A STOCHASTIC-DIFFERENCE-EQUATION MODEL FOR HEDGE-FUND RELATIVE RETURNS

by

Emanuel Derman, Kun Soo Park, and Ward Whitt

Department of Industrial Engineering and Operations Research
Columbia University, New York, NY 10027-6699
{ed2114, kp2143, ww2040}@columbia.edu

April 2008; Revisions April 2009; July 21, 2009

Abstract

We propose a stochastic difference equation of the form $X_n = A_n X_{n-1} + B_n$ to model the annual returns X_n of a hedge fund relative to other funds in the same strategy group in year n. We fit this model to data from the TASS database over the period 2000 to 2005. We let $\{A_n\}$ and $\{B_n\}$ be independent sequences of independent and identically distributed random variables, allowing general distributions, with A_n and B_n independent of X_{n-1}, where $E[B_n] = 0$. This model is appealing because it can involve relatively few parameters, can be analyzed, and can be fit to the limited and somewhat unreliable data reasonably well. The key model parameters are the year-to-year persistence factor $\gamma \equiv E[A_n]$ and the noise variance $\sigma_b^2 \equiv Var(B_n)$. The model was chosen primarily to capture the observed persistence, which ranges from 0.11 to 0.49 across eleven different hedge-fund strategies, according to regression analysis. The constant-persistence normal-noise special case with $A_n = \gamma$ and B_n (and thus X_n) normal provides a good fit for some strategies, but not for others, largely because in those other cases the observed relative-return distribution has a heavy tail. We show that the heavy-tail case can also be successfully modelled within the same general framework. The model is evaluated by comparing model predictions with observed values of (i) the relative-return distribution, (ii) the lag-1 auto-correlation and (iii) the hitting probabilities of high and low thresholds within the five-year period.

1. Introduction

Despite the abundance of stochastic models for stocks, commodities and market indices, relatively few stochastic models have been developed for hedge funds. That is not entirely surprising since hedge funds are not too transparent; there are only a few sources of data, with infrequent voluntary reporting. We contribute by developing a stochastic-process model of the relative annual returns of a hedge fund, exploiting data from the Tremont Advisory Shareholders Services (TASS) hedge-fund database for the period 2000-2005.

1.1. Relative Annual Returns Within the Fund Strategy

The TASS database archives monthly returns and the managed asset value for each hedge fund. In addition, TASS also archives various fund-specific data, such as the strategy of the fund. The eleven strategies and the sample size for each are given in the first and second columns of Table 1; we will explain the rest of Table 1 later. (The appendixes of Hasanhodzic and Lo (2007) and Chan et al. (2006) describe the hedge-fund strategies.)

Table 1: Estimated persistence γ and auto-correlation ρ for the eleven strategies.

Strategy	Sample Size	γ from regression[1]	γ from ratio of exp. returns[2]	ρ auto-correlation[3]
Convertible	238	0.44±0.10	0.39	0.49+0.09/-0.11
Dedicated Short	29	0.49±0.38	0.44	0.16+0.25/-0.35
Emerging Market	315	0.36±0.10	0.36	0.32+0.09/-0.10
Equity Marcro	268	0.09±0.10	0.12	0.12±0.12
Event Driven	533	0.24±0.08	0.16	0.13±0.08
Fixed Income	193	0.29±0.14	0.38	0.37+0.12/-0.14
Fund of Fund	986	0.33±0.05	0.31	0.31+0.05/-0.06
Global Macro	166	0.13±0.15	0.14	0.06±0.15
Long-short Equity	1658	0.15±0.04	0.11	0.07±0.05
Managed Future	235	0.22±0.13	0.17	0.21+0.12/-0.13
Other	167	0.41±0.15	0.38	0.39+0.12/-0.13

1. 95% confidence interval for the regression coefficient
2. Ratio of expected relative returns from the previous to current year for pairs of two successive years whose return values are both above the average.
3. confidence interval of correlation coefficient from 95% confidence interval of Fisher-Z statistic in (4.1).

In order to highlight differences in hedge fund performance within its strategy and to approach a stationary environment, we focus on the *relative* annual returns. We use geometric compounding to convert the twelve reported monthly returns into one annual return, i.e.,

$$r_{annual} = (1 + r_1) \cdot (1 + r_2) \cdots (1 + r_{12}) - 1 \ .$$

We then obtain the relative annual returns by subtracting the average for the strategy for that year.

We think of the TASS relative return data as being observations from a *stationary* discrete-time stochastic process $\{X_n : n \geq 0\}$, with X_n representing the relative annual return from year n. Assuming that the process $\{X_n\}$ is indeed approximately stationary (which is made more plausible by our focus on relative returns), we combine all the data for each category to estimate the distribution of the single-year relative return for each strategy. For each strategy, we seek a stochastic-process model that matches both the observed single-year relative-return distribution and the observed dependence structure. To have a model useful for prediction, it is desirable that the stochastic process be a Markov process, with a state that is as simple as possible.

Since we focus on relative returns, the relative-return distribution necessarily has mean 0, so a key parameter of the distribution to be matched is the *variance* $\sigma^2 \equiv Var(X_n)$, but we also want to match the entire distribution as much as possible. Indeed, in some cases we find that the return distribution has a heavy tail, consistent with an infinite variance.

For a stationary stochastic process, a key parameter describing the dependence structure is the *autocorrelation* $\rho \equiv Cor(X_n, X_{n+1}) \equiv Cov(X_n, X_{n+1})/\sigma^2$. Estimates of the auto-correlation ρ appear in the final column of Table 1. However, we also want to match the full time-dependent behavior of the stochastic process as much as possible. To partially test the time-dependent behavior beyond the auto-correlation ρ, we evaluate the probability that the relative returns will ever hit specified levels within a five-year period. That also illustrates how the model can be applied.

1.2. Persistence of Hedge-Fund Returns

Our modelling approach is motivated by our observation of persistence in the relative returns. Broadly, persistence in hedge-fund returns is a tendency for a fund which generates relatively high (or low) returns in a period to continue generating relatively high (or low) returns again in the next period.

Persistence has been studied quite extensively within the hedge-fund literature, but it remains a highly controversial topic. A consensus has not yet been reached on the degree of persistence in hedge-fund returns, or even whether it exists at all. Indeed, some studies did not find significant persistence; e.g., Brown et al. (1999), Capocci and Hüber (2004), and Boyson and Cooper (2004). However, several studies have found evidence of strictly positive persistence, depending on the time period measured; Agarwal and Naik (2000) found significant persistence for quarterly returns, while Edwards and Caglayan (2001) found significant persistence over one to two years, and Jagannathan et al. (2006) found significant persistence over three years of returns. For hedge-fund indexes, Amenc et al. (2003) found statistically meaningful persistence for most of the strategies.

In this paper, we consider persistence in the (relative) returns. It is important to note that others have looked for persistence in different ways; e.g., Jagannathan et al. (2006) is about alpha persistence. We say that there is a *persistence factor* of γ if for every 1 percentage point the fund makes above the average in the current year, it is expected to earn γ percentage point above the average in the next year. For the stochastic process $\{X_n : n \geq 0\}$, the persistence implies that we should have the following relation between the conditional expected relative return at the end of the current year, given the previous relative return, and the previous relative return itself:

$$E[X_n|X_{n-1}] = \gamma X_{n-1} \tag{1.1}$$

for all n and all values of X_{n-1}. We estimate the persistence factors by performing a regression analysis. In particular, we combine the relative-return data for all pairs (X_n, X_{n+1}) and perform a standard linear regression. Our estimated persistence factors for the eleven hedge-fund strategies ranged from 0.11 to 0.49; estimates by two different methods appear in the third and fourth columns of Table 1. The 95% confidence intervals show that positive persistence is confirmed statistically for all but two strategies; See §4 for more on our data-selection and analysis procedure.

In our statistical analysis we do find strong evidence for persistence, but we hasten to admit that the issue remains controversial. The voluntary reporting has led to questions about the reliability of the data. As Getsmansky et al. (2004) pointed out, under the voluntary reporting system, a hedge fund manager may choose to report smoothed returns intentionally, which causes serial correlation of returns. Possible biases in reported hedge-fund returns are discussed by Fung and Hsieh (2000) and Boyson and Cooper (2004). As we explain in §4.1, in our data selection procedure, we attempt to reduce the bias, but the TASS data should be regarded as somewhat unreliable. We emphasize that *our primary goal is not to make a case for persistence, but instead is to show how persistence can be exploited, if it is there, in order to create a flexible and tractable stochastic-process model of hedge-fund returns.* Our approach should also have other useful applications, where persistence may exist. We introduce the model in the next section. In subsequent sections, we elaborate on the appealing mathematical structure of the model, we describe our data analysis methods and results, and we show that the model provides a flexible framework for fitting.

2. The Proposed Stochastic Difference Equation Model

In order to capture the observed persistence in the performance of hedge-fund relative returns, we first propose the simple *stochastic difference equation* (SDE)

$$X_n = \gamma X_{n-1} + B_n, \quad n \geq 1, \tag{2.1}$$

where γ is a constant with $0 < \gamma < 1$, B_n is independent of X_{n-1} and $\{B_n : n \geq 1\}$ is a sequence of independent and identically distributed (i.i.d.) random variables, each distributed as $N(0, \sigma_b^2)$, where $N(a, b)$ denotes a normally distributed random variable with mean a and variance b.

The SDE in (2.1) is a linear, recursive Markov process; it is also a first-order autoregressive process. Moreover, the SDE in (2.1) is a natural discrete-time analog of the familiar continuous-time *stochastic differential equation*

$$dX(t) = -\nu X(t) + \sigma_c dB(t), \tag{2.2}$$

where $\{B(t) : t \geq 0\}$ is a standard Brownian motion, commonly used in finance, as can be seen by subtracting X_{n-1} from both sides in (2.1) to get

$$X_n - X_{n-1} = -(1 - \gamma) X_{n-1} + B_n, \quad n \geq 1. \tag{2.3}$$

We choose the discrete-time process in (2.1) instead of the continuous-time process in (2.2) because hedge-fund returns are reported much less frequently than stock prices.

The initial SDE model in (2.1) is very appealing because, first, it clearly matches the persistence as specified in (1.1) with the same parameter γ and, second, one need to choose only one remaining model parameter σ_b^2 in order to match the steady-state variance σ^2. That is easily done, because for the model (2.1) it turns out that one variance must be a constant multiple of the other:

$$\sigma^2 = \frac{\sigma_b^2}{1 - \gamma^2}, \tag{2.4}$$

Moreover, as a consequence of (2.1), the distribution of X_n (assuming stationarity) must itself be normal, distributed as $N(0, \sigma_b^2/(1 - \gamma^2))$. Both these conclusions are demonstrated in §3.

This is a beautiful simple story when it works. Clearly, it works (from this preliminary checking) if indeed the two variances are related by (2.4) and the steady-state distribution of the relative returns is approximately normal. Fortunately, for some hedge fund strategies, we find that both conditions are satisfied reasonably well. Moreover, we can go beyond the distribution of relative annual returns to check the time-dependent behavior. In §3 we show that in steady-state, the SDE in (2.1) necessarily has autocorrelation equal to the persistence:

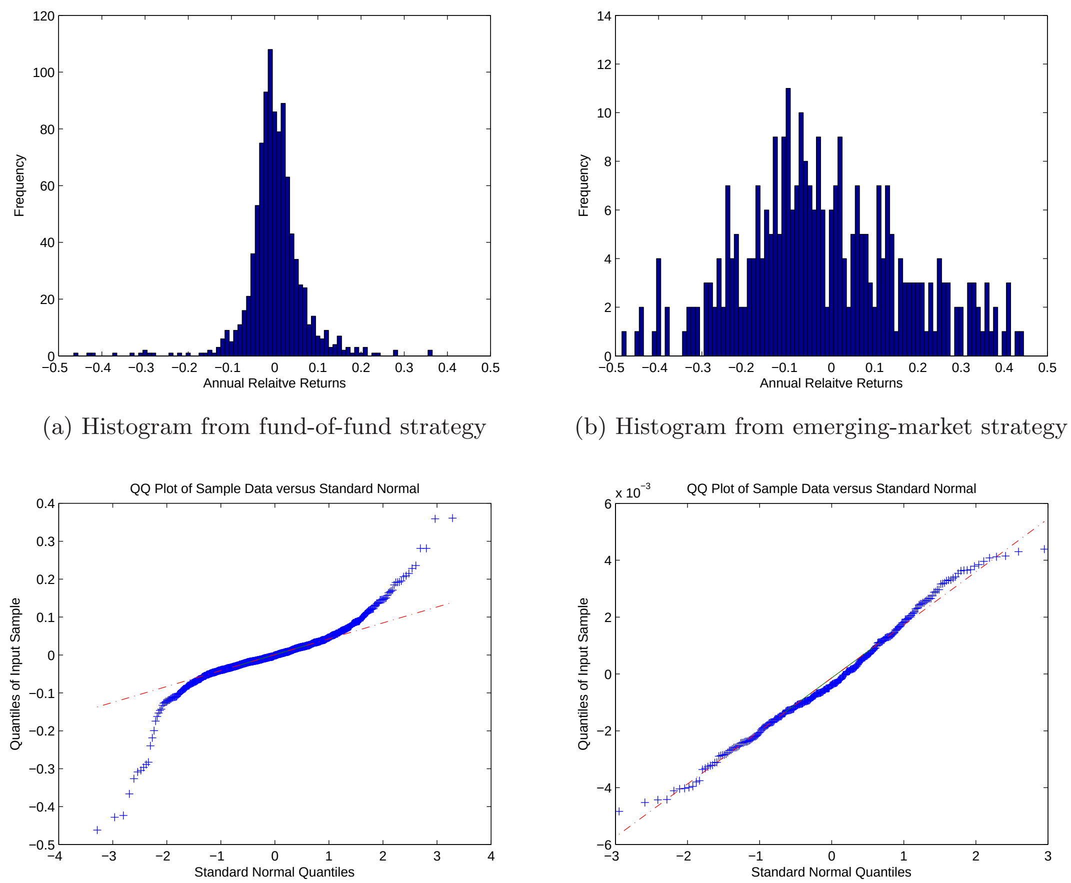
$$\text{autocorrelation} \equiv \rho = \gamma \equiv \text{persistence factor}. \tag{2.5}$$

This special relation in (2.5) turns out to match the TASS data remarkably well, given the limited data, as shown in Table 1, which displays estimates of both ρ and γ.

We find that the simple SDE model in (2.1) provides a remarkably good fit for some of the hedge-fund strategies, e.g., for the emerging-market strategy. However, it does not provide a good fit for all strategies; e.g., for the fund-of-fund and event-driven strategies, largely because for those other strategies the empirical distribution of the relative annual returns is quite far from normal, having a heavy tail. Figure 1 substantiates this claim, showing the histogram and Q-Q plots of the relative annual returns of hedge funds within the fund-of-fund and emerging-market strategies. (The units are chosen so that a relative annual return of 0.10 corresponds to 10 *percentage points* above average.)

We selected these two strategies for three reasons: (i) because these strategies have relatively large numbers of observations (ii) because they have relatively high persistence factors and (iii) because the return distributions exhibit very different tail behavior. Figure 1 shows that the distributions for those two strategies differ significantly. The Q-Q plots in Figure 1 (c) and (d) show that the distribution of the relative returns for the emerging-market strategy is close to normal, whereas for the fund-of-fund strategy it is not.

The fund-of-fund strategy is somewhat special, involving investments in other strategies. It might be considered surprising that the relative returns from the fund-of-fund strategy are less normal, since they tend to be more diversified, but correlations among the returns from different strategies may possibly explain this phenomenon. Understanding the observed tail behavior of different strategies remains a problem for future research. We do emphasize that heavy tails are also observed in other strategies, such as the event-driven strategy, as we show in Appendix §J. Corresponding figures for other strategies appear in Appendix §C.



(a) Histogram from fund-of-fund strategy

(b) Histogram from emerging-market strategy

(c) Q-Q plot for fund-of-fund strategy

(d) Q-Q plot for emerging-market strategy

Figure 1: (a)(b) Histograms of 986 relative returns within the fund-of-fund strategy and 315 relative returns within the emerging-market strategy from the TASS database. (A relative return of 0.15 means 15 percentage points above the average.) (c)(d) Q-Q plots comparing the model to the normal distribution.

Just as for performance persistence, the distribution and other statistical properties of hedge-fund returns are not yet well understood, despite the importance (Lhabitant, 2004; Kassberger and Kiesel, 2006; Tran, 2006). Several authors have reported that the normal distribution may not approximate hedge-funds returns well, primarily because of heavy tails (Lo, 2001; Lhabitant, 2004; Tran, 2006; Geman and Kharoubi, 2003; Eling and Schuhmacher, 2007). It should thus not be surprising that we find that the relative returns are reasonably well approximated by the normal distribution for some strategies, but not for all strategies. Consistent with our analysis, Amo et al. (2007) pointed out that autocorrelation, high-peak, and heavy-tail may be observed from the distributions of hedge-fund returns.

Kassberger and Kiesel (2006) studied the distribution of daily hedge-fund indices within each strategy. Based on the daily indices data from March 2003 to June 2006, they show that the distributions of indices have heavy-tails by Q-Q plots. They claimed that a Normal Inverse Gaussian (NIG) distribution fits the distribution of indices well, since it may have heavy-tail and skewness depending on parameter values.

2.1. A More General SDE Model

The non-normal distribution shown in Figure 1 (c), and in other return distributions, leads us to look for other models. Fortunately, we find that a natural generalization of the simple SDE in (2.1) provides a robust and tractable model for capturing different behavior observed in the TASS data. As a generalization of the simple SDE in (2.1), we propose the SDE

$$X_n = A_n X_{n-1} + B_n \ , \ n \ \geq \ 1 \ , \tag{2.6}$$

where A_n and B_n are independent of X_{n-1} and $\{A_n : n \geq 1\}$ and $\{B_n : n \geq 1\}$ are independent sequences of i.i.d. random variables with general distributions, satisfying

$$E[A_n] = \gamma \quad \text{for} \quad 0 < \gamma < 1, \quad \text{and} \quad E[B_n] = 0. \tag{2.7}$$

In going from (2.1) to (2.6), we have replaced the constant persistence factor γ by the random persistence A_n, but the moment conditions in (2.7) imply that the basic persistence relation (1.1) still holds. Moreover, the autocorrelation still satisfies (2.5), as we show in §3. By allowing A_n and B_n to have general distributions, we have produced a much more flexible class of models. Fortunately, this class of models is also remarkably tractable, as was shown by Vervaat (1979), where many additional references can be found.

We classify the specific models we consider by the assumptions we make about the distributions of A_n and B_n. When $P(A_n = \gamma) = 1$, we have a *constant-persistence* model; when A_n has a non-degenerate distribution, we have a *stochastic-persistence* model. When B_n is normally distributed, we have a *normal-noise* model. To capture the heavier tails we see in the data, we also consider as distributions for B_n the Student-t distribution, a mixture of two distributions, an empirical distribution and a stable distribution.

2.2. The Constant-Persistence Stable-Noise Model

We highlight the constant-persistence stable-noise model, because it is now common to use stable distributions to represent heavy-tailed distributions, building on early work by Mandelbrot (1963), Fama (1965) and others; see Embrechts et al. (1997), Samorodnitsky and Taqqu (1994) and §4.5 of Whitt (2002) for general background. Indeed, there is now a vast literature on heavy tails in financial data; e.g., see Lux (1996), Rachev and Mittnik (2000), Cont (2001) and Gabaix et al. (2007).

