Online Appendix

to

ESTIMATING MUTUAL FUND SKILL: A NEW APPROACH

Angie Andrikogiannopoulou
London School of Economics

Filippos Papakonstantinou
Imperial College London

August 2016

C.1 Hierarchical mixture model

In Figure C.1, we present the directed acyclic graph representation of the hierarchical version of our baseline mixture model described in Section 2 of the paper. In this graph, squares represent quantities that are fixed or observed, e.g., prior parameters and data, while circles represent unknown model parameters that need to be estimated.

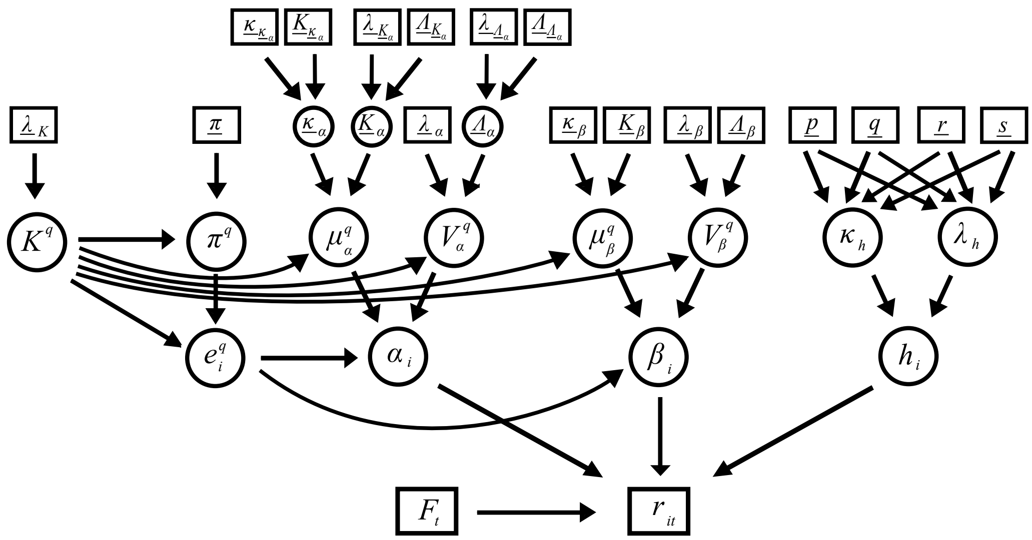


Figure C.1: Representation of the hierarchical mixture model of fund returns as a directed acyclic graph. Squares represent quantities that are fixed or observed, e.g., prior parameters and data, while circles represent unknown model parameters that need to be estimated.

Comparing the graph in this figure with that for the non-hierarchical mixture model in Figure 1 in the paper, we see that the difference between the two versions is that the hierarchical one takes as given the prior distributions for the population parameters. First, this is necessary, since classical estimation would be intractable. Second, we generally use weak priors, and furthermore we perform a prior sensitivity analysis which shows that our posteriors are quite robust to varying the priors (see Section 8.1 in the paper for a brief summary and Section C.12 in this appendix for details).

C.2 Simulations

Here, we present additional results on the simulations we perform in Section 3 of the paper.

First, we present results relating to the first set of our simulations (see Table 1 in the paper), in which we generate alphas from mixed distributions and compare our estimated proportions of skill types with those obtained from fund-level hypothesis tests, with and without the FDR correction. In Table C.1, we present the true percentiles of each simulated alpha distribution, as well as point and interval estimates — the posterior mean and 90% Highest Posterior Density Interval (HPDI) — of these percentiles using our methodology. These results show that our methodology is flexible enough to estimate well not only the proportions of skill types but also the entire alpha distribution even, e.g., in cases in which nonzero alphas are discrete or normal.

Next, we present results relating to the second set of our simulations (see Figure 2 in the paper), in which we generate alphas from continuous distributions and compare our estimated distribution of alpha with that of the hierarchical normal model. In Table C.2, we present the true percentiles of each simulated alpha distribution, as well as point and interval estimates — the posterior mean and 90% HPDI — of these percentiles using our methodology. The results in this table show that our model is flexible enough to estimate reasonably well the entire alpha distribution in the case of both a normal distribution as well as a skewed and fat-tailed distribution without a point mass at zero. In Table C.3, we present the true percentiles of each simulated alpha distribution, as well as point and interval estimates — the posterior mean and 90% HPDI — of these percentiles using the hierarchical normal model. As expected, the results in this table show that the normal model can accurately estimate the distribution if alphas are drawn from a normal, but grossly mis-estimates it if alphas are drawn from a skewed and fat-tailed distribution.

Table C.1: Simulations with Mixed Distributions — True and Estimated Percentiles of Fund Skill Distribution

Results from simulations in which alphas (expressed as annualized percentages) are generated from mixed distributions with a point mass at zero and with nonzero alphas drawn from a discrete distribution (in Panel A), a normal distribution (in Panel B), and a log-normal distribution (in Panel C). The data generating processes (DGPs) within each panel differ in the proportions π^0, π^-, π^+ of funds with zero, negative, and positive alpha, respectively, and/or in the distance of nonzero alphas from zero. In Panel A, $\alpha \sim \pi^0\delta_0 + \pi^-\delta_{x^-} + \pi^+\delta_{x^+}$, with large nonzero alphas $x^- = -3.2$ and $x^+ = 3.8$ and unequal proportions $\pi^0 = 0.75$, $\pi^- = 0.23$, $\pi^+ = 0.02$ (DGP D-1), and with small nonzero alphas $x^- = -1.2$ and $x^+ = 1.8$ and equal proportions $\pi^0 = 0.34$, $\pi^- = 0.33$, $\pi^+ = 0.33$ (DGP D-2). In Panel B, $\alpha \sim \pi^0\delta_0 + \pi^{-,+} f_{\mathcal{N}}\left(\alpha \,\middle|\, -1.45, \sigma^2\right)$ with large variance $\sigma^2 = 72$ and a large point mass $\pi^0 = 0.90$ (DGP N-1), and with small variance $\sigma^2 = 7.2$ and a smaller point mass $\pi^0 = 0.35$ (DGP N-2). In Panel C, $\alpha \sim \pi^0\delta_0 + \pi^- f_{\ln\mathcal{N}}\left(|\alpha| \,\middle|\, \mu, \sigma^2\right) + \pi^+ f_{\ln\mathcal{N}}\left(|\alpha| \,\middle|\, \mu, \sigma^2\right)$ with nonzero alphas far from zero i.e. $\mu = 2$ and $\sigma^2 = 0.2$ (DGP L-1), and close to zero i.e. $\mu = 1$ and $\sigma^2 = 0.35$ (DGP L-2), and with proportions $\pi^0 = 0.45$, $\pi^- = 0.28$, $\pi^+ = 0.27$ in both cases. For each DGP, we report the true percentiles of the alpha distribution and their posterior mean and 90% HPDI estimated using our methodology.

		Percentiles														
		0.5^{th}	**1^{st}**	**5^{th}**	**10^{th}**	**20^{th}**	**30^{th}**	**40^{th}**	**50^{th}**	**60^{th}**	**70^{th}**	**80^{th}**	**90^{th}**	**95^{th}**	**99^{th}**	**99.5^{th}**
Panel A: Discrete nonzero alphas																
DGP D-1: *Large alphas*	True	-3.20	-3.20	-3.20	-3.20	-3.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.80	3.80
	Posterior Mean	-3.66	-3.57	-3.35	-3.22	-2.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.73	3.97
	5%	-3.78	-3.69	-3.42	-3.26	-3.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.51	3.78
	95%	-3.51	-3.46	-3.30	-3.17	-2.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.91	4.16
DGP D-2: *Small alphas*	True	-1.20	-1.20	-1.20	-1.20	-1.20	-1.20	0.00	0.00	0.00	1.80	1.80	1.80	1.80	1.80	1.80
	Posterior Mean	-1.52	-1.46	-1.30	-1.23	-1.12	-0.97	0.00	0.00	0.00	1.58	1.75	1.88	1.97	2.14	2.21
	5%	-1.66	-1.58	-1.37	-1.27	-1.16	-1.05	0.00	0.00	0.00	1.50	1.70	1.83	1.91	2.02	2.06
	95%	-1.37	-1.34	-1.24	-1.18	-1.06	-0.84	0.00	0.00	0.00	1.66	1.79	1.92	2.03	2.24	2.32
Panel B: Normal nonzero alphas																
DGP N-1: *Large variance*	True	-15.48	-12.08	-1.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	9.65	12.81
	Posterior Mean	-15.95	-11.75	-1.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	9.03	12.50
	5%	-18.19	-13.24	-3.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.85	10.92
	95%	-13.97	-10.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	10.21	14.39
DGP N-2: *Small variance*	True	-7.76	-7.13	-5.24	-4.18	-2.78	-1.65	-0.66	0.00	0.00	0.00	0.00	1.39	2.42	4.51	5.25
	Posterior Mean	-8.02	-7.15	-5.10	-4.11	-2.81	-1.60	-0.75	0.00	0.00	0.00	0.06	1.28	2.26	4.32	5.25
	5%	-8.60	-7.56	-5.28	-4.26	-2.97	-1.78	-0.95	0.00	0.00	0.00	0.00	1.06	2.05	3.98	4.76
	95%	-7.54	-6.81	-4.93	-3.95	-2.65	-1.43	0.00	0.00	0.00	0.04	0.38	1.45	2.46	4.73	5.89
Panel C: Log-normal nonzero alphas																
DGP L-1: *Far from zero*	True	-19.92	-17.76	-12.38	-9.80	-6.32	0.00	0.00	0.00	0.00	0.00	5.87	9.30	11.97	18.15	21.65
	Posterior Mean	-20.93	-18.30	-12.29	-9.56	-6.25	0.00	0.00	0.00	0.00	0.00	6.11	9.46	12.18	18.18	20.81
	5%	-22.04	-19.14	-12.71	-9.92	-6.60	0.00	0.00	0.00	0.00	0.00	5.75	9.15	11.78	17.36	19.76
	95%	-19.87	-17.56	-11.87	-9.25	-5.94	0.00	0.00	0.00	0.00	0.00	6.40	9.77	12.63	19.01	21.90
DGP L-2: *Close to zero*	True	-9.81	-8.42	-5.23	-3.84	-2.15	0.00	0.00	0.00	0.00	0.00	1.95	3.58	5.00	8.67	10.95
	Posterior Mean	-10.19	-8.59	-5.17	-3.75	-2.15	0.00	0.00	0.00	0.00	0.00	2.05	3.68	5.16	8.78	10.51
	5%	-11.01	-9.22	-5.43	-3.94	-2.31	0.00	0.00	0.00	0.00	0.00	1.88	3.51	4.92	8.24	9.78
	95%	-9.45	-8.06	-4.95	-3.57	-1.97	0.00	0.00	0.00	0.00	0.00	2.23	3.86	5.39	9.34	11.31

Table C.2: Simulations with Continuous Distributions — True and Estimated Percentiles of Skill Distribution from Our Model

Results from simulations in which alphas (expressed as annualized percentages) are generated from continuous-distribution data generating processes (DGPs) and are estimated using our methodology. In Panel A we present results for alphas simulated from a normal distribution, and in Panel B we present results for alphas simulated from a negatively-skewed and fat-tailed distribution. Specifically, in Panel A, we present results for DGP C-1, i.e., $\alpha \sim f_{\mathcal{N}}\left(\cdot \,\middle|\, \mu, \sigma^2\right)$, with $\mu = -2.5$ and $\sigma^2 = 4$. In Panel B, we present results for DGP C-2, i.e., $\alpha \sim \pi^N f^N + \pi^- f^- + \pi^+ f^+$, with $\pi^N = 0.10$, $\pi^- = 0.80$, $\pi^+ = 0.10$, $f^N(\alpha) = f_{\mathcal{N}}(\alpha \,|0, 0.1)$, $f^-(\alpha) = f_{\ln\mathcal{N}}(|\alpha| \,|0.1, 0.5)$ for $\alpha < 0$, and $f^+(\alpha) = f_{\ln\mathcal{N}}(|\alpha| \,|0.1, 0.5)$ for $\alpha > 0$. For each DGP, we report the true percentiles of the alpha distribution and their posterior mean and 90% HPDI estimated using our methodology. The 90% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.90.

		Percentiles														
		0.5$^{\text{th}}$	**1$^{\text{st}}$**	**5$^{\text{th}}$**	**10$^{\text{th}}$**	**20$^{\text{th}}$**	**30$^{\text{th}}$**	**40$^{\text{th}}$**	**50$^{\text{th}}$**	**60$^{\text{th}}$**	**70$^{\text{th}}$**	**80$^{\text{th}}$**	**90$^{\text{th}}$**	**95$^{\text{th}}$**	**99$^{\text{th}}$**	**99.5$^{\text{th}}$**
Panel A: Normal alphas																
DGP C-1	True	-7.56	-6.95	-5.70	-5.01	-4.15	-3.51	-2.99	-2.47	-1.96	-1.46	-0.80	0.11	0.76	1.96	2.57
	Posterior Mean	-8.34	-7.55	-5.74	-4.92	-4.02	-3.41	-2.91	-2.44	-1.94	-1.41	-0.94	0.00	1.04	1.94	2.31
	5%	-9.02	-8.07	-5.93	-5.06	-4.14	-3.52	-3.02	-2.56	-2.07	-1.59	-1.11	0.00	0.86	1.73	2.01
	95%	-7.72	-7.13	-5.56	-4.78	-3.88	-3.28	-2.79	-2.32	-1.78	-1.26	-0.78	0.56	1.21	2.17	2.65
Panel B: Negatively skewed and fat-tailed alphas																
DGP C-2	True	-6.53	-5.43	-3.21	-2.47	-1.74	-1.35	-1.10	-0.88	-0.67	-0.50	-0.27	0.42	1.11	2.75	3.42
	Posterior Mean	-6.31	-5.21	-3.11	-2.36	-1.69	-1.32	-1.06	-0.85	-0.67	-0.50	-0.27	0.29	1.13	2.78	3.59
	5%	-6.97	-5.67	-3.29	-2.48	-1.76	-1.39	-1.12	-0.92	-0.74	-0.56	-0.36	0.00	0.91	2.45	3.11
	95%	-5.69	-4.78	-2.92	-2.24	-1.61	-1.25	-1.00	-0.79	-0.62	-0.43	0.00	0.57	1.31	3.17	4.23

Table C.3: Simulations with Continuous Distributions — True and Estimated Percentiles of Skill Distribution from Hierarchical Normal

Results from simulations in which alphas (expressed as annualized percentages) are generated from continuous-distribution data generating processes (DGPs) and are estimated using the hierarchical normal model. In Panel A we present results for alphas simulated from a normal distribution, and in Panel B we present results for alphas simulated from a negatively-skewed and fat-tailed distribution. Specifically, in Panel A, we present results for DGP C-1, i.e., $\alpha \sim f_{\mathcal{N}}\left(\cdot \left| \mu, \sigma^2\right.\right)$, with $\mu = -2.5$ and $\sigma^2 = 4$. In Panel B, we present results for DGP C-2, i.e., $\alpha \sim \pi^N f^N + \pi^- f^- + \pi^+ f^+$, with $\pi^N = 0.10$, $\pi^- = 0.80$, $\pi^+ = 0.10$, $f^N(\alpha) = f_{\mathcal{N}}(\alpha\,|0, 0.1)$, $f^-(\alpha) = f_{\ln\mathcal{N}}(|\alpha|\,|0.1, 0.5)$ for $\alpha < 0$, and $f^+(\alpha) = f_{\ln\mathcal{N}}(|\alpha|\,|0.1, 0.5)$ for $\alpha > 0$. For each DGP, we report the true percentiles of the alpha distribution and their posterior mean and 90% HPDI estimated using the hierarchical normal model. The 90% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.90.

		Percentiles														
		0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
Panel A: Normal alphas																
DGP C-1	True	-7.56	-6.95	-5.70	-5.01	-4.15	-3.51	-2.99	-2.47	-1.96	-1.46	-0.80	0.11	0.76	1.96	2.57
	Posterior Mean	-7.60	-7.10	-5.73	-5.00	-4.12	-3.48	-2.94	-2.43	-1.92	-1.38	-0.75	0.14	0.87	2.23	2.73
	5%	-7.75	-7.24	-5.84	-5.10	-4.20	-3.56	-3.02	-2.51	-2.00	-1.46	-0.83	0.04	0.74	2.07	2.56
	95%	-7.43	-6.95	-5.62	-4.90	-4.04	-3.41	-2.87	-2.37	-1.86	-1.31	-0.67	0.23	0.97	2.38	2.89
Panel B: Negatively skewed and fat-tailed alphas																
DGP C-2	True	-6.53	-5.43	-3.21	-2.47	-1.74	-1.35	-1.10	-0.88	-0.67	-0.50	-0.27	0.42	1.11	2.75	3.42
	Posterior Mean	-4.48	-4.14	-3.20	-2.71	-2.10	-1.67	-1.30	-0.95	-0.60	-0.23	0.20	0.81	1.30	2.24	2.58
	5%	-4.63	-4.27	-3.31	-2.79	-2.17	-1.73	-1.35	-1.01	-0.66	-0.29	0.14	0.73	1.22	2.13	2.46
	95%	-4.35	-4.02	-3.11	-2.62	-2.04	-1.61	-1.24	-0.89	-0.55	-0.17	0.27	0.89	1.40	2.36	2.71

C.3 Summary statistics

Here, we present summary information for the funds in the two samples of actively managed open-end US equity funds that we use in our analyses in Sections 5 through 8 of the paper: the baseline sample of 3,497 funds and the restricted sample with reliable investment objective data for 1,865 funds.

