	1.089
Gross return	5.373	4.649	11.736
Turnover ratio	0.514	0.749	0.505
Risk	4.336	4.215	3.265
12b-1 fee	0.362	0.362	0.362
Max front load	5.145	5.145	5.145
Redemption fee	1.640	1.640	1.640

Notes: This Table reports the mean of variables over five years intervals. Number of funds started is the number of new funds started in a given year. Number of categories represents the different investment objectives within a family. Turnover is the weighted average turnover across all funds in a family. Risk is measured as the weighted average standard deviation of monthly return. Gross expense ratio, gross return, turnover ratio, risk, 12b-1 fee, max front load, and redemption fee are the weighted average values across all funds in a family.

Based on Morningstar classification, there are 25 fund categories in our sample after being reviewed for errors and outliers. More specifically, according to size, these types consist of large blend, mid-cap blend, small blend, foreign large blend, large growth, mid-cap growth, small growth, foreign large growth, foreign small/mid growth, large value, mid-cap value, and small value. In terms of sector, these categories include real estate, global real estate, technology, equity energy, financial, consumer cyclical,

health, utilities, natural resources, communications, consumer defensive, industrials, and world stock.

Regarding variable selection for mutual fund, there is not a universal accepted approach, apart from the return variable that is rather basic. To account for the impact of various determinants we include in addition variables such as loads, fee and turnover. In particular, 12b-1 fee refers to marketing and distribution fee related to money paid to selling agents (Collins and Mack, 1997) and marketing expenses (Khorana and Servaes, 2012). In our sample, there are 7958 funds having a 12b-1 plan. As 12b-1 fee is supposed to be a driving factor in raising fund's assets, there could be two possibilities. On the one hand, economies of scale may exist, which provides funds with the benefit of passing on the fee to both existing and new investors (Khorana and Servaes, 2012). On the other hand, 12b-1 fee can raise expenses as it is a component of a fund's expense (Ferris and Chance, 1987; Latzko, 1999). Front-end load and back-end load have also attracted research interests in their impact on funds' expenses (Daraio and Simar, 2006; Khorana and Servaes, 2012; Latzko, 1999). Front-end loads are an initially one-off sales charge as a reduction to the investment to the fund and are used to incur the cost of financial advisors in attracting new investors. Back-end loads, often known as deferred loads or redemption fee are levied when investors redeem their shares. When redemption fee is high, it may also hinder fund shareholders from leaving the fund, especially underperforming ones.

In subsequent analyses, we also include variables such as total expenses (including loads), risk, turnover ratio, and number of funds. The literature has suggested that family's diversification across investment styles would benefit investors in terms of

fewer restrictions imposed on their asset allocation (Mamaysky and Spiegel, 2002). This would also denote the presence of risk hedging improvement in contrast to the economies of scale arising from '*learning-by-doing externality*', which exists in more focused families (Massa, 2000). Furthermore, risk is more likely to affect fund family's market share through its relationship with competition and performance (Basak and Makarov, 2012; Huang et al., 2011; Huang et al., 2007; Spiegel and Zhang, 2013; Vidal-García and Vidal, 2014).

The number of funds offered by the fund family is also included to observe whether there is a presence of cost sharing between funds. Put differently, a fund family may enjoy greater economies of scale as the expenses could be reduced for a group of funds (Malhotra et al., 2007). As indicated in Khorana and Servaes (2012), the number of fund started could signify additional business lines, product differentiation, or simply the incentive to increase the likelihood of having funds on the top 5% best-performing classification. Its squared value gives an indication for the outstanding impact (if any) in case there is a considerable number of new funds.

Based on Morningstar's definition, turnover ratio conveys the fund's trading activity. Funds report this figure by taking the lesser of purchases or sales of all securities with maturities from one year and dividing by average monthly net assets. The lower the turnover ratio, the more the fund is in favour of the buy-and-hold strategy. Stated differently, high turnover ratio indicates active portfolio management strategies (Daraio and Simar, 2006; Khorana and Servaes, 2012). As a result, active fund managers classified based on high turnover ratios may impose more transaction costs on fund shareholders, in turn raising the fund total expenses.

We also include the Fama-French 5 factors. Moreover the factors are: factor 1 the SMB (Small Minus Big) defined as the average return on the nine small stock portfolios minus the average return on the nine big stock portfolios, factor 2 the HML (High Minus Low) that is the average return on the two value portfolios minus the average return on the two growth portfolios, factor 3 the RMW (Robust Minus Weak) the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios, factor 4 the CMA (Conservative Minus Aggressive) the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios, and the last factor $R_m - R_f$, the excess return on the market (see Fama and French, 2014).⁹

4. EMPIRICAL RESULTS

In this section, we run the baseline model and then adopt different priors as in Table 1 to examine how the results change so that we have some sense of the mapping from the prior to the posterior. We do not report results for autocorrelation or correlation across funds and statistics related to errors-in-variables to save space but these are available on request.

For prior sensitivity analysis, we simulate 10,000 different priors from the last column of Table 1 and we repeat posterior analysis using SMC/PF) techniques (Appendix I).

4.1 Selecting the model

⁹ Moreover, the 5 factors come from Fama and French (2014) and defined as $SMB(B/M) = 1/3$ (Small Value + Small Neutral + Small Growth) - $1/3$ (Big Value + Big Neutral + Big Growth), $SMB(OP) = 1/3$ (Small Robust + Small Neutral + Small Weak) - $1/3$ (Big Robust + Big Neutral + Big Weak), $SMB(INV) = 1/3$ (Small Conservative + Small Neutral + Small Aggressive) - $1/3$ (Big Conservative + Big Neutral + Big Aggressive) and thus $SMB = 1/3$ ($SMB(B/M) + SMB(OP) + SMB(INV)$). $HML = 1/2$ (Small Value + Big Value) - $1/2$ (Small Growth + Big Growth). $RMW = 1/2$ (Small Robust + Big Robust) - $1/2$ (Small Weak + Big Weak). $CMA = 1/2$ (Small Conservative + Big Conservative) - $1/2$ (Small Aggressive + Big Aggressive). Data include all NYSE, AMEX, and NASDAQ firms.

To get an idea of what is a ‘good’ model we use the baseline prior and we present in Table 3 values of the marginal likelihood (converted into Bayes factors, BF). For a posterior distribution¹⁰:

$$p(\theta|Y) = \frac{L(\theta;Y)p(\theta)}{p(Y)}, \quad (16)$$

the denominator is the marginal likelihood, $p(Y) = \int L(\theta;Y)p(\theta)d\theta$ (17)

Table 3: The model comparison: Bayes factors.