A random variable Y is said to have a (strictly) *stable law* if, for any positive numbers a_1 and a_2, there is a positive number $c \equiv c(a_1, a_2)$ such that

$$a_1 Y_1 + a_2 Y_2 \stackrel{\mathrm{d}}{=} cY, \tag{2.8}$$

where Y_1 and Y_2 are independent copies of Y and $\stackrel{\mathrm{d}}{=}$ means equality in distribution. It turns out that the constant c must be related to the constants a_1 and a_2 by

$$a_1^\alpha + a_2^\alpha = c^\alpha \tag{2.9}$$

for some constant α with $0 < \alpha \le 2$, called the *index* of the stable law. A random variable Y_α with stable distribution having index α with $0 < \alpha < 2$ satisfies $P(Y_\alpha > x)/x^{-\alpha} \to c_+$ and $P(Y_\alpha < -x)/x^{-\alpha} \to c_-$ as $x \to \infty$ for some positive constants c_+ and c_-. Consequently, $E[|Y_\alpha|^p] < \infty$ for all $p < \alpha$, but $E[|Y_\alpha|^p] = \infty$ for all $p > \alpha$. We will be considering α with $1 < \alpha < 2$, so that our stable distributions will have infinite variance but finite mean, which we take to be zero.

Just as for the normal distribution (which can be regarded as a special stable distribution), the structure of the SDE in (2.1) implies that the stochastic structure of the distribution of B_n is inherited by the distributions of X_n for the constant-persistence models; i.e., the distribution of X_n is again stable with the same index and skewness parameter; that is, we have

$$X_n \stackrel{\mathrm{d}}{=} \left(\frac{1}{1-\gamma^\alpha}\right)^{1/\alpha} B_n, \tag{2.10}$$

as we prove in §3. We use this relation (2.10) in what we think are novel ways: *We use* (2.10) *to test both the constant-persistence stable-noise model and the stable index* α (using the persistence factor γ already estimated); see §7.

For the constant-persistence stable-noise model, the SDE in (2.1) also has the continuous-time analog in (2.2), but where now $\{B(t) : t \ge 0\}$ is a non-Gaussian stable Lévy motion, as in Samorodnitsky and Taqqu (1994). More generally, when the random variable B_n has a non-normal distribution, (2.1) has continuous analog (2.2) where $\{B(t) : t \ge 0\}$ is a Lévy process; see

Wolfe (1982). In §5 of Wolfe (1982), he shows how to construct the continuous-time analog from the discrete-time SDE if it is desired. By now, there is a substantial literature on non-standard stochastic differential equations in finance; e.g., see Barndorff-Nielsen and Shephard (2001) and Borland (2002).

We will show that the constant-persistence stable-noise model is remarkably effective for the fund-of-fund strategy. Nevertheless, other versions of the model in (2.6) are worth considering as well, in part because they have finite variance, which allows us to use the observed variance σ^2 to calibrate the model.

2.3. Previous Models of Hedge-Fund Returns

A conventional assumption is that a firm's net asset value evolves in continuous time as a geometric Brownian motion. Following that convention, a log-normal distribution was used to model hedge fund net asset value by Atlan et al. (2006) and the risky investment the hedge fund holds by Hodder and Jackwerth (2007). However, the log-normal assumption is not empirically tested in those papers.

Others have previously used Markov process models to model hedge-fund returns. Hayes (2006) used discrete-time birth-and-death process to calculate the maximum drawdown in hedge-fund returns, and used the autocorrelation condition to calibrate the model. In Derman et al. (2008) we used three-state Markov chain models to estimate the premium from extended hedge-fund lockup. We used the same TASS data to calibrate that model.

Several econometric models have been proposed as well. A seminal paper is Amin and Kat (2003), which sought a trading strategy with cash and a market portfolio such as S&P 500 to *replicate* the distribution of a hedge-fund's returns. If a replicating portfolio can be found, by considering the required initial investment in the replicating portfolio and the hedge-fund management fee, then it may be possible to evaluate whether or not an investment in the hedge fund is justifiable or not. A similar replicating approach is also found in Hasanhodzic and Lo (2007). They tried to replicate hedge-fund returns with six common risk factors such as the S&P 500, US Dollar Indexes, Bond index, etc, by means of linear regression analysis. Chan et al. (2006) is a paper closely related to Hasanhodzic and Lo (2007). However, the purpose of Chan et al. (2006) was somewhat different; they wanted to decompose the risk factors underlying the hedge fund in order to compare the systematic risks of hedge funds to that of other traditional asset classes.

2.4. Applications of the Stochastic Model

As usual, a stochastic-process model allows us to go far beyond a direct examination of historical data to ask various "what if" questions. There are many ways to apply the model to answer questions, which cannot easily be answered from the data directly. We might simply want to know the probability distribution of the relative return for a particular hedge fund over the following year, given all available past data. From the past data, we can observe the most recent relative return, say $X_0 = c$. We would then apply the model in (2.6) to conclude that the relative return next year should be distributed as $A_1 c + B_1$, where A_1 and B_1 are the independent stochastic persistence and noise, respectively, for that hedge-fund strategy, whose distributions can be determined by data fitting, as described in this paper. We could go further and calculate the discounted present value of the return stream over many years; see (3.11) - (3.12).

We might want to invest in that particular hedge fund because we believe that it will be especially well managed. We could use the model to provide a "measurement-based" quantification of what we mean by good management. In particular, we may postulate that a good fund manager improves the fund performance in one or more of three possible ways: increasing the expected persistence $\gamma \equiv E[A_n]$, reducing the standard deviation of the persistence $\sigma_a \equiv \sqrt{Var(A_n)}$, or reducing the standard deviation of the additive noise $\sigma_b \equiv \sqrt{Var(B_n)}$. With the model, we can quantify the impact of such effects. We first fit the model to the data for that hedge-fund strategy in order to obtain random variables A_n and B_n. We then produce new random variables A'_n and B'_n consistent with the postulated consequences of good management. We then calculate future relative returns, both with the original model and with the revised model. In that way, we can estimate the value added by the good management.

We illustrate with a concrete example: Suppose that the relative returns for a specific fund in the last year are $X_0 = c$. We start by quantifying what it mean for a "good" manager to be effective. Suppose that we conclude that the impact of superior management should increase its nominal estimated expected persistence from γ to 1.5γ, reduce the estimated standard deviation of the persistence from σ_a to $0.8\sigma_a$, and reduce the estimated standard deviation of the noise from from σ_b to $0.5\sigma_b$. As a numerical example, we choose the beta-persistence t-noise model developed in §6.2 for the fund-of-fund strategy (which has parameter values $\gamma = 0.33, \sigma_a = 0.0381, \sigma_b = 0.0642$, and $\alpha = 50, \beta = 101.52$). We then choose new random variable A'_n and B'_n with $\gamma' = 1.5\gamma, \sigma'_a = 0.8\sigma_a, \sigma'_b = 0.5\sigma_b$ and define X'_n based on the new parameter values. Then, algebraic manipulation yields $\alpha' = 84.75$ and $\beta' = 86.46$. It is then immediate that $\mathbb{E}[X_1'|X_0' = c] - \mathbb{E}[X_1|X_0 = c] = (\gamma' - \gamma)c = 0.1650c$, $Var(X_1|X_0 = c) - Var(X_1'|X_0' = c) = c^2(0.36\sigma_a^2) + 0.75\sigma_b^2 = 0.0005c^2 + 0.0031$. We have thus shown how the model can be applied to quantify the impact of good management.

3. Background on the General SDE

The behavior of the general SDE in (2.6) is well described in Vervaat (1979); we will be stating implications from the general results there. We will be considering the standard (good) case in which the expectation $E[\log(A_n)]$ is well defined (at least one of the positive part or the negative part has finite expectation) and the following (minimal) logarithmic-moment conditions are satisfied:

$$-\infty \leq E[\log(A_n)] < 0 \quad \text{and} \quad E[\log^+(B_n)|] < \infty \ , \tag{3.1}$$

where $\log^+(x) \equiv \max\{0, \log(x)\}$. Note that $\log(A_n) = -\infty$ occurs if $A_n = 0$, which is a possibility we want to allow. That corresponds to no persistence at all.

Under condition (3.1), Vervaat shows that we have convergence in distribution $X_n \Rightarrow X_\infty$ as $n \to \infty$, where the distribution of X_∞ is independent of the initial conditions and is characterized as the unique solution to the *stochastic fixed-point equation*

$$X_\infty \stackrel{\mathrm{d}}{=} A_n X_\infty + B_n, \tag{3.2}$$

where the random vector (A_n, B_n) is independent of X_∞ on the right. There is thus a unique stationary version of the process $\{X_n : n \geq 0\}$, obtained by letting the initial value X_0 be distributed as X_∞, while being independent of A_1 and B_1. With our notion of persistence in mind, it is natural to go beyond condition (3.1) and assume in addition that $P(0 \leq A_n < 1) = 1$. That will immediately imply extra moment conditions we make for A_n below. But that extra assumption is actually not required.

Moreover, we actually do not need to assume that A_n is independent of B_n, as we have done, but the strong results in Vervaat (1979) do require that the sequence $\{(A_n, B_n)\}$ be a sequence of i.i.d. random vectors. It is worth noting, though, that the general model in (2.6) has been further generalized beyond Vervaat (1979). First, Brandt (1986) established results for the case in which independence for the sequence $\{(A_n, B_n) : n \geq 1\}$ is dropped; he assumes only that it is a stationary sequence. Next Horst (2001) considers the time-dependent version, allowing the distribution of (A_n, B_n) to depend on n. Finally, Horst (2003) embeds the model in a game-theoretic setting, letting the values of (A_n, B_n) depend on the strategic decisions of multiple players. These extensions are significantly less tractable than (2.6) here, but they open the way to interesting new applications.

Given (3.1), we can also characterize the distribution of X_∞ via an *infinite-series representation*

$$X_\infty \stackrel{\mathrm{d}}{=} \sum_{k=1}^{\infty} A_1 A_2 \cdot A_{k-1} B_k, \tag{3.3}$$

where the series on the right converges with probability 1 (w.p.1). It is thus easy to approximately generate samples from the distribution of X_∞ by considering a truncated version of the series. If $|A_n|$ tends to be relatively small, as with our persistence estimates, then relatively few terms are required.

Moreover, it is easy to apply the stochastic fixed-point equation (3.2) in order to deduce that the steady-state value X_∞ is distributed simply as a constant multiple of B_n, as given in (2.10), when B_n has a stable law. We have the following elementary proposition:

Proposition 1. *For the simple SDE in* (2.1)*, if B_n has a stable law with index α, i.e., if* (2.8) *and* (2.9) *hold for $0 < \alpha \le 2$ (with $\alpha = 2$ being the case of a normal distribution), then*

$$X_\infty \stackrel{\mathrm{d}}{=} \left(\frac{1}{1-\gamma^\alpha}\right)^{1/\alpha} B_n; \tag{3.4}$$

i.e., (2.10) *is valid.*

Proof. First, since we are considering the simple SDE in (2.1), we have $A_n \equiv \gamma$. Since the distribution of X_∞ is the unique solution to the stochastic fixed-point equation (3.2), it suffices to show that $X_\infty \equiv cB_n$ satisfies equation (3.2) for some constant c, i.e., it suffices to show that

$$cB \stackrel{\mathrm{d}}{=} \gamma(cB) + B_n, \tag{3.5}$$

where B and B_n are independent random variables with the common distribution of B_n. Since B_n has a stable law with index α, we can apply (2.9) to get the equation $c^\alpha = (\gamma c)^\alpha + 1^\alpha$, which has the desired value for c as its unique solution. ∎

Important moment properties of the SDE in (2.6) are given in §5 of Vervaat (1979), but these require extra conditions on the moments of the model elements. Prior to the moment conditions made in (2.7), in addition to the conditions above, we assume the technical regularity conditions

$$E[|A_n|] < 1, \quad E[|B_n|] < \infty \quad \text{and} \quad E[|X_0|] < \infty. \tag{3.6}$$

Under these conditions, it follows that $E[|X_\infty|] < \infty$ and $E[|X_n|] < \infty$ for all n. By 5.2.1 of Vervaat (1979), if (3.6) holds, then in general

$$E[X_\infty] = \frac{E[B_n]}{1 - E[A_n]} \quad \text{and} \quad E[X_n] \to E[X_\infty] \quad \text{as} \quad n \to \infty. \tag{3.7}$$

Since we assume condition (2.7) in addition to conditions (3.1) and (3.6), we can conclude that $E[X_\infty] = 0$ and $E[X_n|] \to 0$ as $n \to \infty$.

We will not want to go beyond these first-moment conditions for B_n in (3.6) when we consider stable noise, because then B_n will have infinite variance. However, for the finite-variance case, we also assume that $E[A_n^2] < 1$, and $E[B_n^2] < \infty$ and $E[X_0^2] < \infty$. Then 5.2.2 of Vervaat (1979) provides the following important expression for the variance of the steady-state distribution:

$$\sigma^2 \equiv Var(X_\infty) = \frac{E[B_n^2]}{1 - E[A_n^2]} = \frac{Var(B_n)}{1 - E[A_n^2]} \equiv \frac{\sigma_b^2}{1 - \sigma_a^2 - \gamma^2}, \tag{3.8}$$

where we have introduced the new notation $\sigma_a^2 \equiv Var(A_n)$ and used the assumption that $E[A_n] = \gamma$ in the final expression. Paralleling (3.7), it also implies the convergence $Var(X_n) \to Var(X_\infty)$ as $n \to \infty$. When $P(A_n = \gamma) = 1$, then (3.8) reduces to (2.4).

We now exploit the variance limit above under the the moment conditions in order to characterize the auto-correlation of the stationary version of the stochastic process $\{X_n\}$. We will characterize the asymptotic behavior, with a non-stationary initial condition. For that purpose, assume that $E[X_0] = 0$ along with the moment conditions, so that we have $E[X_n] = 0$ for all n. Then the time-dependent auto-covariance is simply

$$Cov(X_{n+1}, X_n) = E[X_{n+1}X_n] = \gamma E[X_n^2] = \gamma Var(X_n), \tag{3.9}$$

which implies that the associated auto-correlations satisfy

$$\rho_n \equiv Cor(X_{n+1}, X_n) = \frac{Cov(X_{n+1}, X_n)}{\sqrt{Var(X_{n+1})Var(X_n)}} = \gamma\sqrt{\frac{Var(X_n)}{Var(X_{n+1})}} \to \gamma \quad \text{as} \quad n \to \infty. \tag{3.10}$$

We have thus shown for the general SDE model in (2.6) that $\rho = \gamma$, where $\rho \equiv \rho_\infty$ is the auto-correlation for the stationary version of $\{X_n\}$, obtained by letting X_0 be distributed as X_∞, just as claimed in (2.5) for the simple SDE in (2.1).

In our hedge-fund context it is natural to be interested in the discounted present value of a return stream. It is thus convenient that the discounting can be incorporated into our current framework. First, if we postulate a constant rate of interest r compounded continuously, so that the annual discounting factor is e^{-r}, then the (random) present value of the entire relative-return stream and its conditional expected value are

$$V(r) = \sum_{n=1}^{\infty} e^{-nr} X_n \quad \text{and} \quad E[V(r)|X_0] = \frac{X_0}{1 - \gamma e^{-r}}. \tag{3.11}$$

More generally, we may have random annual interest rate R_n in year n, so that the present value is

$$V = \sum_{n=1}^{\infty} (\prod_{k=1}^{n} R_k) X_n \ . \tag{3.12}$$

Given our model with specified distributions for A_n and B_n, a well-defined stochastic process $\{R_n : n \geq 1\}$, which could be (but need not be) a sequence of i.i.d. random variables with specified distribution, and the initial value X_0, we can easily determine the distribution of V by simulation. We can first generate a segment of the process $\{X_n\}$ recursively, and then do the same for the sum in (3.12). Given typical discounting processes $\{R_n\}$, the series will converge quickly, so that truncated versions will yield good approximations.

4. Empirical Observations from the TASS Data

4.1. Hedge-Fund Data Selection and Analysis

We first explain how we try to remove biases in the TASS data. We then describe the regression procedure to estimate the persistence factor.

TASS differentiates between the date the fund starts reporting and the date the fund starts operating. When a fund starts reporting returns after operating for several months or years, the fund may simultaneously report several monthly returns at the time its first return is reported. It is then possible for the fund manager to report only good returns. Otherwise, if the returns are bad, the manager may choose not to report them. This phenomenon creates the so-called *backfill bias*, since the backfilled returns tend to be higher due to the missing bad returns. Fung and Hsieh (2000) calculate that the difference from actual returns and reported returns is about 3.6% per year from this reason. In order to at least partially address this problem, we consider monthly returns only after the fund's first reporting date. Similarly, if a fund's monthly returns are reported less than six times a year, we exclude these data, due to the possibility of hiding bad returns.

We also considered the asset value managed by a fund. We treat all funds equally, without regard to the asset value, so we present a "fund view" as opposed to a "dollar view." However, we did start by removing very small funds from our sample. Specifically, we consider monthly returns only if the fund's asset value managed has reached 25 million dollars at least once, at which point we assume that the fund becomes mature, so that it can produce relatively stable returns. A similar data selection strategy was used by Boyson and Cooper (2004). To better understand this issue, we computed the average asset value managed for each fund and plotted the distribution of the values; it is shown in the Appendix §B. As might be expected, the distribution of the sizes has a heavy tail.

After selecting the monthly returns based on the above criteria, we proceeded to estimate the persistence factor by regression. In particular, we made pairs of two successive annual returns for each hedge fund from 2000 to 2005. Thus, there are possibly five pairs of annual returns of a fund, if it does not cease reporting during that period. (Thus, our sample sizes in Table 1 are the number of pairs in the strategy.) The monthly returns are annualized to measure the yearly persistence of returns, using geometric compounding. We next calculate relative annual returns for each fund by subtracting the average annual returns of the funds in the same strategy. The relative returns for two successive years are then coupled as a pair to estimate yearly persistence factor. In order to make meaningful pairs of relative returns for two successive years, the averages of annual returns for the first year and each strategy of the funds are calculated first. When calculating the average annual returns and the associated relative returns for the next consecutive year, we only include returns from the funds which existed and were not dropped from the TASS database during the previous year. Thus, the average annual return for any given year depends on whether that year is treated as an initial year or a next year. They are not necessarily equal, since some funds may start reporting to TASS in the next consecutive year. In this way, we finally construct pairs of two consecutive relative returns from 2000 to 2005 for each strategy of the fund.

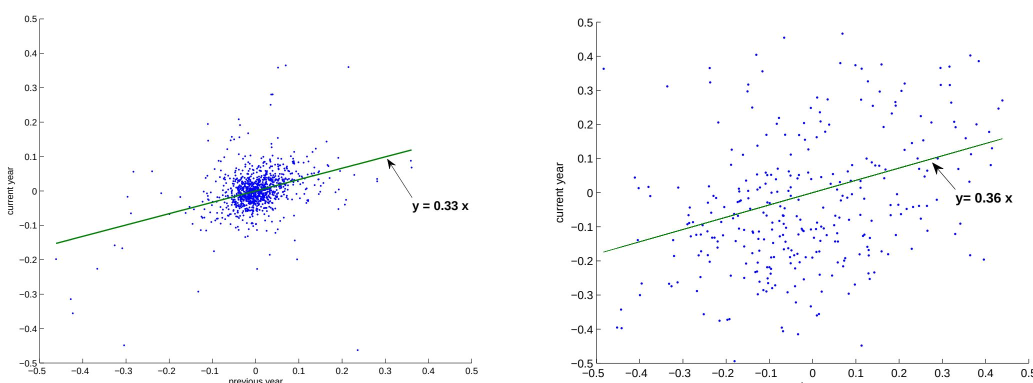
Before conducting regression, we also exclude pairs of returns with extreme values, depending on the distribution of the pairs of returns for each strategy category. Even one or two outliers can seriously affect the regression, especially if we do not have a large number of observations. Specifically, we excluded pairs of relative returns when one absolute relative return exceeds $\pm$ 30 % for the fixed-income and equity-macro and $\pm$ 40% for the convertible, dedicated-short-bias, and global-macro strategies. We also excludes pairs of relative returns exceeding $\pm$ 50% for the emerging-market, event-driven, fund-of-fund, long/short-equity, managed-future, and others strategies. (These percentages were chosen to be appropriate by visual inspection. The percentages are roughly equivalent relative to the overall standard deviation of the return distribution for the strategy.) On the positive side, this data-selection procedure helps us avoid data errors. On the other hand, this data-selection procedure might lead us to underestimate heavy tails. As a consequence, our heavy-tail findings should be even more convincing.