Table C.4: Summary Statistics of Fund Characteristics

Summary statistics of fund characteristics for the two samples of actively managed open-end US equity funds used in the empirical analyses in Sections 5 through 8 of the paper. In Panel A, we present summary statistics for the baseline sample of 3,497 funds, and in Panel B for the restricted sample of 1,865 funds with reliable investment objective information; both samples span the period January 1975 through December 2011. Fund age is the number of years since the fund's establishment. Total net asset value (TNAV) is measured in millions of dollars. Expense ratio is defined as total annual management, administrative, and 12b-1 fees and expenses divided by year-end TNAV, and is expressed as a percentage. Turnover ratio is defined as the minimum of aggregate purchases and sales of securities divided by the average TNAV over the calendar year, and is expressed as a percentage. Fund inflows are defined as the net fund flows into the mutual fund over the calendar year, divided by the TNAV at the end of the previous calendar year, and they are expressed as a percentage; negative values indicate net outflows. The summary statistics reported are calculated across all fund-months in each sample.

Panel A: Baseline sample

			Percentiles						
	Mean	**Std.Dev.**	**5th**	**10th**	**25th**	**50th**	**75th**	**90th**	**95th**
Fund age	12.56	13.20	1	2	4	8	16	30	42
Total net asset value	942	3,992	4	10	37	143	551	1,747	3,634
Expense ratio	1.31%	1.00%	0.27%	0.66%	0.95%	1.24%	1.58%	1.97%	2.24%
Turnover ratio	96%	161%	10%	17%	34%	66%	116%	183%	249%
Fund inflows	46%	159%	−37%	−27%	−14%	1.4%	33%	124%	273%

Panel B: Restricted sample

			Percentiles						
	Mean	**Std.Dev.**	**5th**	**10th**	**25th**	**50th**	**75th**	**90th**	**95th**
Fund age	15.25	14.26	1	2	5	11	20	35	47
Total net asset value	1,245	4,804	7	14	53	202	765	2,403	4,928
Expense ratio	1.29%	0.96%	0.12%	0.65%	0.94%	1.22%	1.54%	1.95%	2.21%
Turnover ratio	88%	109%	10%	17%	34%	66%	113%	177%	237%
Fund inflows	37%	142%	−36%	−27%	−14%	−0.5%	27%	101%	217%

Table C.5: Assignment of Funds to Investment Strategies

The number and fraction of funds allocated to each investment objective — Growth & Income, Growth, Aggressive Growth — in the restricted sample of 1,865 funds with reliable investment objective information from the Thomson database.

Investment Objective	**# of Funds**	**%age of Funds**
Growth & Income	405	21.7%
Growth	1,230	66.0%
Aggressive Growth	230	12.3%

C.4 Fund fees

Here, we present a plot of the empirical density of annual fund fees and expenses, expressed as a percent of total net asset value. This empirical density is constructed from the average (across the lifetime of each fund) annual fees and expenses for the 3,497 funds in our sample. Fees and expenses are reported annually in the CRSP Survivorship-Bias-Free US Mutual Fund Database, and they include annual management, administrative, and 12b-1 fees, and expenses.

The empirical density of fees and expenses shown in Figure C.2 has a mode at 0.95%. The mean, median, and standard deviation of fees and expenses are 1.16%, 1.09%, and 0.68%, respectively.

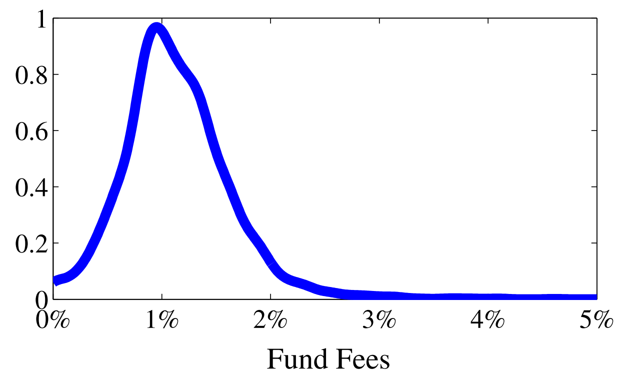


Figure C.2: Plot of the empirical density of annual fund fees and expenses (expressed as a percent of total net asset value) across 3,497 funds.

C.5 The distribution of skill

In this section, we present some additional figures and tables regarding the estimation of the baseline model presented in Section 2 of the paper using returns net of expenses for 3,497 funds. These results supplement those presented in Section 5 of the paper.

First, we present results on the posterior distributions of the population mean and standard deviation of alpha and the factor loadings (in Table C.6), and of the population correlations between factor loadings (in Table C.7). We present these results conditional on $K^- = 2$, $K^+ = 1$, rather than presenting tables for each of the 16 possible models.

Table C.6: Population Mean and Standard Deviation of Alpha and Factor Loadings

Results on the posterior distributions of the population mean and standard deviation of annualized alpha (expressed as a percent) and the factor loadings, estimated using our baseline model presented in Section 2 with returns net of expenses, conditional on the model with the highest posterior probability, i.e., $K^- = 2$ and $K^+ = 1$. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter. The population mean and variance of alpha for the zero-alpha funds is constrained to equal zero.

	Means					Standard Deviations				
	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**	**Mean**	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
α^0	0	0	0	[0 , 0]	0	0	0	0	[0 , 0]	0
α_1^-	−1.11	−1.11	0.16	[−1.39 , −0.75]	0.01	0.66	0.64	0.19	[0.36 , 1.16]	0.01
α_2^-	−2.11	−1.50	1.68	[−7.39 , −0.93]	0.07	3.04	2.06	3.05	[0.81 , 9.90]	0.08
α^+	1.04	1.00	0.30	[0.54 , 1.77]	0.02	1.29	1.22	0.37	[0.77 , 2.18]	0.02
β_{M}	0.95	0.95	0.00	[0.94 , 0.96]	0.00	0.21	0.21	0.00	[0.21 , 0.22]	0.00
β_{SMB}	0.19	0.19	0.01	[0.18 , 0.20]	0.00	0.31	0.31	0.00	[0.30 , 0.31]	0.00
β_{HML}	0.02	0.02	0.01	[0.01 , 0.03]	0.00	0.34	0.34	0.00	[0.33 , 0.35]	0.00
β_{UMD}	0.01	0.01	0.00	[0.01 , 0.01]	0.00	0.10	0.10	0.00	[0.10 , 0.11]	0.00

Table C.7: Population Correlation Matrix

Means and standard deviations (in parentheses) of the posteriors of population correlations between the factor loadings, for our baseline model presented in Section 2 with returns net of expenses.

	β_{SMB}	β_{HML}	β_{UMD}
β_{M}	0.30	−0.47	0.11
	(0.01)	(0.01)	(0.01)
β_{SMB}		−0.12	0.18
		(0.01)	(0.01)
β_{HML}			−0.44
			(0.01)

As explained in Section 2.4 of the paper, to estimate our model we need to derive the joint posterior distribution of the model parameters conditional on the data. Since this joint posterior cannot be calculated analytically, we obtain information about it by drawing from it using a Markov chain Monte Carlo (MCMC) algorithm. Section 2.4 of the paper and Section B of the paper's appendix provide details about the MCMC algorithm we employ.

Using this algorithm, we make 5 million draws from which we discard the first 10% as burn-in and retain every 50$^{\text{th}}$ after that to mitigate serial correlation. These draws form a Markov chain with stationary distribution equal to the joint posterior. In Figure C.3, we present trace plots (plots of the retained draws against the iteration number) of the proportions of funds with zero, negative, and positive alpha. In Figure C.4, we present trace plots of the population means and variances of the distributions of alpha for negative- and positive-alpha funds, and of the factor loadings, conditional on the highest-posterior-probability model, i.e., $K^- = 2$, $K^+ = 1$. These plots indicate no convergence problems.

We note that, in mixture models, the posterior distribution of parameters is invariant to permutations of the components' labels. As a result, inference regarding parameters that are not invariant to component relabeling in the MCMC draws is problematic. We circumvent this issue in two ways. First, we focus on inferences that are invariant to label switching, i.e., inference on: the numbers of components K^-, K^+; the population proportions π^0, π^-, π^+; the population mean μ_β and variance V_β of the loadings; the population shape κ_h and scale λ_h of the error distribution; the individual-level alpha α_i, loadings β_i, and error precisions h_i; the individual-level latent allocations to groups e_i^0, $\sum_{1\leq k\leq K^-} e_{i,k}^-$, $\sum_{1\leq k\leq K^+} e_{i,k}^+$; and the density of alpha and the loadings. Second, to conduct inferences that are not invariant to label switching, i.e., on component-specific probabilities $\{\pi_k^q\}$ and distribution parameters $\{(\mu_{\alpha,k}^q, V_{\alpha,k}^q)\}$, we retrospectively relabel components in the MCMC draws so the estimated marginal posteriors of parameters of interest are close to unimodality (see Stephens, 1997). This achieves a unique labeling throughout the draws, so we obtain point estimates through averaging over the draws.[1] We see in Figure C.4 that we have successfully removed the label-switching behavior from the means and variances.

[1] We do not impose artificial identifiability restrictions through the priors, because they do not guarantee a unique labeling and can produce biased estimates (see Celeux, 1998). Also, see Jasra, Holmes and Stephens (2005) for a review of the various methods that have been proposed to solve the label switching problem.

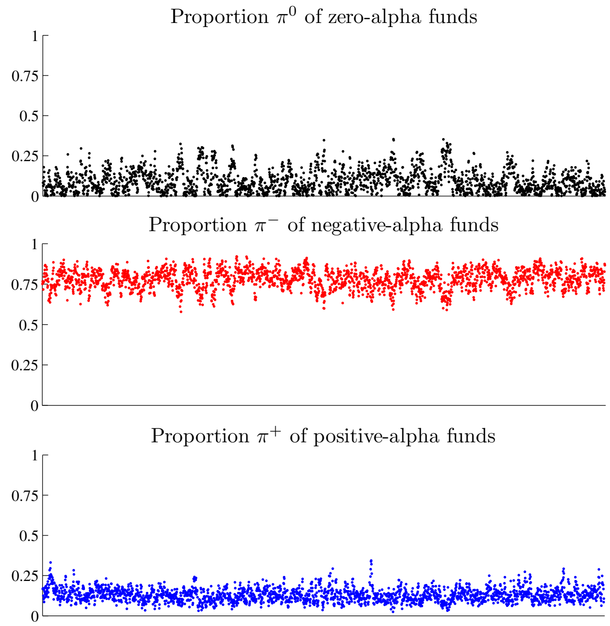


Figure C.3: Trace plots of the MCMC draws for the population proportions of zero-alpha funds (in the top panel, using black dots), negative-alpha funds (in the middle panel, using red dots), and positive-alpha funds (in the bottom panel, using blue dots).

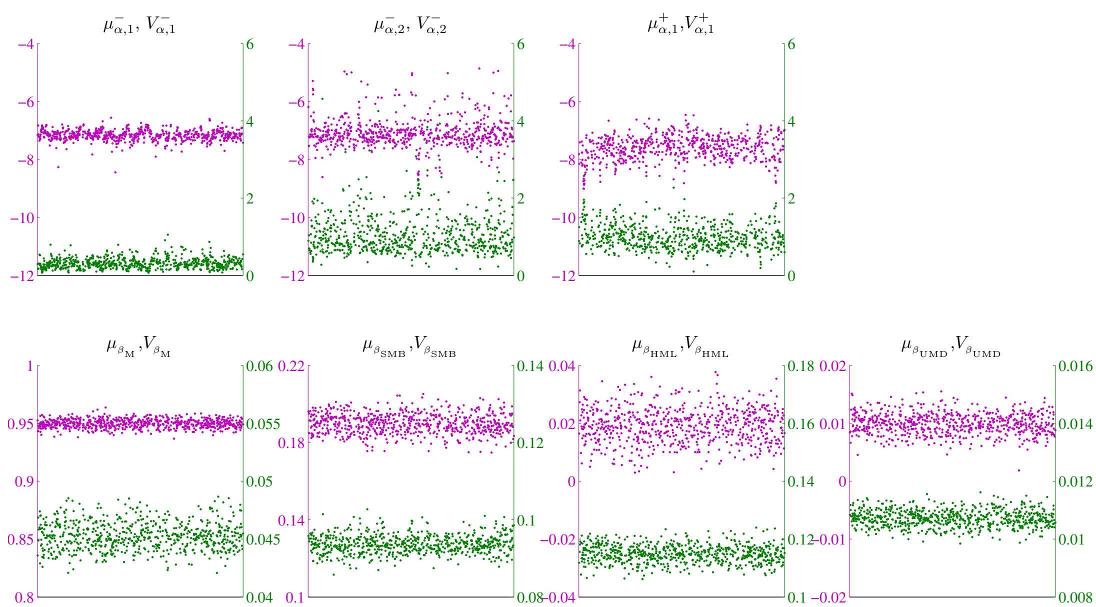


Figure C.4: Trace plots of the MCMC draws for the population means (purple dots toward the top of each panel, with values associated with the left vertical axes) and the population variances (green dots toward the bottom of each panel, with values associated with the right vertical axes) of alpha and the factor loadings, conditional on the model with the highest posterior probability, i.e., $K^- = 2$, $K^+ = 1$. The mean and variance of alpha are those of the underlying normal distribution.

C.6 Robust skewness, tail weight, and distance for standard distributions

In this section, we present robust quantile-based measures of skewness and tail weight for various well-known distributions, to provide context for the measures we calculate for the alpha distribution we estimate in Section 5 of the paper. We also present distance measures between the standard normal and various well-known distributions, again to provide context for the distance we calculate between the alpha distribution estimated from our model and from the hierarchical normal model in Section 5 of the paper.

The robust measure of skewness for our estimated alpha distribution is -0.20, its left tail weight is 0.34, and its right tail weight is 0.27 (all quoted in excess of the values corresponding to the normal; see Table 5 in the paper). In Table C.8, we see that the robust skewness is similar (in absolute value) to that of a χ^2 distribution that has between 30 and 50 degrees of freedom, at 0.22 and 0.17 respectively, and the left and right tail weight measures are similar to those of a $t\,(2)$ and a $t\,(3)$ distribution, at 0.36 and 0.28 respectively.

The Hellinger distance between our estimated alpha distribution and the one estimated from the normal model is $H^2 = 0.11$ (see Table 6 in the paper). As we can see in Table 5, this is close to i) the Hellinger distance ($H^2 = 0.11$) between two normals that have the same mean but one has twice the standard deviation of the other, ii) the Hellinger distance ($H^2 = 0.08$) between the standard normal $\mathcal{N}\,(0, 1)$ and a $\chi^2\,(3)$ distribution that is scaled to have the same mean and variance as the standard normal, and iii) the Hellinger distance ($H^2 = 0.11$) between the standard normal $\mathcal{N}\,(0, 1)$ and the $t\,(1)$ distribution. The Wasserstein distance between our estimated distribution and the one estimated from the normal model is $W = 0.22$ (see Table 6 in the paper). From 5, we can also see that this distance is close to i) the Wasserstein distance ($W = 0.19$) between a standard normal and a normal that has the same mean but 25% smaller/greater standard deviation, ii) the Wasserstein distance ($W = 0.21$) between the standard normal $\mathcal{N}\,(0, 1)$ and a $\chi^2\,(4)$ distribution that is scaled to have the same mean and variance as the standard normal, and iii) the Wasserstein distance ($W = 0.25$) between the standard normal $\mathcal{N}\,(0, 1)$ and the $t\,(3)$ distribution.

Table C.8: Robust Measures of Skewness and Tail Weight for the normal, χ^2, and t distributions

Robust quantile-based measures of skewness and tail weight that rely on 99% of the range of each distribution, for the normal distribution (in Panel A), and for the χ^2 (in Panel B) and t distributions (in Panel C) with various degrees of freedom. The measure of skewness is as in Groeneveld and Meeden (1984) — $S := \left[Q(1-p)+Q(p)-2Q(0.5)\right]/\left[Q(1-p)-Q(p)\right]$ — and the measures of left and right tail weight are as in Brys, Hubert, and Struyf (2006) — $LTW := -\left[Q\left(\frac{1-p}{2}\right)+Q\left(\frac{p}{2}\right)-2Q(0.25)\right]/\left[Q\left(\frac{1-p}{2}\right)-Q\left(\frac{p}{2}\right)\right]$ and $RTW := \left[Q\left(\frac{1+q}{2}\right)+Q\left(1-\frac{q}{2}\right)-2Q(0.75)\right]/\left[Q\left(\frac{1+q}{2}\right)-Q\left(1-\frac{q}{2}\right)\right]$ — where $Q(x)$ is the x^{th} quantile of the distribution, and we use $p=0.005$ and $q=0.995$. The measures are reported as deviations from the corresponding values for the normal distribution (0 for the skewness and 0.52 for the left and right tail weight measures).