	Bayes Factors
Formulation in (12), spline-I	
No restrictions	1.000
$b_t = b_o$ are the same	0.0061
$A = 0$, no stochastic volatility	0.0056
$b_t = b_o$ and $A = 0$	0.0011
No cross-sectional correlation	0.0003
No measurement error	0.0001
Formulation in (13), spline-II	
No restrictions	7.3221
$b_t = b_o$ are the same	0.0414
$A = 0$, no stochastic volatility	0.0032
$b_t = b_o$ and $A = 0$	0.0017
No cross-sectional correlation	0.0002
No measurement error	0.0000

Note: The table presents Bayes factors relative to the full model without any restrictions. The Bayes factors are computed using the marginal likelihood from the Sequential Monte Carlo (particle filtering) approach described in Appendix.

¹⁰ θ denotes the parameter vector, Y the data and L is the likelihood. Also, $p(\theta)$ is the prior.

Clearly, no simplification is possible, at least using the baseline prior as the Bayes factors for alternative models are inferior to the model without restrictions. It also turns out that spline-II behaves much better compared to spline-I.

Our fundamental objective is to evaluate the performance of the mutual funds. From (7) the elements of g_t are the ‘generalized Jensen’s alphas’ for all funds for date t . From (2) coefficients a_i denote the persistence. To take off, we would like first of all to present a measure of overall performance, which is

$$PERF_t = n^{-1} \sum_{i=1}^n \gamma_{it}, \quad (25)$$

that is the average performance in the fund industry for a given date.

As the measure depends on all other parameters, in standard Bayesian fashion we take an average across all Monte Carlo draws. Table 4 reports results for model selection, reporting means and standard deviations, including the five factors of Fama and French across specifications. It is perhaps important to point out that our likelihood-based procedures avoid resorting to asymptotic inferences, which are questionable in finite samples, especially when the econometric model is complicated. Since we use Bayesian techniques we examine thoroughly sensitivity to the prior assumptions as we mentioned before in connection to Table 1.

Table 4: Model selection, posterior means and posterior standard deviations.

	full model		without measurement error		without time-varying b_i		without cross-sectional correlation	
	Post. mean	Post. SD	Post. mean	Post. SD	Post. mean	Post. SD	Post. mean	Post. SD
SD of returns	0.0171	0.0015	0.0031	0.0021	-0.0022	0.0010	0.0134	0.0019
Skewness	0.0043	0.0035	-0.0030	0.0011	-0.0044	0.0014	0.0035	0.0023
Kurtosis	0.0031	0.00012	0.00132	0.0254	0.00441	0.0165	0.012	0.00024
Expense ratio	-0.0014	0.00032	-0.0133	0.0043	-0.0033	0.0017	-0.0144	0.00017
Loads & Turnover	-0.0015	0.00013	-0.0126	0.0032	-0.013	0.0083	-0.0033	0.00012

12b-1 fees	0.0045	0.00017	0.0023	0.0020	0.0033	0.0432	0.0011	0.00044
Net asset value	0.0032	0.00021	0.0011	0.0017	0.0025	0.0019	0.0015	0.00035
Fund size	0.0015	0.00017	-0.0013	0.0015	-0.0022	0.0017	0.0036	0.00026
Net-exp. ratio	-0.0033	0.00024	0.0045	0.0032	0.00210	0.0013	0.0032	0.00323
Turnover ratio	0.0017	0.00015	0.0045	0.0033	0.0037	0.0022	0.0034	0.00021
F&F 1	0.281	0.032	0.171	0.0117	0.128	0.035	0.155	0.054
F&F 2	0.048	0.032	0.023	0.024	0.032	0.0156	0.0245	0.0272
F&F 3	-0.015	0.011	0.0056	0.0137	0.0044	0.0031	-0.0221	0.0217
F&F 4	0.129	0.0522	0.0171	0.0081	0.0225	0.0152	0.2513	0.0351
F&F 5	0.083	0.0714	0.0331	0.0266	0.0742	0.0553	0.0484	0.0493
Bayes R ²	0.9345	0.0454	0.8716	0.0325	0.8652	0.0352	0.8751	0.0255

Note: We have used the prior that out of 10,000 alternative priors in the last column of Table 3 yields the highest value for the marginal likelihood. So, we use the prior that best ‘fits’ the data. The five Fama-French (F&F) factors are also included: F&F 1 the SMB (Small Minus Big), F&F 2 the HML (High Minus Low), F&F 3 the RMW (Robust Minus Weak), F&F 4 the CMA (Conservative Minus Aggressive), and F&F 5 $R_m - R_f$. We also compute a Bayesian variant of R^2 , which is computed, using the correlation of actual and predicted values, averaged over all posterior draws. SD implies standard deviation. Post. implies posterior. We have used 15000 iterations of the Sequential Monte Carlo (particle filtering) approach discarding the first 5000 to mitigate possible start up effects. We used 10000 particles per iteration.

Table 4 reports that risk asserts a positive and significant impact on performance across specifications, but the model without time-varying β_t . The positive sign comes in line with Basak and Makarov, (2012), Huang et al., (2011), Vidal-García and Vidal, (2014). The convex flow-performance nexus (Chevalier and Ellison, 1997) implies the incentives for fund managers to increase the underlying risk of their funds over time. However, in the model specification where there are no time-varying effects the impact of risk turns negative. Similar results are reported for the third moment, insinuating the importance of allowing for time varying β_t for correctly identifying the impact of underlying moments on performance. Fama and French five factors show strong positive and significant effect on performance. The exception is the RMW factor where negative, but not significant, effects are reported.

Whether size matters for fund performance has attracted much research interest (Chen et al., 2004; Ferreira et al., 2012; Chen et al., 2013; Khorana and Servaes, 2012). Large funds may benefit from economies of scale which they can pass on to investors

(Khorana and Servaes, 2012), from investment opportunities which may not be available for small funds, and from better spreads thanks to large positions and trading volumes (Ferreira et al., 2012).¹¹ This advantage and economies of scope may also be present at the family level (Chen et al., 2004; Chen et al., 2013). Herein we provide statistical significant evidence, under the full model, that size, indeed, matters.¹² The coefficient of fund size takes a positive sign for the full model and the model without cross sectional variation in line with Pollet and Wilson (2008) and Jordan and Riley (2015).

Table 4 also reports that expense ratio asserts a significant negative impact on performance across specifications. This is line with Ferreira et al. (2012). Carhart (1997) finds also a negative relationship between fees and net-fee performance. Gil-Bazo and Ruiz-Verdú (2009) opt for pooled ordinary least squares to estimate the effect of funds' expense ratio on before-fee risk-adjusted performance. The results report that before-fee performance is inversely related to fees. According to fund's strategic behaviour, they set fees based on past or expected performance. One underlying rationale is that underperforming funds could have investors who are less sensitive to performance (Christoffersen and Musto, 2002). Therefore, if money does not flow from worse performing funds to the better counterparts, the remained investors would be

¹¹ However, large funds may encounter some disadvantages in terms of liquidity and management (Chen et al., 2004; Pollet and Wilson, 2008). According to the organisational diseconomies hypothesis (Chen et al., 2004), fund size is inversely related to performance. This could be because of hierarchical costs, or management dilution when the fund expands. One manager can easily manage small assets, but it needs a team to manage a large asset base. Large funds would eventually trade large volumes, which may cause difficulties for them in opening and closing their positions. Hence, this liquidity constraint also explains why large funds could be associated with lower performance.