We conducted a linear auto-regression analysis with pairs of two successive years of annual relative returns. The coefficient from this linear regression, i.e., the least square fit is the calculated persistence. The regression analysis results in very low intercept for all strategy category. Thus, we finally conduct a auto-regression without intercept and consider only the coefficient term. The results are shown in the third column of Table 1.

An alternative way to estimate the persistence factor is to consider the ratio of the next-year average returns to the current-year average return, restricting attention to the returns that are positive in the current year. The fourth column of Table 1 shows the ratio of two successive average returns restricting attention to the returns that are positive and negative in the current year, respectively. We observe that these alternative persistence estimates tend to be similar to the regression estimates.

4.2. Persistence of Relative Returns

We started by constructing scatter plots of the relative returns for each hedge-fund strategy, using all pairs (X_n, X_{n+1}), and performed auto-regression analysis in that setting in order to estimate the persistence factor, which thus becomes the the regression coefficient. Figure 2 shows the scatter plots of the relative annual returns for the fund-of-fund and emerging-market strategies. A linear



(a) Fund-of-fund strategy

(b) Emerging-market strategy

Figure 2: Scatter plots and auto-regression lines for relative returns from two successive years within (a) the fund-of-fund strategy and (b) the emerging-market strategy.

relationship is not overwhelmingly clear in Figure 2. Nevertheless, we do observe more pairs of returns in the lower left and higher right sides of the scatter plot, indicating the existence of persistence. We mention that the persistence factor may also be derived in another way. We can also estimate the persistence factor from the ratio of the two successive years' expected relative returns, when those values are both above the average. This directly measures the ratio of current year's expected relative returns to the previous year's expected relative returns, but we have yet to develop the statistical properties of this estimator. The estimated persistence factors by both these methods are given in Table 1.

4.3. Distribution of Relative Returns

We now turn to the distribution of the relative annual returns. As illustrated by Figure 1, we constructed histograms showing the empirical distribution and constructed Q-Q plots to test for normality. As we have indicated before, the emerging-market strategy relative-return distribution seems to be approximately normal, but the fund-of-fund relative-return distribution does not. The distributions and Q-Q plots for the other strategies are given in §C of the Appendix. The Q-Q plots there show that the relative-return distribution for the global-macro strategy also is well approximated by the normal distributions, but all others have significant departures from normality in the tails. We also performed the Lilliefors test in Appendix §C, from which we conclude, statistically, that the relative returns from most of the strategies are not consistent with the normal distribution. (See Lilliefors (1967) for the details of the test.) In order to facilitate visual comparison with the normal distribution, we also plotted histograms from a simulation of i.i.d. normal random variable with the same sample sizes; see, Appendix §D. Finally, we note that the fund-of-fund relative-return distribution has a relatively high peak in the center.

4.4. Autocorrelation of Relative Returns

In §3 we showed that the auto-correlation is equal to the persistence for the general SDE model in (2.6). Thus we want to see if that is true for the TASS data. To examine this issue, we estimate the auto-correlations in the data, using the sample correlation coefficient estimator, denoted by r. In order to estimate the 95% confidence intervals for the auto-correlation correlation, we use the well-known result that the Fisher Z statistic, defined by

$$Z = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right) , \tag{4.1}$$

is approximately normally distributed with mean zero and standard deviation $\sigma_z = 1/\sqrt{n-3}$, where n is the sample size; e.g., see Serfling (1980) or Lin (1989).

From (4.1), we derive the confidence interval of the correlation coefficient ρ from the confidence interval of Z. The confidence interval is not symmetric around the observed sample autocorrelation coefficient r because r is a non-symmetric function of Z in (4.1). The last column in Table 1 summarizes the results. Table 1 shows that the two 95% confidence intervals – for the persistence γ and the auto-correlation ρ – overlap significantly for most strategies. Thus we conclude that γ and ρ coincide with each other and regard this as support for the validity of the SDE model in (2.6). Figure 3 adds by providing a graphical comparison of these confidence intervals.

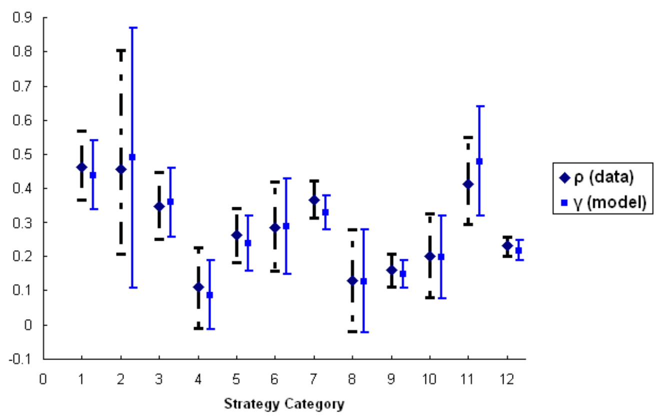


Figure 3: A comparison of estimates of the auto-correlation ρ and the persistence γ, showing the 95% confidence intervals for both. As before, the horizontal axis represents the strategy: 1. Convertible 2. Dedicated Short 3. Emerging Market 4. Equity Marcro 5. Event Driven 6. Fixed Income 7. Fund of Fund 8. Global Macro 9. Long-short Equity 10. Managed Future 11. Other 12. All

5. Testing the Constant-Persistence Normal-Noise Model

We now describe how we evaluated the fit of the constant-persistence normal-noise model. This model has only two parameters γ and $\sigma_b \equiv SD(B_n) \equiv \sqrt{Var(B_n)}$, so the fit to the observed persistence γ and standard deviation $\sigma \equiv SD(X_n)$ is immediate. If we use only those two parameters, we obtain a perfect fit by applying (2.4) and letting $\sigma_b^2 = (1-\gamma^2)\sigma^2$. Such a fit seems to provide a reasonable rough model in all cases.

In this section we want to evaluate the quality of that fit more closely. One test is the autocorrelation; the predicted relation between the autocorrelation and persistence in (2.5) holds more generally, and was just discussed above; Table 1 shows that the fit is pretty good, given the limited data. There are two principal remaining issues: (i) Is the relative-return distribution approximately normal? and (ii) Are the standard deviations (or variances) actually related by (2.4)? We have already addressed the first question in §4.3, finding that the return distribution is approximately normal in some cases, but not all. Now we turn to the one remaining question.

As indicated before, assuming stationarity, we combine all the relative-return data to estimate the one-year relative-return distribution. The standard deviation of that distribution is denoted by σ; it is estimated directly by the sample standard deviation once the data have been combined.

Testing is possible because we can also directly observe the values of the noise variables B_n. We estimate $\sigma_b \equiv SD(B_n) \equiv \sqrt{Var(B_n)}$ by acting as if the model is valid, implying that $B_n \equiv X_{n+1} - \gamma X_n$ would be i.i.d random variables, using the previously estimated value of the persistence γ. We thus estimate σ_b directly by the sample standard deviation as well, but we are here assuming the model to get the i.i.d. structure and we are using our estimate of the persistence γ. From (2.4), the constant-persistence normal-noise model (and other finite-variance-noise models) predict that $\sigma/\sigma_b = \sqrt{1/(1-\gamma^2)}$. Since we have already estimated γ from the data, we can compare σ/σ_b and $\sqrt{1/(1-\gamma^2)}$ in order to test the validity of the model.

Table 2 shows the results. From the last two columns in the table, we observe that σ/σ_b and

Table 2: Estimation of the standard deviations σ and σ_b to test the constant-persistence model.

Strategy	σ return[1]	σ_b noise[2]	ratio data[3]	ratio model[4]
1. Convertible	0.0686+0.0068/-0.0056	0.0579+0.0057/-0.0048	1.18	1.11+0.07/-0.05
2. Dedicated Short	0.1393+0.0480/-0.0284	0.1353+0.0466/-0.0275	1.03	1.15+0.88/-0.14
3. Emerging Market	0.1903+0.0161/-0.0138	0.1797+0.0152/-0.0130	1.06	1.07+0.05/-0.04
4. Equity Macro	0.0801+0.0074/-0.0062	0.0655+0.0061/-0.0051	1.22	1.00+0.01/-0.00
5. Event Driven	0.1007+0.0064/-0.0057	0.0884+0.0056/-0.0050	1.14	1.03+0.03/-0.02
6. Fixed Income	0.0693+0.0077/-0.0063	0.0661+0.0073/-0.0060	1.05	1.04+0.06/-0.03
7. Fund of Fund	0.0681+0.0031/-0.0029	0.0565+0.0026/-0.0024	1.21	1.06+0.02/-0.02
8. Global Macro	0.1070+0.0129/-0.0104	0.1027+0.0124/-0.0100	1.04	1.01+0.03/-0.01
9. Long-short Equity	0.1520+0.0054/-0.0050	0.1376+0.0048/-0.0045	1.10	1.01+0.01/-0.01
10. Managed Future	0.1265+0.0126/-0.0105	0.1214+0.0121/-0.0101	1.04	1.02+0.03/-0.02
11. Other	0.1003+0.0120/-0.0097	0.0976+0.0117/-0.0094	1.03	1.14+0.16/-0.08

1. σ: Standard deviation and 95 % confidence interval of the relative annual return
2. σ_b: Standard deviation and 95 % confidence interval of $B_n \equiv X_n - \gamma X_{n-1}$.
3. Ratio: σ/σ_b observed from the data.
4. Ratio: $\sqrt{1/(1-\gamma^2)}$, ratio σ/σ_b from the constant-persistence normal-noise model; see (2.4). 95% confidence interval of the ratio is obtained from 95% confidence interval of γ in Table 1.

$\sqrt{1/(1-\gamma^2)}$ are quite close for some fund strategies, but not for others. In particular, we see a good match for the emerging-market, fixed-income, global-macro, and managed-future strategies, but we see a poor match, in various degrees, for the others; the worst being the equity-macro and fund-of-fund strategies.

Where the match is good, we need to also test the normal-distribution property, which we have done, and discussed in §4.3. Where the match is poor, we see right away that we need to consider a different model, which is what much of the rest of this paper is about.

6. Stochastic-Persistence Models

In this section, we consider the stochastic-persistence models with various stochastic noise distributions as an alternative to the constant-persistence normal-noise model. Our analysis here illustrates the great model flexibility for fitting to data. Our goal in this section is to remedy both deficiencies found in the constant-persistence normal-noise model for some strategies: With the extra flexibility, we obtain a perfect match of the variance σ^2, remedying the problems observed in the last two columns of Table 2, and in addition seek a good match in the overall distribution.

6.1. Beta Persistence

In order to achieve this new flexibility in a controlled way, we assume that A_n has a *beta distribution*, which is a probability distribution that concentrates on the open unit interval $(0, 1)$. The beta distribution has two parameters, α and β, with mean $\alpha/(\alpha+\beta)$ and variance $\alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$. We can choose α and β to match the mean $E[A_n]$ and the variance $Var(A_n)$, provided that the variance is not too large. We remark that the beta distribution arises naturally in Bayesian frameworks when focusing on an unknown parameter lying in a fixed interval; e.g., see Browne and Whitt (1996). However, other persistence distributions can be used in essentially the same way.

By introducing beta persistence, we have thus increased the parameters associated with the persistence from only one (γ) in the deterministic case to two with this beta distribution. We can fit the beta parameters α and β to the mean and variance by

$$\gamma = \frac{\alpha/\beta}{1+\alpha/\beta} \quad \text{and} \quad c_a^2 \equiv \frac{\sigma_a^2}{\gamma^2} = \frac{\beta}{\alpha(\alpha+\beta+1)} = \frac{1}{\frac{\alpha}{\beta}\left(\frac{\alpha}{\beta}+1+\frac{1}{\beta}\right)} \ . \tag{6.1}$$

From (6.1), we see that the mean γ depends on α and β only through their ratio, while c_a^2, the squared coefficient of variation (SCV, variance divided by the square of the mean), is strictly increasing in both α and β for any given ratio α/β.

The full beta-persistence stochastic-noise model has three basic parameters: σ_b^2, γ and σ_a^2, but we only directly observe γ and σ^2. We have used γ to specify the mean $E[A_n]$. We thus have only σ^2 to use in order to determine the two model variances σ_a^2 and σ_b^2. Hence, there is one extra degree of freedom.

We apply the variance formula (3.8) to determine a relation that all these variances must satisfy. Formula (3.8) implies that we must have

$$0 \le \sigma_b^2 \le (1-\gamma^2)\sigma^2 \quad \text{and} \quad 0 \le \sigma_a^2 \le 1-\gamma^2. \tag{6.2}$$

Given both σ^2 and σ_b^2, formula (3.8) gives a formula for σ_a^2. In summary, there is a one-parameter family of variance pairs (σ_a^2, σ_b^2) consistent with our data.

We can draw some initial conclusions. First, if $\sigma_a^2 = 0$, so that $A_n = \gamma$ w.p.1, then we can estimate σ_b^2 directly by looking at $X_n - \gamma X_{n-1}$, as we already did. By formula (2.4) or (3.8), we then should have $\sigma_b^2 = (1-\gamma^2)\sigma^2$, but that is inconsistent with the results in Table 2. Hence we conclude that we do need to have stochastic persistence; i.e., we should consider some non-degenerate beta distribution for A_n.

One way to proceed at this point is to exploit what we have done in the previous section, and assume that we have already fit the variance σ_b^2 by acting as if the persistence A_n were constant. In other words, we let σ_b^2 be the estimated variance of $X_n - \gamma X_{n-1}$, using our estimate of the persistence γ.

Given that we start with an estimate of σ^2 and have already estimated γ and σ_b^2 by the methods already described, we can choose the variance $\sigma_a^2 \equiv Var(A_n)$ to satisfy (3.8). For the fund-of-fund return data, we have $\gamma = 0.33$ from Table 1, while $\sigma = 0.0681$ and $\sigma_b = 0.0565$ from Table 2, so that our estimated beta parameters are, first, $\sigma_a^2 = 0.2028$ and then $\alpha = 0.03$ and $\beta = 0.06$. However, the result is not plausible, because these small values of α and β produce a strongly U-shaped density for A_n; see Appendix §E.

We deduce that we should consider larger values of α and β, and thus smaller values for the variance σ_a^2 and larger values for σ_b^2. For given α, β is determined to match γ. From visual inspection, we estimate that $\alpha = 50$ should be reasonable; see Appendix §G.

Once we have chosen α, that determines β and thus σ_a^2, which in turn determines σ_b^2 by (3.8). For $\alpha = 50$, we get $\beta = 101.51$, $\sigma_a^2 = 0.0014$ and $\sigma_b = 0.9369\sigma = 0.0642$. Having calibrated the model parameter values, we then approximate the random variable X_∞ by taking a truncated version of the infinite series in (3.3). In our context, where we always have $\gamma < 1/2$, fewer than 10 terms suffices. We use only 5 for the fund-of-fund data with $\gamma = 0.33$. That yields the approximation

$$X_\infty \;\approx\; B_1 + A_1 B_2 + A_1 A_2 B_3 + A_1 A_2 A_3 B_4 + A_1 A_2 A_3 A_4 B_5 \; . \tag{6.3}$$

We get one realization from X_∞ by generating four independent copies of A_n and five independent copies of B_n.

6.2. The Beta-Persistence Normal-Noise and t-Noise Models

So far, by this rather involved process, we have specified only the variance of the noise $\sigma_b^2 \equiv Var(B_n)$. A simple specific noise distribution with that variance is the normal distribution that we have been considering; we get it by simply assuming that $B_n \stackrel{\mathrm{d}}{=} N(0, \sigma_b^2)$. For that special noise distribution, the single parameter σ_b^2 fully specifies the noise distribution. We call this the *beta-persistence normal-noise* model. However, when we apply this procedure and apply simulation to estimate the relative-return distribution, we see that the return distribution remains too close to the normal distribution. That remains the case for a wide range of α values; See, Appendix §E. Thus we rule out the beta-persistence normal-noise model. Our analysis leads us to conclude that this beta-noise feature, by itself, does not address the heavy tails seen in the data for the fund-of-fund strategy.

In order to capture the heavy tails in the observed relative-return distribution, we consider non-normal noise distributions. In doing so, we build on our previous analysis. As before, we aim to match the estimated values of γ and σ. We exploit the beta persistence we have already constructed, with $\alpha = 50$, $\sigma_a^2 = 0.0133$ and $\sigma_b = 0.0638$.

As a new candidate noise distribution, we propose the (Student)-t distribution, which is known to have a heavier tail than the normal distribution. Specifically, we assume that $B_n \stackrel{\mathrm{d}}{=} \kappa T(\nu)$ where $T(\nu)$ denotes a random variable with the standard t-distribution having parameter ν, which is commonly referred to as the degrees of freedom, and κ is a constant scale factor. Since we keep the beta persistence, we call the overall model the *beta-persistence t-noise model*.

For $\nu > 2$, the variance of a t-distributed random variable T is $\nu/(\nu - 2)$. Since $E[B_n] = 0$, we can match the given variance via

$$\sigma_b^2 \equiv Var(B_n) = E[B_n^2] = E[\kappa^2 T^2] = \frac{\kappa^2 \nu}{\nu - 2} \; . \tag{6.4}$$

We first use ν as a parameter to choose in order to select the desired shape of the distribution of X_n, consistent with a fixed first two moments of B_n (mean 0 and variance σ_b^2). We then use κ to match the observed variance. Thus, for any given ν, κ is determined by $\kappa = \sigma_b\sqrt{(\nu-2)/\nu}$.

Figure 4 shows the simulated distribution of the relative return X_n from the beta-persistence t-noise model with $A_n \stackrel{\text{d}}{=} Beta(50, 101.51)$ and $B_n \stackrel{\text{d}}{=} 0.0278 \cdot T(2.4)$ compared to the observed relative-return distribution for the fund-of-fund strategy. Comparing Figures 1 and 4, we see that the beta-persistence t-noise model approximates the observed relative-return distribution much better than the constant-persistence normal-noise model does. The two-sample Kolmogorov-Smirnov test also statistically shows that we cannot reject the hypothesis that the simulated data and empirical data come from the same distribution, with p value of 0.3080 (The high p value indicates that we cannot reject the hypothesis that the two random variables are drawn from the same distribution; e.g., see Massey (1951).)

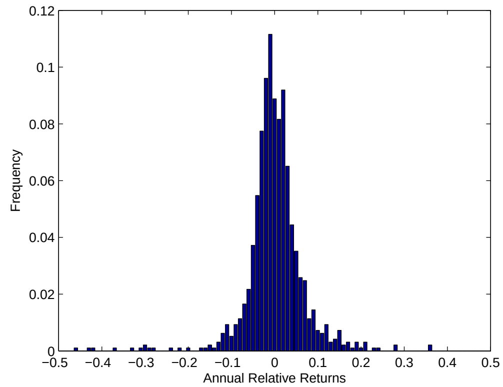
However, looking closely at Figure 4, we see that the observed relative-return distribution still has heavier tails than predicted by the model, especially in the left tail. That conclusion is confirmed by the Q-Q plot in Figure 4(c).

6.3. The Beta-Persistence Empirical-Noise Model

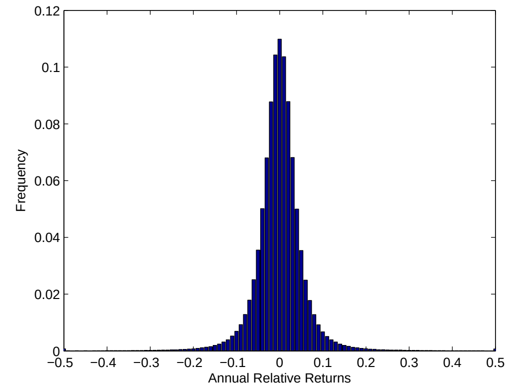
A relatively simple way to obtain a better fit to the data within the beta-persistence class of models is to let B_n have the observed empirical distribution for $X_n - \gamma X_{n-1}$, using the estimated value of γ. This automatically gives B_n and its estimated variance σ_b^2. It now goes further to directly match the shape. This procedure works quite well, as we show in Appendix §G. Overall, the approach works well if we are content to use the model for simulation. However, we might want a parametric model, with not too many parameters, so we consider further refinements.