Panel A: $\mathcal{N}\left(\mu,\sigma^2\right)$ distribution

	Quantile Skewness (S)	Left Tail Weight (LTW)	Right Tail Weight (RTW)
$\forall \mu, \sigma^2$	0	0	0

Panel B: $\chi^2(k)$ distributions

	Quantile Skewness (S)	Left Tail Weight (LTW)	Right Tail Weight (RTW)
$k=3$	0.64	−0.51	0.19
$k=4$	0.57	−0.41	0.17
$k=5$	0.52	−0.34	0.16
$k=10$	0.37	−0.21	0.12
$k=20$	0.27	−0.13	0.09
$k=30$	0.22	−0.10	0.08
$k=50$	0.17	−0.08	0.06

Panel C: $t(k)$ distributions

	Quantile Skewness (S)	Left Tail Weight (LTW)	Right Tail Weight (RTW)
$k=1$	0	0.46	0.46
$k=2$	0	0.36	0.36
$k=3$	0	0.28	0.28
$k=4$	0	0.22	0.22
$k=5$	0	0.18	0.18
$k=10$	0	0.09	0.09
$k=50$	0	0.02	0.02

Table C.9: Measures of Distance Between the Standard Normal and Other Distributions

Distance measures between the standard normal distribution and various normal, χ^2, and t distributions. The Hellinger distance between densities $f_X,\ f_Y$ is $H^2 := 1 - \int \sqrt{f_X(s)\, f_Y(s)} ds$, and takes values in $[0, 1]$. The Wasserstein distance between densities $f_X,\ f_Y$ is $W := \inf_{f_{XY}} E[\|X - Y\|]$ where f_{XY} is any joint density with marginals $f_X,\ f_Y$, and takes values in $[0, +\infty)$. For the Wasserstein distance, we present values that rely on 99% of the range of the distribution, i.e., we exclude the extreme tails to make the distance measure robust. In Panel A, we present the distances between $\mathcal{N}(0, 1)$ and normal distributions with the same mean but different standard deviation, as indicated in each row of the panel. In Panel B, we present the distances between $\mathcal{N}(0, 1)$ and χ^2 distributions with various degrees of freedom as indicated in each row of the panel; these distributions are scaled to have the same mean (0) and variance (1) as the standard normal. In Panel C, we present the distances between $\mathcal{N}(0, 1)$ and t distributions with various degrees of freedom as indicated in each row of the panel.

Panel A: Distance between $\mathcal{N}(0, 1)$ and $\sigma \cdot \mathcal{N}(0, 1)$ distributions

	Hellinger Distance (H^2)	Wasserstein Distance (W)
$\sigma = 0.25$	0.31	0.58
$\sigma = 0.50$	0.11	0.39
$\sigma = 0.75$	0.02	0.19
$\sigma = 1.25$	0.02	0.19
$\sigma = 1.50$	0.04	0.39
$\sigma = 1.75$	0.07	0.58
$\sigma = 2.00$	0.11	0.78

Panel B: Distance between $\mathcal{N}(0, 1)$ and $[\chi^2(k) - k]/\sqrt{2k}$ distributions

	Hellinger Distance (H^2)	Wasserstein Distance (W)
$k = 3$	0.08	0.24
$k = 4$	0.06	0.21
$k = 5$	0.05	0.19
$k = 10$	0.02	0.13
$k = 20$	0.01	0.09
$k = 30$	0.01	0.08
$k = 50$	0.00	0.06

Panel C: Distance between $\mathcal{N}(0, 1)$ and $t(k)$ distributions

	Hellinger Distance (H^2)	Wasserstein Distance (W)
$k = 1$	0.11	1.89
$k = 2$	0.05	0.45
$k = 3$	0.03	0.25
$k = 4$	0.02	0.17
$k = 5$	0.01	0.13
$k = 10$	0.00	0.06
$k = 50$	0.00	0.01

C.7 Portfolio performance

In this section, we present additional results regarding the out-of-sample performance of portfolios that select top-performing funds using i) the FDR methodology, ii) a hierarchical model in which fund alphas are drawn from one normal, iii) a hierarchical model in which fund alphas are drawn from two normals, and iv) our estimation methodology.

Our baseline portfolio formation rule described in detail in Section 6.1 of the paper is the following: At the beginning of each month in the period 1980–2011, we use the preceding 60 months of fund returns to estimate the 4-factor model using each methodology, and we form and hold until the end of the month a portfolio of funds with high estimated probability of having a positive alpha; if all funds have a low probability of having a positive alpha, we select funds whose posterior mean alpha (for the Bayesian methodologies) or OLS t-statistic (for the FDR methodology) is in the top 1% among all funds in the data set for the preceding 5 years.

In Table C.10, we present results on portfolio performance under alternative portfolio formation rules: portfolios formed using a 36-month (instead of a 60-month) rolling estimation window, portfolios that are left empty and portfolios that keep the top 2% (instead of the top 1%) of funds in months in which all funds have a low probability of having a positive alpha, and portfolios that always keep the top 1% of funds sorted by their posterior mean alpha. In particular, for each portfolio we construct, we use its monthly portfolio returns for the period 1980–2011 to estimate its annualized OLS 4-factor alpha, $\hat{\alpha}$, and the associated $\hat{\alpha}$ t-statistic and residual standard deviation, its information ratio, the mean and standard deviation of its return in excess of the risk-free return, and its Sharpe ratio. We see that, as with our baseline portfolio formation rule used in Section 6.1 in the paper, for portfolios constructed using these alternative formation rules, those based on our methodology yield higher performance than those based on the other methodologies. This is true not only in terms of estimated alpha, but also in terms of the information ratio and even in terms of the Sharpe ratio. The exception to this finding is that the conservative portfolio based on our methodology yields a lower Sharpe ratio than those based on the other methodologies (see Panel B of the table). However, we note that, by construction, the conservative portfolios are not active in all months, and indeed conservative portfolios based on different methodologies are active in different months, therefore their performance is not directly comparable.

Furthermore, our methodology estimates *alpha*, therefore it is not surprising that its advantage for the Sharpe ratio is smaller than it is for the estimated alpha or for the information ratio.

In Table C.11, we present results on portfolio performance — using the measures described above — for each of the two halves of our sample period (1980–1995 and 1995–2011). We find that our portfolio exhibits superior performance in both subperiods.

In Table C.12, we present results on the performance of quantile-based portfolios. That is, at the beginning of each month in the period 1980–2011, we use the preceding 60 months of fund returns to estimate the 4-factor model using each methodology — the hierarchical model in which fund alphas are drawn from one normal or from two normals, and our methodology — then we sort funds into ten quantiles based on the posterior mean alpha, and we hold these quantile-based portfolios until the end of the month. As before, for each portfolio we construct, we use its monthly portfolio returns for the period 1980–2011 to estimate its annualized OLS 4-factor alpha, $\hat{\alpha}$, and the associated $\hat{\alpha}$ t-statistic and residual standard deviation, its information ratio, the mean and standard deviation of its return in excess of the risk-free return, and its Sharpe ratio. We see that the slope of returns going from the bottom quantile to the top quantile is steeper for the portfolios constructed using our methodology than for those constructed using the alternatives. In particular, as we see in the column labeled 'Q10−Q1', the portfolio that buys the funds in the top quantile and sells the funds in the bottom quantile has $\hat{\alpha} = 3.23\%$ per year for the hierarchical model with one normal, $\hat{\alpha} = 3.61\%$ for the hierarchical model with two normals, and $\hat{\alpha} = 4.29\%$ for our methodology. The difference in $\hat{\alpha}$ between the portfolio constructed using our methodology and the one constructed using the hierarchical model with one normal (two normals) is 1.06% (0.68%) and is statistically significant at the 1% level, with a t-statistic of 4.54 (3.59). These results show that our methodology can better identify funds at the tails (both the right and the left tail) of the skill distribution, which is consistent with our theoretical argument that our more flexible semi-parametric model can better capture the tails of the distribution.

Table C.10: Out-of-sample Portfolio Performance — Alternative Portfolio Construction Rules

Out-of-sample performance measures for portfolios that use alternative portfolio construction rules to select funds using the FDR methodology, hierarchical models in which fund αs are drawn from one normal or from two normals, and our estimation methodology. At the beginning of each month in the period 1980–2011, we use the preceding 36 or 60 months of fund returns to estimate the 4-factor model using each methodology, and we form and hold until the end of the month equal-weighted portfolios of funds that are estimated to have high performance. In Panels A and B, we construct portfolios using 60-month rolling estimation windows, and we select funds with high estimated probability of having a positive α. During months in which all funds have a low probability of having a positive α, in Panel A we select funds whose posterior mean α or OLS α t-statistic is in the top 2% among all funds in the ranking period (the 'aggressive' portfolio), while in Panel B, we leave the portfolio empty (the 'conservative' portfolio). In Panel C, we present results on the aggressive portfolio constructed using 36-month rolling estimation periods and keeping the top 1% instead of the top 2% during months in which all funds have a low probability of having a positive α. In Panel D, we use a 60-month rolling estimation window, and in all months we select funds whose posterior mean α is in the top 1% among all funds in the ranking period. For each portfolio we construct, we use its monthly returns from 1980 through 2011 to estimate its annualized OLS 4-factor alpha $\hat{\alpha}$ and residual standard deviation $\hat{\sigma}_\varepsilon$ (both expressed as percents), $\hat{\alpha}$ t-statistic, Information Ratio ($\hat{\alpha}/\hat{\sigma}_\varepsilon$), mean and standard deviation (both expressed as percents) of its return in excess of the risk-free return, and its Sharpe Ratio (mean/std. dev. of excess return).

Panel A: Aggressive Portfolio with top 2%

	FDR	**1 Normal**	**2 Normals**	**Our Model**
$\hat{\alpha}$	1.63	1.62	1.35	2.38
$\hat{\alpha}$ t-statistic	2.13	1.87	1.55	3.27
$\hat{\sigma}_\varepsilon$	4.14	4.42	4.60	3.93
Information Ratio	0.39	0.36	0.29	0.61
Mean Return	7.27	7.42	7.25	8.36
Std. dev. Return	16.69	13.64	13.81	15.00
Sharpe Ratio	0.44	0.54	0.53	0.56

Panel B: Conservative Portfolio

	FDR	**1 Normal**	**2 Normals**	**Our Model**
$\hat{\alpha}$	1.86	2.01	1.68	2.84
$\hat{\alpha}$ t-statistic	2.22	2.10	1.77	3.14
$\hat{\sigma}_\varepsilon$	3.27	4.58	4.75	3.80
Information Ratio	0.57	0.44	0.35	0.75
Mean Return	9.75	7.79	7.83	7.65
Std. dev. Return	15.80	13.95	14.19	16.30
Sharpe Ratio	0.62	0.56	0.55	0.47

Panel C: Aggressive Portfolio with 36-month window

	FDR	**1 Normal**	**2 Normals**	**Our Model**
$\hat{\alpha}$	1.24	1.36	1.86	2.10
$\hat{\alpha}$ t-statistic	1.58	1.68	1.89	2.02
$\hat{\sigma}_\varepsilon$	4.22	4.34	6.19	6.41
Information Ratio	0.29	0.31	0.30	0.33
Mean Return	7.49	7.06	8.17	9.07
Std. dev. Return	17.05	14.57	16.46	17.34
Sharpe Ratio	0.44	0.48	0.50	0.52

Panel D: Alpha-sorted Portfolio

	1 Normal	**2 Normals**	**Our Model**
$\hat{\alpha}$	1.45	1.69	2.35
$\hat{\alpha}$ t-statistic	2.00	2.19	2.87
$\hat{\sigma}_\varepsilon$	3.85	4.18	4.59
Information Ratio	0.38	0.40	0.51
Mean Return	7.19	7.62	8.55
Std. dev. Return	14.61	14.93	15.78
Sharpe Ratio	0.49	0.51	0.54

Table C.11: Out-of-sample Portfolio Performance — Sub-samples

Out-of-sample performance measures for two non-overlapping sub-samples, for portfolios that select funds using the FDR methodology, hierarchical models in which fund αs are drawn from one normal or from two normals, and our estimation methodology. At the beginning of each month in the period 1980–2011, we use the preceding 60 months of fund returns to estimate the 4-factor model using each methodology, and we form and hold until the end of the month equal-weighted portfolios of funds that are estimated to have high probability of having a positive α (see Section 6.1 of the paper for more details). During months in which all funds have a low probability of having a positive α, we select funds whose posterior mean α (for the three hierarchical methodologies) or OLS t-statistic (for the FDR methodology) is in the top 1% among all funds in the data set for the preceding 60 months. For each portfolio we construct, we use its monthly portfolio returns from 1980 to 1995 (in Panel A) and from 1995 to 2011 (in Panel B) to estimate its annualized OLS 4-factor alpha $\hat{\alpha}$ and residual standard deviation $\hat{\sigma}_{\varepsilon}$ (both expressed as percents), $\hat{\alpha}$ t-statistic, Information Ratio ($\hat{\alpha}/\hat{\sigma}_{\varepsilon}$), mean and standard deviation (both expressed as percents) of its return in excess of the risk-free return, and its Sharpe Ratio (mean/std. dev. of excess return).

Panel A: 1^{st}–half Sub-sample

	FDR	**1 Normal**	**2 Normals**	**Our Model**
$\hat{\alpha}$	2.45	2.22	1.80	3.36
$\hat{\alpha}$ t-statistic	2.73	2.74	2.22	3.51
$\hat{\sigma}_{\varepsilon}$	3.05	3.03	3.07	3.73
Information Ratio	0.80	0.73	0.59	0.90
Mean Return	8.52	8.49	8.07	10.59
Std. dev. Return	15.90	14.83	14.91	15.47
Sharpe Ratio	0.54	0.57	0.54	0.68

Panel B: 2^{nd}–half Sub-sample

	FDR	**1 Normal**	**2 Normals**	**Our Model**
$\hat{\alpha}$	1.54	1.92	1.62	2.24
$\hat{\alpha}$ t-statistic	1.27	1.70	1.33	2.28
$\hat{\sigma}_{\varepsilon}$	4.99	4.20	4.60	3.74
Information Ratio	0.31	0.46	0.35	0.60
Mean Return	7.04	6.51	6.47	7.21
Std. dev. Return	17.48	12.41	12.82	14.46
Sharpe Ratio	0.40	0.52	0.50	0.50

Table C.12: Out-of-sample Portfolio Performance — Quantile-based Portfolios

Out-of-sample performance measures for portfolios that select funds using hierarchical models in which fund αs are drawn from one normal (in Panel A) or from two normals (in Panel B), and our estimation methodology (in Panel C). At the beginning of each month in the period 1980–2011, we use the preceding 60 months of fund returns to estimate the 4-factor model using each methodology, we sort funds into ten quantiles (Q1 through Q10) based on the posterior mean α, and we hold these quantile-based portfolios until the end of the month. We also form the portfolio (labeled 'Q10−Q1') which buys the funds belonging to the top quantile and sells the funds belonging to the bottom quantile. For each portfolio, we use its monthly returns from 1980 through 2011 to estimate its annualized OLS 4-factor alpha $\hat{\alpha}$ and residual standard deviation $\hat{\sigma}_{\varepsilon}$ (both expressed as percents), $\hat{\alpha}$ t-statistic, Information Ratio ($\hat{\alpha}/\hat{\sigma}_{\varepsilon}$), mean and standard deviation (both expressed as percents) of its return in excess of the risk-free return, and its Sharpe Ratio (mean/std. dev. of excess return).

Panel A: 1 Normal

	Quantiles										
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q10−Q1**
$\hat{\alpha}$	−2.86	−2.14	−0.94	−0.86	−1.18	−1.19	−0.94	0.26	0.30	0.37	3.23
$\hat{\alpha}$ t-statistic	−6.02	−5.47	−2.02	−1.57	−1.87	−1.86	−1.86	0.52	0.60	0.76	5.94
$\hat{\sigma}_{\varepsilon}$	2.48	2.20	2.45	2.94	3.30	3.17	2.61	2.50	2.62	2.69	2.84
Information Ratio	−1.15	−0.97	−0.38	−0.29	−0.36	−0.38	−0.36	0.10	0.11	0.14	1.14
Mean Return	3.33	4.18	5.28	5.68	5.25	5.33	5.64	6.60	6.46	6.42	3.10
Std dev Return	14.96	15.32	15.50	15.71	15.69	15.77	16.19	16.24	16.26	15.29	2.91
Sharpe Ratio	0.22	0.27	0.34	0.36	0.33	0.34	0.35	0.41	0.40	0.42	1.06

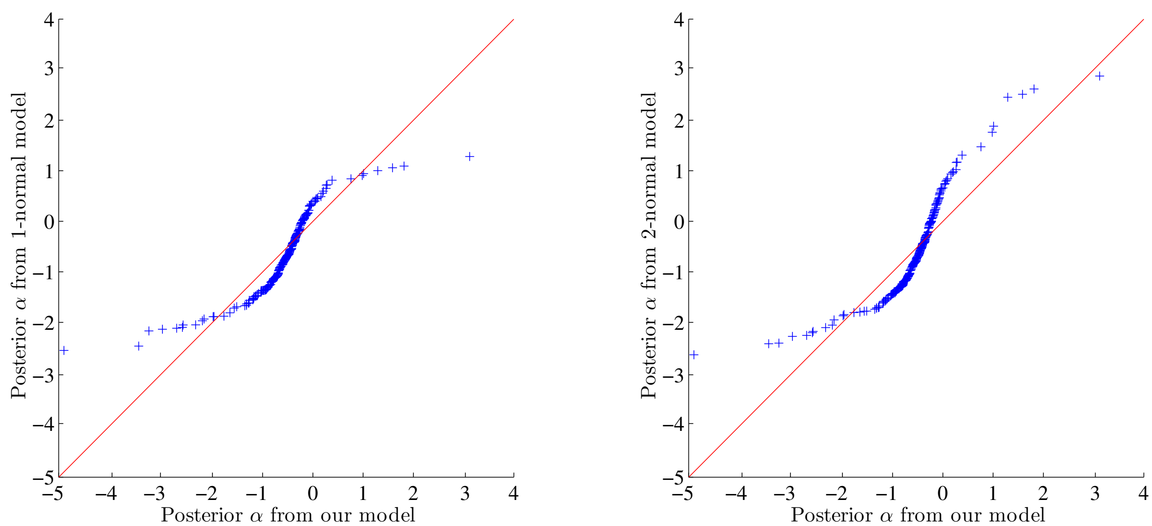
Panel B: 2 Normals

	Quantiles										
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q10−Q1**
$\hat{\alpha}$	−3.09	−1.84	−1.02	−1.19	−1.00	−1.23	−0.49	−0.09	0.24	0.52	3.61
$\hat{\alpha}$ t-statistic	−6.35	−4.63	−2.24	−2.29	−1.64	−2.09	−0.96	−0.17	0.49	1.02	6.42
$\hat{\sigma}_{\varepsilon}$	2.56	2.22	2.39	2.80	3.15	2.93	2.73	2.49	2.61	2.81	3.01
Information Ratio	−1.21	−0.83	−0.43	−0.43	−0.32	−0.42	−0.18	−0.03	0.09	0.19	1.20
Mean Return	3.14	4.45	5.22	5.29	5.47	5.26	6.07	6.27	6.42	6.56	3.42
Std dev Return	15.10	15.31	15.40	15.60	15.59	15.85	16.16	16.24	16.24	15.37	3.10
Sharpe Ratio	0.21	0.29	0.34	0.34	0.35	0.33	0.38	0.39	0.40	0.43	1.10

