THE JOURNAL OF FINANCE • VOL. LXIV, NO. 2 • APRIL 2009

Hedge Fund Risk Dynamics: Implications for Performance Appraisal

NICOLAS P.B. BOLLEN and ROBERT E. WHALEY*

ABSTRACT

Accurate appraisal of hedge fund performance must recognize the freedom with which managers shift asset classes, strategies, and leverage in response to changing market conditions and arbitrage opportunities. The standard measure of performance is the abnormal return defined by a hedge fund's exposure to risk factors. If exposures are assumed constant when, in fact, they vary through time, estimated abnormal returns may be incorrect. We employ an optimal changepoint regression that allows risk exposures to shift, and illustrate the impact on performance appraisal using a sample of live and dead funds during the period January 1994 through December 2005.

HEDGE FUND MANAGERS are generally free to change trading strategies, leverage, and allocations to different asset classes.[1] These changes can be in response to macro-economic conditions. During the weeks surrounding the Long Term Capital Management crisis, for example, hedge fund managers reduced leverage and reallocated portfolios to less volatile instruments. The changes can also be in response to arbitrage opportunities, as illustrated by the cycles of M&A activity and the corresponding level of risk arbitrage conducted by hedge funds. From an investor's perspective, the dynamic nature of hedge funds complicates due diligence activities including risk management and performance appraisal.

The standard approach to evaluate fund performance is to regress fund returns on risk factors that proxy for different trading strategies. Fung and Hsieh (2001), for example, develop factors to represent the payoffs of trend-following strategies, Mitchell and Pulvino (2001) generate a return series to represent a risk arbitrage strategy, and Agarwal and Naik (2004) use options-based returns to provide a flexible functional form to represent unspecified nonlinear equity strategies. In general, the coefficients in these analyses are assumed to be constant, and therefore do not allow for shifts in trading strategies, leverage, or allocations to different asset classes. If coefficients are assumed constant when,

*The Owen Graduate School of Management, Vanderbilt University. The authors thank seminar participants at the Commodity Futures Trading Commission, Emory University, Indiana University, Vanderbilt University, and the 2008 American Finance Association annual meeting for helpful suggestions. Cliff Ball, Bruce Cooil, Chris Kirby, Andrew Patton, and Jay Shanken provided useful comments. Research support from the Managed Funds Association is gratefully acknowledged.

[1] We use the term "hedge fund" when referring to all managed portfolios. In the empirical analysis, we make a distinction between hedge funds, funds of funds, commodity trading advisors, and commodity pool operators.

in fact, they are time varying, parameter estimates will be unreliable and thus will shed no light on the skill with which a manager adjusts his strategy mix.

The purpose of this study is to investigate hedge fund risk dynamics. We have two goals. First, we would like to determine the most effective econometric technique for measuring changes in hedge fund risk. We find that a robust, simple regression-based method works best. Second, we would like to gauge the economic significance of hedge fund risk dynamics. We characterize the extent to which hedge funds of different styles engage in strategy shifts, and measure the error in performance appraisal that occurs when regression parameters are assumed constant.

We apply two existing empirical methods that allow for time-series variation in risk exposures, and determine which is superior in the hedge fund context. The first is an optimal changepoint regression that allows for a discrete number of shifts in parameter values. The dates on which shifts occur are selected along with parameter values in order to maximize the regression's explanatory power. We limit our analysis to one shift in parameters for each fund. The second is a stochastic beta model that uses an autoregressive process for risk exposures. We study the two models' power to reject the null hypothesis of constant risk exposures by generating data under both alternatives. Since the optimal changepoint regression performs better overall, we use it to study the prevalence and economic significance of changes in hedge fund risk exposures.

Using data covering the period 1994 to 2005, we find that approximately 40% of the hedge funds in our sample feature a statistically significant shift in risk exposures. The probability of observing a significant shift increases with the number of observations available for a given fund—a consequence of increased statistical power and a greater likelihood that changes in market conditions will warrant a shift in risk exposures. An analysis of the duration of strategies indicates that strategy shifts are more common early in a fund's history. Furthermore, strategy shifts in live funds appear to be associated with higher Sharpe ratios, suggesting a relation between the ability of a manager to change strategies and fund performance.