6.4. The Beta-Persistence Mixed-Noise Model

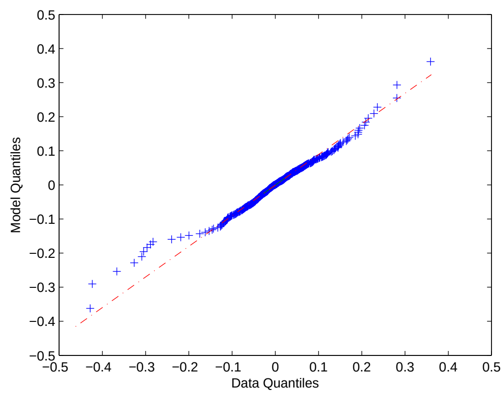
Since the beta-persistence t-noise model did not adequately capture the heavy left tail of the observed relative-return distribution for the fund-of-fund strategy, we continue to search for a better parametric model. In order to better match this feature, we consider a mixture of two distributions for our noise distribution. We do this both to illustrate the flexibility of our general modelling framework and to obtain a better fit.



(a) Histogram of relative returns from the data



(b) Beta-persistence t-noise model



(c) Q-Q plot comparing the model to the data

Figure 4: (a) The relative-return distribution from the data within the fund-of-fund strategy (986 observations). (b) Simulation estimate of the relative-return distribution (sample size 10^6) using the beta-persistence t-noise model, with the sample size of 10^6, with $\alpha = 50$, $\beta = 101.51$, $\nu = 2.4$, $k = 0.0278$, $\gamma = 0.33$ and $\sigma = 0.0681$. (c) Q-Q plot of the beta-persistence t-noise model to the data.

Again building upon our previous fitting, we let the distribution of B_n be a mixture of an exceptional normal distribution with some small probability p and the t distribution with probability $1 - p$. We start with the beta stochastic persistence in order to calibrate the two variances σ^2 and σ_b^2, and then we introduce the t-noise distribution in order to capture the main shape of the return distribution. In addition, we now add a small normal component to capture the heavy left tail. We call this overall construction our *beta-persistence mixed-noise model.*

The noise random variable B_n in this model can be defined explicitly by

$$B_n = \begin{cases} Z_1 \stackrel{\mathrm{d}}{=} \mu_1 + \kappa T(\nu) \text{ with probability } 1 - p \\ Z_2 \stackrel{\mathrm{d}}{=} N(\mu_2, \sigma_2^2) \text{ with probability } p \ . \end{cases} \tag{6.5}$$

Here it is understood that Z_1 represents the regular returns, while Z_2 represents the exceptional low returns. We intend to make the probability p small.

From (6.5), we have six parameters to fit: $p, \kappa, \nu, \mu_1, \mu_2$ and σ_2. We start by controlling the overall shape. That is done by choosing the t parameter ν, in the method just described. We then calibrate p by counting the number of relative returns less than -2σ. Then it remains to fit the four remaining parameters κ, μ_1, μ_2 and σ_2. But now we can write down expressions for the mean and variance of B_n:

$$\begin{aligned} \mathbb{E}[B_n] &= (1-p)\mu_1 + p\mu_2 = 0, \\ \sigma_b^2 &= \mathbb{E}[B_n^2] \;=\; (1-p)\left(\kappa^2 \frac{\nu}{\nu-2} + \mu_1^2\right) + p(\mu_2^2 + \sigma_2^2) \;. \end{aligned} \tag{6.6}$$

Since we have two equations in four parameter values, we have two degrees of freedom. Thus, we fit μ_2 and σ_2 directly from the data. We directly fit the mean and standard deviation of the relative returns counted for estimating p. In this way, we can fit p, μ_2 and σ_2 at the same time. Then, from (6.6), we can obtain explicit representations for μ_1 and κ, namely,

$$\mu_1 = -p\mu_2/(1-p) \quad \text{and} \quad \kappa = \sqrt{[(\nu-2)/\nu(1-p)]\left(\sigma_b^2 - p(\mu_2^2+\sigma_2^2) - (1-p)\mu_1^2\right)} \;. \tag{6.7}$$

For the fund-of-fund relative returns, out of 986 data points in our sample, we find 18 relative returns below $-2\sigma = -0.1363$. Thus our estimate for p is $p = 18/986 = 0.0183$. As indicated above, in this step we also select the mean and standard deviation of this "exceptional distribution." We find that the mean and standard deviation of those 18 returns are $\mu_2 = -0.2746$ and $\sigma_2 = 0.0717$. Finally, we fit the remaining parameters, getting $\mu_1 = -0.0051$ and $\kappa = 0.0232$. Again, after calibrating parameters for X_n, we use (6.3) to generate realizations of the modelled stationary return X_∞.

Figure 5 (a) and (b) shows the simulated return distribution for this beta-persistence mixed-noise model. We now do see a heavier left tail in the model, just like that in the data, but unfortunately now the left tail of the return distribution generated by the model now is heavier than the left tail of the observed distribution from the data. This actually should not be surprising because our model exaggerates the probability of a return below -2σ, including the t-variable as well as the exceptional normal component.

In order to reduce the gap between the model and the data in the left tail, we consider a new parameter fitting procedure that reduces p while keeping μ_2 and σ_2 as specified. The new procedure starts from the given parameter values $p, \mu_2, \sigma_2, \mu_1, \kappa$ and the simulation obtained from the fitting procedure stated above. We first calculate the probability of relative returns falling below the threshold in the model, denoted by f. Since $\mu_2 \ll -2\sigma$ and $\sigma_2 \ll (-2\sigma + \mu_2)$, we ignore the probability of exceptional random variables exceeding the threshold. Let t be the probability that t-distribution falls below the threshold (which we do not evaluate directly). From the definition of t and the observed f, we obtain $p + (1-p)t \approx f$, which yields $t \approx (f-p)/(1-p)$. To obtain a corrected model, we replace f by p and p by p', and have $p' + (1-p')t \approx p$ for $t \approx (f-p)/(1-p)$. Combining these two equations, we get the following expression for p' (which is to replace p):

$$p' = \frac{2p - p^2 - f}{1-f}. \tag{6.8}$$

Our revised model is (6.5) with p replaced with p' in (6.8). We assume that μ_2 and σ_2 remain unchanged. We thus need to calculate new values of μ_1' and κ' via (6.7), using p' instead of p.

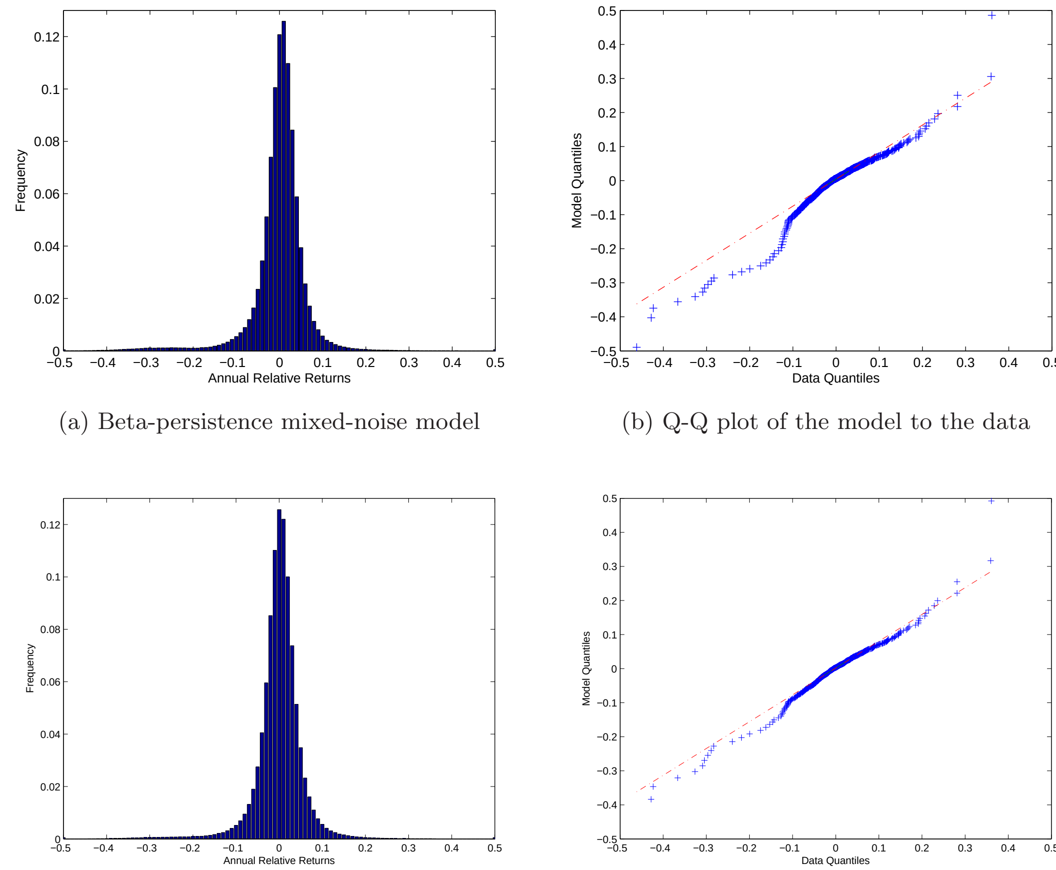
Then, we perform simulation once more with new parameters. Since the first simulation has $f = 0.0284$, we obtain $p' = 0.0081, \mu_1' = 0.0022$ and $\kappa' = 0.0236$ from the new procedure. We found that this procedure significantly improves the fitting. As shown in Figure 5, the left tail from the new procedure matches the data much better than before.

7. The Constant-Persistence Stable-Noise Model

The procedures in §6 introduced more and more complexity in order to obtain a better and better fit. A more parsimonious alternative is to directly address the heavy-tail property at the outset by using a stable distribution. In doing so, we have to abandon the information provided by the variance σ^2 and the other variances, because the stable distribution has infinite variance. We thus lose a convenient model parameter when we take this step.

However, we gain simplicity, because we can use the constant-persistence model and avoid any representation of the distribution of A_n. Moreover, the stable distribution has the advantage of providing additional tractability. In particular, with constant persistence, stable noise provides the nice relation between the distribution of X_n and the distribution of B_n given in (2.10) and Proposition 1. That relation says that X_n will be distributed the same as a constant multiple of B_n.

Indeed, Proposition 1 provides an ideal way to test whether the constant-persistence stable-noise might be appropriate. A simple test is to plot the distributions of X_n and B_n and see if they



(a) Beta-persistence mixed-noise model

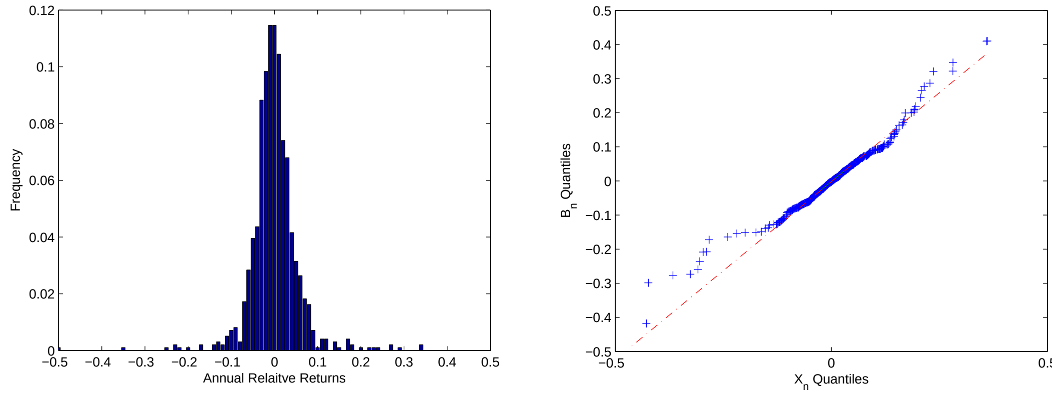
(b) Q-Q plot of the model to the data

(c) Beta-persistence mixed-noise model re-calibrated

(d) Q-Q plot of the model to the data

Figure 5: (a) Simulation estimate of the relative-return distribution (sample size 10^6) using the beta-persistence mixed-noise model with $\gamma = 0.33, \alpha = 50, \nu = 2.4, \sigma = 0.0681, \mu_2 = -0.2746, \sigma_2 = 0.0717, p = 0.0186, \mu_1 = -0.0051$ and $\kappa = 0.0232$. (to be compared to Figure 4 (a)). (b) Q-Q plot comparing the model to the data. (c) (d) Simulation estimate of the relative-return distribution and Q-Q plot for the same model re-calibrated with $p' = 0.0098, \mu_1' = 0.0027, \kappa' = 0.0237$ in (6.8).

look similar. As noted before, we obtain B_n directly from $X_n - \gamma X_{n-1}$, using the previous estimate for the persistence γ. Figures 4 (a) and 6 (a) show the empirical distributions of X_n (stationary version) and B_n obtained from the fund-of-fund data. Clearly, these distributions look remarkably similar, although the Q-Q plot in Figure 6 (b) shows some discrepancy in the tails. Moreover, the relationship is further substantiated by Table 3, where the ratio of the quantile differences of these distributions are calculated at different levels. These quantile ratios constitute estimates of the proportionality constant c. These quantile ratios are consistently around 1.2, with some discrepancy again in the tails. Thus, Figure 6 and Table 3 suggest that $X_n \stackrel{\mathrm{d}}{=} cB_n$ approximately, where c is a constant whose value is about 1.2. We also performed the two sample Kolmogorov-



(a) $B_n \equiv X_n - \gamma X_{n-1}$ from the data

(b) Q-Q plot comparing X_n to cB_n

Figure 6: (a) Distribution of $B_n \equiv X_n - \gamma X_{n-1}$ for the fund-of-fund relative returns, to be compared to Figure 4 (a), and (b) Q-Q plot comparing the distributions of X_n and cB_n with $c = 1.2$.

Smirnov test to compare the distributions, and obtained a p value of 0.5196, which provides further support.

Table 3: The Quantile Differences of X_n and B_n and Their Ratios

Quantile Difference[1]	X_n	B_n	Ratio [2]
55% − 45%	0.0111	0.0085	1.3096
60% − 40%	0.0210	0.0170	1.2321
65% − 35%	0.0327	0.0265	1.2342
70% − 30%	0.0425	0.0364	1.1683
75% − 25%	0.0566	0.0492	1.1506
80% − 20%	0.0709	0.0609	1.1633
85% − 15%	0.0907	0.0778	1.1656
90% − 10%	0.1211	0.1053	1.1509
95% − 5%	0.1887	0.1430	1.3194

1. Difference between two quantile values.
2. Ratio: Quantile Difference for X /Quantile Difference for B.

Recall from our discussion in §1 that the index α of a stable law coincides with its tail-decay parameter (of the form $Cx^{-\alpha}$ for some constant C). The conventional elementary way to investigate power tails and estimate the index α is to directly construct a log-log plot of the tails of the distributions. Figure 7 shows the log-log plots of the two distribution tails for the fund-of-fund relative-return data. (Figure 7 also shows corresponding plots for a model, to be discussed below.) We observe that the left tail of the return distribution is approximated quite well by the linear slope of −1.6, which implies that there is approximately a power tail and that $\alpha \approx 1.6$. As we have observed before, the heavy-tail behavior is more evident in the left tail than in the right tail. The two sample Kolmogorov-Smirnov test result also shows high p value (0.1446), which statistically shows that the two samples could be drawn from the same distribution. (In Appendix §F we provide log-log plots of the tails of the simulated distributions from the other models for contrast.)

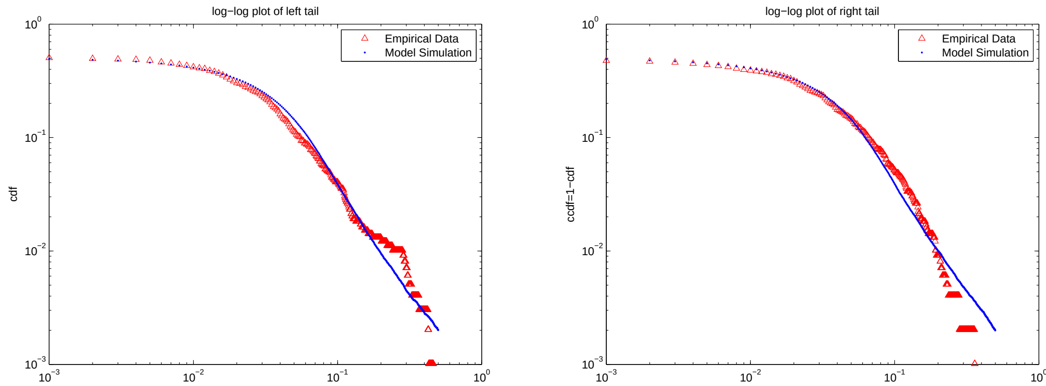


Figure 7: Log-log plots of the left and right tails of the fund-of-fund relative-return distribution, from the TASS data and the constant-persistence stable-noise model with parameters $\gamma = 0.33$, $\alpha = 1.6, \beta = 0, k = 0.029$.

We now combine the last two observations to develop a test for the constant-persistence stable-noise model. On the one hand, we have directly estimated the stable index α from the log-log plots of the distribution tails (getting $\alpha \approx 1.6$), but on the other hand, for the constant-persistence stable-noise model, the observed quantile ratio $c \approx 1.2$ also provides an estimate of the index α. That is true because, given the quantile ratio c and the persistence γ, we can solve for α in the equation

$$c^{\alpha} = \frac{1}{1 - \gamma^{\alpha}}, \tag{7.1}$$

obtained from (2.10). We see that the observed value $c = 1.20$ is consistent with the other parameter values: $\gamma \approx 0.33$ and $\alpha \approx 1.6$. Thus the constant-persistence stable-noise model passes this test.

Non-Gaussian stable laws actually have four parameters, and are commonly referred to by $S_{\alpha}(\kappa, \beta, \mu)$; see Samorodnitsky and Taqqu (1994). (We use κ instead of the conventional σ to avoid confusion with the standard deviation considered previously.) As before, α is the index, which ranges in $0 < \alpha < 2$. The other three parameters are: the scale κ, the skewness β and the location parameter μ. When the stable law has a finite mean, μ is that mean. Since we are considering stable laws with finite mean, where that mean is zero, we always have $\mu = 0$. For $\alpha > 1$ and $\mu = 0$, we have the scaling relation

$$cS_{\alpha}(\kappa, \beta, 0) \stackrel{\mathrm{d}}{=} S_{\alpha}(c\kappa, \beta, 0) \quad \text{for all} \quad c > 0 \tag{7.2}$$

for all model parameters. Choosing the scale parameter κ is like choosing the measuring units. In addition to the index, the shape is determined by the skewness parameter β which ranges in $-1 \leq \beta \leq 1$. From (7.2), we see that the scale has no effect on the index or the skewness.