Panel C: Our Model

	Quantiles										
	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q10−Q1**
$\hat{\alpha}$	−3.48	−1.62	−0.65	−1.50	−0.85	−1.06	−0.78	−0.06	−0.04	0.80	4.29
$\hat{\alpha}$ t-statistic	−7.17	−3.90	−1.22	−2.85	−1.63	−1.79	−1.66	−0.11	−0.07	1.48	6.89
$\hat{\sigma}_{\varepsilon}$	2.63	2.23	2.82	2.79	2.68	2.84	2.46	2.63	2.68	3.05	3.41
Information Ratio	−1.32	−0.72	−0.23	−0.54	−0.32	−0.37	−0.32	−0.02	−0.01	0.26	1.26
Mean Return	2.83	4.68	5.69	5.04	5.48	5.37	5.83	6.25	6.19	6.77	3.96
Std dev Return	15.36	15.42	15.50	15.53	15.43	15.61	16.01	16.05	16.20	15.73	3.53
Sharpe Ratio	0.18	0.30	0.37	0.32	0.36	0.34	0.36	0.39	0.38	0.43	1.12

Observing Tables C.10 and C.11 above, it is interesting that the performance of the portfolio constructed using the hierarchical model with two normals is in some cases worse than that of the portfolio constructed using the hierarchical model with one normal. This may be due to noise, since funds are allocated to portfolios using only a few years of data, but it could also be explained as follows. The hierarchical model with two normals attempts to estimate the fat tails of the alpha distribution and is therefore more aggressive about placing funds in the right tail, but due to its limited flexibility it may be unable to do so accurately. This intuition is consistent with the evidence in Figure C.5, which presents representative Quantile-Quantile plots of the posterior alphas from the hierarchical model with one normal (in Panel *a*) and with two normals (in Panel *b*) versus the posterior alphas from our model, for one of the estimation periods used in the construction of the portfolios. We see that, while the model with two normals does a better job than the model with one normal in estimating the largest alpha at more than 3% annualized, it also overestimates (relative to our model) a number of large alphas, and therefore would over-aggressively include them in the portfolio of the best-performing funds.



(a) 1-normal model vs. our model.

(b) 2-normal model vs. our model

Figure C.5: Representative Quantile-Quantile plots of posterior mean alphas estimated from the hierarchical model with one normal (in Panel *a*) and with two normals (in Panel *b*) versus posterior mean alphas estimated using our methodology, for one of the 60-month periods used in the rolling estimation employed to construct the portfolios in Section 6.1 of the paper. The blue cross marks plot the quantiles, and the solid red line plots the 45° line.

C.8 Posterior predictive densities

In this section, we explain in detail how we make draws from the posterior predictive densities of benchmark portfolio returns and of fund returns. These draws are necessary for our calculation of optimal portfolios in Section 6.2 of the paper. The posterior predictive density $p\left(r_{i,T+1}, F_{T+1}\middle| r, F\right)$ is $\int p\left(r_{i,T+1}, F_{T+1}, \chi_i, \chi_F\middle| r, F\right) d\chi_i \chi_F$, where χ_i and χ_F denote the parameters of the distribution of r_i and F, respectively, and where r and F collect all fund and benchmark portfolio returns, respectively. Using simple rules of probability, we can rewrite $p\left(r_{i,T+1}, F_{T+1}, \chi_i, \chi_F\middle| r, F\right)$ as $p\left(r_{i,T+1}, F_{T+1}\middle| \chi_i, \chi_F\right)$ times $p\left(\chi_i, \chi_F| r, F\right)$ and $p\left(r_{i,T+1}, F_{T+1}\middle| \chi_i, \chi_F\right)$ as $p\left(r_{i,T+1} \left|F_{T+1}, \chi_i\right.\right) p\left(F_{T+1}| \chi_F\right)$, while $p\left(\chi_i, \chi_F| r, F\right)$ is proportional to $p\left(\chi_i \left|r, F\right.\right) p\left(\chi_F \left|F\right.\right)$. That is,

$$p\left(r_{i,T+1}, F_{T+1}\middle| r, F\right) \propto \int p\left(r_{i,T+1} \left|F_{T+1}, \chi_i\right.\right) p\left(F_{T+1}\middle| \chi_F\right) p\left(\chi_i| r, F\right) p\left(\chi_F\middle| F\right) d\chi_i \chi_F.$$

Thus, to make draws from the posterior predictive density $p\left(r_{i,T+1}, F_{T+1}\middle| r, F\right)$, we make draws from $p(\chi_F \left|F\right.)$ and $p(\chi_i \left|r, F\right.)$, which we use to make draws from $p\left(F_{T+1}| \chi_F\right)$ and subsequently from $p\left(r_{i,T+1} \left|F_{T+1}, \chi_i\right.\right)$. To make draws from $p\left(\chi_i \left|r, F\right.\right)$, we work as in Section 2.4 of the paper, and to make draws from $p\left(r_{i,T+1} \left|F_{T+1}, \chi_i\right.\right)$ we use the linear factor model in Equation 1 of the paper. Below, we describe how we make draws from $p\left(\chi_F \left|F\right.\right)$ and from $p\left(F_{T+1}| \chi_F\right)$.

For the factor returns, we assume that they are i.i.d. normal, that is $F_t \left|\mu_F, \Sigma_F\right. \sim \mathcal{N}\left(\mu_F, \Sigma_F\right)$, and that the distribution parameters $\left(\mu_F, \Sigma_F\right)$ follow the conjugate Normal-inverse-Wishart prior given by $\mu_F, \Sigma_F \left|\underline{\mu}_F, \underline{\kappa}_F, \underline{\nu}_F, \underline{\Lambda}_F\right. \sim NIW\left(\underline{\mu}_F, \underline{\kappa}_F, \underline{\nu}_F, \underline{\Lambda}_F\right)$, i.e.,

$$\begin{aligned} \mu_F \left|\underline{\mu}_F, \underline{\kappa}_F, \Sigma_F\right. &\sim \mathcal{N}\left(\underline{\mu}_F, \frac{1}{\underline{\kappa}_F} \Sigma_F\right) \\ \Sigma_F^{-1} \left|\underline{\nu}_F, \underline{\Lambda}_F\right. &\sim \mathcal{W}\left(\underline{\nu}_F, \underline{\Lambda}_F^{-1}\right), \end{aligned}$$

where $\underline{\mu}_F$, $\underline{\kappa}_F$, $\underline{\nu}_F$, and $\underline{\Lambda}_F$ are prior parameters. In particular, using the Jeffrey's prior $p\left(\mu_F, \Sigma_F\right) \propto |\Sigma_F|^{-\frac{k_F+1}{2}}$ (with k_F the number of factors) and observing data $F := \{F_t\}_{t=1}^T$, the posterior of $\left(\mu_F, \Sigma_F\right)$ is Normal-inverse-Wishart $\mu_F, \Sigma_F \left|F\right. \sim NIW\left(\hat{\mu}_F, T, T-1, T\hat{\Sigma}_F\right)$, i.e.,

$$\begin{aligned} \mu_F \left|\Sigma_F, F\right. &\sim \mathcal{N}\left(\hat{\mu}_F, \frac{1}{T} \Sigma_F\right) \\ \Sigma_F^{-1} \left|F\right. &\sim \mathcal{W}\left(T-1, \left(T\hat{\Sigma}_F\right)^{-1}\right), \end{aligned}$$

where

$$
\begin{aligned}
\hat{\mu}_F &:= \frac{1}{T}\sum_{t=1}^{T} F_t \\
\hat{\Sigma}_F &:= \frac{1}{T}\sum_{t=1}^{T} \left(F_t - \hat{\mu}_F\right)\left(F_t - \hat{\mu}_F\right)'.
\end{aligned}
$$

Thus, to generate $m = 1, \ldots, M$ draws for the benchmark portfolio returns from the posterior predictive density, we generate draw $\Sigma_F^{(m)}$ from $\Sigma_F^{-1} \,|F$ above, we use this to generate draw $\mu_F^{(m)}$ from $\mu_F \,|\Sigma_F, F$ above, and we use both to generate draw $F_{T+1}^{(m)}$ from $\mathcal{N}\left(\mu_F^{(m)}, \Sigma_F^{(m)}\right)$.

To generate $m = 1, \ldots, M$ draws for fund i's returns from the posterior predictive density, first we randomly pick $m = 1, \ldots, M$ draws from our MCMC draws for α_i, β_i, and h_i, whose joint distribution converges to their joint posterior distribution. Then, we generate $m = 1, \ldots, M$ draws $\varepsilon_{i,T+1}^{(m)} \sim \mathcal{N}\left(0, \left(h_i^{(m)}\right)^{-1}\right)$. Finally, we combine draws $\alpha_i^{(m)}$, $\beta_i^{(m)}$, $\varepsilon_{i,T+1}^{(m)}$ with draw $F_{T+1}^{(m)}$ for the benchmark portfolio returns whose generation is described above, and substitute them in the linear factor model equation to calculate the draw $r_{i,T+1}^{(m)} = \alpha_i^{(m)} + \left(F_{T+1}^{(m)}\right)' \beta_i^{(m)} + \varepsilon_{i,T+1}^{(m)}$.

C.9 The distribution of skill by fund investment objective

In this section, we present additional results regarding the estimation of the $K^- = 2$, $K^+ = 1$ model separately for funds classified to each of the three investment objectives (Growth & Income, Growth, and Aggressive Growth); see Section 7.1 of the paper for details. In particular, in Table C.13, we present the percentiles of the estimated distributions of alpha and the factor loadings.

Table C.13: Percentiles of Estimated Distributions — By Investment Objective

Percentiles of the estimated population distributions of annualized alpha (expressed as a percent) and factor loadings, estimated with returns net of expenses using the $K^- = 2$, $K^+ = 1$ model separately for funds classified to the three investment objective categories: Growth & Income (Panel A), Growth (Panel B), and Aggressive Growth (Panel C).

Panel A: Growth & Income Objective

	Percentiles														
	0.5^{th}	**1^{st}**	**5^{th}**	**10^{th}**	**20^{th}**	**30^{th}**	**40^{th}**	**50^{th}**	**60^{th}**	**70^{th}**	**80^{th}**	**90^{th}**	**95^{th}**	**99^{th}**	**99.5^{th}**
α	-3.75	-3.15	-2.04	-1.66	-1.34	-1.14	-0.96	-0.80	-0.63	-0.40	0.00	0.00	0.47	1.32	1.71
$\beta_{\mathbf{M}}$	0.29	0.34	0.47	0.55	0.63	0.70	0.75	0.80	0.85	0.90	0.97	1.06	1.13	1.27	1.31
$\beta_{\mathbf{SMB}}$	-0.46	-0.42	-0.30	-0.24	-0.17	-0.12	-0.07	-0.03	0.01	0.06	0.11	0.18	0.24	0.36	0.40
$\beta_{\mathbf{HML}}$	-0.34	-0.29	-0.16	-0.09	-0.00	0.06	0.11	0.16	0.21	0.26	0.32	0.40	0.47	0.60	0.65
$\beta_{\mathbf{UMD}}$	-0.25	-0.23	-0.17	-0.14	-0.10	-0.07	-0.04	-0.02	0.00	0.02	0.05	0.09	0.12	0.18	0.21

Panel B: Growth Objective

	Percentiles														
	0.5^{th}	**1^{st}**	**5^{th}**	**10^{th}**	**20^{th}**	**30^{th}**	**40^{th}**	**50^{th}**	**60^{th}**	**70^{th}**	**80^{th}**	**90^{th}**	**95^{th}**	**99^{th}**	**99.5^{th}**
α	-4.90	-4.10	-2.56	-1.99	-1.45	-1.12	-0.87	-0.64	-0.40	0.00	0.00	0.10	0.48	2.28	3.62
$\beta_{\mathbf{M}}$	0.53	0.58	0.69	0.75	0.83	0.88	0.93	0.97	1.02	1.06	1.11	1.19	1.25	1.36	1.40
$\beta_{\mathbf{SMB}}$	-0.58	-0.50	-0.29	-0.18	-0.05	0.04	0.13	0.20	0.28	0.36	0.45	0.59	0.69	0.90	0.98
$\beta_{\mathbf{HML}}$	-0.90	-0.81	-0.56	-0.43	-0.27	-0.15	-0.06	0.03	0.12	0.22	0.34	0.50	0.63	0.87	0.97
$\beta_{\mathbf{UMD}}$	-0.28	-0.25	-0.18	-0.14	-0.08	-0.05	-0.02	0.01	0.04	0.07	0.11	0.16	0.20	0.28	0.31

Panel C: Aggressive Growth Objective

	Percentiles														
	0.5^{th}	**1^{st}**	**5^{th}**	**10^{th}**	**20^{th}**	**30^{th}**	**40^{th}**	**50^{th}**	**60^{th}**	**70^{th}**	**80^{th}**	**90^{th}**	**95^{th}**	**99^{th}**	**99.5^{th}**
α	-17.81	-12.26	-4.35	-2.36	-0.96	-0.36	0.00	0.00	0.00	0.00	0.30	1.12	1.99	4.63	6.25
$\beta_{\mathbf{M}}$	0.65	0.69	0.80	0.86	0.93	0.98	1.03	1.07	1.11	1.15	1.20	1.27	1.33	1.44	1.48
$\beta_{\mathbf{SMB}}$	-0.28	-0.21	-0.02	0.08	0.20	0.29	0.37	0.43	0.50	0.58	0.67	0.79	0.89	1.07	1.14
$\beta_{\mathbf{HML}}$	-1.12	-1.03	-0.79	-0.66	-0.50	-0.39	-0.29	-0.20	-0.11	-0.01	0.10	0.26	0.39	0.64	0.72
$\beta_{\mathbf{UMD}}$	-0.30	-0.27	-0.17	-0.12	-0.06	-0.01	0.03	0.06	0.10	0.14	0.19	0.25	0.30	0.40	0.43

C.10 The prevalence of short-term skill and its evolution

Here, we present additional tables pertaining to our analysis of short-term skill and its evolution over time (see Section 7.2 of the paper) using returns net of expenses for 3,497 funds.

Table C.14 presents the evolution over time of the posterior means of the population proportions of zero-, negative-, and positive-alpha funds, while Table C.15 presents the evolution over time of the percentiles of the estimated distribution of alpha. Tables C.14 and C.15 correspond to Figures 9a and 9b of the paper, respectively.

Table C.14: Proportions of Fund Types — Short-term Skill

Evolution over time of posterior means of population proportions of zero-, negative-, and positive-alpha funds, in a model with short-term skill. Posterior means are estimated at the end of each year using data from the preceding 60 months. All estimations use returns net of expenses in the $K^{-} = 2$, $K^{+} = 1$ model with two and one components for the alpha distribution of negative-alpha and positive-alpha funds, respectively.

	π^0	π^-	π^+
1975 – 1979	0.29	0.58	0.13
1976 – 1980	0.40	0.29	0.32
1977 – 1981	0.29	0.42	0.29
1978 – 1982	0.37	0.22	0.41
1979 – 1983	0.49	0.15	0.36
1980 – 1984	0.33	0.19	0.48
1981 – 1985	0.40	0.30	0.30
1982 – 1986	0.34	0.17	0.48
1983 – 1987	0.46	0.18	0.37
1984 – 1988	0.43	0.26	0.31
1985 – 1989	0.39	0.34	0.27
1986 – 1990	0.54	0.14	0.32
1987 – 1991	0.26	0.55	0.19
1988 – 1992	0.23	0.58	0.19
1989 – 1993	0.49	0.32	0.18
1990 – 1994	0.42	0.36	0.23
1991 – 1995	0.20	0.73	0.07
1992 – 1996	0.22	0.70	0.08
1993 – 1997	0.24	0.65	0.11
1994 – 1998	0.06	0.81	0.12
1995 – 1999	0.05	0.91	0.04
1996 – 2000	0.43	0.30	0.27
1997 – 2001	0.40	0.39	0.21
1998 – 2002	0.49	0.37	0.14
1999 – 2003	0.50	0.33	0.17
2000 – 2004	0.35	0.55	0.09
2001 – 2005	0.09	0.88	0.03
2002 – 2006	0.04	0.95	0.01
2003 – 2007	0.17	0.73	0.10
2004 – 2008	0.17	0.71	0.12
2005 – 2009	0.22	0.57	0.21
2006 – 2010	0.17	0.72	0.11
2007 – 2011	0.22	0.62	0.16

Table C.15: Percentiles of Estimated Distribution of Alpha — Short-term Skill

Evolution over time of various percentiles of the estimated distribution of annualized alpha (expressed as a percent), in a model with short-term skill. The distributions are estimated at the end of each year using data from the preceding 60 months. All estimations use returns net of expenses in the $K^- = 2$, $K^+ = 1$ model with two components for the alpha distribution of negative-alpha funds and one component for that of positive-alpha funds.