¹² Empirical findings for the size-performance relationship are somewhat mixed. While Chen et al. (2004), Huang et al. (2011) and Ferreira et al. (2012) find a presence of diseconomies of scale, Pollet and Wilson (2008) report that large funds tend to diversify their portfolios which in turn increases their performance. Jordan and Riley (2015) report a positive effect of small size on fund's future alpha. Regressing future alpha on past alpha and size, Elton et al. (2012) do not find a statistically significant predictability of size on future alpha. However, they document that as size increases, expense ratios and management fees decrease.

charged more. Another explanation is that those performance-insensitive investors are target clients of funds with low expected performance (Gil-Bazo and Ruiz-Verdú, 2008). By doing so they are able to charge higher fees, also to compensate for their inability to compete with better performing funds.

Regarding the turnover ratio, we find that it asserts a positive impact on performance across specification. As turnover ratio indicates the fund family's trading activities, increasing turnover would imply active underlying portfolio management tactics that would also increase performance (Daraio and Simar, 2006; Khorana and Servaes, 2012).¹³ Similarly, 12b-1 fees also have a positive and significant effect on performance across specification. 12b-1 fees convey information regarding fund's assets. In this respect, higher fees would imply larger funds that, in turn, we report that have higher performance. In line with economies of scale we show that funds would pass fees to both existing and new investors as performance rises in line with Khorana and Servaes (2012).¹⁴

4.2 Funds Performance

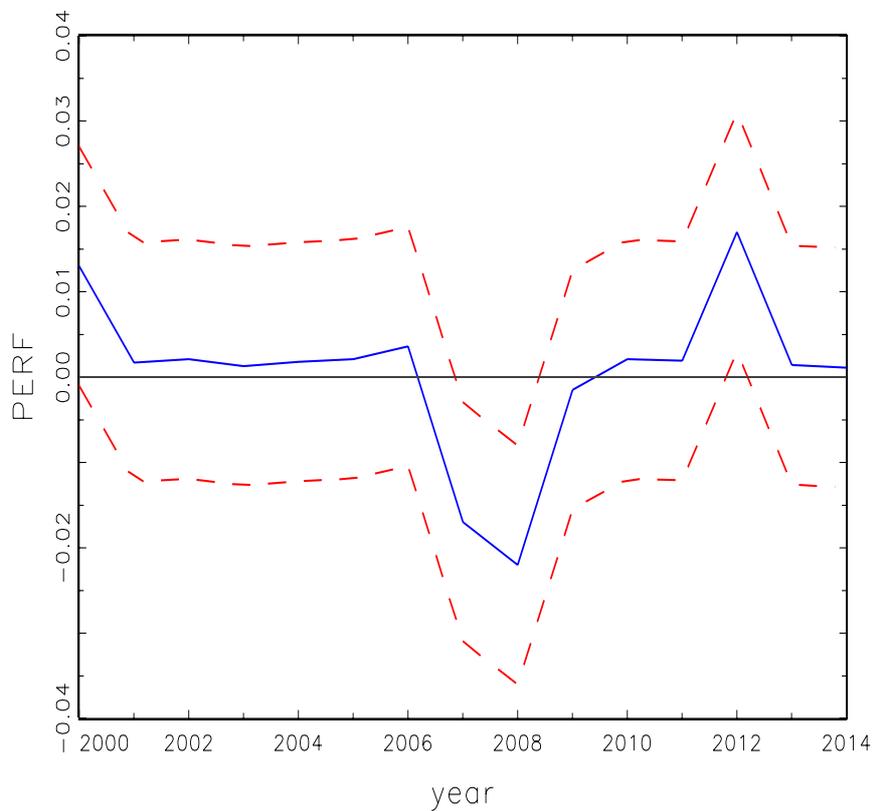
The performance results are drawn in Figure 1, based on the full model specification as selected by the results presented in Table 3, along with plus or minus two posterior standard deviations. Clearly, there is a dive in performance during the financial crisis, though the starting point of the downfall is reported as early as in 2006. Thereafter, there is a pick in recovery up till 2012, but that is short lived as a drop is next detected for the remaining of the sample period up till 2014. In this Figure, we detect that funds'

¹³ Funds report turnover ratio by taking the lesser of purchases or sales of all securities with maturities from one year and dividing it by the average monthly net assets. The lower the turnover ratio, the more the fund is in favor of the buy-and-hold strategy. Stated differently, high turnover ratio indicates active portfolio management strategies (Daraio and Simar, 2006; Khorana and Servaes, 2012). As a result, active fund managers classified based on high turnover ratios may induce higher performance.

¹⁴ However, there is some evidence that shows that 12b-1 fee can raise expenses that would eventually compromise performance.

performance are following a quite long financial cycle, over ten years period. The slow performance starts as early as in 2001 with the recovery being recorded in 2012. The financial crisis made aggravate things, yet it is evident that funds' financial cycle is elongated. Persistence, therefore, might be of importance here. Next, we report persistence.

Figure 1: The posterior average overall funds' performance over time.



Notes: The figure presents the posterior mean of the overall performance as in equation (25) over time. We have used Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear dashed lines.

Table 5 reports the average performance indicator across funds categories over the whole sample period. It is striking that some 12 funds report negative performance. The highest performer appears funds in utilities with the lowest performer the real estate funds. Overall performance, as identified also in Figure 1, has been rather subdued during the sample period. There is a dive reported during the financial crisis, but worryingly despite some recovery up till 2012, there is a further decline in recent years.

Table 5: Average performance indicator across funds' categories.