Given the index α, we also have available the two parameters κ and β. As α increases, the shape of the distribution is more centered. As β increases, the distribution is skewed more to the left. Thus we formulate the constant-persistence stable-noise model by letting $B_n \stackrel{\mathrm{d}}{=} \kappa \cdot S_\alpha(1, \beta, 0)$. Using Proposition 1 and the scaling relation (7.2) for the constant-persistence stable-noise model (2.1), we have

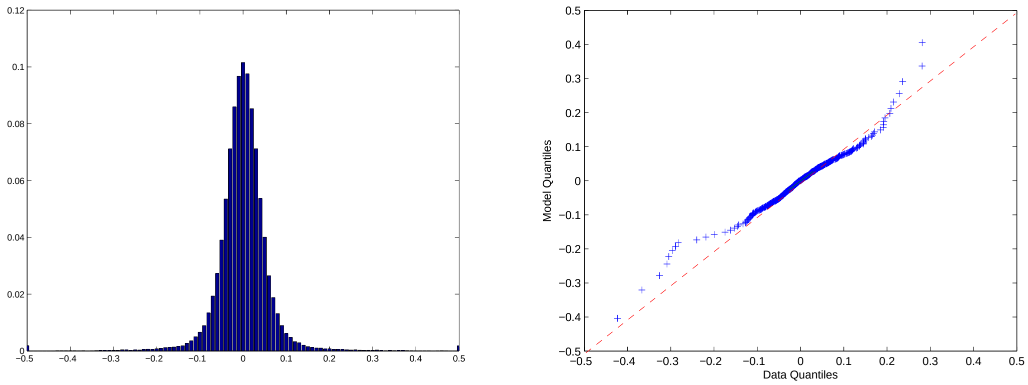
$$X_\infty \stackrel{\mathrm{d}}{=} \left(\frac{1}{1-\gamma^\alpha}\right)^{1/\alpha} \kappa \cdot S_\alpha(1, \beta, 0) \ . \tag{7.3}$$

We emphasize that this characterization of the limiting distribution in the constant-persistence stable-noise model simplifies further analysis and simulation; e.g., we do not need the approximation formula in (6.3).

We are now ready to consider specific parameter values for our constant-persistence stable-noise model. We can select the index from the slope of the log-log plots, as in Figure 7. We then can set the scale parameter κ by looking at the quantile ratios. We have chosen the value $\kappa = 0.029$. We can choose the skewness to match the shape. We compare plots of the distribution of either B_n or X_n to plots of stable distributions as a function of the skewness parameter β. In this informal way, we picked $\beta = 0$; see, Appendix §H for the details.

Figure 8 (a) shows the estimated relative-return distribution from the calibrated constant-persistence stable-noise model. Note that the chosen value of $\kappa = 0.029$ matches the peak of the distribution from the data and model reasonably well; see Figure 4 (a) for comparison. Figure 8 shows that the model approximates the empirical distribution reasonably well. However, Figure 8 (b) shows that the tails of the simulated distribution from the model fits the tails of the distribution from the data only roughly, not as good as Figure 5 (d).

Now we further test the validity of the model by comparing the quantile ratio in Table 3 and c in (7.3). Since the quantile ratio is estimated from the data and c is predicted by the model, if they coincide, the validity of the model is verified. It turns out that the model with calibrated $\alpha = 1.6$ and $\gamma = 0.33$, $\kappa = 0.029$ from the data generates $c = 1.1232$ which is consistent with Table 3. This provides solid support for the constant-persistence stable-noise model.



(a) Constant-persistence stable-noise model

(b) Q-Q plot comparing model and data

Figure 8: (a) A simulation estimate of the relative-return distribution (sample size 10^6) of the constant-persistence stable-noise model with $\alpha = 1.6, \beta = 0, \kappa = 0.029$ (to be compared to Figure 4 (a)). (b) Q-Q plot comparing the predicted relative-return distribution based on the constant-persistence stable-noise model to the empirical distribution from the fund-of-fund TASS data.

8. An Additional Model Test: Hitting Probabilities

In this section, we consider the probability that the hedge-fund relative return ever exceeds some level during the 5-year time period. Such hitting probabilities are important for risk management. We consider high or low levels of relative returns, measured in units of (sample) standard deviation σ. By simply counting the number of hedge funds whose relative returns have ever reached the level during 5-year period (2000-2004), we calculate the hitting probability from the data.

Table 4 shows the hitting probabilities of each level for five years from the data within the fund-of-fund strategy and the corresponding beta-persistence t-noise, beta-persistence mixed noise and constant-persistence stable-noise models. The probability estimate from the data is the observed proportion of funds whose relative returns had ever hit the level during the entire five-year period, among the 92 total number of funds within fund-of-fund strategy in 2000. The initial relative return in the model simulation is set to have the stationary limiting distribution of each model, i.e., X_∞.

We perform two different simulation estimates. First, in order to estimate the true hitting probabilities, we generate 10,000 independent values of X_∞ for initial relative returns, using (6.3) and (7.3) and then use the recursion $X_n = A_n X_{n-1} + B_n$ to calculate 95% confidence interval of hitting probability throughout five years. Second, in order to assess whether the model is consistent with the data, given the small sample size, we simulate 92 independent values of the X_∞ random variables as the initial relative returns in 2000 and then use the recursion formula of

$X_n = A_n X_{n-1} + B_n$ to determine the hitting probability within 5 years. We repeat 20 of these simulations and record the maximum and minimum hitting probability observed and investigate if the range of hitting probabilities includes the probability from the data. It is observed that the hitting probabilities for the high level fit the probability from the data relatively well. However, all the first estimates predict higher hitting probabilities for the low levels than are predicted from the data estimates. Nevertheless, the range of probabilities from the 20 simulations includes the hitting-probability estimates from data in most cases. (See, also Appendix G for corresponding results for the Beta-persistence empirical-noise model.)

Table 4: Hitting probabilities of thresholds over a five-year period (2000-2004)

Level[1]	data[2]	t-noise		Mixed noise		Stable noise	
		$N = 92^3$	$N = 10,000^4$	$N = 92^3$	$N = 10,000^4$	$N = 92^3$	$N = 10,000^4$
3 σ	0.0326	[0,0.0435]	0.0280±0.0032	[0,0.0543]	0.0174±0.0026	[0,0.0870]	0.0326±0.0035
2 σ	0.0761	[0.0326,0.1087]	0.0696±0.0050	[0.0217,0.0761]	0.0464±0.0041	[0.0217,0.1630]	0.0712±0.0050
1 σ	0.2363	[0.1739,0.3478]	0.2569±0.0086	[0.1304,0.2717]	0.2012± 0.0079	[0.1304,0.3696]	0.2593±0.0086
-1 σ	0.2391	[0.1848,0.3043]	0.2603±0.0086	[0.1196,0.3152]	0.2028± 0.0079	[0.1739,0.3587]	0.2590±0.0086
-2 σ	0.0542	[0.0326,0.1413]	0.0718±0.0051	[0.0217,0.1522]	0.0797± 0.0053	[0.0217,0.1087]	0.0670±0.0049
-3 σ	0.0326	[0,0.0543]	0.0273±0.0032	[0.0109,0.0978]	0.0516±0.0043	[0,0.0652]	0.0328±0.0035

1. $\sigma = 0.0681$, the observed standard deviation of the fund-of-fund relative returns.
2. Number of funds that have ever hit the level for 2000-2004 divided by 92, the total number in 2000.
3. Minimum and maximum of the probabilities from 20 simulations with sample size of 92 initially.
4. 95 % confidence interval of hitting probability from simulation with sample size of 10,000 initially

9. Conclusion

In this paper, we proposed a stochastic difference equation (SDE) of the form $X_n = A_n X_{n-1} + B_n$ to model the relative returns of hedge funds. In §2-§3 we showed that the model is remarkably tractable, with many convenient analytical properties. Afterwards, we showed that the model is remarkably flexible for model fitting by showing how it can be calibrated to the data from the TASS database from 2000 to 2005. The foundation of our approach is persistence. It is quantified in the model via $\gamma \equiv \mathbb{E}[A_n]$. We presented a strong case for basing the model on persistence by showing that the observed persistence estimated from the data by regression is statistically significant for all but two strategies (see Table 1). The persistence was found to range from 0.11 to 0.49 across the eleven fund strategies.

For the emerging-market strategy, the parsimonious (two parameter) constant-persistence normal-noise model with $A_n = \gamma$ and $B_n \stackrel{d}{=} N(0, \sigma_b^2)$ provides an excellent fit, with σ_b^2 fit to the estimated relative-returns variance σ^2 directly by (2.4). However, the constant-persistence normal-noise model is not suitable for the fund-of-fund strategy, and most other strategies, largely because the relative-return distribution has heavy tails. However, we find that some strategies are well approximated by the beta-persistence normal-noise model. In particular, that is the case for the long-short equity strategy, as we show in the Appendix §I. We do a complete fitting for that strategy there.

For the heavy-tailed distributions, we demonstrated the SDE model flexibility by showing that a good fit can be obtained for the fund-of-fund relative-return process by choosing variables A_n and B_n in different ways. The beta-persistence mixed-noise model in §6.4, the constant-persistence stable-noise model in §7 and the beta-persistence empirical-noise model in Appendix §G all produced remarkably good fits, given the limited and unreliable data. Each of these models has advantages and disadvantages: The empirical-noise model is evidently most accurate, but it is a complicated non-parametric model, which may only be useful in simulation studies. The stable-noise model has the most appealing mathematical form, but it is not as accurate and it cannot exploit the variance for fitting (since it implies infinite variance). The mixed-noise model falls in between: it has good accuracy and it is a parametric model that can use the variance for fitting, but the parametric structure is complicated, making it harder to use in mathematical analysis. But these three models are just a sample of what could be considered. They illustrate that our SDE model offers a flexible model for fitting.

We paid special attention to matching the (assumed stationary) single-year relative-return distribution, but we also evaluated the fit of the stochastic-process model over time. As shown in (3.10), the SDE model predicts that the autocorrelation coefficient should coincide with the persistence factor γ. Table 1 shows that is consistent with the data. In §8 we also showed that the model predicted 5-year hitting probabilities of high (or low) thresholds reasonably well too. The fit here was especially good for the beta-persistence empirical-noise model, as shown in Appendix §G. In this test, our conclusions were not as strong as we would like because of the relatively small sample sizes and the somewhat unreliable data. We think that there is the potential for even better fitting with better data.

Overall, we contend that the value of our proposed modelling approach has been demonstrated. It should be useful in other financial contexts as well, wherever persistence may exist. As we explained in §2, our SDE is a discrete-time analog of the common stochastic differential equation, which should be regarded as an attractive alternative when time is naturally regarded as discrete. §2.4 contains a numerical example illustrating how our model can be applied to go beyond data description to answer various "what if" questions. There we briefly considered how the model might be applied to quantify the value of good fund management.

Acknowledgments. Ward Whitt was partially supported by NSF grant DMI-0457095.

References

Agarwal, V., N. Y. Naik. 2000. Multi-period performance persistence analysis of hedge funds. *Journal of Financial and Quantitative Analysis* **35**(3) 327–342.

Amenc, N., S. E. Bied, L. Martellini. 2003. Predictability in hedge fund returns. *Financial Analysts Journal* **59**(5) 32–46.

Amin, G. S., H. M. Kat. 2003. Hedge fund performance 1990-2000: Do the "money machines" really add value? *Journal of Financial and Quantitative Analysis* **38**(2) 251–274.

Amo, A., H. Harasty, P. Hillion. 2007. Diversification benefits of funds of hedge funds: Identifying the optimal number of hedge funds. *The Journal of Alternative Investments* **10**(2) 10–22.

Atlan, M., H. Geman, M. Yor. 2006. Options on hedge funds under the high-water mark rule. *Working Paper, Université Pierre et Marie Curie and Université Paris Dauphine* .

Barndorff-Nielsen, O. E., N. Shephard. 2001. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B* **63**(2) 167–241.

Borland, L. 2002. A theory of non-Gaussian option pricing. *Quantitative Finance* **2**(6) 415–431.

Boyson, N. M, M. J. Cooper. 2004. Do hedge funds exhibit performance persistence? A new approach. *Working Paper, Northeastern and Purdue University* .

Brandt, A. 1986. The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advances in Applied Probability* **18**(1) 211–220.

Brown, S. J., W. N. Goetzmann, R. G. Ibbotson. 1999. Offshore hedge funds survival and performance 1989-1995. *Journal of Business* **72**(1) 91–117.

Browne, S., W. Whitt. 1996. Portfolio choice and the Bayesian Kelly criterion. *Advances in Applied Probability* **28**(4) 1145–1176.

Capocci, D., G. Hüber. 2004. Analysis of hedge fund performance. *Journal of Empirical Finance* **11**(1) 55–89.

Chan, N., M. Getmansky, S. M. Hass, A. W. Lo. 2006. Do hedge fund increase systematic risk? *Economic Review - Federal Reserve Bank of Atlanta* **91**(4) 49–80.

Cont, R. 2001. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* **1**(2) 223–236.

Derman, E., K. S. Park, W. Whitt. 2008. Markov chain models to estimate the premium for extended hedge fund lockups. *Working Paper, Columbia University* .

Edwards, F. R., M. O. Caglayan. 2001. Hedge fund performance and manager skill. *Journal of Futures Markets* **21**(11) 1003–1028.

Eling, M., F. Schuhmacher. 2007. Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking and Finance* **31**(9) 2632–2647.

Embrechts, P., C. Klüppelberg, T. Mikosch. 1997. *Modeling Extremal Events*. Springer.

Fama, E. 1965. The behavior of stock-market prices. *Journal of Business* **38**(1) 34–105.

Fung, W., D. A. Hsieh. 2000. Performance characteristics of hedge funds and commodity funds: Natural versus spurious biases. *Journal of Financial and Quantitative Analysis* **35**(3) 291–307.

Gabaix, X., P. Gopikrishnan, V. Plerou, E. Stanley. 2007. A unified econophysics explanation for the power-law exponents of stock market activity. *Physica A: Statistical Mechanics and its Applications* **382**(1) 81–88.

Geman, H., C. Kharoubi. 2003. Hedge funds revisited: Distributional characteristics, dependence structure and diversification. *Journal of Risk* **5**(4) 55–73.

Getsmansky, M., A. W. Lo, I. Makarov. 2004. An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics* **74**(3) 529–609.

Hasanhodzic, J., A. W. Lo. 2007. Can hedge-fund returns be replicated?: The linear case. *Journal of Investment Management* **5**(2) 5–45.

Hayes, B. T. 2006. Maximum drawdowns of hedge funds with serial correlation. *Journal of Alternative Investments* **8**(4) 26–38.

Hodder, J. E., J. C. Jackwerth. 2007. Incentive contracts and hedge fund management. *Journal of Financial and Quantitative Analysis* **42**(4) 811–826.

Horst, U. 2001. The stochastic equation $Y_{t+1} = A_t Y_t + B_t$ with non-stationary coefficients. *Journal of Applied Probability* **38**(1) 80–94.

Horst, U. 2003. Stability of linear stochastic difference equations in strategically controlled random environments. *Advances in Applied Probability* **35**(4) 961–981.

Jagannathan, R., A. Malakhov, D. Nonikov. 2006. Do hot hands persist among hedge fund managers. *NBER Working Paper* .

Kassberger, S., R. Kiesel. 2006. A fully parametric approach to return modelling and risk management of hedge funds. *Financial Markets and Portfolio Management* **20**(4) 472–491.

Lhabitant, F. 2004. *Hedge Funds: Quantitative Insights*. Wiley.

Lilliefors, H. W. 1967. On the Kolmogorov-Smirnov test for normality with mean and variances unknown. *Journal of the American Statistical Association* **62**(318) 399–402.

Lin, L. I. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**(1) 255–268.

Lo, A. 2001. Risk management for hedge funds: Introduction and overview. *Financial Analysts Journal* **57**(6) 16–33.

Lux, T. 1996. The stable Paretian hypothesis and the frequence of large returns: an examination of major German stocks. *Applied Financial Economics* **6**(6) 463–475.

Mandelbrot, B. 1963. The variation of certain speculative prices. *Journal of Business* **36**(4) 394–419.

Massey, F. J. 1951. The Kolmogorov-Smirnov test for goodness of fit. *Journal of American Statistical Association* **46**(253) 68–78.

Rachev, S., S. Mittnik. 2000. *Stable Paretian Models in Finance*. Wiley.

Samorodnitsky, G., M. S. Taqqu. 1994. *Stable Non-Gaussian Random Processes*. Chapman & Hall.

Serfling, R. J. 1980. *Approximation Theorems of Mathematical Statistics*. Wiley.

Tran, V. 2006. *Evaluating Hedge Fund Performance*. Wiley Finance.

Vervaat, W. 1979. On a stochastic difference equation and a representation of non-negative infinitely divisible random variable. *Advanced Applied Probability* **11**(4) 750–783.

Whitt, W. 2002. *Stochastic-Process Limits*. Springer.

Wolfe, S. J. 1982. On a continuous analogue of the stochastic difference equation $X_{n+1} = \rho X_n + B_n$. *Stochastic Processes and Their Applications* **12**(3) 301–312.

APPENDIX
TO
A STOCHASTIC MODEL FOR
HEDGE FUND RELATIVE RETURNS

by

Emanuel Derman, Kun Soo Park, and Ward Whitt

Department of Industrial Engineering and Operations Research
Columbia University, New York, NY 10027-6699
{ed2114, kp2143, ww2040}@columbia.edu

A. Introduction and Summary

This appendix has nine more sections. In §B we display plots of the sizes of the managed assets of the funds in our sample. In §C, we provide the relative-return distributions of hedge funds across 10 strategies in the TASS database from 2000-2005. It is observed that the relative-return distributions for some strategies are approximately normal, while others have high peaks or heavy tails, which is not fit by the normal distribution. In §D, we show how the relative-return distribution in the constant-persistence normal-noise model depends on the sample size of the simulation. We compare simulations with the sample size of the data to larger simulations with sample size of 10^6.

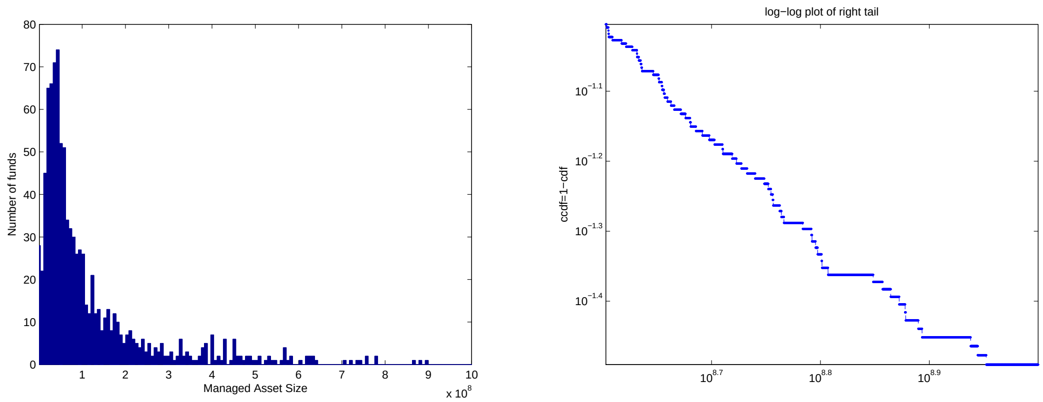
We supplement the analysis of the other models in the remaining sections. In §E, we show how the beta-persistence model depends on the beta-distribution parameters α and β. It is shown that the shape of the estimated relative-return distribution is insensitive to α and β. In §F, we show that the heavy-tail and light-tail distributions behave differently in log-log scale. In §G, we show that the beta-persistence empirical-noise model provides a good fit the the data and reasonable estimates to the hitting probabilities. In §H, we show how the tails of the relative-return distribution in the constant-persistence stable-noise model behave, depending on the parameter β in the stable distribution. It is observed that the estimated relative-return distribution fits the data reasonably well for fund-of-fund strategy when $\beta = 0$. In §I, we provide a fitting for long-short equity strategy, which has the largest sample size in the data. We conclude that the beta-persistence normal-noise model fits the data well. Finally, in §J, we provide a fitting for the event-driven strategy whose relative-return distribution has heavy tails. It is observed that the beta-persistence t-noise model and constant-persistence stable-noise model provides a good fit to the data.