	Percentiles										
	5th	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**
1975 – 1979	-2.59	-2.07	-1.55	-1.22	-0.96	-0.70	0.00	0.00	0.00	0.92	1.81
1976 – 1980	-1.71	-1.17	-0.63	0.00	0.00	0.00	0.00	0.34	0.74	1.21	1.66
1977 – 1981	-2.07	-1.64	-1.18	-0.87	-0.49	0.00	0.00	0.00	1.40	2.29	3.07
1978 – 1982	-3.03	-1.86	-0.59	0.00	0.00	0.00	0.42	1.05	1.65	2.65	3.70
1979 – 1983	-3.17	-1.49	0.00	0.00	0.00	0.00	0.00	0.64	1.30	2.45	3.78
1980 – 1984	-3.97	-2.34	0.00	0.00	0.00	0.00	0.74	1.20	1.79	2.81	3.89
1981 – 1985	-2.50	-1.34	-0.49	0.00	0.00	0.00	0.00	0.00	1.05	2.08	3.17
1982 – 1986	-4.58	-2.14	0.00	0.00	0.00	0.00	0.96	1.37	1.83	2.54	3.25
1983 – 1987	-4.64	-2.43	0.00	0.00	0.00	0.00	0.00	1.20	1.83	2.68	3.47
1984 – 1988	-4.04	-2.12	-0.67	0.00	0.00	0.00	0.00	0.49	1.35	2.42	3.51
1985 – 1989	-3.67	-2.10	-0.91	-0.32	0.00	0.00	0.00	0.00	1.24	2.35	3.39
1986 – 1990	-3.27	-1.24	0.00	0.00	0.00	0.00	0.00	0.42	1.08	2.03	3.05
1987 – 1991	-2.84	-1.86	-1.08	-0.68	-0.42	-0.20	0.00	0.00	0.00	1.73	2.76
1988 – 1992	-2.92	-1.84	-1.01	-0.62	-0.38	-0.20	0.00	0.00	0.00	1.45	2.50
1989 – 1993	-2.54	-1.47	-0.63	-0.18	0.00	0.00	0.00	0.00	0.00	1.36	2.44
1990 – 1994	-3.40	-2.02	-0.94	-0.39	0.00	0.00	0.00	0.00	0.45	1.42	2.40
1991 – 1995	-3.75	-2.69	-1.76	-1.28	-0.95	-0.70	-0.48	-0.25	0.00	0.00	0.33
1992 – 1996	-3.93	-2.80	-1.84	-1.33	-0.98	-0.71	-0.46	0.00	0.00	0.00	0.31
1993 – 1997	-4.39	-3.31	-2.31	-1.75	-1.33	-0.98	-0.61	0.00	0.00	0.20	0.53
1994 – 1998	-4.29	-3.51	-2.73	-2.26	-1.91	-1.62	-1.35	-1.06	-0.61	0.00	0.00
1995 – 1999	-3.47	-3.05	-2.60	-2.32	-2.10	-1.91	-1.73	-1.54	-1.32	-0.88	0.00
1996 – 2000	-2.85	-2.04	-1.20	0.00	0.00	0.00	0.00	0.00	0.68	1.20	1.67
1997 – 2001	-3.74	-2.92	-2.05	-1.43	0.00	0.00	0.00	0.00	0.15	0.48	0.76
1998 – 2002	-3.16	-2.26	-1.37	-0.81	0.00	0.00	0.00	0.00	0.00	0.78	2.20
1999 – 2003	-2.36	-1.75	-1.10	-0.59	0.00	0.00	0.00	0.00	0.00	1.56	2.44
2000 – 2004	-3.19	-2.52	-1.85	-1.44	-1.10	-0.74	0.00	0.00	0.00	0.00	1.88
2001 – 2005	-5.01	-4.01	-3.04	-2.48	-2.08	-1.75	-1.46	-1.19	-0.88	0.00	0.00
2002 – 2006	-4.45	-3.59	-2.77	-2.30	-1.95	-1.68	-1.43	-1.21	-0.97	-0.68	0.00
2003 – 2007	-2.86	-2.38	-1.90	-1.59	-1.36	-1.15	-0.94	-0.65	0.00	0.00	2.25
2004 – 2008	-2.01	-1.74	-1.44	-1.25	-1.10	-0.95	-0.80	-0.54	0.00	1.13	1.85
2005 – 2009	-2.59	-2.03	-1.48	-1.14	-0.88	-0.61	0.00	0.00	0.29	0.90	1.31
2006 – 2010	-2.64	-2.22	-1.80	-1.52	-1.31	-1.11	-0.91	-0.60	0.00	0.30	0.64
2007 – 2011	-2.62	-2.21	-1.77	-1.49	-1.25	-1.02	-0.68	0.00	0.00	0.12	0.20

C.11 Fund flow analysis

In this section, we present additional results on our analysis in Section 7.4 of the paper, for the relation between fund flows and past fund performance as well as the relation between fund flows and subsequent fund performance.

In Table C.16 (C.17), we present the average past (future) performance across all funds that belong to each flow quintile for each 5-year non-overlapping period in our sample. These tables are similar, respectively, to Panels A and B of Table 16, which presents results averaged across all 5-year non-overlapping periods in our sample. They show the same effects as those shown in Table 16 and discussed in detail in Section 7.4 of the paper.

In Table C.18, we analyze these effects in a regression framework. In Panel A of the table, we examine the relation between fund flows and past performance using the specification

$$F_y^q = \alpha_0 + \alpha_1 Perf^q_{y-5,y-1} + \varepsilon_y^q,$$

where F_y^q is the flow in year y averaged across all funds in flow quintile q, and $Perf^q_{y-5,y-1}$ is the posterior performance (alpha relative to the 4-factor model) estimated using our methodology over the previous 5 years (from $y-5$ to $y-1$) averaged across all funds in flow quintile q. In Panel B of the table, we examine the relation between fund flows and future performance using the specification

$$Perf^q_{y+1,y+5} - Perf^q_{y-5,y-1} = \beta_0 + \beta_1 F_y^q + u_y^q,$$

where F_y^q is as above and the dependent variable is the difference between performance in the 5-year period after and the 5-year period before year y, averaged across all funds in flow quintile q. To eliminate the effect of time, performance measures in both specifications are de-meaned by subtracting the mean performance across all funds operating contemporaneously.

The effects estimated from these regressions are consistent with those calculated from the quantile-based analysis. For example, in Panel A of Table C.18, we see that an increase of 1% in the annualized posterior mean of alpha in the 5-year period prior to flow measurement corresponds to an increase of 128% in the measured flows. In Panel B, we see that an increase of 100% in capital flows corresponds to a decrease of 0.25% in the difference between the annualized posterior mean of alpha in the subsequent and the preceding 5-year period.

Table C.16: Fund Flows and Past Performance — Quantile Analysis

The relation between fund flows and past fund performance. At the end of each non-overlapping 5-year period from 1975 to 2010, we sort funds into quintiles (Q1 through Q5) based on their flows (expressed as a percent of beginning-of-year total net asset value) in the subsequent year. In each 5-year period, we measure fund performance using i) the posterior probability (expressed as a percent) of having a positive alpha, ii) the posterior probability (expressed as a percent) of having a negative alpha, and iii) the posterior mean of alpha (expressed as an annualized percent), all estimated from our model with 4-factors. For each period, we present average performance measures and average flows across all funds in each flow quintile. We also present averages across all periods. In the columns labeled 'Q5−Q1', we report the difference between the top and the bottom flow quintile. */**/*** indicate significance of this difference at the 10%/5%/1% levels. In the columns labeled 'N', we write the number of funds used in the calculations for each 5-year period.

		Panel A: Positive-α probability (as a %)						**Panel B: Negative-α probability (as a %)**					
		Flow Quintiles						**Flow Quintiles**					
Period	**N**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q5−Q1**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q5−Q1**
1975 – 1980	244	8.23	9.28	10.31	13.47	15.70	7.47 ***	67.33	65.28	61.59	58.24	55.54	−11.81 ***
1980 – 1985	266	40.53	45.67	46.90	48.73	60.93	20.40 ***	27.13	22.08	20.56	21.33	11.56	−15.57 ***
1985 – 1990	422	31.74	37.04	36.19	46.06	49.83	18.09 ***	29.47	21.87	21.24	17.20	15.63	−13.84 ***
1990 – 1995	679	13.58	13.98	16.73	20.74	23.04	9.46 ***	59.03	58.38	52.39	47.90	46.52	−12.51 ***
1995 – 2000	1187	3.29	4.26	4.73	5.91	5.48	2.19 ***	89.43	86.91	85.36	82.43	83.88	−5.55 ***
2000 – 2005	1332	8.59	9.03	11.02	12.28	14.01	5.42 ***	53.44	51.10	45.69	43.58	41.78	−11.66 ***
2005 – 2010	1032	12.24	13.98	14.78	18.83	19.23	6.99 ***	56.79	53.13	51.01	44.28	43.40	−13.39 ***
Average		16.88	19.04	20.08	23.69	26.89	10.01 ***	54.64	51.26	48.29	44.99	42.62	−12.03 ***

		Panel C: α (as a %/year)						**Panel D: Flows (as a %/year)**					
		Flow Quintiles						**Flow Quintiles**					
Period	**N**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q5−Q1**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q5−Q1**
1975 – 1980	244	−0.84	−0.72	−0.60	−0.48	−0.36	0.48 ***	−35.43	−24.19	−19.68	−10.45	46.59	82.02 ***
1980 – 1985	266	−0.12	0.24	0.48	0.36	1.32	1.44 ***	−21.68	−9.50	−2.04	10.46	85.10	106.78 ***
1985 – 1990	422	−0.24	0.00	0.24	0.72	0.60	0.84 ***	−20.56	−8.34	−2.32	6.69	39.05	59.61 ***
1990 – 1995	679	−0.72	−0.84	−0.48	−0.24	−0.24	0.48 ***	−30.25	−11.56	0.34	21.65	173.56	203.81 ***
1995 – 2000	1187	−1.92	−1.80	−1.80	−1.68	−1.68	0.24 ***	−33.26	−14.93	−5.49	9.58	157.08	190.34 ***
2000 – 2005	1332	−1.08	−0.96	−0.72	−0.72	−0.60	0.48 ***	−37.19	−19.03	−8.55	9.34	159.23	196.42 ***
2005 – 2010	1032	−0.84	−0.72	−0.60	−0.36	−0.36	0.48 ***	−30.65	−13.34	−4.03	8.77	154.76	185.41 ***
Average		−0.82	−0.69	−0.50	−0.34	−0.19	0.63 ***	−30.03	−14.15	−6.46	8.25	116.38	146.41 ***

Table C.17: Fund Flows and Future Performance — Quantile Analysis

The relation between fund flows and subsequent fund performance. At the beginning of each non-overlapping 5-year period from 1980 to 2010, we sort funds into quintiles (Q1 through Q5) based on their flows (expressed as a percent of beginning-of-year total net asset value) in the previous year. In each 5-year period, we measure fund performance using i) the posterior probability (expressed as a percent) of having a positive alpha, ii) the posterior probability (expressed as a percent) of having a negative alpha, and iii) the posterior mean of alpha (expressed as an annualized percent), all estimated from our model with 4-factors. For each period, we present average performance measures and average flows across all funds in each flow quintile. We also present averages across all periods. In the columns labeled 'Q5−Q1', we report the difference between the top and the bottom flow quintile. */**/*** indicate significance of this difference at the 10%/5%/1% levels. In the columns labeled 'N', we write the number of funds used in the calculations for each 5-year period.

		Panel A: Positive-α probability (as a %)						**Panel B: Negative-α probability (as a %)**					
		Flow Quintiles						**Flow Quintiles**					
Period	**N**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q5−Q1**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q5−Q1**
1980 – 1985	243	45.12	45.95	51.23	51.21	45.58	0.46	24.52	20.56	17.99	18.33	24.25	−0.27
1985 – 1990	324	37.09	36.04	37.45	46.78	38.75	1.65	21.93	22.56	23.38	16.32	22.38	0.45
1990 – 1995	615	18.81	14.53	16.19	18.06	16.82	−2.01	51.87	58.13	53.89	51.90	54.54	2.67
1995 – 2000	924	4.63	4.86	4.06	4.43	4.42	−0.21	85.98	85.08	87.06	86.43	86.44	0.45
2000 – 2005	1353	14.18	11.06	10.50	9.28	8.59	−5.59 ***	40.07	46.94	47.56	50.10	53.34	13.27 ***
2005 – 2010	1064	16.74	15.04	15.33	15.17	14.88	−1.86 **	48.61	50.78	50.64	51.19	52.09	3.48 *
Average		22.76	21.25	22.48	24.16	21.51	−1.25 ***	45.50	47.34	46.75	45.70	48.84	3.34 ***

		Panel C: α (as a %/year)						**Panel D: Flows (as a %/year)**					
		Flow Quintiles						**Flow Quintiles**					
Period	**N**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q5−Q1**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q5−Q1**
1980 – 1985	243	0.24	0.36	0.72	0.60	0.24	0.00	−26.21	−19.45	−14.34	−7.23	67.23	93.44 ***
1985 – 1990	324	0.24	0.12	−0.12	0.72	0.36	0.12	−15.33	−3.23	2.49	12.68	111.89	127.22 ***
1990 – 1995	615	−0.48	−0.72	−0.60	−0.48	−0.60	−0.12	−37.10	−15.87	−4.86	17.57	188.39	225.49 ***
1995 – 2000	924	−1.80	−1.80	−1.80	−1.80	−1.80	0.00	−26.43	−6.40	6.35	30.05	224.75	251.18 ***
2000 – 2005	1353	−0.60	−0.84	−0.84	−0.96	−1.08	−0.48 ***	−36.78	−13.18	1.02	31.91	293.88	330.66 ***
2005 – 2010	1064	−0.48	−0.60	−0.60	−0.60	−0.72	−0.24 **	−26.56	−11.97	−1.34	17.39	183.39	209.95 ***
Average		−0.48	−0.58	−0.54	−0.42	−0.60	−0.12 ***	−28.04	−12.67	−2.34	17.59	178.43	206.47 ***

Table C.18: Fund Flows and Fund Performance — Regression Analysis

Regression analysis of the relation between fund flows and past and future fund performance. In Panel A, we present estimates from an OLS regression of annual fund flows on fund performance over the previous 5 years. In Panel B, we present estimates of an OLS regression in which the explanatory variable is the annual fund flow and the dependent variable is the difference between fund performance in the subsequent 5-year period and fund performance in the previous 5-year period. Measures of fund flows and fund performance are averaged across all funds that belong to the same flow quintile each year. In both panels, we present results from three specifications: In specification (1), the performance measure is the posterior probability (expressed as a percent) of having a positive alpha; in specification (2), it is the posterior probability (expressed as a percent) of having a negative alpha; and in specification (3), it is the posterior mean of alpha (expressed as an annualized percent). All performance measures are estimated from our model with 4 factors, and are de-meaned by subtracting the mean performance across all funds operating contemporaneously. Fund flows are expressed as a percent of beginning-of-year total net asset value; in Panel B, 'Flows $\times 100$' means that the quoted coefficients correspond to an increase of 100% in flows. t-statistics are reported below the coefficients. $^{*}/^{**}/^{***}$ indicate significance at the 10%/5%/1% levels.

Panel A: Flows and past performance

	(1)	(2)	(3)
Intercept	19.23 ***	19.37 ***	19.26 ***
	4.76	5.09	4.82
Positive-α probability	7.60 ***		
	8.95		
Negative-α probability		−8.73 ***	
		−10.64	
α			128.17 ***
			9.30
Adj R^2	0.34	0.43	0.36
Number of observations	155	155	155

Panel B: Flows and future performance

	(1)	(2)	(3)
Intercept	1.22 ***	−1.07 ***	0.05 ***
	2.57	−3.04	2.70
Flows $\times 100$	−3.96 ***	5.13 ***	−0.25 ***
	−7.87	9.33	−8.27
Adj R^2	0.33	0.40	0.35
Number of observations	130	130	130

C.12 Prior sensitivity analysis

Here, we present some additional results for the prior sensitivity analysis discussed in Section 8.1 of the paper.

First, we present more detailed results from the sensitivity analysis that replaces the symmetric $\mathcal{D}(1)$ prior with the asymmetric $\mathcal{D}(1, 3, 3)$ prior which overweights component probability vectors with lower values for the proportion of zero-alpha funds. In Table C.19, we present statistics for the posterior of the population proportions of zero-, negative-, and positive-alpha funds, and in Table C.20 we present various percentiles of the estimated distribution of annualized alpha.

Table C.19: Proportions of Fund Types — Sensitivity to Asymmetric Prior

Analysis of the sensitivity of the posterior distribution of the population proportions of zero-alpha, negative-alpha, and positive-alpha funds, to an asymmetric $\mathcal{D}(1, 3, 3)$ prior over these population proportions. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter.