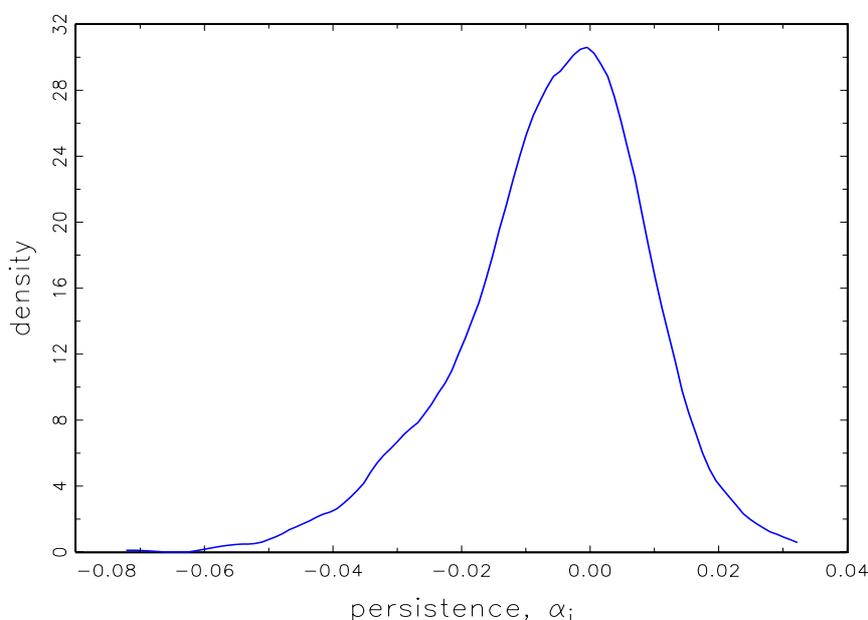
Category	Performance	Category	Performance
Large Blend	0.007 (0.034)	Technology	-0.003 (0.005)
Mid-Cap Blend	0.001 (0.021)	Financial	-0.004 (0.002)
Small Blend	-0.003 (0.012)	Consumer Cyclical	-0.007 (0.012)
Large Growth	0.005 (0.013)	Equity Energy	0.005 (0.012)
Mid-Cap Growth	0.002 (0.007)	World Stock	-0.003 (0.003)
Small Growth	-0.003 (0.004)	Global Real Estate	-0.075 (0.120)
Foreign Large Growth	0.005 (0.002)	Consumer Defensive	0.003 (0.004)
Foreign Small/Mid Growth	-0.003 (0.001)	Real Estate	-0.082 (0.118)
Foreign Large Blend	0.002 (0.004)	Communications	-0.005 (0.017)
Large Value	0.009 (0.012)	Health	0.003 (0.002)
Mid-Cap Value	0.004 (0.08)	Industrials	0.005 (0.015)
Small Value	-0.001 (0.02)	Natural Resources	-0.003 (0.002)
Utilities	0.0151 (0.01)		

Notes: For each category, the Sequential Monte Carlo / Particle Filtering draws for the performance indicator (PERF) are averaged. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear in parentheses.

The corresponding density function of persistence according to parameter α_i is depicted by Figure 2a. Parameters α_i are fund-specific but time-invariant so we present the posterior distribution of (averaged across Monte Carlo draws) estimates of these

parameters. As expected there is fat tail to the left of the density towards negative values. This is of some interest, as negative performance would persist.¹⁵

Figure 2a: The sample distribution of posterior means of parameters α_i , the persistence.



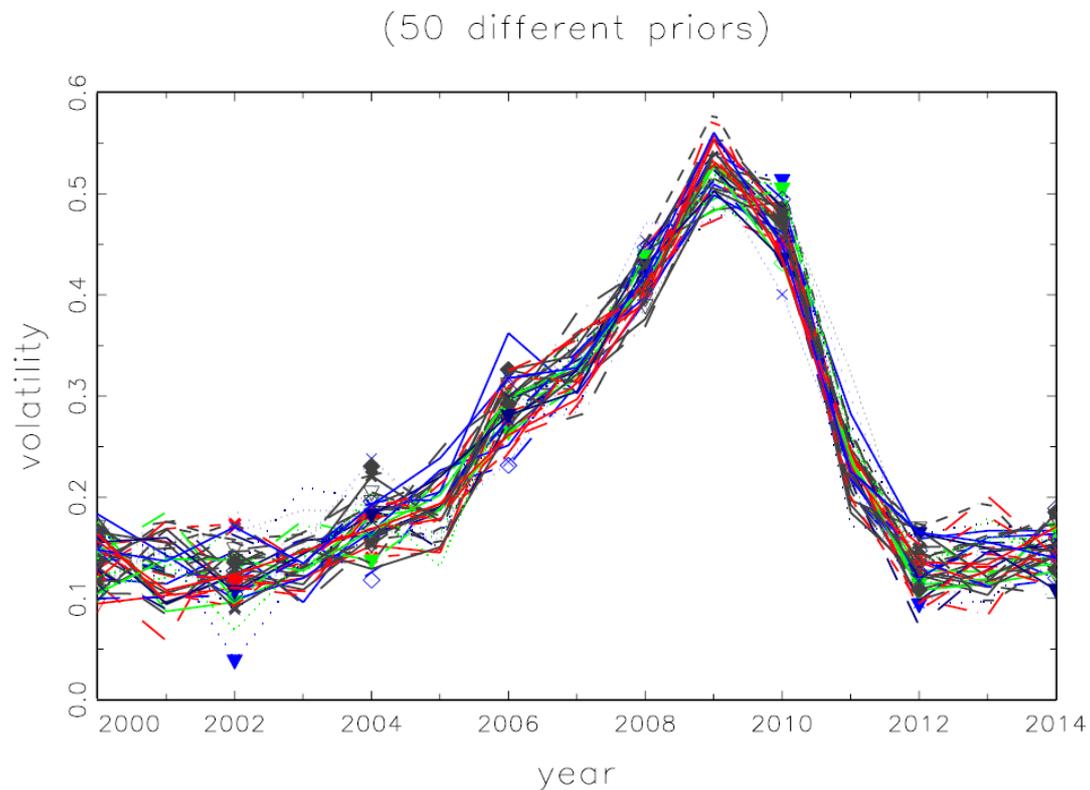
Notes: The figure sample distribution of posterior means of parameters α_i , see equation (2). We have used Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear dashed lines.

Funds' posterior mean volatility for 50 different priors, overtime, is presented in Figure 2b. These are filtered estimates from equation (14). In line with the performance results volatility picks in 2009 at the height of the financial crisis and stabilises thereafter. The striking characteristic of our measure of volatility is that it starts picking up as early as

¹⁵ Figure of the posterior distribution of α_{is} conditional on a 'significant' γ_{it} , that is the ratio of posterior mean to posterior standard deviation exceeds 2 in absolute value shows that persistence predominantly leans towards negative values (Figure available under request).

2003-2004, well before the financial crisis. Effectively, our modelling allows measuring risk in early stages and as such could act as early warning.

Figure 2b: The posterior mean funds' volatility ω_i over time average across all funds for 50 different priors.



Note: The figure presents the posterior mean funds' volatility ω_i over time average across all funds, see equation (14), for 50 different priors. We have used Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear dashed lines.

To further explore persistence over time Table 6 reports the average persistence over time. Note that persistence is negative over the whole sample, but during the financial crisis, that is from 2007 to 2009. Clearly during the financial crisis there is a regime change in funds' performance as reflected by the persistence parameter, which takes positive values. Effectively, we find evidence of a strong negative spiral that further

lowers levels of performance during the financial crisis due to strong persistence, though since 2010 persistence is subdued.

Table 6: The yearly average persistence over time.