To illustrate the economic significance of risk dynamics, we compare ex-post performance rankings based on three alternative models: the optimal changepoint regression, a constant parameter model using all available observations, and a constant parameter rolling-window model using the most recent 24 observations. We limit the analysis to funds that demonstrate a significant switch in risk exposures. When the changepoint rank is compared to the constant parameter rank using all observations, a scatter plot indicates widespread divergence. A regression of changepoint rank on the constant parameter rank yields an adjusted-R^2 of only 28%. When the changepoint rank is compared to a constant parameter rank using the most recent 24 observations, the fit is much tighter, but still yields an adjusted-R^2 of just 59%, suggesting that performance appraisal can be significantly distorted when changes in risk exposure are not appropriately measured.

The paper is organized as follows. In the first section, we review related literature. Section II discusses the two empirical methods that we use to measure changes in risk. Section III describes the data. Section IV shows results of a competition between the two empirical methods. In Section V, we analyze a large database of individual hedge funds to determine the frequency and magnitude of changes in risk. Section VI studies the causes and effects of changes in risk exposure, and Section VII offers concluding remarks.

I. Related Literature

Jensen (1968) is usually credited with the first application of a linear factor model to the problem of performance appraisal of managed portfolios. He measures mutual fund performance using the alpha from a regression of a fund's excess returns on those of the market, motivated by the capital asset pricing model (CAPM) of Sharpe (1964). Subsequent empirical work challenges the ability of the CAPM to capture systematic variation in asset returns. Consequently, mutual fund performance is now typically measured using the Fama–French three-factor model or Carhart's (1997) four-factor model. An important difference between the CAPM and these multifactor extensions is that the additional factors are not explicitly motivated by an asset-pricing model. Carhart notes:

> The 4-factor model is consistent with a model of market equilibrium with four risk factors. Alternately, it may be interpreted as a performance attribution model, where the coefficients and premia on the factor-mimicking portfolios indicate the proportion of mean return attributable to four elementary strategies . . . I employ the model to "explain" returns, and leave risk interpretations to the reader. (p. 61)

In this paper, we use the alpha from various multifactor models to measure the performance of hedge funds, and interpret it as the mean excess return generated by the fund manager beyond that attributable to investment in the strategy-mimicking factors.

In all linear factor models, slope coefficients measure the fund's exposure to the included factors. Jensen (1968) notes, in the context of the CAPM, that a constant slope implies that the portfolio's risk level is stationary over time. He then adds:

> However, we know this need not be strictly true since the portfolio manager can certainly change the risk level of his portfolio very easily . . . Indeed, the portfolio manager may consciously switch his portfolio holdings between equities, bonds and cash in trying to outguess the movements of the market. (p. 394)

Jensen then argues that if the portfolio's risk level changes over time, the alpha of the constant parameter model will reflect the average incremental return generated by a manager's timing ability.

Other research extends the constant parameter model used by Jensen to explicitly allow for a strategy of market timing. Since successful market timers increase market exposure prior to a market advance, and decrease exposure prior to a decline, market timing can be represented by a function relating a fund's market exposure to the level of the market return. Treynor and Mazuy (1966), for example, specify fund returns to be a quadratic function of market returns:

$$R_t = \alpha + \beta R_{m,t} + \gamma R_{m,t}^2 + \varepsilon_t, \tag{1}$$

where R_t is the excess return of a fund and $R_{m,t}$ is the excess return of the market. If a manager is a successful market timer, the expected value of γ is greater than zero. With $\gamma > 0$ and $R_{m,t}^2 > 0$, market timing provides a positive incremental return, independent of the sign of the market return. In a similar vein, Henriksson and Merton (1981) use an indicator variable to allow for distinct levels of exposure:

$$R_t = \alpha + \beta R_{m,t} + \gamma I_t R_{m,t} + \varepsilon_t, \tag{2}$$

where I_t equals one if $R_{m,t} > 0$ and zero otherwise. For a successful market timer, γ is, again, greater than zero. In a market advance, γ is the incremental exposure to market risk. During a decline, the risk exposure reduces to β. Ferson and Schadt (1996) develop conditional versions of both models to control for trading activity based on public information.