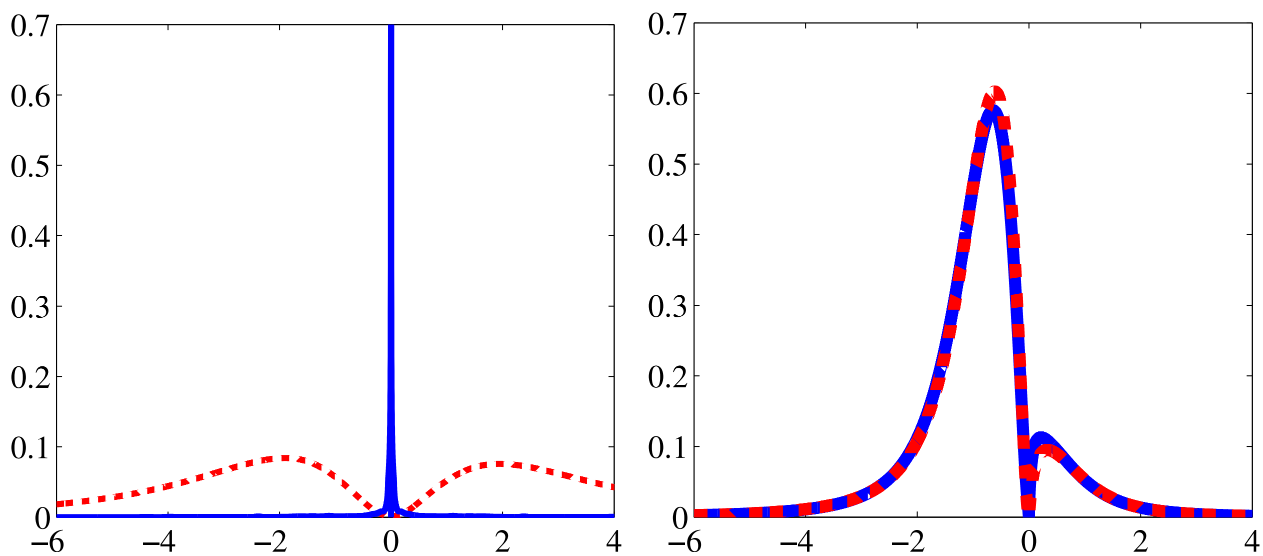
	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
π^0	0.07	0.05	0.06	[0.00 , 0.21]	0.00
π^-	0.80	0.80	0.05	[0.68 , 0.88]	0.00
π^+	0.13	0.13	0.04	[0.07 , 0.21]	0.00

Table C.20: Percentiles of Estimated Skill Distribution — Sensitivity to Asymmetric Prior

Analysis of the sensitivity of the percentiles of the estimated population distribution of annualized alpha (expressed as a percent) to an asymmetric $\mathcal{D}(1, 3, 3)$ prior over the proportions of zero-, negative-, and positive-alpha funds. We report the posterior mean and 90% HPDI of the percentiles of the alpha distribution. The 90% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.90.

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
Posterior Mean	-6.47	-5.03	-2.69	-1.99	-1.41	-1.10	-0.88	-0.70	-0.53	-0.36	0.00	0.29	0.82	2.40	3.37
5%	-8.00	-5.79	-2.99	-2.18	-1.51	-1.20	-1.00	-0.82	-0.65	-0.46	-0.27	0.00	0.51	1.88	2.62
95%	-5.46	-4.34	-2.36	-1.80	-1.30	-1.00	-0.78	-0.60	-0.44	-0.19	0.00	0.55	1.08	2.99	4.22

Next, we present more detailed results for the analysis that checks the sensitivity of the posterior predictive distribution of alpha on its prior predictive distribution. To achieve the desired change in the prior predictive distribution, we replace the $\mathcal{D}(1)$ prior with the $\mathcal{D}(1,3,3)$ prior (as above), and also we vary the baseline hyperprior parameters for the population mean and variance of the negative and positive components of the alpha distribution. In particular, in the baseline priors we use parameters $\underline{\kappa}_{\underline{\kappa}_\alpha} = 0$, $\underline{K}_{\underline{\kappa}_\alpha} = 100$, $\underline{\lambda}_{\underline{K}_\alpha} = 1$, and $\underline{\Lambda}_{\underline{K}_\alpha} = \underline{\lambda}_{\underline{\Lambda}_\alpha} = \underline{\Lambda}_{\underline{\Lambda}_\alpha} = 1$, and in the alternative specification for which we present results here we use $\underline{\kappa}_{\underline{\kappa}_\alpha} = -6$, $\underline{K}_{\underline{\kappa}_\alpha} = 0.5$, $\underline{\lambda}_{\underline{K}_\alpha} = 4$, and (as in the baseline) $\underline{\Lambda}_{\underline{K}_\alpha} = \underline{\lambda}_{\underline{\Lambda}_\alpha} = \underline{\Lambda}_{\underline{\Lambda}_\alpha} = 1$. The former parameter values (combined with our other prior parameters) imply that the prior predictive density is almost flat for all nonzero values of alpha except those very close to zero, while the latter imply that the prior predictive density is concentrated over a small range of values away from zero. As we show in Figure C.6, the change in the prior predictive density is drastic but the effect on the posterior predictive density is very small. In Tables C.21 and C.22, we also present statistics for the posteriors of the population proportions of zero-alpha, negative-alpha, and positive-alpha funds, and various percentiles of the estimated distribution of annualized alpha when using the alternative prior parameters.



(a) Prior predictive density, baseline vs alternative.

(b) Posterior predictive density, baseline vs alternative.

Figure C.6: Plots of the prior and posterior predictive density of α_i at $\alpha_i \neq 0$, for the baseline prior specification ($\pi \sim \mathcal{D}(1)$, $\underline{\kappa}_{\underline{\kappa}_\alpha} = 0$, $\underline{K}_{\underline{\kappa}_\alpha} = 100$, $\underline{\lambda}_{\underline{K}_\alpha} = 1$, $\underline{\Lambda}_{\underline{K}_\alpha} = \underline{\lambda}_{\underline{\Lambda}_\alpha} = \underline{\Lambda}_{\underline{\Lambda}_\alpha} = 1$) and for an alternative ($\pi \sim \mathcal{D}(1,3,3)$, $\underline{\kappa}_{\underline{\kappa}_\alpha} = -6$, $\underline{K}_{\underline{\kappa}_\alpha} = 0.5$, $\underline{\lambda}_{\underline{K}_\alpha} = 4$, $\underline{\Lambda}_{\underline{K}_\alpha} = \underline{\lambda}_{\underline{\Lambda}_\alpha} = \underline{\Lambda}_{\underline{\Lambda}_\alpha} = 1$). In Panel (a), we plot the prior predictive density for the baseline prior specification (in solid blue) and for the alternative prior specification (in dotted red). In Panel (b), we plot the posterior predictive density for the baseline prior specification (in solid blue) and for the alternative prior specification (in dotted red). For clarity, we do not represent the point mass at zero alpha, which has probability 0.33 for the baseline and 0.14 for the alternative prior predictive density, and probability 0.09 for the baseline and 0.08 for the alternative posterior predictive density.

Table C.21: Proportions of Fund Types — Sensitivity to Prior Predictive Distribution

Analysis of the sensitivity of the posterior distribution of the population proportions of zero-alpha, negative-alpha, and positive-alpha funds, to the alternative specification for the prior predictive distribution implied by replacing $\pi \sim \mathcal{D}(1)$ with $\pi \sim \mathcal{D}(1, 3, 3)$ and the baseline values for hyperparameters $\underline{\kappa}_{\underline{\kappa}_\alpha}$, $\underline{K}_{\underline{\kappa}_\alpha}$, and $\underline{\lambda}_{\underline{K}_\alpha}$ with -6, 0.5, and 4, respectively. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter.

	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
π^0	0.08	0.06	0.06	[0.00 , 0.21]	0.00
π^-	0.80	0.82	0.06	[0.68 , 0.88]	0.00
π^+	0.12	0.11	0.03	[0.06 , 0.20]	0.00

Table C.22: Estimated Percentiles of Skill — Sensitivity to Prior Predictive Distribution

Analysis of the sensitivity of the percentiles of the estimated population distribution of annualized alpha (expressed as a percent) to the alternative specification for the prior predictive distribution implied by replacing $\pi \sim \mathcal{D}(1)$ with $\pi \sim \mathcal{D}(1, 3, 3)$ and the baseline values for hyperparameters $\underline{\kappa}_{\underline{\kappa}_\alpha}$, $\underline{K}_{\underline{\kappa}_\alpha}$, and $\underline{\lambda}_{\underline{K}_\alpha}$ with -6, 0.5, and 4, respectively. We report the posterior mean and 90% HPDI of the percentiles of the alpha distribution. The 90% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.90.

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
Posterior Mean	-6.55	-5.05	-2.67	-1.98	-1.41	-1.10	-0.89	-0.71	-0.54	-0.37	0.00	0.29	0.92	2.39	3.26
5%	-8.05	-5.92	-3.02	-2.19	-1.52	-1.21	-1.00	-0.83	-0.66	-0.49	-0.29	0.00	0.51	1.85	2.42
95%	-5.43	-4.25	-2.33	-1.78	-1.30	-1.00	-0.78	-0.60	-0.43	0.00	0.00	0.63	1.22	2.98	4.15

C.13 Dependence between model parameters

In this section, we present results from the estimation of the model that allows for a type-specific distribution for all model parameters — skill α_i, the factor loadings β_i, and fund return error precision h_i — as well as a full correlation matrix between them, as presented in Section 8.2 of the paper. In Table C.23 we present results on the posterior distribution of population correlations between alpha, the factor loadings, and the error precision, in Table C.24, we present results on the posterior distribution of the population proportions of zero-, negative-, and positive-alpha funds, and in Table C.25 we present the percentiles of the estimated distributions of alpha and the factor loadings (separately for zero-, negative-, and positive-alpha funds).

We generally see that the correlations between alpha and the factor loadings and alpha and the error precision are quite small in magnitude. Furthermore, though the estimated distributions of the factor loadings differ between zero-, negative-, and positive-alpha funds, importantly, the skill distribution we estimate here is similar to the one we estimate in the baseline model. This shows that the baseline model is robust to the presence of weak correlation between alpha and the other model parameters.

Table C.23: Correlation Matrix — Specification with Full Correlation

Means and standard deviations (in parentheses) of the posterior distributions of population correlations between alpha, the factor loadings, and the error precision, estimated with returns net of expenses using the $K^- = 2$, $K^+ = 1$ model that allows alpha, the factor loadings, and the error precision to be correlated.

	Panel A: Zero-alpha funds					Panel B: Negative-alpha funds					Panel C: Positive-alpha funds				
	β_{M}	β_{SMB}	β_{HML}	β_{UMD}	h	β_{M}	β_{SMB}	β_{HML}	β_{UMD}	h	β_{M}	β_{SMB}	β_{HML}	β_{UMD}	h
α	0	0	0	0	0	−0.13	−0.15	−0.01	−0.11	0.09	0.03	−0.09	−0.17	0.09	0.16
	(0)	(0)	(0)	(0)	(0)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.12)	(0.11)	(0.12)	(0.06)	(0.17)
β_{M}		0.02	0.24	−0.26	−0.08		0.36	−0.56	0.47	−0.25		0.07	−0.69	−0.09	−0.10
		(0.05)	(0.06)	(0.05)	(0.04)		(0.02)	(0.01)	(0.02)	(0.02)		(0.04)	(0.02)	(0.03)	(0.03)
β_{SMB}			−0.05	−0.09	−0.03			−0.06	0.33	−0.35			−0.07	−0.10	−0.20
			(0.06)	(0.06)	(0.03)			(0.02)	(0.02)	(0.01)			(0.04)	(0.05)	(0.04)
β_{HML}				−0.30	−0.16				−0.55	0.07				−0.25	0.07
				(0.06)	(0.03)				(0.01)	(0.01)				(0.03)	(0.02)
β_{UMD}					0.02					−0.19					−0.09
					(0.02)					(0.02)					(0.03)

Table C.24: Proportions of Fund Types — Specification with Full Correlation

Results on the posterior distributions of the population proportions of funds with zero, negative, and positive alpha, estimated with net returns using the $K^- = 2$, $K^+ = 1$ model that allows for full correlation between alpha, the factor loadings, and the error precision. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter.

	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
π^0	0.15	0.15	0.01	[0.12 , 0.17]	0.00
π^-	0.65	0.65	0.02	[0.62 , 0.69]	0.00
π^+	0.20	0.20	0.01	[0.18 , 0.23]	0.00

Table C.25: Percentiles of Estimated Distributions — Specification with Full Correlation

Percentiles of the estimated distributions of annualized alpha (expressed as a percent) and factor loadings, estimated with net returns in the K^-=2, K^+=1 model that allows for full correlation between alpha, factor loadings, and error precision.

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
α	-6.37	-5.11	-2.77	-1.97	-1.28	-0.91	-0.66	-0.46	-0.25	0.00	0.05	0.52	1.01	2.68	3.73
Zero-Alpha Funds															
$\beta_{\mathbf{M}}$	-0.10	-0.03	0.18	0.29	0.43	0.52	0.61	0.68	0.76	0.84	0.94	1.07	1.19	1.39	1.47
$\beta_{\mathbf{SMB}}$	-0.38	-0.34	-0.24	-0.19	-0.12	-0.07	-0.03	0.00	0.04	0.08	0.13	0.20	0.25	0.35	0.39
$\beta_{\mathbf{HML}}$	-0.36	-0.32	-0.18	-0.11	-0.03	0.04	0.09	0.14	0.19	0.24	0.30	0.39	0.46	0.59	0.64
$\beta_{\mathbf{UMD}}$	-0.25	-0.23	-0.16	-0.13	-0.09	-0.06	-0.03	-0.01	0.01	0.04	0.07	0.11	0.15	0.21	0.23
Negative-Alpha Funds															
$\beta_{\mathbf{M}}$	0.69	0.72	0.80	0.84	0.89	0.92	0.95	0.98	1.01	1.04	1.08	1.12	1.16	1.24	1.27
$\beta_{\mathbf{SMB}}$	-0.41	-0.37	-0.25	-0.19	-0.11	-0.04	0.02	0.08	0.17	0.38	0.55	0.67	0.75	0.89	0.93
$\beta_{\mathbf{HML}}$	-0.70	-0.63	-0.44	-0.34	-0.22	-0.13	-0.05	0.02	0.09	0.16	0.25	0.38	0.48	0.67	0.74
$\beta_{\mathbf{UMD}}$	-0.24	-0.22	-0.15	-0.11	-0.06	-0.03	-0.00	0.02	0.05	0.08	0.11	0.15	0.19	0.26	0.29
Positive-Alpha Funds															
$\beta_{\mathbf{M}}$	0.36	0.43	0.60	0.70	0.81	0.89	0.96	1.03	1.09	1.16	1.24	1.36	1.45	1.63	1.69
$\beta_{\mathbf{SMB}}$	-0.19	-0.14	0.01	0.09	0.19	0.26	0.32	0.38	0.43	0.49	0.56	0.66	0.74	0.89	0.95
$\beta_{\mathbf{HML}}$	-1.54	-1.40	-1.00	-0.80	-0.54	-0.36	-0.21	-0.06	0.09	0.24	0.42	0.68	0.89	1.28	1.42
$\beta_{\mathbf{UMD}}$	-0.44	-0.40	-0.29	-0.23	-0.16	-0.11	-0.06	-0.03	0.01	0.06	0.11	0.18	0.24	0.34	0.38

C.14 Conditional asset-pricing model

In this section, we present results from the estimation of the model that allows for funds' market factor loadings to be time-varying, by allowing them to depend in a fund-specific manner on the T-bill rate, the dividend yield, the term spread, and the default spread; see Section 8.3 of the paper for details. In particular, in Table C.26, we present results on the posterior distributions of the population proportions of funds with zero, negative, and positive alpha, in Table C.27 we present the percentiles of the estimated distribution of alpha, the factor loadings, and the coefficients on the conditioning variables, and in Figure C.7 we compare the estimated distributions of alpha and the factor loadings from the two versions of our model: the baseline, in which each fund's market factor loading is constant, and the conditional one, in which this loading is allowed to be time-varying in a fund-specific manner.

Table C.26: Proportions of Fund Types — Conditional Model

Results on the posterior distributions of the proportions of funds with zero, negative, and positive alpha, estimated with returns net of expenses using the conditional model presented in Section 8.3 of the paper. Each fund's market loading at month t may depend in a fund-specific manner on the following quantities at month $t-1$: the 1-month T-bill rate; the dividend yield of the CRSP value-weighted index; the term spread, proxied by the yield difference between constant-maturity 10-year Treasury bonds and 3-month T-bills; and the default spread, proxied by the yield difference between Moody's Baa- and Aaa-rated corporate bonds. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for posterior mean estimates.

	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
π^0	0.08	0.06	0.06	[0.00 , 0.22]	0.00
π^-	0.81	0.82	0.05	[0.69 , 0.90]	0.00
π^+	0.11	0.11	0.03	[0.05 , 0.18]	0.00

Table C.27: Percentiles of Estimated Distributions — Conditional Model

Percentiles of the estimated distributions of annualized alpha (expressed as a percent), the factor loadings, and the coefficients on the conditioning variables, estimated with returns net of expenses using the conditional model presented in Section 8.3 of the paper. Each fund's market factor loading at month t may depend in a fund-specific manner on the following quantities at month $t-1$: i) the one-month T-bill rate (TB); ii) the dividend yield (DY) of the CRSP value-weighted index; iii) the term spread (TS), proxied by the yield difference between constant-maturity 10-year Treasury bonds and three-month T-bills; and iv) the default spread (DS), proxied by the yield difference between Moody's Baa-rated and Aaa-rated corporate bonds.