Year	Mean	Standard deviation	Min	Max
2000	-0.042	0.017	-0.064	0.0032
2001	-0.051	0.022	-0.071	0.0049
2002	-0.055	0.021	-0.075	0.0083
2003	-0.042	0.025	-0.077	0.015
2004	-0.038	0.020	-0.082	0.017
2005	-0.021	0.019	-0.079	0.025
2006	-0.035	0.017	-0.081	0.027
2007	0.032	0.021	-0.087	0.012
2008	0.043	0.025	-0.091	0.007
2009	0.023	0.027	-0.088	0.003
2010	-0.032	0.031	-0.085	0.019
2011	-0.036	0.034	-0.082	0.022
2012	-0.039	0.035	-0.077	0.025
2013	-0.041	0.031	-0.078	0.021
2014	-0.043	0.027	-0.081	0.019
Average	-0.0251	0.0248	-0.0799	0.0152

Note: The persistence parameters α_i are fund-specific but time-invariant so we present the posterior distribution of (averaged across Monte Carlo draws) estimates of these parameters.

Table 7 further reports the average persistence across categories. Once more the dominant fund refers to the utilities with a positive persistence of 0.031 compared to small blend of -0.032.

Table 7: The average persistence indicator across funds categories.

Category	Persistence	Category	Persistence
Large Blend	0.017 (0.009)	Technology	0.022 (0.017)
Mid-Cap Blend	0.005 (0.002)	Financial	-0.032 (0.011)
Small Blend	-0.032 (0.014)	Consumer Cyclical	0.006 (0.003)
Large Growth	0.012 (0.008)	Equity Energy	-0.005 (0.008)
Mid-Cap Growth	-0.005 (0.003)	World Stock	-0.004 (0.031)
Small Growth	-0.009 (0.002)	Global Real Estate	-0.017 (0.003)
Foreign Large Growth	0.017 (0.003)	Consumer Defensive	0.007 (0.004)
Foreign Small/Mid Growth	0.005 (0.003)	Real Estate	-0.015 (0.005)
Foreign Large Blend	0.017 (0.008)	Communications	-0.043 (0.032)
Large Value	0.012 (0.007)	Health	0.005 (0.004)
Mid-Cap Value	0.005 (0.004)	Industrials	-0.032 (0.025)
Small Value	0.001 (0.001)	Natural Resources	0.005 (0.004)
Utilities	0.031 (0.017)		

Notes: For each category, the Sequential Monte Carlo / Particle Filtering draws for the persistence indicator (α_i) are averaged. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration. Posterior standard deviations appear in parentheses.

An interesting question is what happens at the tails of the distribution. A valid but messy way would be to produce posteriors for each fund across all Monte Carlo draws. Instead we provide posterior distributions of posterior mean persistence parameters a_i at different values of the right or left tail (see Figures A3a and A3b in Appendix II). Fixing, for example, the posterior mean at $a = 0.01$ we take all posterior draws in excess of this value and present the posteriors averaged across all Monte Carlo draws. On the other hand, fixing the posterior mean at $a = -0.01$ we take all posterior draws below this value. Once more, we observe that for negative persistence the densities lean towards negative values while there characterized as leptokurtic compared to positive values for persistence.

4.3 Best and worst fund performers.

Having derived the performance over time and across the main funds categories we turn to the ten best and worst fund performers for an equally weighted portfolio. In a recent Rossello (2015) provide a ranking of investment funds using nonparametric modelling. The results herein complemented previous rankings and extend the modelling techniques. Moreover, Table 8 reports the ten best and worst performance indicators over the sample period. During the financial crisis even the best performers turn to negative values, and this lasted till 2011. Since then there is a remarkable recovery, beyond the previous pick of the highest performance in 2005. Similarly, the 10 worst performers were hit by the financial crisis, much more harshly compared to the best 10 performers. However, when it comes to the worst performers, it appears that the recovery in 2014 is much stronger than the best performers though for the former volatility is clearly an issue.

Table 8: The Yearly average 10 best and worst performance indicators over time.

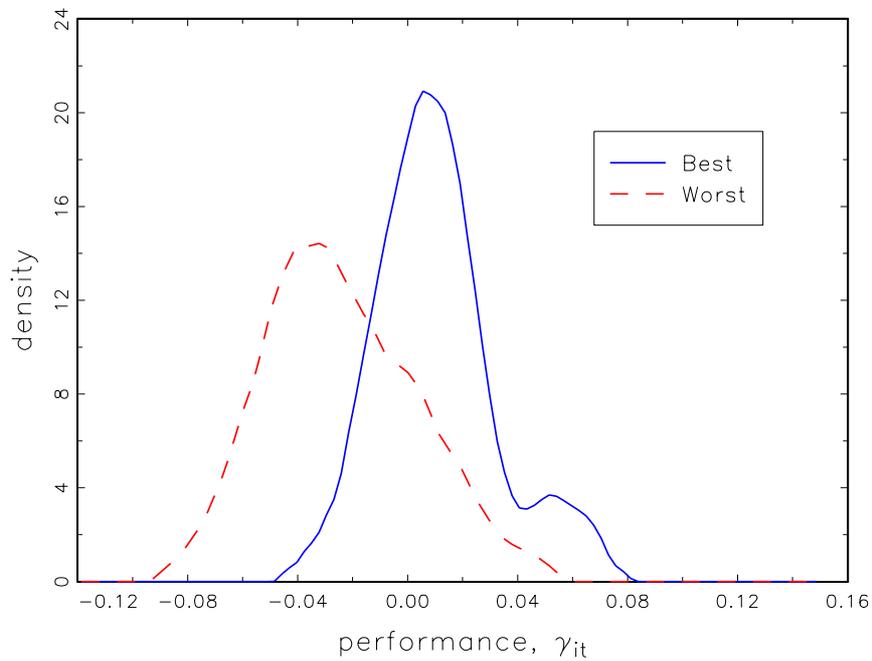
Year	Best				Worst			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
2000	0.085	0.077	-0.062	0.251	0.003	0.127	-0.241	0.192
2001	0.072	0.055	-0.031	0.182	0.005	0.125	-0.239	0.233
2002	0.079	0.082	-0.071	0.191	0.007	0.119	-0.255	0.241
2003	0.081	0.093	-0.115	0.257	0.0155	0.118	-0.217	0.255
2004	0.067	0.085	-0.101	0.229	0.0221	0.121	-0.221	0.26
2005	0.083	0.079	-0.082	0.212	0.0203	0.031	-0.052	0.087
2006	0.081	0.072	-0.061	0.225	0.0151	0.029	-0.036	0.019
2007	0.065	0.082	-0.105	0.203	-0.0915	0.103	-0.331	0.085
2008	-0.055	0.095	-0.132	0.103	-0.127	0.155	-0.567	0.061
2009	-0.062	0.084	-0.135	0.102	-0.0941	0.161	-0.515	0.211
2010	-0.032	0.065	-0.126	0.082	-0.0833	0.165	-0.485	0.174
2011	-0.005	0.032	-0.045	0.061	-0.104	0.132	-0.101	0.155
2012	0.0121	0.044	-0.072	0.06	0.005	0.144	-0.252	0.254
2013	0.0125	0.039	-0.065	0.065	0.009	0.152	-0.303	0.281
2014	0.0313	0.03	-0.031	0.0414	0.101	0.177	-0.241	0.211
Average	0.0343	0.0676	-0.0823	0.151	-0.0198	0.1239	-0.2704	0.1813

Notes: The table present the yearly PERF measures as in equation (25). Performance is derived from Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects. We use 10,000 particles per Monte Carlo iteration.