Exploiting an insight of Jagannathan and Korajczyk (1986), Fung and Hsieh (1997) argue that hedge fund returns feature option-like payoffs relative to the return of underlying assets, consistent with dynamic trading strategies qualitatively identical to market timing. To capture the option-like payoffs of dynamic trading, Fung and Hsieh (2001) create "style factors" that mimic the time-series properties of trend-following strategies. They create five factors with returns created by using combinations of exchange-traded put and call options in stock, bond, interest rate, currency, and commodity markets. Similarly, Agarwal and Naik (2004) use the returns of buy-and-hold investments in at-the-money and out-of-the-money index put and call options to generate a flexible functional form that captures the dynamic trading of a hedge fund manager. In both cases, the authors have eliminated the need to specify functional forms in factor model regressions as in equations (1) and (2). Instead, a hedge fund's trading strategy can be characterized by a constant exposure to the appropriate style factors:

$$R_t = \alpha + \beta^{\mathrm{T}} F_t + \varepsilon_t, \tag{3}$$

where β^{T} is the transpose of a vector of factor loadings and the vector F contains observations of the factors.

In our empirical analysis, we employ two sets of factors in parallel. The first set consists of the five trend-following factors used in Fung and Hsieh (2001), the three Fama–French factors, the change in the 10-year Treasury yield, and the change in the credit spread. The latter two are found to be relevant for some funds in Fung and Hsieh (2004). While the trend-following factors are constructed to model dynamic trading in the underlying assets, the Fama–French factors represent constant exposure to the market portfolio, size premium, and value premium. To allow for time variation in exposure to the size and value premia, we add squared terms for these factors, analogous to the Treynor and

Mazuy (1966) quadratic regression. We do not add a squared market term since the trend-following stock factor already captures dynamic exposure to the market.

The first set of factors have been widely adopted in the hedge fund literature, although it is open to debate whether they represent trading strategies that can be mimicked in the spirit of Carhart (1997) or Sharpe (1992). Thus, as a robustness check, we create a second set of factors that represent trading strategies that can be mimicked, that is, long positions in liquid futures contracts. As with the size and value premia, we add a squared term when using the futures contract factors, that is,

$$R_t = \alpha + \beta^{\mathrm{T}} F_t + \gamma^{\mathrm{T}} F_t^2 + \varepsilon_t, \tag{4}$$

where the vector F contains observations of the returns of a buy-and-hold investment in futures contracts on different assets, where the positions roll as maturities near as described in Section III.

The style factors and the futures contract factors described above reflect specific strategies consisting of a particular type of time variation in the exposure of a hedge fund to the price movements of underlying assets. Several studies, however, present evidence that hedge funds' strategies themselves vary over time. In the context of the style factors, this can be expressed as:

$$R_t = \alpha + \beta_t^{\mathrm{T}} F_t + \varepsilon_t. \tag{5}$$

Fung and Hsieh (1997) use principal components analysis to determine the dominant styles in hedge funds. They investigate the stationarity of the identified style factors by conducting the analysis over two subperiods of the data. They report that style factors vary somewhat across the subperiods and suggest that this may reflect aggregate changes in trading strategies. Fung and Hsieh (2004) analyze the time-series properties of regression residuals to identify two breakpoints in aggregate exposure to style factors: September 1998, which they attribute to the LTCM debacle, and March 2000, which they attribute to the end of the internet bubble. They find that a hedge fund index's factor loading on the S&P 500 index drops by half over the resulting subperiods, consistent with a reduction in equity exposure during the bear market. These aggregate breakpoints are also used in Agarwal et al. (2006) and Fung et al. (2008). While these breakpoints are important evidence that hedge fund strategies change over time, they do not provide much insight into the time variation of individual hedge fund exposures, which is the topic we turn to next.