B. The Values of Managed Assets

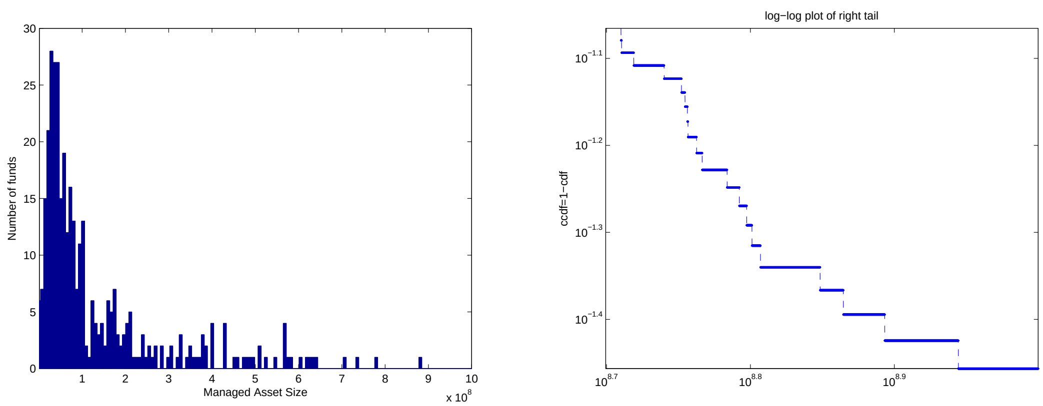
As described in §4 of the main paper, we started by examining the TASS data. We followed the previous researchers, such as Boyson and Cooper (2004), in our data selection procedure. For each strategy, in order to avoid very small funds, which might have different characteristics, we first removed all funds from the data for which the managed asset value never reaches our 25 million dollar threshold. For the fund-of-fund strategy, we first removed 407 fund pairs from the data; that left the 986 fund pairs in our sample. (A pair is the relative annual returns for two successive years.)

To further explore the data, we considered the distribution of the average asset values managed by the fund. In Figure 1 (a) below, we plot the histogram of the average managed asset value among the the 986 funds in the fund-of-fund strategy. These 986 observations are taken only from the funds exceeding the 25 million dollar threshold. We see that the largest managed asset values are of order \$$10^8$. We also show a corresponding log-log plot in Figure 1 (b), which shows that the size distribution has a heavy tail.

We also measure the total value of asset managed by the larger and smaller funds (in terms of managed asset values) in Table 5. We first study the total value of asset managed for all 986 returns observed for fund-of-fund strategy. Since the relative returns from 2000 to 2004 are included at the same time for all 986 observations, asset values of some funds are counted multiple times for their life during the period. Thus, we also choose one specific year, namely, 2004, and take a snapshot of that year in terms of asset size such that we can see how the asset size of each fund, not the returns over the years, is distributed in one year.



(a) Histogram and log-log plot of the asset value managed by funds (all)



(b) Histogram and log-log plot of the asset value managed by funds (2004)

Figure 1: Histogram and log-log plot of the value of managed assets for funds under the fund-of-fund strategy.

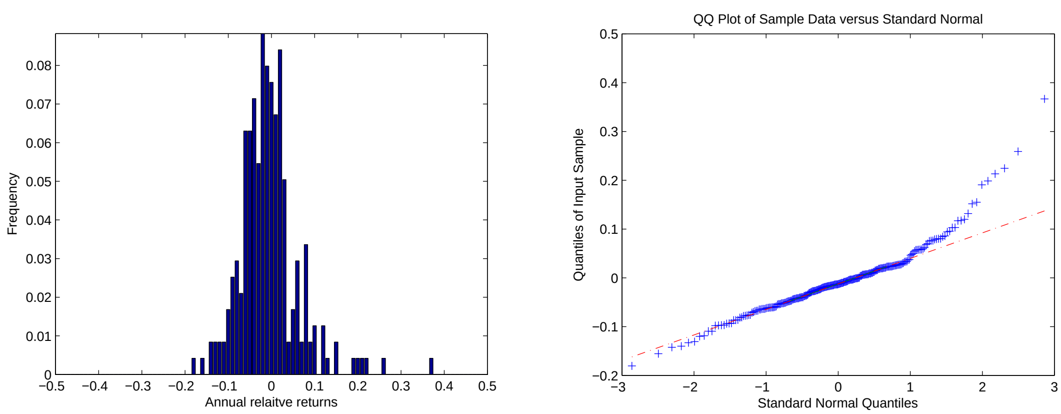
The table shows that top 10% funds constitute large portion of total asset values, up to 65%. It also shows that the percentage of total asset values in two methods are not significantly different. Although the 65% is not small, we believe that this is not an extreme value such that we need some other measure to analyze the relative returns under the same strategy.

Table 5: Managed asset values for fund-of-fund strategy

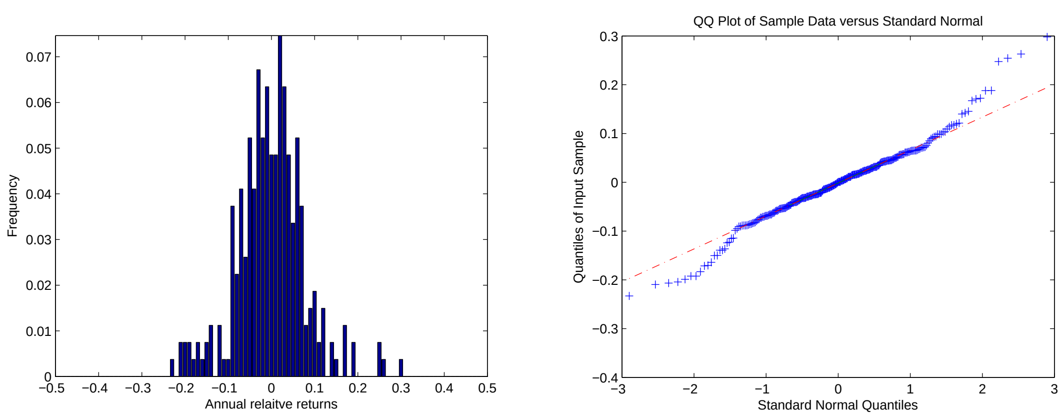
Ranking	Managed asset	Manages asset for 2004
Top 1%	33 %	38 %
Top 5%	55 %	58 %
Top 10%	65 %	67 %
Bottom 10%	0.5 %	1 %
Bottom 5%	0.1 %	0.2 %
Bottom 1%	0 %	0 %
Total Managed Asset	$\$1 \times 10^{11}$	$\$2 \times 10^{11}$

C. Distribution of Relative Returns from the Data

In this section, we carry out the analysis of Figure 1 in the main paper for the other hedge-fund strategies. In particular, we display histograms of the relative returns within each of these strategies and provide Q-Q plots comparing the empirical distribution to the normal distribution. It is pointed out by Lhabitant (2004), Tran (2006), Geman and Kharoubi (2003), Eling and Schuhmacher (2007), Kassberger and Kiesel (2006) that hedge fund returns or indexes have heavy-tails, which are not fitted by normal distribution. In contrast, although most returns do indeed show heavy tails, we find that relative returns within the global-macro and emerging-market strategies can be fit to the normal distribution; see Figure 1 in the main paper and Figure 2 below. (We omit dedicated-short-biased strategy since we only have 29 observations.)

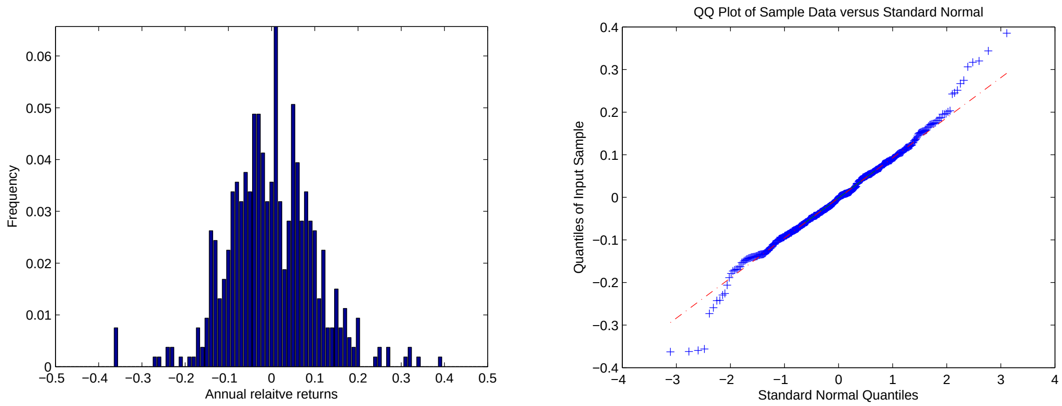


(a) Convertible ($\gamma = 0.44$)

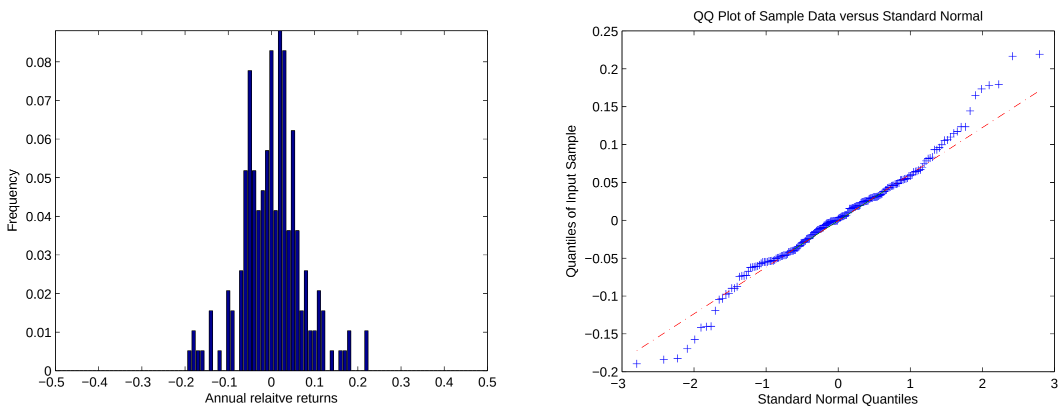


(b) Equity macro ($\gamma = 0.09$)

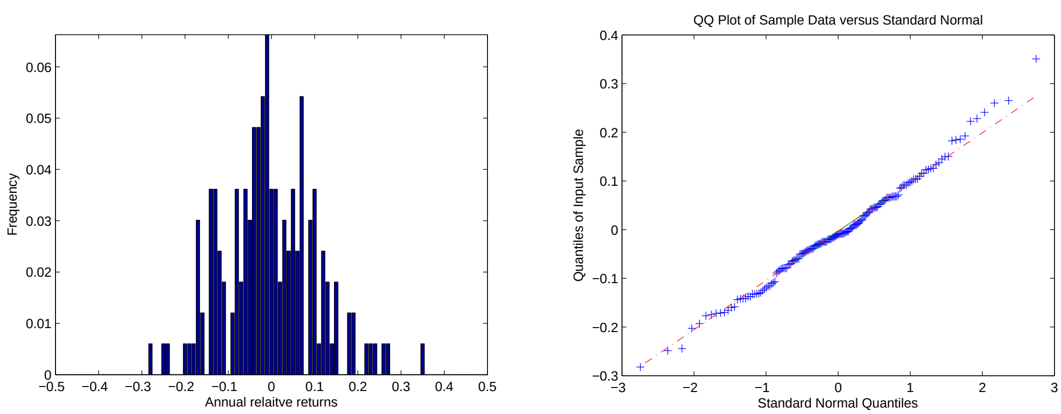
Figure 2: Relative-return distributions and Q-Q plots comparing the empirical distribution to the normal distribution for 10 strategies in TASS database from 2000-2004.



(c) Event driven ($\gamma = 0.24$)

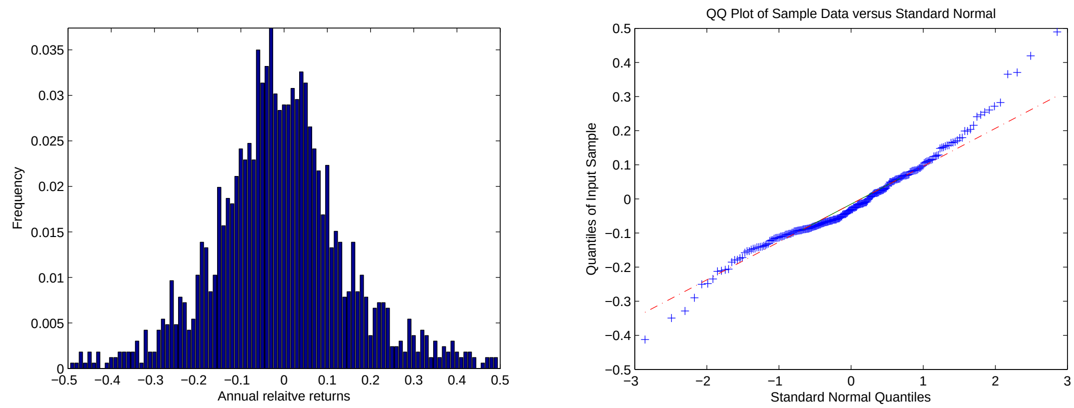


(d) Fixed income ($\gamma = 0.29$)

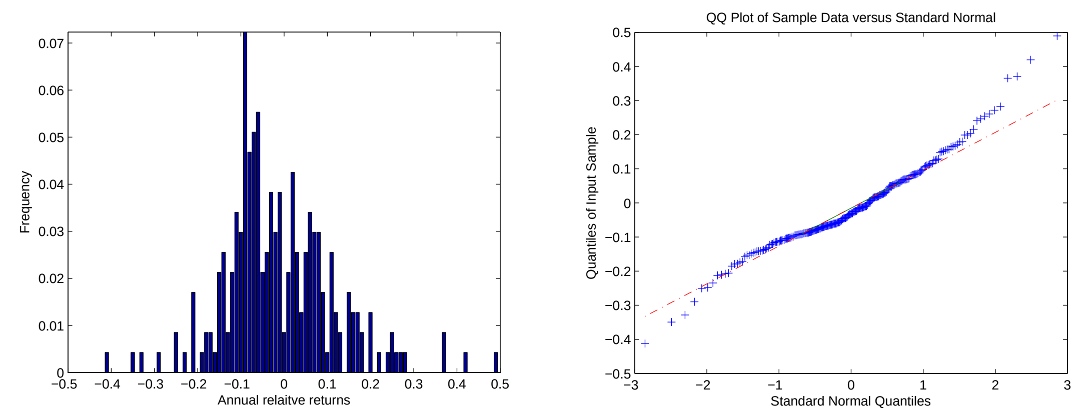


(e) Global macro ($\gamma = 0.13$)

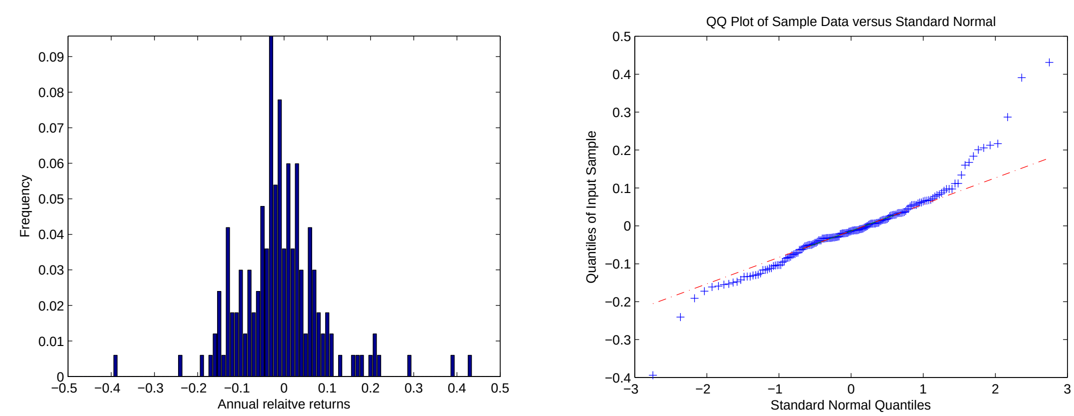
Figure 2: (Continue) Relative-return distributions and Q-Q plots comparing the empirical distribution to the normal distribution for 10 strategies in TASS database from 2000-2004.



(f) Long-short equity ($\gamma = 0.15$)



(g) managed future ($\gamma = 0.20$)



(h) Others ($\gamma = 0.48$)

Figure 2: (Continue) Relative-return distributions and Q-Q plots comparing the empirical distribution to the normal distribution for 10 strategies in TASS database from 2000-2004.

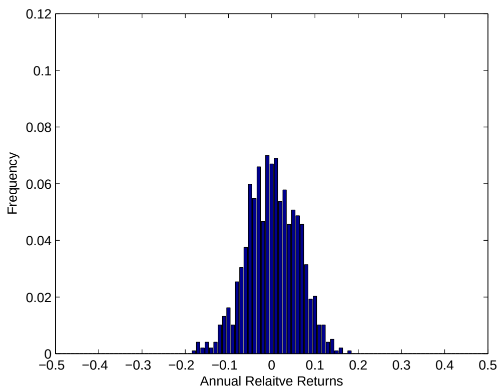
The table below shows results for the Lilliefors test. It tests the hypothesis that the sample comes from a normal distribution. The two distributions with relatively high p-values (greater that 0.05) from the Lilliefors test have distributions the look like the normal distribution in Figure 2, both directly and in the Q-Q plot .

Table 6: Lilliefors test results with 95 % significance level

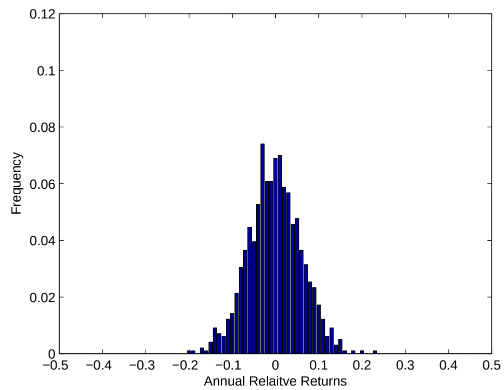
Strategies	Result	p-value
Convertible	Reject	0.0001
Equity Macro	Reject	0.0071
Event Driven	Accept	0.1204
Fixed Income	Reject	0.0424
Global Macro	Accept	0.3002
Long-short Equity	Reject	0.0001
Managed Future	Reject	0.0021
Other	Reject	0.0001

D. Constant-Persistence Normal-Noise Model Simulation

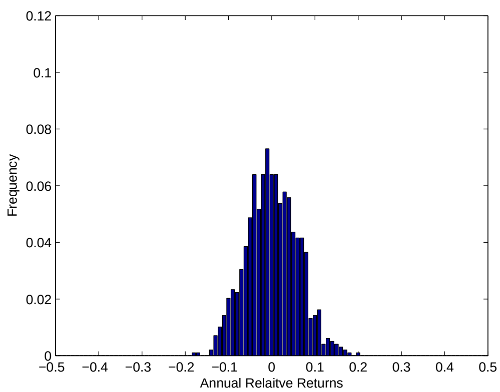
In this section, we show how the relative-return distribution in the constant-persistence normal-noise model depends on the sample size of the simulation. Since the relative returns we have from the data is limited, when fitting the relative-return distribution, it might be helpful to compare the empirical distribution to the estimated distribution with the sample size of the data. Figure 3 (a)-(c) illustrate estimated distributions, each with the same size of the data, 986, for the fund-of-fund strategy. We then do the same for the emerging-market strategy in Figure 3 (e)-(g) with sample size of 315. We also provide the estimated relative-return distribution with sample size of 10^6 in Figure 3 (d) and (h) in order to see how the shape of the estimated relative-return distribution changes as the sample size increases.