In Figure 3 we present the posterior distribution of g_{it} for an equally weighted portfolio of the ten best and ten worst funds (in terms of simple average returns in the sample).¹⁶ Clearly the worst performers are shifted to the left of the density of the best performers.

¹⁶ In addition, results are available under request of posterior distributions obtained through 20 different priors from the last column in Table 1 to examine prior sensitivity. These results provide evidence of the performance for the ten best (worst) funds ad are in line with the evidence herein.

Figure 3: The posterior distribution of g_{it}



Notes: The figure presents the posterior distribution of g_{it} for the best and the worst performers in the sample. The posterior distribution is drawn with Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects.

4.4 The fund of funds portfolio

Following from Basak and Makarov (2014) on the manager's portfolio choice where they argue that managers either win or lose we propose to model a fund of funds portfolio as robustness for the above findings. Moreover, we propose to solve a simple quadratic programming problem a la Markowitz:

$$\min_{w \geq 0} : w' m - A w' S w, \text{ s.t. } w \geq 0, w' \mathbf{1} = 1$$

where, in familiar notation, μ and Σ are taken from parameter estimates in (9).

We consider a fund-of-funds portfolio P that is formed by using the optimal weights w assuming that $A = 3$. Table 8 presents the performance and persistence of the optimal portfolio P .¹⁷ It is striking that performance in a la Markowitz portfolio (see Table 9) is well below the ten best fund performers, whilst persistence is also subdued compared to above results.

Table 9: The yearly average Markowitz performance and persistence indicators over time.

Year	Performance				Persistence			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
2000	0.017	0.011	-0.0051	0.022	-0.005	0.012	-0.032	0.014
2001	0.005	0.013	-0.0049	0.021	-0.007	0.011	-0.035	0.021
2002	0.009	0.011	-0.0061	0.024	-0.006	0.01	-0.036	0.023
2003	0.005	0.012	-0.0044	0.026	-0.002	0.009	-0.038	0.018
2004	0.007	0.013	-0.0032	0.032	-0.004	0.015	-0.034	0.028
2005	0.012	0.011	-0.0055	0.029	-0.005	0.019	-0.033	0.036
2006	0.013	0.012	-0.0071	0.029	-0.003	0.021	-0.027	0.039
2007	-0.015	0.013	-0.0265	0.035	-0.002	0.023	-0.038	0.042
2008	-0.022	0.022	-0.0277	0.021	-0.12	0.035	-0.226	0.085
2009	-0.008	0.035	-0.0414	0.009	-0.132	0.042	-0.185	0.072
2010	0.007	0.027	-0.0303	0.013	-0.181	0.037	-0.255	0.072
2011	0.009	0.025	-0.005	0.016	0.004	0.015	-0.044	0.022
2012	0.011	0.022	-0.007	0.015	0.003	0.014	-0.032	0.019
2013	0.012	0.025	-0.006	0.017	0.002	0.012	-0.033	0.015
2014	0.012	0.024	-0.006	0.019	-0.001	0.011	-0.025	0.013
Average	0.0049	0.0184	-0.0124	0.0219	-0.0306	0.0191	-0.0715	0.0346

Note: The table presents yearly average Markowitz performance and persistence indicators over time. The posterior distribution is drawn with Sequential Monte Carlo / Particle Filtering draws. We have used 15,000 draws the first 5,000 of which are discarded to mitigate possible start up effects.

¹⁷ The posterior distribution of optimal portfolio P is presented in Figure A3a and the posterior distribution of average persistence of its component returns is presented in Figure A3b, for 50 different priors from the last column in Table 1. In Figure A3c, we report the posterior distribution of its persistence only when average persistence of its component returns is ‘significant’ (viz. the ratio of posterior mean to posterior S.D. exceeds 2 in absolute value).

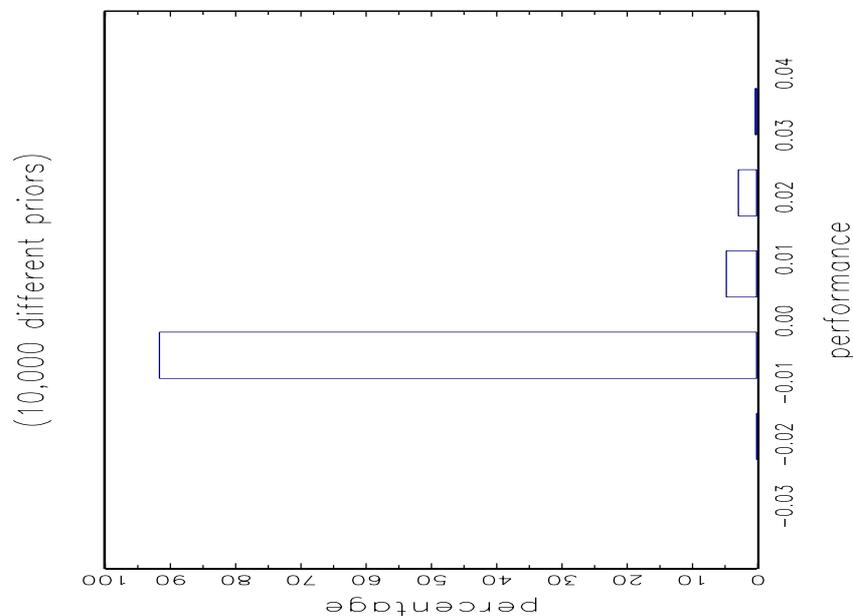
5. PRIOR SENSITIVITY ANALYSIS

An interesting question is what different priors imply about the funds' ability to perform better than the market. In Figure 4, we consider all 10,000 different priors reported in the last column of Table 1, and we present posterior estimates of temporal averages of

g_{it} whose $\tau = \frac{E(\bar{\gamma}_i)}{\sqrt{\text{var}(\bar{\gamma}_i)}}$, $\bar{\gamma}_i = T^{-1} \sum_{t=1}^T \gamma_{it}$ exceeds 2 or is lower than -2, for at least ten

funds. This presents direct evidence as to whether there are priors, which support the idea that funds perform better than the market.