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
α	-6.55	-5.05	-2.77	-2.11	-1.56	-1.25	-1.03	-0.84	-0.67	-0.48	-0.17	0.21	0.80	2.38	3.32
β_{M}	0.14	0.22	0.45	0.57	0.72	0.82	0.91	1.00	1.08	1.17	1.28	1.43	1.55	1.78	1.87
β_{SMB}	-0.61	-0.53	-0.32	-0.21	-0.07	0.03	0.11	0.19	0.27	0.35	0.45	0.59	0.70	0.91	0.99
β_{HML}	-0.84	-0.76	-0.53	-0.41	-0.27	-0.16	-0.07	0.01	0.10	0.19	0.30	0.44	0.56	0.79	0.87
β_{UMD}	-0.27	-0.24	-0.17	-0.13	-0.08	-0.05	-0.02	0.01	0.03	0.06	0.10	0.14	0.18	0.26	0.28
γ_{TB}	-130.94	-119.26	-87.49	-70.49	-49.86	-35.07	-22.44	-10.69	1.09	13.67	28.46	49.00	65.92	98.21	109.42
γ_{DY}	-37.65	-33.82	-23.55	-18.13	-11.52	-6.76	-2.67	1.12	4.94	8.99	13.77	20.36	25.84	36.03	39.72
γ_{TS}	-0.17	-0.15	-0.11	-0.08	-0.05	-0.03	-0.02	0.00	0.02	0.03	0.05	0.08	0.10	0.15	0.16
γ_{DS}	-0.44	-0.40	-0.29	-0.24	-0.17	-0.12	-0.08	-0.04	0.00	0.04	0.09	0.16	0.22	0.32	0.36

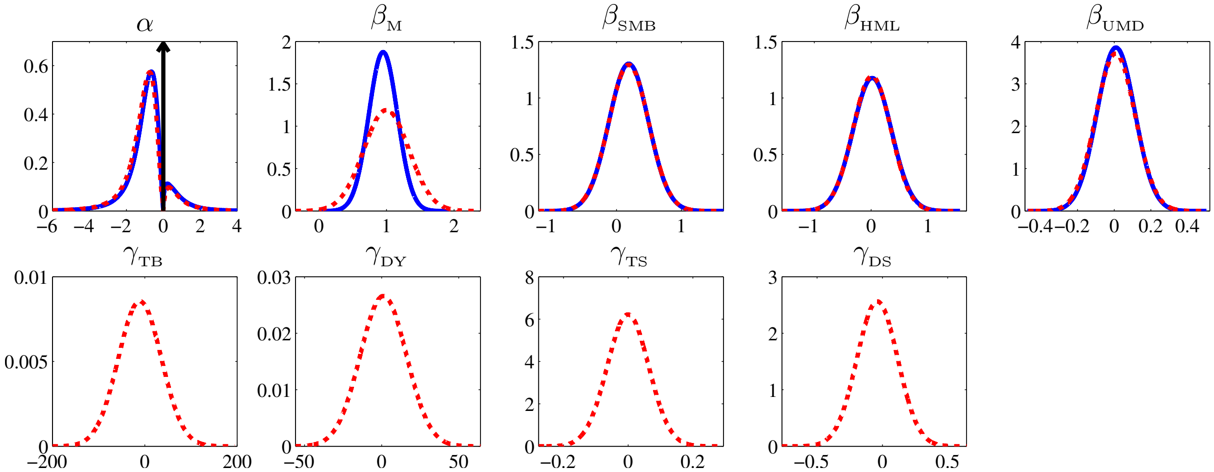


Figure C.7: Comparison of the estimated densities of annualized alpha (expressed as a percent) and factor loadings for two models: blue solid lines plot densities from a model in which funds' market loadings are constant, red dashed lines plot densities from a model in which they are time-varying. Both sets of densities are estimated using our model with an unknown number of components for the alpha distribution of negative-alpha and positive-alpha funds. The black vertical arrow at zero alpha represents a point mass for both estimated distributions; the point mass for the model with constant market loadings has probability 0.09 and the other model's point mass has probability 0.08. Both models are estimated using returns net of expenses for 3,497 funds.

C.15 Distributional assumptions for errors, factor loadings

In this section, we present tables and figures relating to our analysis in Section 8.4 in the paper, of the sensitivity of our results to distributional assumptions about the fund return errors and about the factor loadings in Equation 1 in the paper.

First, we present results from an estimation in which we exclude, for each fund, the observations whose residuals most deviate from the values of a normal random sample of size equal to the number of observations for that fund; we remove no more than 8 observations per fund, and in total we remove 1.8% of all observations. In this restricted data set, our assumption that each fund's errors in Equation 1 are normal should be even more accurate than in the whole data set, therefore this analysis can help determine whether the results we present using the whole data set are driven by non-normality in the errors. In Table C.28, we present results on the posterior distribution of the population proportions of funds with zero, negative, and positive alpha, and in Table C.29 we present various percentiles of the estimated distribution of alpha. In Figure C.8, we plot the estimated skill densities from the estimation that uses the whole data set and from the one that uses the restricted data set. In Table C.28, we see that the proportions we estimate are largely unchanged. In Table C.29 and in Figure C.8, we see that the estimated skill distribution is very similar in both cases, with the most noteworthy difference being that the distribution estimated on the restricted data set has slightly fatter tails, especially in the left. The reason for this is that the excluded observations correspond to residuals that are large, so by excluding them the posterior estimates of fund-level αs have lower variance. As a result, there is less shrinkage toward the mean, therefore not only do we place more fund-level αs in the tails, but we are also more confident that they should be in the tails, so we estimate fatter tails. The effect is stronger on the left than on the right tail, because more of the funds for which the exclusion of observations leads to a big reduction in their error variance are located in the left than in the right tail.

Table C.28: Proportions of Fund Types — Excluding non-Normal Residuals

Results on the posterior distributions of the population proportions of funds with zero, negative, and positive alpha, estimated using our baseline model with net returns for 3,497 funds, but excluding, for each fund, the observations whose residuals most deviate from the values of an equal-sized normal random sample. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter.

	Mean	Median	Std.Dev.	95% HPDI	NSE
π^0	0.08	0.08	0.06	[0.00 , 0.22]	0.00
π^-	0.78	0.78	0.05	[0.68 , 0.86]	0.00
π^+	0.14	0.13	0.04	[0.07 , 0.21]	0.00

Table C.29: Percentiles of Estimated Skill Distribution — Excluding non-Normal Residuals

Percentiles of estimated distributions of annualized alpha (expressed as a percent) from our estimation on the whole data set (in Panel A) and from the estimation on a restricted data set (in Panel B), which excludes, for each fund, the observations whose residuals most deviate from the values of a normal random sample of size equal to the number of observations for that fund. Both estimations use returns net of expenses for 3,497 funds. For each estimation, we report the posterior mean and 90% HPDI of the percentiles of the alpha distribution. The 90% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.90.

Panel A: Estimation using the whole data set

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
Posterior Mean	-6.60	-5.01	-2.62	-1.96	-1.42	-1.12	-0.96	-0.72	-0.55	-0.36	0.00	0.30	0.87	2.35	3.24
5%	-8.48	-5.95	-2.97	-2.17	-1.54	-1.25	-1.04	-0.86	-0.69	-0.50	-0.29	0.00	0.49	1.77	2.40
95%	-5.36	-4.04	-2.25	-1.77	-1.30	-1.00	-0.78	-0.60	-0.42	0.00	0.02	0.61	1.20	2.98	4.24

Panel B: Estimation using the restricted data set

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
Posterior Mean	-8.97	-6.41	-2.98	-2.20	-1.61	-1.30	-1.07	-0.87	-0.66	-0.42	0.00	0.46	1.09	2.51	3.30
5%	-11.46	-7.74	-3.46	-2.47	-1.73	-1.41	-1.20	-1.00	-0.81	-0.57	-0.28	0.00	0.81	2.04	2.61
95%	-7.07	-5.23	-2.56	-1.98	-1.51	-1.19	-0.95	-0.74	-0.54	0.00	0.04	0.74	1.37	3.01	4.08

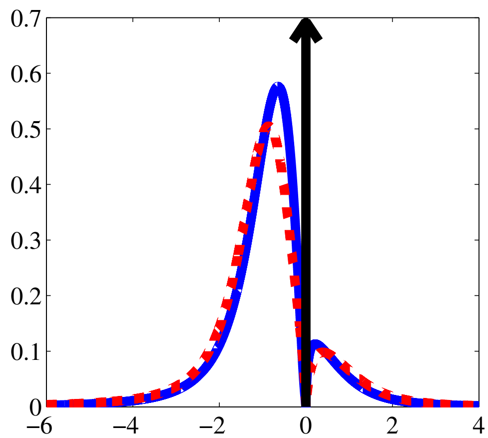


Figure C.8: Plots of the estimated density of annualized 4-factor alphas, expressed as percents. In solid blue, we plot the density estimated using all observations of the 3,497 funds in our data, i.e., the density from our baseline estimation. In dotted red, we plot the density estimated using the same data but excluding the 1.8% of all observations whose residuals most deviate from the normal distribution. The black vertical arrow at zero alpha represents a point mass for both estimated densities; the point mass has probability 0.09 for the baseline estimation and probability 0.08 for the estimation on the restricted data set.

Now, we turn our attention to our assumption that the factor loadings β are normal. In Figure C.9, we present Q-Q plots of posterior mean βs estimated using our methodology versus OLS estimates of the βs. The plots lie quite close to the 45° line, showing that the quantiles of the two sets of distributions are very similar, and therefore that shrinkage in our posterior estimates is very limited; only the posteriors for β_{UMD} exhibit (a little) more than a non-trivial degree of shrinkage. This implies that our distributional assumption is unlikely to have a significant effect on our fund-level estimates of alpha, and therefore on our estimated distribution of alpha.

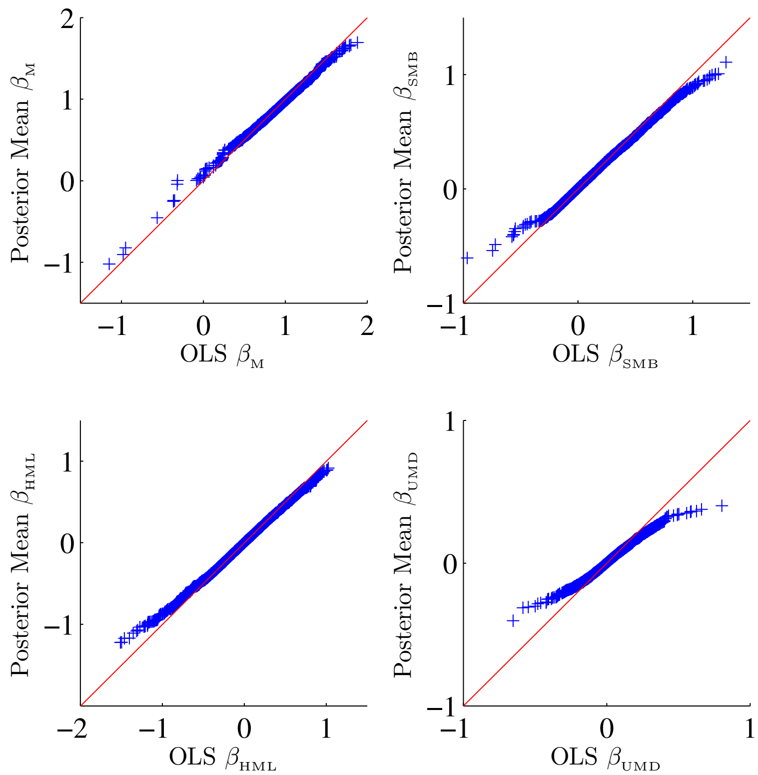


Figure C.9: Quantile-Quantile plots of posterior mean βs estimated using our methodology versus OLS estimates of βs. The blue cross marks plot the quantiles, and the solid red line plots the 45° line.

Next, we present results from an estimation in which the scale $\underline{\Lambda}_\beta$ of the inverse Wishart prior distribution for the population variance V_β of β is large ($\underline{\Lambda}_\beta = 10^3 I$ instead of $\underline{\Lambda}_\beta = I$ as in our baseline estimation), which effectively eliminates shrinkage in the posterior estimates of fund-level βs. In Figure C.10, we present Q-Q plots for the posterior mean βs estimated using our methodology with "no shrinkage" priors for the βs versus OLS estimates of the βs; comparing these Q-Q plots with those in Figure C.9, we see that now all plots lie almost exactly on the 45° line, verifying that shrinkage in our estimates for the βs has effectively been eliminated. In Table C.30, we present results on the posterior distribution of the population proportions of funds with zero, negative, and positive alpha, and in Table C.31 we present various percentiles of the estimated distribution of alpha, from this estimation. In Figure C.11, we plot the estimated skill densities from the baseline estimation and from the estimation that imposes no shrinkage. These tables and figures show that the results on skill are almost identical for the two estimations: the one with the baseline priors, which induces little shrinkage, and the one with these alternative priors, which induces no shrinkage. Thus, we conclude that our baseline distributional assumption for the βs does not drive our main results in the paper.

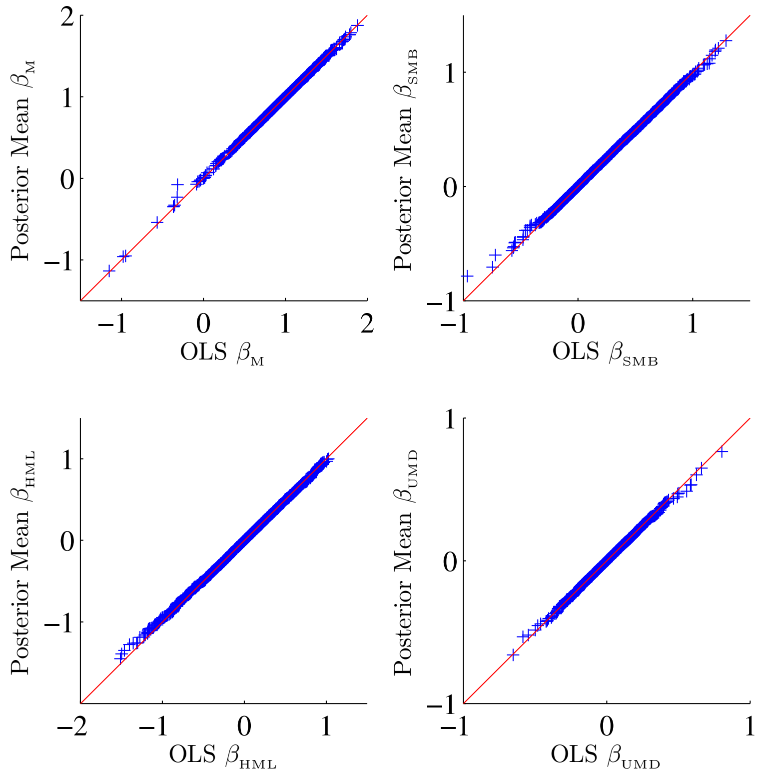


Figure C.10: Quantile-Quantile plots of posterior mean βs estimated using our methodology — but with priors that impose very little shrinkage on β — versus OLS estimates of βs. The blue cross marks plot the quantiles, and the solid red line plots the 45° line.

Table C.30: Proportions of Fund Types — No Shrinkage for βs

Results on the posterior distributions of the population proportions of funds with zero, negative, and positive alpha, estimated using our model presented in Section 2 with returns net of expenses for 3,497 funds, but with $\underline{\Lambda}_\beta = 10^3 I$ instead of $\underline{\Lambda}_\beta = I$ as in the baseline estimation. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter.

	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
π^0	0.08	0.07	0.07	[0.00 , 0.25]	0.00
π^-	0.79	0.79	0.06	[0.65 , 0.88]	0.00
π^+	0.13	0.13	0.04	[0.06 , 0.23]	0.00

Table C.31: Estimated Percentiles of Skill — Baseline vs. Model with no Shrinkage for βs

Percentiles of estimated distributions of annualized alpha (expressed as a percent) from our baseline specification (in Panel A) and from the specification with no shrinkage for βs, i.e., with prior parameter $\underline{\Lambda}_\beta = 10^3 I$ instead of $\underline{\Lambda}_\beta = I$ (in Panel B). Both specifications are estimated with returns net of expenses for 3,497 funds. For each specification, we report the posterior mean and 90% HPDI of the percentiles of the alpha distribution. The 90% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.90.