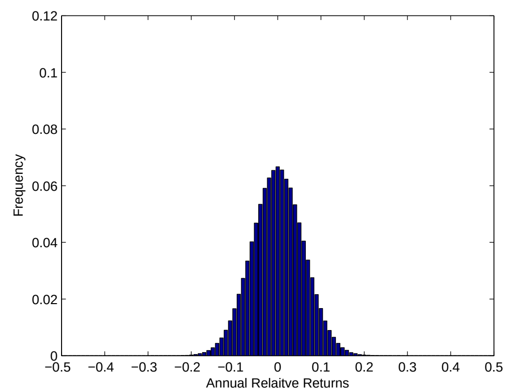
(a) 986 simulation of the model



(b) Relative-return distribution from 986 samples

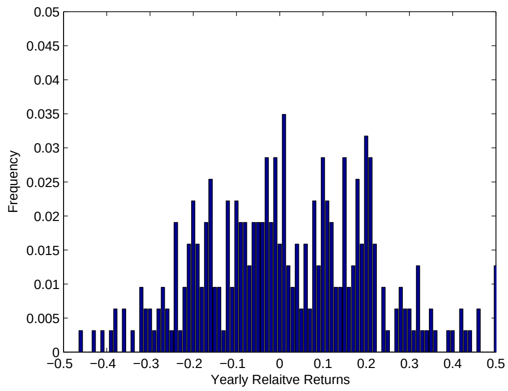


(c) Relative-return distribution from 986 samples

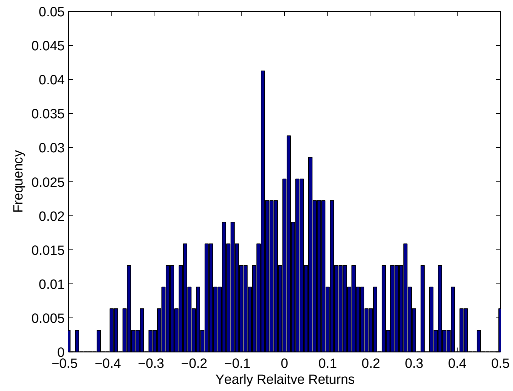


(d) Relative-return distribution from 10^6 samples

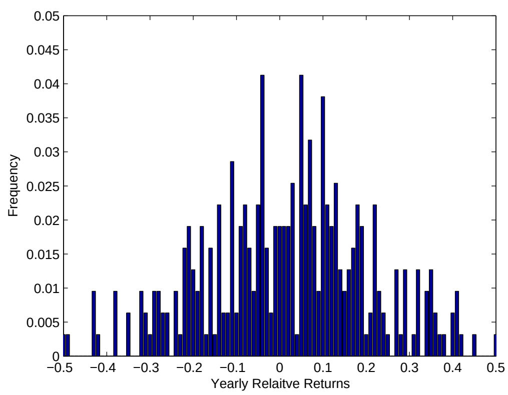
Figure 3: (a)(b)(c) The estimated relative-return distribution with the sample size of 986 in the constant-persistence normal-noise model with $\gamma = 0.33, \sigma_b = 0.0565$ for fund-of-fund strategy. (d) The estimated relative-return distribution with the sample size of 10^6.



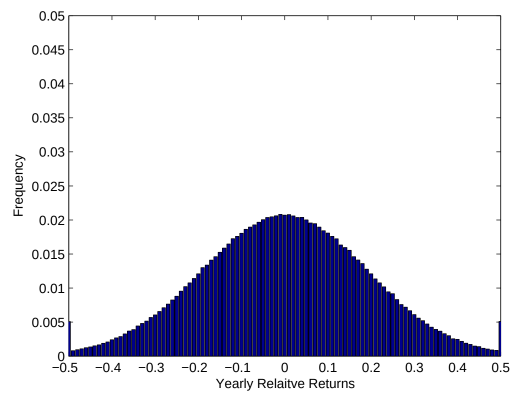
(e) 315 simulation of the model



(f) 315 simulation of the model



(g) 315 simulation of the model

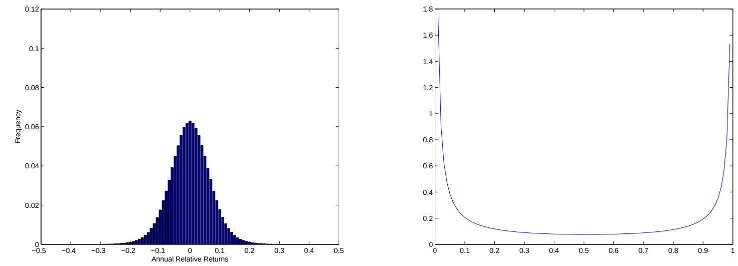


(h) $1,000,000$ simulation of the model

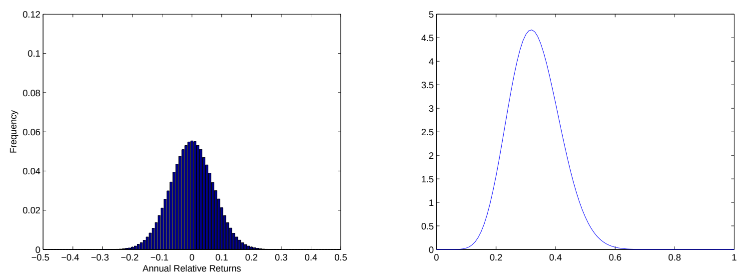
Figure 3: (Continue)(e)(f)(g) The estimated relative-return distribution with the sample size of 315 in the constant-persistence normal-noise model with $\gamma = 0.36, \sigma_b = 0.1797$ for emerging-market strategy (h) The estimated relative-return distribution with the sample size of 10^6.

E. Beta-Persistence Model Simulations

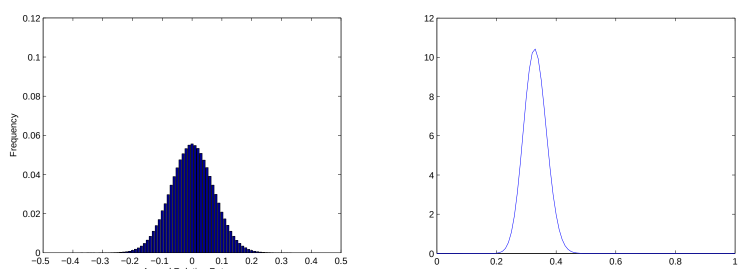
In this section, we illustrate how the beta-persistence model depends on the beta-distribution parameters α and β. It is observed that the overall relative-return distribution predicted by the model does not depend much on beta-distribution parameters. See, Figure 4 for the beta-persistence normal-noise model. The observation also holds for the other beta-persistence models with t and mixture noise.



(a) The beta-persistence normal-noise model with $\alpha = 0.03$ and corresponding beta PDF



(b) The beta-persistence normal-noise model with $\alpha = 10$ and corresponding beta PDF

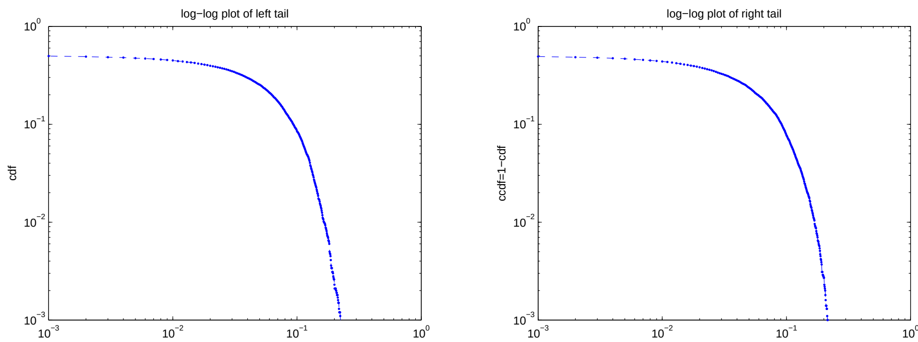


(c) The beta-persistence normal-noise model with $\alpha = 50$ and corresponding beta PDF

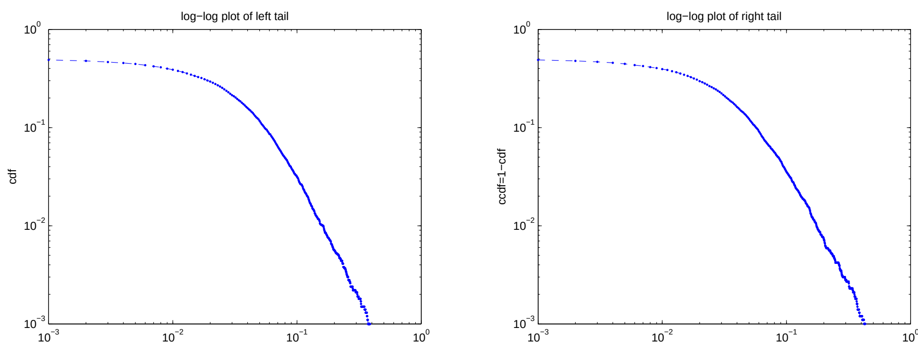
Figure 4: Simulation estimate of the relative-return distribution and the associated beta pdf from the beta-persistence normal-noise model for the fund-of-fund strategy with $\gamma = 0.33$, $\sigma = 0.0681$ and (a) $\alpha = 0.03$ and $\beta = 0.06$, (b) $\alpha = 10$ and $\beta = 20.30$, (c) $\alpha = 50$ and $\beta = 101.51$.

F. Log-Log Plots of Distribution Tails in Different Models

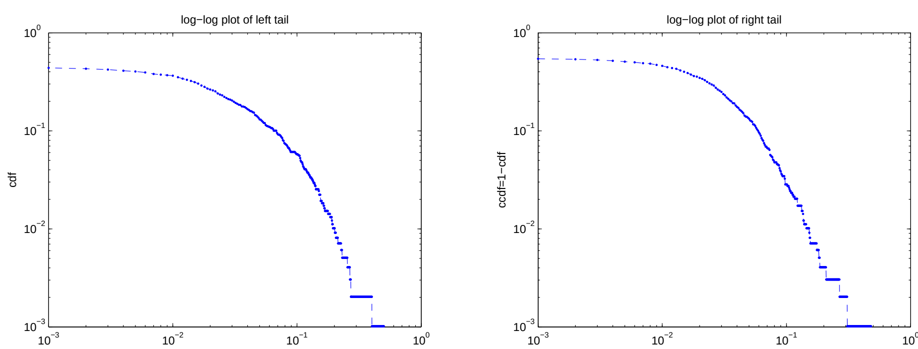
In this section, we plots the distribution tails for the normal, t, and mixture noise model in order to show the differences in their tail behavior. All except the normal have heavy tails, which is shown as linear behavior for larger values (at the right in each plot) in Figure 5.



(a) 10,000 simulation of constant-persistence normal-noise model



(b) 10,000 simulation of constant-persistence t-noise model

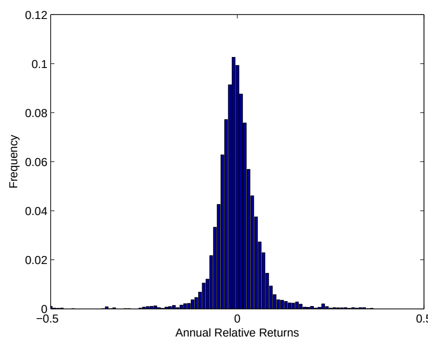


(c) 10,000 simulation of beta-persistence mixed-noise model

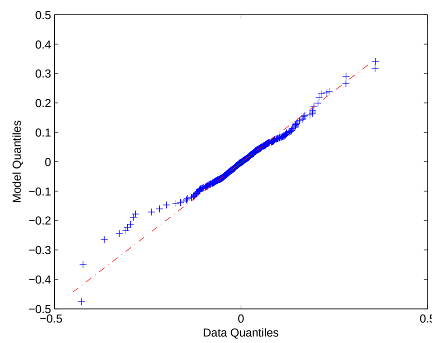
Figure 5: Log-log plots of the estimated relative-return distributions with sample size of 10^4 in the (a) constant-persistence normal-noise model, (b) constant-persistence t-noise model, (c) constant-persistence mixed-noise model.

G. The Beta-Persistence Empirical-Noise Model

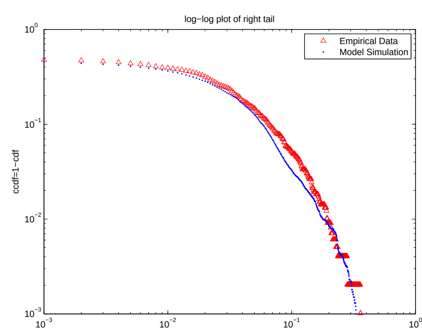
To seek a still better fit to the data within the beta-persistence class of models, we can let B_n have the observed empirical distribution for $X_n - \gamma X_{n-1}$, using the estimated value of γ. This automatically gives B_n and its estimated variance σ_b^2. It now goes further to directly match the shape, but sacrifices the explicit parametric form. In order to simulate B following the same distribution of B_n obtained from the data, we construct distribution function of B_n numerically. This is done by splitting the support of relative returns, $[-0.5, 0.5]$ equally and cumulatively count the number of returns falling each interval, from left to the right. As a numerical example, we construct distribution function of B_n from the relative returns within fund-of-fund strategy. Given the distribution function, we can generate B using inverse transform method; we generate uniform random variable and find the inverse value of given distribution function numerically. Figure 6 shows the distribution function of X based on the simulation of B constructed from empirically obtained B_n. As we see from the figure, the beta-persistence empirical-noise model also provides a good fit to the data.



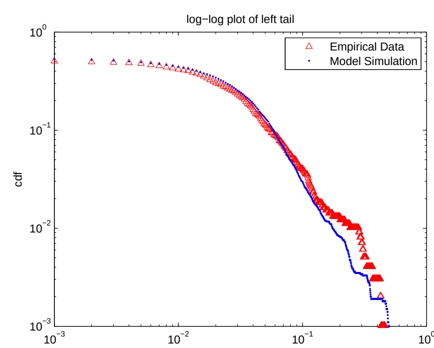
(a) Relative-return distribution



(b) Q-Q plot comparing the model to the data



(c) Left-tail log-log plot



(d) Right-tail log-log plot

Figure 6: Simulated samples from the beta-persistence empirical-noise model with $\gamma = 0.33$, $\alpha = 50$, $\sigma = 0.0681$ and the empirical relative-return distribution for the fund-of-fund strategy from the data.

Table 7 below shows the hitting probabilities from the beta-persistence empirical-noise model. It is observed that the maximum and minimum of 20 simulations of hitting probabilities cover the empirically observed hitting probabilities from the data. The large number (10^4) of simulation results in the fourth column of Table 7 also suggests that the beta-persistence empirical-noise model provides reasonable estimates of the hitting probabilities.

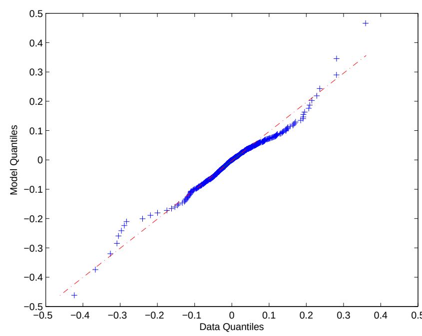
Table 7: Hitting probabilities of thresholds over a five-year period (2000-2004)

Level[1]	data[2]	empirical-noise	
		$N = 92^3$	$N = 10,000^4$
3 σ	0.0326	[0,0.0652]	0.0313±0.0034
2 σ	0.0761	[0.0217,0.1196]	0.0659±0.0049
1 σ	0.2363	[0.1630,0.3261]	0.2226±0.0082
-1 σ	0.2391	[0.1413,0.2826]	0.2021±0.0079
-2 σ	0.0542	[0.0109,0.0870]	0.0477±0.0042
-3 σ	0.0326	[0,0.0543]	0.0271±0.0032

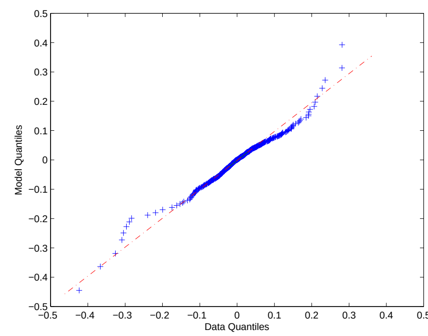
1. $\sigma = 0.0681$, the observed standard deviation of the fund-of-fund relative returns.
2. Number of funds that have ever hit the level for 2000-2004 divided by total 92 funds in 2000.
3. Minimum and maximum of the probabilities from 20 simulations with sample size of 92 initially.
4. 95 % confidence interval of hitting probability from simulation with sample size of 10,000 initially

H. Constant-Persistence Stable-Noise Model Simulations

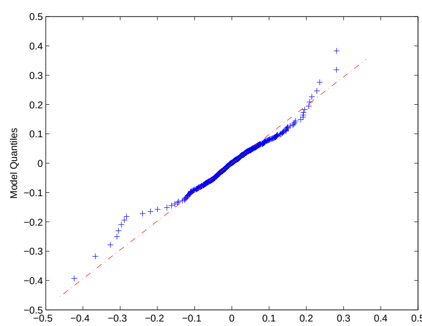
In this section, we show how the relative-return distribution in the constant-persistence stable-noise model depends on parameter β in the stable distribution. Figure 7 shows Q-Q plots and log-log plots of the left and right tails of the estimated distributions for $\beta = -0.2, -0.1, 0,$ and 0.1. It is observed that the constant-persistence stable-noise model with $\beta = -0.1$ fits the Q-Q plot well whereas the left and right tails of the distribution are approximated well with $\beta = 0.1$. Overall, $\beta = 0$ fits both the Q-Q plot and the left and right tails relatively well at the same time. It is hard to find stable random variable parameters that can fit both Q-Q plot and log-log figures at the same time. It is because the shape of the stable distribution cannot match the observed distribution exactly. However, the constant-persistence stable-noise model still provides a reasonably good fit to the data with fewer parameters than the other models, such as the beta-persistence mixed-noise model.



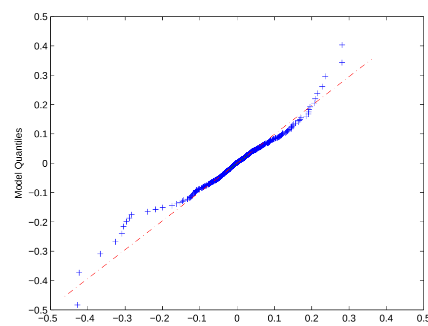
(a) Q-Q plot with $\beta = -0.2$



(b) Q-Q plot with $\beta = -0.1$

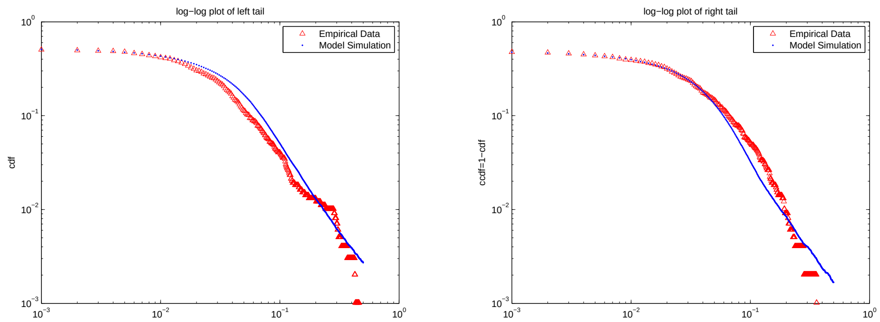


(c) Q-Q plot with $\beta = 0$

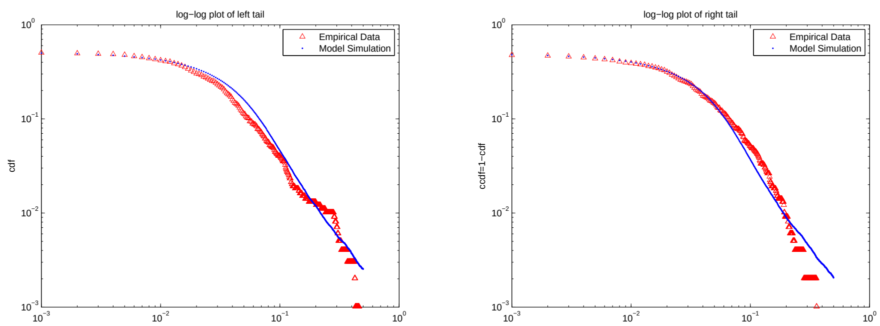


(d) Q-Q plot with $\beta = 0.1$

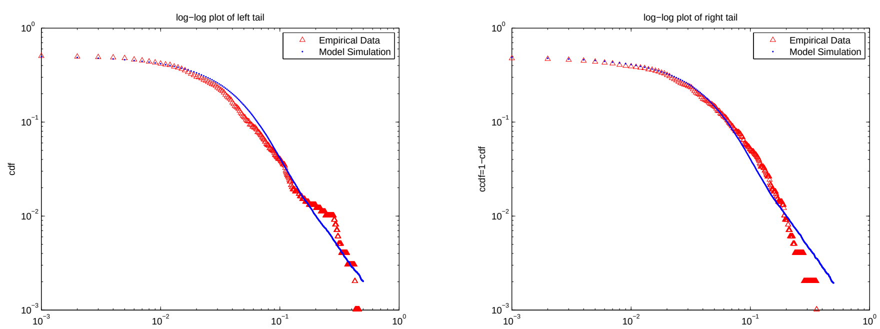
Figure 7: Q-Q plots and Log-log plots of left and right tails of the relative-return distributions from the constant-persistence stable-noise model with $\alpha = 1.6, k = 0.0029$ for $\beta = -0.2, -0.1, 0,$ and 0.1.