Figure 4: funds' performance with all priors.



Note: all 10,000 different priors reported in the last column of Table 1.

Overall, the empirical results show that the baseline model and the adoption of different priors from Table 1 show that results are stable as reported from the mapping of the prior to the posterior.

7. CONCLUSIONS

This paper proposes a novel panel data model so as to capture time-varying heteroskedasticity, time-varying covariances of performance of US funds. Such modelling also permits general autocorrelation plus errors in underlying variables. Our results show that there has been striking variability across funds categories in terms of performance and persistence. Risk asserts a positive and significant impact on performance across different specifications. All Fama and French five factors show strong positive and significant effect on funds' performance. The exception is the RMW factor that turns negative, but it is not significant. There has been striking variability in terms of performance and persistence across funds categories and over time, and in particular through the financial crisis. Persistence in performance during the financial crisis is strong, insinuating a negative spiral to further lower levels of performance. The reported stochastic volatility exhibits a rising trend as early as 2003-2004 and could act as an early warning of future crisis. Finally, we show that our results are stable across different priors as reported from the mapping of the prior to the posterior of the Bayesian baseline model with the adoption of different priors.

The reported results have some implications for investors, professional managers and regulators. The revealed relationships could be part of investors' information set when select a fund whereas fund managers could benefit from the knowledge of components that enhance their portfolio performance. Finally, regulators and supervisory

authorities whose task is to safeguard a secure and well-functioning financial system may take into account that risk improves performance.

REFERENCES

- Andrieu, C., and G.O. Roberts. (2009) The pseudo-marginal approach for efficient computation. *Ann. Statistics*, 37, 697–725.
- Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle Markov chain Monte Carlo methods, *Journal of the Royal Statistical Society: Series B*, 72, 269–342.
- Annaert, J., J. Van Den BROECK, R. Vander Vennet, (2003). Determinants of mutual fund underperformance: a Bayesian stochastic frontier approach. *European Journal of Operational Research* 151, 617-632.
- Basak, S., D. Makarov, (2014). Strategic asset allocation in money management. *The Journal of Finance* 69, 179-217.
- Blake, D., Caulfield, T., Ioannidis, C. and Tonks, I., (2016). New Evidence on Mutual Fund Performance: A Comparison of Alternative Bootstrap Methods. *Journal of Financial and Quantitative Analysis*, forthcoming.
- Blake, D., Caulfield, T., Ioannidis, C. and Tonks, I. P., (2014). Improved inference in the evaluation of mutual fund performance using panel bootstrap methods. *Journal of Econometrics*, 183 (2), pp. 202-210.
- Berk, J.B., R.C. Green, (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy* 112, 1269-1295.
- Brown, K.C., W.V. Harlow, L.T. Starks, (1996). Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance*, 85-110.
- Brown, S.J., W.N. Goetzmann, J. Park, (2001). Careers and survival: Competition and risk in the hedge fund and CTA industry. *The Journal of Finance* 56, 1869-1886.
- Busse, J.A., (2001). Another look at mutual fund tournaments. *Journal of Financial and Quantitative Analysis* 36, 53-73.
- Cabello, J. M., Ruiz, F., Pérez-Gladish, B., & Méndez-Rodríguez, P. (2014). Synthetic indicators of mutual funds' environmental responsibility: An application of the reference point method. *European Journal of Operational Research*, 236(1), 313–325.
- Carhart, M.M., (1997). On persistence in mutual fund performance. *The Journal of finance* 52, 57-82.
- Cappé, O., E. Moulines, E., and T. Rydén (2005). *Inference in Hidden Markov Models*. Springer, Berlin.
- Casarin, R., J.-M. Marin (2007). Online data processing: Comparison of Bayesian regularized particle filters. University of Brescia, Department of Economics. Working Paper n. 0704.
- Chen, H.-l., G.G. Pennacchi, (2009). Does prior performance affect a mutual fund's choice of risk? Theory and further empirical evidence. *Journal of Financial and Quantitative Analysis* 44, 745-775.
- Chen, J., H. Hong, M. Huang, J.D. Kubik, (2004). Does Fund Size Erode Performance? Organizational Diseconomies and Active Money Management. *American Economic Review* 94, 1276-1302.
- Chen, J., H. Hong, W. Jiang, J.D. Kubik, (2013). Outsourcing mutual fund management: firm boundaries, incentives, and performance. *The Journal of Finance* 68, 523-558.
- Chevalier, J., G. Ellison, (1997). Risk Taking by Mutual Funds as a Response to Incentives. *The Journal of Political Economy* 105, 1167-1200.
- Christoffersen, S.E., D.K. Musto, (2002). Demand curves and the pricing of money management. *Review of Financial Studies* 15, 1499-1524.

Chopin, N. (2002). A sequential particle filter method for static models, *Biometrika* 89 (3) 539-551.

Chopin, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Annals of Statistics* 32, 2385-2411.

Cremers, K.M., A. Petajisto, (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies* 22, 3329-3365.

Cullen, G., D. Gasbarro, G.S. Monroe, J.K. Zumwalt, (2012). Changes to mutual fund risk: Intentional or mean reverting? *Journal of Banking & Finance* 36, 112-120.

Daraio, C., L. Simar, (2006). A robust nonparametric approach to evaluate and explain the performance of mutual funds. *European Journal of Operational Research* 175, 516-542.

Doucet, A., N. de Freitas, and N. Gordon (2001). *Sequential Monte Carlo Methods in Practice*. New York: Springer.

Durham, G., and J. Geweke, (2013). Adaptive sequential posterior simulators for massively parallel computing environments. Working Paper.

Fearnhead, P. (2007). Computational methods for complex stochastic systems: a review of some alternatives to MCMC. *Statistics and Computing*, 18(2):151-171.

Flury, T., and N. Shephard, (2011) Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models, *Econometric Theory* 27, 933-956.

Ferson, W. and Haitao Mo. (2016). Performance Measurement with Selectivity, Market and Volatility Timing, 2016), *Journal of Financial Economics*, forthcoming.

Ferson, W. and Jerchern Lin, (2014). Alpha and Performance Measurement: The Effects of Investor Disagreement and Heterogeneity, *Journal of Finance*, 69(4), 1565-1596.

Gilks, W.R., and C. Berzuini, (2001). Following a moving target: Monte Carlo inference for dynamic Bayesian models, *Journal of the Royal Statistical Society B* 63, 127-146.

De Jonghe, O., R. Vander Venet, (2008). Competition versus efficiency: What drives franchise values in European banking? *Journal of Banking & Finance* 32, 1820-1835.

Elton, E.J., M.J. Gruber, C.R. Blake, (2012). Does mutual fund size matter? The relationship between size and performance. *Review of Asset Pricing Studies*, ras001.