II. Empirical Methodology

Our central challenge is to determine whether and to what extent individual hedge funds feature dynamic exposure to underlying factors. This section reviews two frameworks within which we can conduct such an analysis. Subsection A introduces a model of structural change featuring factor loadings that can shift discretely across different time periods. Subsection B describes a model in which factor loadings are state variables governed by a first-order autoregressive process. Subsection C discusses pros and cons of both approaches to modeling hedge fund risk dynamics.

A. Discrete Structural Change

One way to model changes in a hedge fund's exposure to underlying factors is to assume that exposures undergo discrete shifts, where the timing of the shifts is not known and must be inferred from the data.[2] Several empirical methods to test for discrete shifts are available. Brown, Durbin, and Evans (1975), for example, develop a graphical analysis that seeks to identify departures from constancy in regression parameters and establish significance tests by measuring properties of regression residuals. Time variation in the properties of regression residuals indicates that underlying regression parameters have changed. In particular, Brown, Durbin, and Evans compute the cumulative sum (CUSUM) of recursive residuals and the cumulative sum of squared recursive residuals, and establish bounds for these sums under the null hypothesis of constant regression parameters. When the CUSUM breaches these bounds, one can reject the null hypothesis. Fung and Hsieh (2004) use a CUSUM test to determine that the Hedge Fund Research (HFR) Fund of Funds Index has time-varying exposure to seven underlying factors.

Andrews, Lee, and Ploberger (1996, hereafter ALP) develop an alternative to the CUSUM, which they label the *changepoint regression*. ALP show that under certain conditions, similar to the ones we face in hedge fund research, the changepoint regression delivers power superior to the CUSUM test. In particular, they study performance of the tests in simulations with 120 datapoints and two regressors, which approximate the number of monthly observations and number of factors that we use to model hedge fund returns. They report that for the majority of parameterizations studied, the changepoint regression correctly rejects the null substantially more often than the CUSUM test. For this reason, we choose to use the changepoint regression to study changes in hedge fund risk exposures.

The changepoint regression with a single changepoint can be written as

$$\begin{aligned} R_t &= \alpha_0 + \beta_0^{\mathrm{T}} F_t + \varepsilon_t && \text{for } t = 1, \ldots, T\pi \\ R_t &= \alpha_0 + \alpha_1 + \left(\beta_0^{\mathrm{T}} + \beta_1^{\mathrm{T}}\right) F_t + \varepsilon_t && \text{for } t = T\pi + 1, \ldots, T, \end{aligned} \tag{6}$$

where the unknown changepoint π satisfies $0 < \pi < 1$, $T\pi$ is an integer, and $\varepsilon_t \sim N(0, \sigma^2)$. The methodology developed by ALP allows for an arbitrary number of changepoints; however, we implement their approach using just one. While this constraint is likely binding for some hedge funds, the limited histories of most funds make estimation with multiple changepoints infeasible.

[2] This is a situation in which the econometrician may have less information than the investor. Large institutional investors may learn of strategy shifts directly from conversations with fund managers.

When all parameters are allowed to undergo change, as in equation (6), the situation is known as *pure structural change*. When some parameters are fixed, the situation is known as *partial structural change*. Since there is no reason a priori to restrict some exposures and not others, we focus on the former. The null and alternative hypotheses are

$$\begin{aligned} &\mathrm{H}_0\text{:} \quad \alpha_1 = \beta_1 = 0 \\ &\mathrm{H}_1\text{:} \quad \alpha_1 \neq 0 \text{ or } \beta_1 \neq 0. \end{aligned} \tag{7}$$

The test derived in ALP is an average of Wald tests (in the case of known residual variance) or an average of Likelihood Ratio tests (in the more realistic case of unknown residual variance), where the averages are computed over all permissible change dates. For the case of unknown residual variance, which is the situation faced in analysis of hedge fund returns, ALP define the regression F-statistic for a single change date π as

$$F(\pi) = \frac{[Q^* - Q(\pi)](T - 2\nu)}{Q(\pi)\nu}, \tag{8}$$

where $Q(\pi)$ is the sum of squared errors for the unrestricted case allowing for a parameter shift at date π; Q^* is the restricted case, that is, the standard least squares regression; and $\nu - 1$ is the number of factors. The test for a significant parameter shift is a weighted average of the F-statistics for each permissible change date. The implementation of the test requires specification of a scale parameter, which directs power against alternative hypotheses with parameter changes of varying magnitudes, and a weighting function $J(\pi)$. As in ALP, we equally weight the F-statistics, and use the limiting case representing small changes in parameter coefficients. This is defined by ALP as

$$Avg - F = \sum_{\pi} F(\pi)J(\pi). \tag{9}$$

The test methodology therefore requires estimating least squares parameters for all possible change dates to construct the test statistic in (9). Different change dates generate different parameter estimates. We report parameter estimates for the change date that delivers the global minimum sum of squared errors.