Panel A: Baseline specification

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
Posterior Mean	-6.60	-5.01	-2.62	-1.96	-1.42	-1.12	-0.96	-0.72	-0.55	-0.36	0.00	0.30	0.87	2.35	3.24
5%	-8.48	-5.95	-2.97	-2.17	-1.54	-1.25	-1.04	-0.86	-0.69	-0.50	-0.29	0.00	0.49	1.77	2.40
95%	-5.36	-4.04	-2.25	-1.77	-1.30	-1.00	-0.78	-0.60	-0.42	0.00	0.02	0.61	1.20	2.98	4.24

Panel B: Specification with no shrinkage for βs

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
Posterior Mean	-6.47	-4.99	-2.66	-1.98	-1.42	-1.12	-0.90	-0.71	-0.54	-0.35	0.00	0.33	0.91	2.41	3.31
5%	-8.03	-5.84	-3.01	-2.19	-1.54	-1.24	-1.03	-0.85	-0.67	-0.48	-0.28	0.00	0.53	1.86	2.45
95%	-5.32	-4.19	-2.29	-1.77	-1.31	-1.01	-0.78	-0.60	-0.43	0.00	0.05	0.65	1.21	2.99	4.21

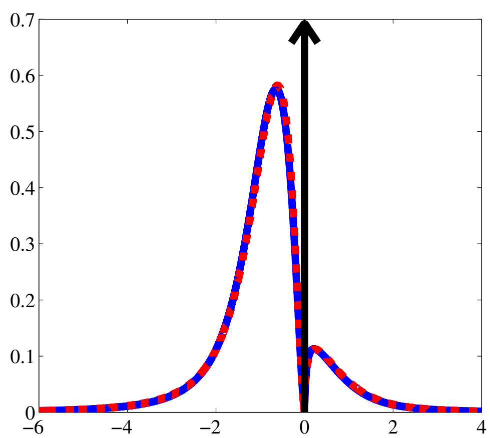


Figure C.11: Estimated density of annualized 4-factor alphas (expressed as a percent). The blue solid line plots the estimated population density of alpha from our baseline specification from Section 2. The red dotted line plots the estimated population density of alpha from an estimation in which the scale $\underline{\Lambda}_\beta$ of the inverse Wishart prior distribution for the population variance V_β of β is large ($\underline{\Lambda}_\beta = 10^3 I$ instead of $\underline{\Lambda}_\beta = I$ as in the baseline), so effectively with no shrinkage for the estimation of fund-level β. The black vertical arrow at zero alpha represents a point mass for both estimated densities; the point mass has probability 0.09 for the baseline model and probability 0.08 for the specification with large $\underline{\Lambda}_\beta$. For both plots, returns net of expenses for 3,497 funds are used.

C.16 Cross-sectional error dependence

In Section 8.5 of the paper, we present a model that allows for cross-sectional dependence in the error terms through linear latent error factors. To estimate this model, in the Gibbs sampler we i) add a block in which we draw from the conditional posterior of the latent factors (see Geweke and Zhou, 1996); ii) we augment the observed factors with the latent factors; and iii) we replace $\rho_i = \left(\alpha_i, \beta_i'\right)'$ with $\rho_i^* := \left(\alpha_i, \beta_i', \delta_i'\right)'$ and h_i with h_i^*, where as we explain in the paper, δ_i are fund-specific error factor loadings; and $\xi_{it} \sim \mathcal{N}\left(0, h_i^{*-1}\right)$ is the cross-sectionally independent part of the error term, with h_i^* a fund-specific precision. In this section, we present results from the estimation of this model. In particular, we present results from the specification with a single latent factor; results from the specification with 4 latent factors are very similar.

In Table C.32 we present results on the posteriors of the population proportions of funds with zero, negative, and positive alpha, in Table C.33 we present the percentiles of the estimated distributions of alpha and the factor loadings, and in Figure C.12 we compare the estimated distributions of alpha and the factor loadings from two versions of our model: one in which errors are assumed to be independent and one in which this is relaxed.

In Figure C.13, we present a sensitivity analysis of the effect of changing the prior for the distribution of the latent error factor coefficients on the estimated population densities of alpha and factor loadings. We find that there is an almost imperceptible effect to changing by several orders of magnitude the scale of the inverse Wishart prior distribution for the population variance V_δ of the latent error factor coefficients δ.

Table C.32: Proportions of Fund Types — Cross-sectional Error Dependence

Results on the posteriors of the population proportions of funds with zero, negative, and positive alpha, estimated with returns net of expenses using the model in which errors are cross-sectionally dependent through the latent error factor model presented in Section 8.5 of the paper. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimates.

	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
$\boldsymbol{\pi^0}$	0.08	0.07	0.06	[0.00 , 0.22]	0.00
$\boldsymbol{\pi^-}$	0.79	0.80	0.05	[0.68 , 0.87]	0.00
$\boldsymbol{\pi^+}$	0.13	0.13	0.04	[0.07 , 0.21]	0.00

Table C.33: Percentiles of Estimated Distributions — Cross-sectional Error Dependence

Percentiles of the estimated population distributions of annualized alpha (expressed as a percent) and factor loadings, estimated with returns net of expenses using the model in which errors are cross-sectionally dependent through the latent error factor model presented in Section 8.5 of the paper.

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
$\boldsymbol{\alpha}$	-6.88	-5.35	-2.94	-2.23	-1.63	-1.31	-1.07	-0.87	-0.67	-0.46	0.00	0.36	0.98	2.53	3.40
$\boldsymbol{\beta}_{\mathbf{M}}$	0.42	0.47	0.61	0.69	0.78	0.85	0.91	0.96	1.01	1.07	1.14	1.23	1.31	1.45	1.51
$\boldsymbol{\beta}_{\mathbf{SMB}}$	-0.63	-0.56	-0.35	-0.24	-0.11	-0.01	0.07	0.15	0.22	0.31	0.40	0.54	0.65	0.85	0.93
$\boldsymbol{\beta}_{\mathbf{HML}}$	-0.76	-0.70	-0.54	-0.45	-0.34	-0.27	-0.20	-0.14	-0.08	-0.02	0.06	0.16	0.25	0.41	0.47
$\boldsymbol{\beta}_{\mathbf{UMD}}$	-0.23	-0.21	-0.14	-0.10	-0.06	-0.03	0.00	0.03	0.05	0.08	0.11	0.16	0.19	0.26	0.29

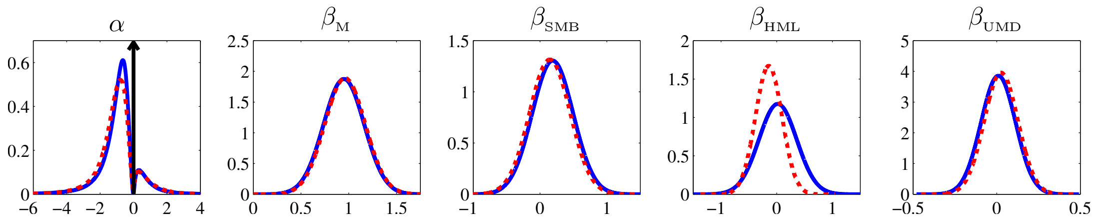


Figure C.12: Comparison of the estimated population densities of annualized alpha (expressed as a percent) and factor loadings under two models: blue solid lines plot densities from a model in which errors are independent across funds, and red dashed lines plot densities from a model in which errors are cross-sectionally dependent through the latent error factor model presented in Section 8.5 of the paper. The black vertical arrow at zero alpha represents a point mass, which has probability 0.09 for the former model and probability 0.08 for the latter model. Both models are estimated using returns net of expenses for 3,497 funds.

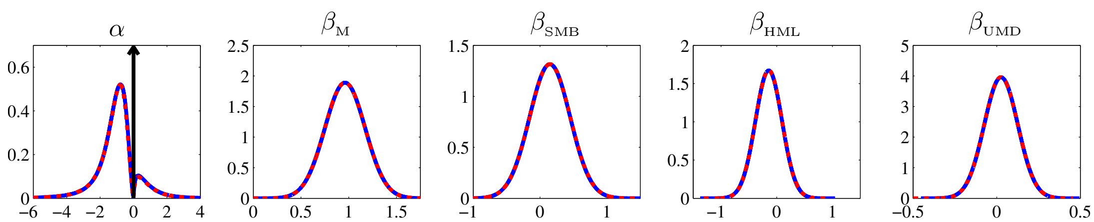


Figure C.13: Sensitivity analysis of the estimated population densities of annualized alpha (expressed as a percent) and factor loadings for the model in which errors are cross-sectionally dependent through the latent linear error factor model presented in Section 8.5 of the paper. The densities plotted in blue solid lines correspond to the estimated densities from an estimation in which the scale of the inverse Wishart prior distribution for the population variance V_δ of the latent error factor coefficients δ is small, while the densities plotted in red dashed lines correspond to the estimated densities from an estimation in which the scale of the inverse Wishart prior distribution for V_δ is large. The black vertical arrow at zero alpha represents a point mass with probability 0.08 for both estimated distributions. Both sets of densities are estimated using returns net of expenses for 3,497 funds.

C.17 Variations to model specification

In this section, we present results from two variations of our specification of the alpha distribution (see Section 8.6 of the paper). First we replace our assumption that nonzero alphas are drawn from two non-overlapping distributions — one for negative-alpha and one for positive-alpha funds — with the assumption that they are drawn from a common distribution, and then we replace the point mass at zero with a narrow normal centered at zero. Table C.34 presents posterior results for the population proportions of zero-alpha and nonzero-alpha funds estimated using the alternative model with a point mass and a normal, and Table C.35 presents various percentiles of the estimated distributions of alpha and the factor loadings using the alternative model with a narrow normal and two non-overlapping mixtures of log-normals.

Table C.34: Proportions of Fund Types — Point Mass and Normal

Results on the posterior distributions of the population proportions of funds with zero and nonzero alpha, estimated with returns net of expenses using a model in which alpha is drawn from a mixture distribution with two components: a point mass at zero and a normal distribution. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter.

	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
π^0	0.08	0.06	0.07	[0.00 , 0.24]	0.00
π^N	0.92	0.94	0.07	[0.76 , 1.00]	0.00

Table C.35: Percentiles of Estimated Distributions — Narrow Normal

Percentiles of the estimated population distributions of annualized alpha (expressed as a percent) and factor loadings estimated from a model in which alpha is drawn from a mixture with three components: a narrow normal centered at zero and two non-overlapping mixtures of log-normal distributions (one for negative- and one for positive-alpha funds). The model is estimated with returns net of expenses for 3,497 funds.

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
α	-6.52	-5.01	-2.65	-1.98	-1.42	-1.12	-0.90	-0.72	-0.55	-0.35	-0.08	0.29	0.87	2.34	3.23
$\beta_{\mathbf{M}}$	0.40	0.46	0.60	0.68	0.77	0.84	0.90	0.95	1.00	1.06	1.13	1.22	1.30	1.44	1.50
$\beta_{\mathbf{SMB}}$	-0.60	-0.52	-0.31	-0.20	-0.07	0.03	0.11	0.19	0.27	0.35	0.45	0.58	0.69	0.90	0.98
$\beta_{\mathbf{HML}}$	-0.85	-0.77	-0.54	-0.42	-0.27	-0.16	-0.07	0.02	0.11	0.20	0.30	0.45	0.58	0.80	0.89
$\beta_{\mathbf{UMD}}$	-0.26	-0.23	-0.16	-0.12	-0.08	-0.04	-0.02	0.01	0.04	0.06	0.10	0.14	0.18	0.25	0.28

C.18 Results excluding first two years per fund

Here, we present results from estimating our baseline model on a restricted data set that excludes returns observations for the first two years of each fund's life. The year of inception for most funds (89%) is provided in the CRSP database. Among the funds for which this information is provided, about 70% report returns in the database from their very first year of existence. As a result, for the 11% of funds for which the year of inception is not provided, we assume that it coincides with the time at which they start reporting their returns. Extrapolating the 70% figure stated earlier to all funds, it appears that this assumption would yield an incorrect inception year for only 3% of all funds. In any case, these mistakes would make our analysis here even more conservative, since by excluding for these funds the first two years of returns observations available in the data, we would effectively be excluding observations for *more* than the first two years since inception.

In Table C.36, we present statistics for the estimated posterior distribution of the population proportions of zero-alpha, negative-alpha, and positive-alpha funds. We see that, as in the baseline estimation, the large majority of funds have negative alpha.

Table C.36: Proportions of Fund Types — Excluding first two years of data

Results on the posterior distributions of the population proportions of funds with zero, negative, and positive alpha, estimated using our baseline model with net returns for 3,497 funds, but excluding the first two years of returns for each fund. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter.

	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
π^0	0.07	0.06	0.06	[0.00 , 0.20]	0.00
π^-	0.81	0.81	0.05	[0.70 , 0.90]	0.00
π^+	0.12	0.11	0.04	[0.05 , 0.20]	0.00

C.19 MCMC without reversible jumps

In this section, we present the results from the model with $K^- = 2$ components for the alpha distribution of negative-alpha funds and $K^+ = 1$ component for the alpha distribution of positive-alpha funds. This is the model with the highest posterior probability according to our baseline estimation, which incorporates model specification uncertainty (see Table 3 in the paper). That is, viewing our analysis as a model selection analysis, this is essentially the model it selects. As we mention in the paper, for computational convenience and/or tractability, we use this model in the construction of portfolios in Section 6.1, in the analysis in Section 7.1 of the distribution of skill by fund investment objective, in the analysis in Section 7.2 of the evolution of skill over time, and in the robustness check in Section 8.2 in which we allow for a full correlation matrix between all model parameters, so its results can be useful as a benchmark for comparison.

In Table C.37, we present results on the posterior distribution of the population proportions of funds with zero, negative, and positive alpha, in Table C.38 we present the percentiles of the estimated densities for alpha and the factor loadings, and in Figure C.14 we present the estimated densities for alpha and the factor loadings and we compare them with those from the baseline model that incorporates model specification uncertainty (i.e., K^-, K^+ are estimated). In short, we find the proportions of funds with zero, negative, and positive alpha to be 14%, 71% and 15%, respectively, which are not very far from the ones from our baseline estimation; and the estimated distributions of factor loadings are almost identical, and the estimated distribution of alpha is quite similar with those from the baseline model.

Table C.37: Proportions of Fund Types — $K^- = 2$, $K^+ = 1$ Model

Results on the posterior distributions of the population proportions of funds with zero, negative, and positive alpha, estimated with returns net of expenses using the model with two negative components and one positive component for the distribution of alpha. The 95% HPDI is the smallest interval such that the posterior probability that a parameter lies in it is 0.95. NSE stands for autocorrelation-adjusted numerical standard errors for the posterior mean estimate of each parameter.

	Mean	**Median**	**Std.Dev.**	95% **HPDI**	**NSE**
π^0	0.14	0.12	0.09	[0.01 , 0.33]	0.00
π^-	0.71	0.72	0.06	[0.60 , 0.83]	0.00
π^+	0.15	0.15	0.05	[0.05 , 0.26]	0.00

Table C.38: Percentiles of Estimated Distributions — $K^{-} = 2, = K^{+} = 1$ Model

Percentiles of estimated population distributions of annualized alpha (expressed as a percent) and factor loadings, for the model with two negative components and one positive component for the alpha distribution, estimated with returns net of expenses.

	Percentiles														
	0.5th	**1st**	**5th**	**10th**	**20th**	**30th**	**40th**	**50th**	**60th**	**70th**	**80th**	**90th**	**95th**	**99th**	**99.5th**
α	-6.87	-4.54	-2.39	-1.90	-1.47	-1.21	-1.01	-0.83	-0.63	-0.18	0.00	0.29	0.78	2.39	3.37
β_{M}	0.40	0.46	0.60	0.68	0.77	0.84	0.90	0.95	1.00	1.06	1.13	1.22	1.30	1.45	1.50
β_{SMB}	-0.60	-0.52	-0.31	-0.20	-0.07	0.03	0.11	0.19	0.27	0.35	0.45	0.58	0.69	0.90	0.97
β_{HML}	-0.85	-0.77	-0.54	-0.41	-0.27	-0.16	-0.07	0.02	0.10	0.20	0.30	0.45	0.58	0.81	0.89
β_{UMD}	-0.26	-0.23	-0.16	-0.12	-0.08	-0.04	-0.02	0.01	0.04	0.06	0.10	0.14	0.18	0.25	0.28

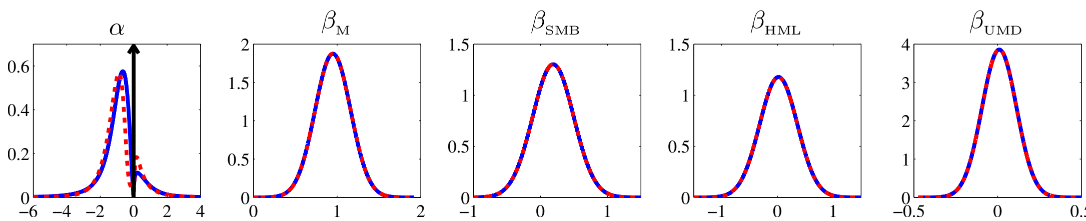


Figure C.14: Comparison of the estimated densities of annualized alpha (expressed as a percent) and factor loadings under two models: blue solid lines plot densities from our baseline model presented in Section 2, red dashed lines plot densities from the $K^{-} = 2, K^{+} = 1$ model with two negative components and one positive component for the distribution of alpha. The black vertical arrow at zero alpha represents a point mass for both estimated distributions; the point mass has probability 0.09 for the baseline model, and probability 0.14 for the $K^{-} = 2, K^{+} = 1$ model. Both models are estimated using returns net of expenses for 3,497 funds.

References

Brys, G, M Hubert, and A Struyf. 2006. "Robust measures of tail weight." *Computational Statistics and Data Analysis*, 50: 733–759.

Celeux, G. 1998. "Bayesian inference for mixture: The label switching problem." In *COMPSTAT*, ed. Roger Payne and Peter Green, 227–232. Physica-Verlag HD.

Geweke, J, and G Zhou. 1996. "Measuring the pricing error of the arbitrage pricing theory." *Review of Financial Studies*, 9(2): 557–587.

Groeneveld, R. A, and G Meeden. 1984. "Measure of skewness and kurtosis." *Journal of Finance*, 33(4): 391–399.

Jasra, A, C. C Holmes, and D. A Stephens. 2005. "Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling." *Statistical Science*, 20(1): 50–67.

Stephens, M. 1997. "Bayesian methods for mixtures of normal distributions." PhD diss. University of Oxford.