Fama and French, (2014). A Five-Factor Asset Pricing Model. Working Paper. Available at SSRN: <http://ssrn.com/abstract=2287202>.

Ferreira, M.A., A. Keswani, A.F. Miguel, S.B. Ramos, (2012). The determinants of mutual fund performance: A cross-country study. *Review of Finance*, rfs013.

Fu, X.M., Y.R. Lin, P. Molyneux, (2014). Bank competition and financial stability in Asia Pacific. *Journal of Banking & Finance* 38, 64-77.

Gaspar, J.M., M. Massa, P. Matos, (2006). Favoritism in mutual fund families? evidence on strategic cross-fund subsidization. *The Journal of Finance* 61, 73-104.

Gilks, W.R., and C. Berzuini, (2001). Following a moving target: Monte Carlo inference for dynamic Bayesian models, *Journal of the Royal Statistical Society B*, 63, 127-146.

Gil-Bazo, J., P. Ruiz-Verdú, (2008). When cheaper is better: Fee determination in the market for equity mutual funds. *Journal of Economic Behavior & Organization* 67, 871-885.

Gil-Bazo, J., P. Ruiz-Verdú, (2009). The relation between price and performance in the mutual fund industry. *The Journal of Finance* 64, 2153-2183.

Glode, Vincent, (2011). Why mutual funds underperform, *Journal of Financial Economics* 93, 546-559.

Goriaev, A., T.E. Nijman, B.J. Werker, (2005). Yet another look at mutual fund tournaments. *Journal of Empirical Finance* 12, 127-137.

Gordon, N.J. (1997). A hybrid bootstrap filter for target tracking in clutter. *IEEE Transactions on Aerospace and Electronic Systems* 33: 353– 358.

Gordon, N.J., D.J. Salmond, and A.F.M. Smith (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE-Proceedings-F* 140: 107–113.

Gruber, M.J., (1996). Another puzzle: The growth in actively managed mutual funds. *The Journal of Finance* 51, 783-810.

Guercio, D.D., J. Reuter, (2014). Mutual fund performance and the incentive to generate alpha. *The Journal of Finance* 69, 1673-1704.

Hall, J., M.K. Pitt, and R. Kohn (2014). Bayesian inference for nonlinear structural time series models. *Journal of Econometrics* 179 (2), 99–111.

Huang, J., C. Sialm, H. Zhang, (2011). Risk shifting and mutual fund performance. *Review of Financial Studies*, hhr001.

Huang, J., K.D. Wei, H. Yan, (2007). Participation costs and the sensitivity of fund flows to past performance. *The Journal of Finance* 62, 1273-1311.

In, F., M. Kim, R.J. Park, S. Kim, T.S. Kim, (2014). Competition of socially responsible and conventional mutual funds and its impact on fund performance. *Journal of Banking & Finance* 44, 160-176.

Jordan, B.D., T.B. Riley, (2015). Volatility and mutual fund manager skill. *Journal of Financial Economics* 118, 289-298.

Kempf, A., S. Ruenzi, (2008). Tournaments in mutual-fund families. *Review of Financial Studies* 21, 1013-1036.

Kempf, A., S. Ruenzi, T. Thiele, (2009). Employment risk, compensation incentives, and managerial risk taking: Evidence from the mutual fund industry. *Journal of Financial Economics* 92, 92-108.

Keswani, A., D. Stolin, (2006). Mutual Fund Performance Persistence and Competition: A Cross-sector Analysis *Journal of Financial Research*, 29, 349-366.

Keswani, A., D. Stolin, (2008). Which money is smart? Mutual fund buys and sells of individual and institutional investors. *The Journal of Finance* 63, 85-118.

Khorana, A., H. Servaes, (2012). What drives market share in the mutual fund industry? *Review of Finance* 16, 81-113.

Koop Gary, Dale J. Poirier, (2004). Bayesian variants of some classical semi-parametric regression techniques, *Journal of Econometrics*, Volume 123, Issue 2, December, Pages 259-282.

Koop, Gary, (2013). Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics* 28, 177-203.

Koski, J.L., J. Pontiff, (1999). How Are Derivatives Used? Evidence from the Mutual Fund Industry. *Journal of Finance* 54, 791-816.

Lou, D., (2012). A flow-based explanation for return predictability. *Review of Financial Studies* 25, 3457-3489.

Liu, J., M. West (2001). Combined parameter and state estimation in simulation-based filtering. In: Doucet, A., de Freitas, N., Gordon, N. (Eds.), *Sequential Monte Carlo Methods in Practice*. Springer-Verlag.

Lunde, A.; A. Timmermann; and D. Blake (1999). The Hazards of Mutual Fund Underperformance: A Cox Regression Analysis. *Journal of Empirical Finance*, 6, 121-52.

Malhotra, D., R. Martin, P. Russel, (2007). Determinants of cost efficiencies in the mutual fund industry. *Review of Financial Economics* 16, 323-334.

- Mamaysky, H., M. Spiegel, (2002). A theory of mutual funds: Optimal fund objectives and industry organization.
- Murthi, B., Y.K. Choi, P. Desai, (1997). Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach. *European Journal of Operational Research* 98, 408-418.
- Nemeth, C., P. Fearnhead (2014). Particle Metropolis adjusted Langevin algorithms for state-space models. Pre-print arXiv:1402.0694v1.
- Pitt, M.K., R.S. Silva, P. Giordani, and R. Kohn (2012). On some properties of Markov chain Monte Carlo simulation methods based on the particle filter, *Journal of Econometrics*, 171(2), 134–151.
- Pollet, J.M., M. Wilson, (2008). How does size affect mutual fund behavior? *The Journal of Finance* 63, 2941-2969.
- Ristic, B., S. Arulampalam, and N. Gordon (2004). *Beyond Kalman Filters: Particle Filters for Applications*. Norwood, MA: Artech House.
- Rossello, D. (2015). Ranking of investment funds: Acceptability versus robustness, *European Journal of Operational Research*, Volume 245, Issue 3, 16 September, Pages 828-836.
- Sato, Y., (2015). Fund Tournaments and Asset Bubbles. *Review of Finance*, rfv044.
- Sharpe, W.F., (1966). Mutual fund performance. *Journal of Business*, 119-138.
- Spiegel, M., H. Zhang, (2013). Mutual fund risk and market share-adjusted fund flows. *Journal of Financial Economics* 108, 506-528.
- Taylor, J., (2003). Risk-taking behavior in mutual fund tournaments. *Journal of Economic Behavior & Organization* 50, 373-383.
- Utz S., M. Wimmer, R. E. Steuer, (2015) Tri-criterion modeling for constructing more-sustainable mutual funds, *European Journal of Operational Research*, Volume 246, Issue 1, 1 October, Pages 331-338
- Wahal, S., A.Y. Wang, (2011). Competition among mutual funds. *Journal of Financial Economics* 99, 40-59.