As described in Section IV, we create bootstrapped critical values for the test statistic in (9), as well as for individual parameters in (6), by generating data under the null and constructing the empirical distribution of the resulting test statistics and parameter estimates.

B. Stochastic Beta

A second approach to modeling dynamic hedge fund risk is to consider a hedge fund's exposure to underlying factors to be an unobserved state variable following a particular stochastic process. We assume that exposures evolve according to a first-order autoregressive (i.e., AR(1)) process as follows:

$$R_t = \alpha + \beta_t^{\mathrm{T}} F_t + \varepsilon_t$$
$$\beta_t = \mu + T\beta_{t-1} + \nu_t, \tag{10}$$

where T (for transition matrix) captures the speed with which the fund's risk exposures revert to a long-run mean.[3] Parameters of the model in equation (10) can be estimated in a variety of ways. We use maximum likelihood and a Kalman filter for which the first line in (10) is the measurement equation and the second is the state equation. See the Appendix for details.

Autoregressive models have been used in this context extensively in prior literature. Ohlson and Rosenberg (1982), for example, use an AR(1) model to measure time variation in the exposure of an equal-weighted stock index to a capitalization-weighted stock index. More recently, Mamaysky, Spiegel, and Zhang (2008) allow for time-varying factor loadings that result from a mutual fund manager's active trading. The manager's trading signal is a state variable that also evolves according to an AR(1).[4] In the context of hedge funds, autoregressive factor loadings seem especially appropriate for capturing the cycles in arbitrage opportunities. LTCM's early success in bond arbitrage, for example, attracted a flood of capital from rival traders that eventually eliminated profit opportunities. In response, LTCM adapted by seeking out investments in different markets. In the words of Lowenstein (2000, p. 97), "...they simply rebooted their computers in virgin terrain."

C. Implementation Issues

The two models described above assume very different stochastic processes for factor loadings. The structural change or changepoint model assumes discrete shifts in factor loadings, whereas the stochastic beta model allows for continuous perturbations to factor loadings and predicts a reversal to a long-run level. Parameters of both models, however, allow us to estimate the same relevant properties of factor loadings. The expected magnitude of changes in factor loadings, for example, is estimated by the difference in their levels in the structural change model and by the volatility of the AR(1) process in the stochastic beta model. Similarly, the expected duration of a particular level of factor loadings is estimated by the time between changepoints in the structural change model and by the transition matrix in the stochastic beta model. Since both models can capture the dynamics of factor loadings, we choose between them in Section IV on the basis of their power to reject the null hypothesis of constant risk exposures.

The structural change model is more parsimonious and its parameters are far easier to estimate than the stochastic beta model. The stochastic beta model, on the other hand, is more elegant, allowing for a smooth transition in a fund's exposure to risk factors over time.[5] Unfortunately, sophisticated assumptions for the evolution of factor loadings run the risk of inferior power if misspecified,

[3] An alternative is to specify a regime-switching model for shifts in exposure.

[4] Stochastic betas are also used in Jostova and Philipov (2005) and Busse and Irvine (2006).

[5] Another alternative is the smooth transition regression of Lin and Terasvirta (1994), which specifies a deterministic but gradual shift between different levels of beta.

as discussed in Ghysels (1998). Consequently, in Section IV, we conduct power tests by generating data under the structural change and stochastic beta models. We then estimate parameters of both models to determine which more accurately identifies changes in factor loadings.