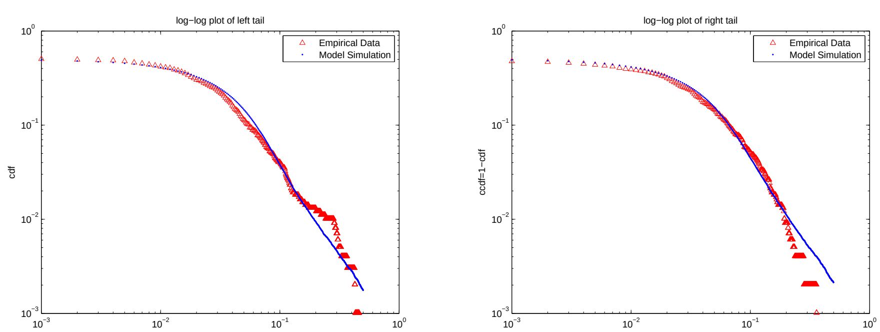
(e) Constant-persistence normal-noise model (Log-log plots) with $\beta = -0.2$



(f) Constant-persistence normal-noise model (Log-log plots) with $\beta = -0.1$



(g) Constant-persistence normal-noise model (Log-log plots) with $\beta = 0$



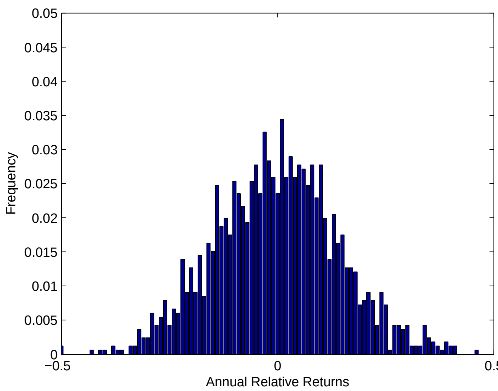
(h) Constant-persistence normal-noise model (Log-log plots) with $\beta = 0.1$

Figure 7: (Continued) Q-Q plots Log-log plots of left and right tails of the relative-return distributions from the constant-persistence stable-noise model with $\alpha = 1.6, k = 0.0029$ for $\beta = -0.2, -0.1, 0$, and 0.1.

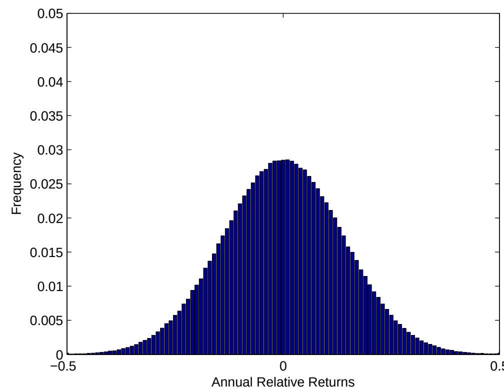
I. Analysis of Relative Returns within the Long-Short Equity Strategy

In this section, we fit the relative returns within the long-short equity strategy. Table 1 in the main paper shows that this strategy has the largest sample size. Thus it is natural to fit our SDE model to the data in this case. Although we observe relative large number of observations from the data for this strategy, we see that the relative returns does not have high performance persistence.

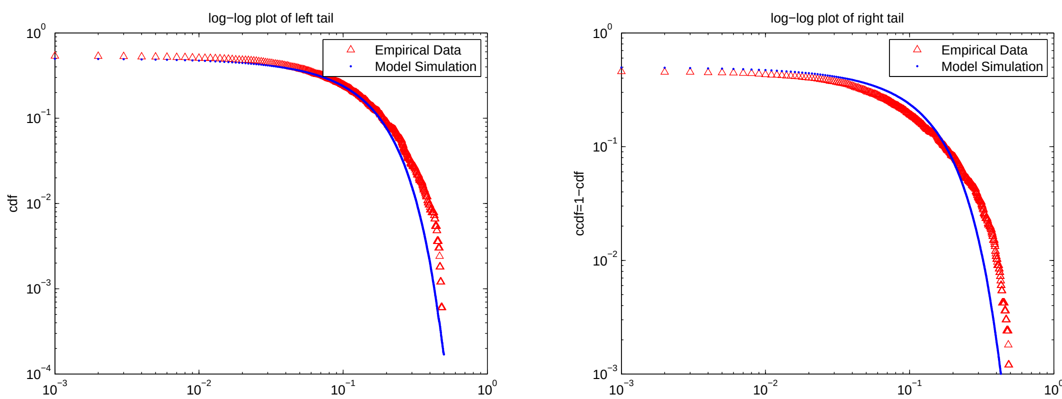
The Q-Q plot of the relative returns in Figure 2 (f) suggests that the distribution does not have heavy tails. That is also supported in the log-log plots of the distribution tails in Figure 8 since both the left and right tails do not end with a linear line and instead decrease quickly in the right side of the Figure 8 (c). Thus, we start from normal-noise model to fit the data. As observed in Table 2 in the main paper, the ratio σ/σ_b from the data and model do not match. Thus, we use the beta-persistence normal-noise model first with $\alpha = 50$. For given $\sigma = 0.1520$ and $\gamma = 0.15$ from the data, we calibrate other parameters β, σ_a and σ_b, following §6 of the main paper. Figure 8 (a) and (b) show the estimated relative-return distribution. It is observed from Figure 8 (d) that the Q-Q plot of the model to the data is close to a linear line with slope 1. Thus, we conclude that the relative-return distribution is approximated well by the beta-persistence normal-noise model for the long-short equity strategy.



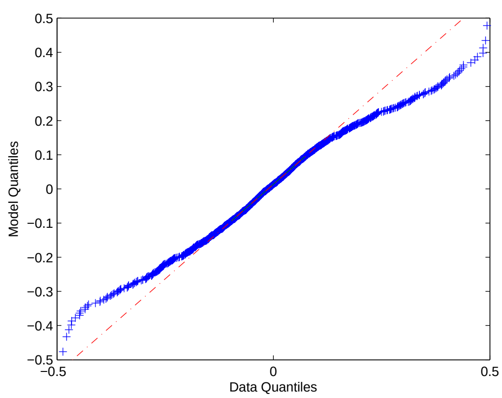
(a) Relative-return distribution with 1658 samples



(b) Relative-return distribution with 10^6 samples



(c) Log-log plot of the left and right tails from the beta-persistence normal-noise model

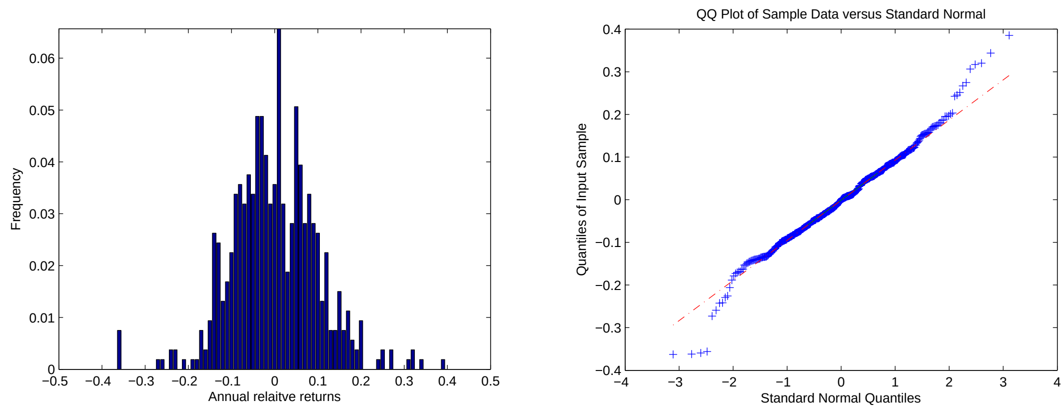


(d) Q-Q plot (simulated distribution to empirical one)

Figure 8: Relative returns simulated from the beta-persistence normal-noise model with $\alpha = 50$, $\sigma = 0.1520$, $\gamma = 0.15$ for the long-short equity strategy.

J. Analysis of Relative Returns within the Event-Driven Strategy

In this section, we analyze another single strategy whose relative-return distribution has heavy tails. In particular, we analyze the event-driven strategy since it has relative big sample size (533) and high persistence factor ($\gamma = 0.24$). The Q-Q plot in Figure 9 shows that the relative-return distribution has heavier tails than a normal distribution. We thus proceed using our beta-persistence t-noise and constant-persistence stable-noise models to fit the data.



(a) Event-driven strategy ($\gamma = 0.24$)

Figure 9: Distribution of relative returns from event-driven strategy and Q-Q plot comparing the distribution to the normal distribution.

J.1. Beta-Persistence t-Noise Model

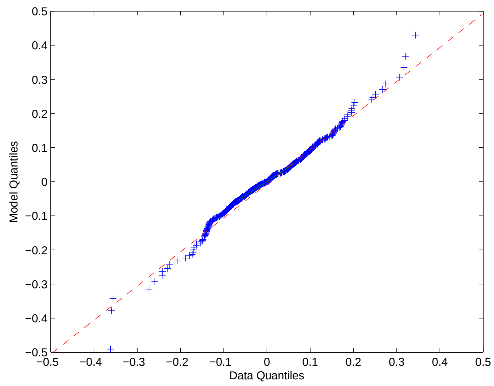
In this section, we test whether the beta-persistence t-noise model can fit the data for the event-driven strategy. Recall that in the beta-persistence t-noise model, once α is set, then the other parameter β in the beta random variable is determined to fit the mean ($\gamma = 0.24$). Just as we did for the fund-of-fund strategy, we set $\alpha = 50$, so that the persistence random variable is relatively narrowly distributed around $\gamma = 0.24$. We then set the degrees of freedom in the t random variable to fit the distribution of relative returns from the data. Another parameter k in the model is determined to fit the standard deviation of X_n ($\sigma = 0.1007$). We find that $v = 3.5$ fits the distribution well.

From Figure 10, we observe that the quantiles in the Q-Q plot comparing the samples from the model to the data coincide reasonably well. We obtain p value of 0.1349 from Kolmogorov-Smirnov two sample test. Thus, we cannot reject the hypothesis that the simulated returns and empirical returns come from the same distribution.

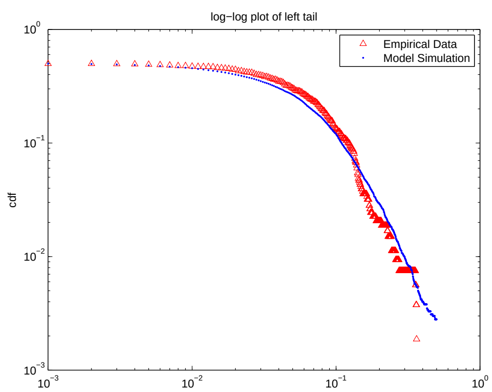
J.2. Constant-Persistence Stable-Noise Model

In this section, we test whether the constant-persistence stable-noise model provides a good fit the data. In order to test that, we measure the quantiles of X_n and B_n that directly come from $X_n - \gamma X_{n-1}$, using previous estimate for the persistence factor γ. Table 8 shows that the ratios of quantiles from X and B are roughly equal to 1.3. We thus proceed the model fitting by assuming that $c = 1.3$.

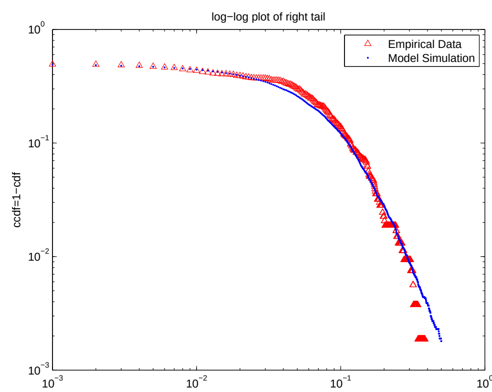
Given $c = 1.3$, we now compare X_n and cB_n from the data for the event-driven strategy. Figure 11 shows the histograms of X_n and cB_n from the data, which look similar. We also conducted Kolmogorov-Smirnov two-sample test and obtained a p-value of 0.2834. Thus we cannot reject the hypothesis that these two sets of samples come from the same distribution. The Q-Q plot also



(a) Q-Q plot comparing the model the the data



(b) Log-log plot of left tails



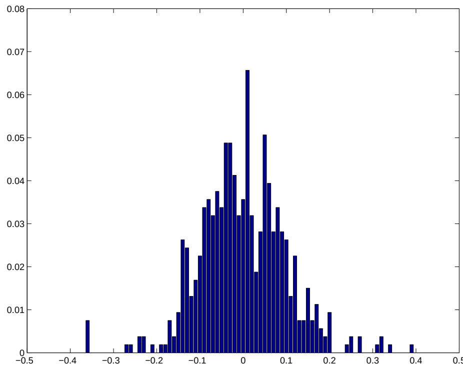
(c) Log-log plot of right tails

Figure 10: The beta-persistence t-noise model 10^4 number of simulation with $\alpha = 50, v = 3.5, \gamma = 0.24$ comparing to the data for event-driven strategy.

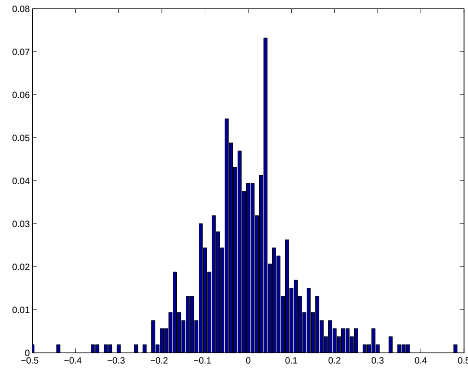
Table 8: The Quantile Differences of X_n and B_n and Their Ratios

Quantile Difference[1]	X_n	B_n	Ratio [2]
$55\% - 45\%$	0.0259	0.0207	1.2533
$60\% - 40\%$	0.0460	0.0372	1.2378
$65\% - 35\%$	0.0783	0.0578	1.3552
$70\% - 30\%$	0.1012	0.0703	1.4396
$75\% - 25\%$	0.1270	0.0921	1.3789
$80\% - 20\%$	0.1580	0.1204	1.3132
$85\% - 15\%$	0.1878	0.1587	1.1832
$90\% - 10\%$	0.2935	0.2067	1.1587
$95\% - 5\%$	0.3051	0.2876	1.0610

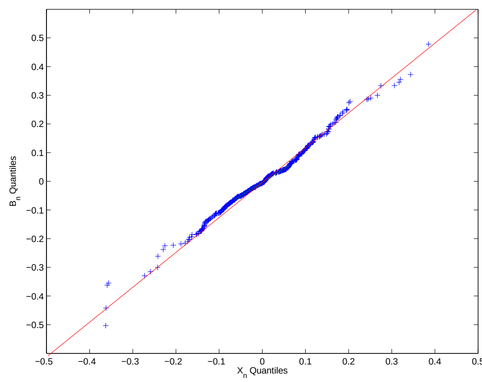
1. Difference between two quantile values.
2. Ratio: Quantile Difference for X /Quantile Difference for B.



(a) Event-driven strategy X_n



(b) Event-driven strategy $cB_n = c(X_n - \gamma X_{n-1})$



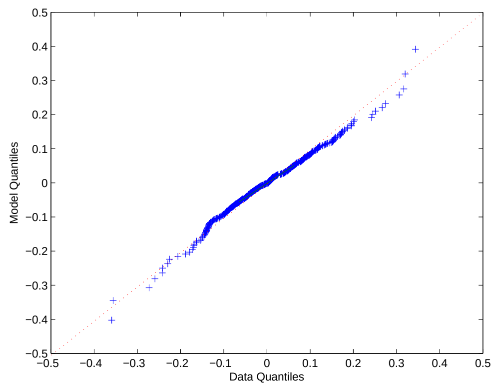
(c) Q-Q plot comparing X_n and cB_n

Figure 11: X_n and cB_n from event-driven strategy and Q-Q plot comparing the distribution of X_n and cB_n with $c = 1.3$ for event-driven strategy.

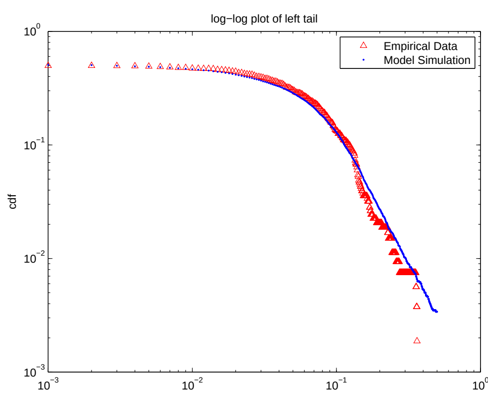
shows that the quantiles from the distributions of the samples from the model and the data coincide with each other remarkably well.

Figure 12 shows that the constant-persistence stable-noise model fits the relative returns within the event-driven strategy reasonably well with stable-distribution parameters $\alpha = 1.75$, $\beta = -0.2$ and $\kappa = 0.055$. The Q-Q plots in the figure show that the quantiles of the distributions of the samples from the model and the data coincide well. Also, log-log plots of the left and right tails show that the tail behaviors of the distribution of the samples from the model approximate the distribution of the samples from the data reasonably well.

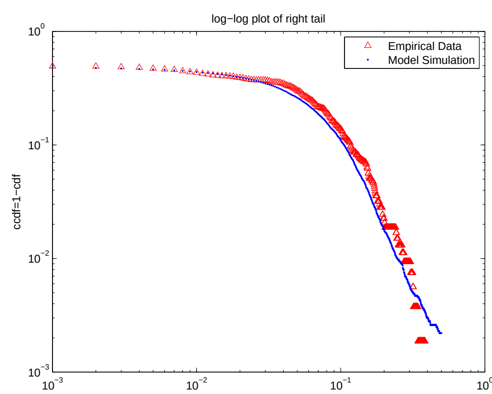
We test if the c and α in Figure 12 and γ reasonably fit (7.1) in the main papr. We observe that $c^\alpha = 1.58$ and $1/(1-\gamma^\alpha) = 1.08$ coincide only roughly. Nevertheless, in summary, we conclude that the fitting to a heavy-tailed distribution works reasonably well, given the limited data.



(a) Q-Q plot comparing the model to the data



(b) Log-log plot of left tails



(c) Log-log plot of right tails

Figure 12: Event-driven strategy Q-Q plot comparing the distribution of 533 samples from the data and 10^4 samples from the constant-persistence stable-noise model with $\alpha = 1.75, \beta = -0.2, \gamma = 0.24, k = 0.055$ for event-driven strategy.