Three other implementation issues warrant discussion. First, reported hedge fund returns often feature significant levels of positive serial correlation. This may bias downward estimates of the exposure of fund returns to contemporaneous factor returns, and may result in an artificial relation between fund returns and lagged factor returns. Lo (2002), Getmanksy, Lo, and Makarov (2004), and Jagannathan, Malakhov, and Novikov (2006) suggest that the serial correlation may be generated by a fund manager's use of stale trade prices, or smoothly evolving model values, to compute fund returns when the fund invests in illiquid assets. Alternatively, Asness, Krail, and Liew (2001) and Bollen and Pool (2008) suggest that the serial correlation may arise from a fund manager artificially smoothing reported returns in order to reduce the measured volatility of the fund. Further, if fund managers follow momentum strategies, fund returns will naturally be related to lagged factor returns, and will likely exhibit positive serial correlation. In any case, serial correlation in hedge fund returns may camouflage shifts in a fund's risk exposures. Our estimates of the prevalence and magnitude of changes in hedge fund risk exposures may then understate their importance.

Second, characteristics of the sample return series—length of history and return frequency—will have a strong influence in determining the appropriate econometric technique. Consider, for example, a fund manager who changes risk exposure according to the stochastic beta model but very slowly over time (i.e., the diagonal elements of the parameter matrix T in the AR(1) process of (10) approach one). With a long history, the fund's risk dynamics would likely be better captured by estimating parameters of the stochastic beta model than the changepoint regression, since only the former model allows for gradual changes in risk exposure. If the available history length is relatively short, however, the data may not provide enough information to estimate parameters of the stochastic beta model. It is conceivable that the changepoint regression could detect a change in exposure over a short horizon in this example, even if the data were generated by a stochastic beta model. Similarly, even if changes are gradual over time as a result of, say, selling a large holding in an illiquid asset, the full liquidation will be completed over the course of a few days, not a few months. If daily data are available, the stochastic beta model would perform better than the changepoint regression as it would capture the smooth transition out of the asset from day to day. If only monthly data are available, the transition will be discrete, as it will have presumably occurred sometime during the month. In this case, even though the stochastic beta model is correct, the data are too crude to permit its estimation, but the changepoint regression may be able to capture the shift. We address both of these issues in the power tests by documenting the relative ability of the two techniques over different horizons.

Third, as is true in all hedge fund research that uses factor models, identification of the factors is problematic. To minimize the number of parameters to estimate, we must select a subset of available factors. Our approach is to first select a subset of factors to maximize the explanatory power of a constant parameter regression, while rewarding parsimony using the Bayesian Information Criterion. Once the subset is chosen, we estimate parameters of the two models for dynamic factor loadings. We recognize that this approach is somewhat ad hoc in that the factor selection is conducted under the null hypothesis that parameters are fixed. If we then find that parameters change, this means that the factor selection was conducted with an erroneous assumption. The benefit of the two-stage approach is that it reduces the dimensionality of the problem. The cost is the potential that the wrong factors are selected due to the assumption of constant parameters. We address this issue by comparing the selection of factors and the identification of the switch date generated by the two-stage procedure we use and by a procedure that does both simultaneously. While it is feasible to select factors and the switch date simultaneously for the purpose of this comparison, generating critical values for parameters is not. In unreported analysis, we find a significant overlap in the selection of factors between the two approaches, and switch dates that are generally within a few months of each other.

III. Data

The primary source of hedge fund data used in our empirical analysis is the Center for International Securities and Derivatives Markets (CISDM) database. The sample period is from January 1994 through December 2005. The CISDM database includes live and dead hedge funds and managed futures funds, as well as indices of both. We focus attention on the individual funds. To avoid survivorship bias, dead funds are included in our analysis.[6] For each fund, we collect observations of returns and record supplemental information including fund type (hedge fund, fund of fund, commodity trading advisor, or commodity pool operator) and fund strategy. Returns are net of all management and performance-based fees, including the fees charged by funds of funds managers. To help ensure reliability in model estimation, we drop funds with less than 24 months of contiguous returns. After we apply our exclusionary criterion, the sample contains 6,158 funds. Of these, 3,013 are live funds and 3,145 are dead funds.

Before describing the attributes of the CISDM sample in greater detail, it is important to note that we test the robustness of our results by performing the same analyses using hedge funds from the Lipper TASS database during the same sample period. After applying the same exclusionary criterion as above, the TASS sample includes 2,751 lives funds and 3,504 dead funds. According to

[6] Backfill is not monitored by the CISDM. A common method for controlling for backfill bias is to drop 12 or 24 observations at the beginning of each return series. We do not drop observations given the relatively short histories of many funds. Unreported analysis indicates that our estimates of changes in risk exposure are not affected (other than a reduced sample size) when the first 12 observations for each fund are dropped.

Table I

Summary Statistics of Reported Monthly Returns of CISDM Funds

See Appendix Table A1 for definitions of fund types. The summary statistics are the number of funds and the equally-weighted averages of the mean monthly return, μ; the standard deviation of monthly returns, σ; the Sharpe ratio, *SR*; the skewness, *Skew*; the excess kurtosis, *Kurt*; the autocorrelation coefficient, AR(1); the percentage of funds with an AR(1) coefficient significantly positive at the 5% probability level, $\% \gg 0$; and the percentage of funds with an AR(1) coefficient significantly negative at the 5% probability level, $\% \ll 0$. Data are from January 1994 through December 2005.

Type	No. of Funds	μ	σ	*SR*	*Skew*	*Kurt*	AR(1)	$\% \gg 0$	$\% \ll 0$
				Panel A: Live Funds					
HF	1,445	0.0115	0.0358	0.3649	0.1533	3.4603	0.1543	30.0%	0.3%
FOF	1,022	0.0070	0.0168	0.3620	−0.2438	2.4814	0.2091	37.7%	0.5%
CTA	302	0.0114	0.0551	0.1661	0.4178	1.9850	0.0027	7.3%	5.6%
CPO	244	0.0084	0.0520	0.1308	0.4002	1.8071	0.0556	10.2%	1.2%
	3,013								
				Panel B: Dead Funds					
HF	1,622	0.0093	0.0540	0.1686	−0.0238	3.7379	0.1257	21.1%	1.3%
FOF	373	0.0056	0.0280	0.1650	−0.3102	3.9531	0.2022	39.4%	0.5%
CTA	513	0.0086	0.0637	0.0551	0.4189	2.9907	−0.0074	7.4%	4.3%
CPO	637	0.0041	0.0525	0.0133	0.1978	2.5492	0.0033	5.3%	3.8%
	3,145								
				Panel C: All Funds					
HF	3,067	0.0103	0.0454	0.2610	0.0596	3.6071	0.1392	25.3%	0.8%
FOF	1,395	0.0067	0.0198	0.3093	−0.2616	2.8749	0.2073	38.1%	0.5%
CTA	815	0.0097	0.0605	0.0962	0.4185	2.6180	−0.0037	7.4%	4.8%
CPO	881	0.0053	0.0524	0.0458	0.2539	2.3437	0.0178	6.7%	3.1%
	6,158								

Agarwal, Daniel, and Naik (2007), only 23% of CISDM funds are also included in the TASS database, hence the TASS data set provides a largely independent sample. In the interest of brevity, we report only the CISDM results, except where the TASS results are important in demonstrating the robustness of our findings.[7]

Table I contains summary statistics of the CISDM funds in the sample. For each fund type, the table lists the number of funds and equally weighted cross-sectional averages of each fund's average monthly return, standard deviation, Sharpe ratio, skewness, and excess kurtosis.[8] Live funds feature substantially higher Sharpe ratios than dead funds in all categories. In the case of hedge

[7] The full set of TASS results can be obtained from the authors upon request or in the Internet Appendix available at http://www.afajof.org/supplements.asp.

[8] To account for the relatively short fund histories, skewness is computed as $\frac{n}{(n-1)(n-2)s^3} \sum_{i=1}^{n} (x_i - m)^3$, where s is sample standard deviation, m is sample mean, and n is number of observations. Similarly, excess kurtosis is computed as $\frac{n(n+1)}{(n-1)(n-2)(n-3)s^4} \sum_{i=1}^{n} (x_i - m)^4 - 3\frac{(n-1)^2}{(n-2)(n-3)}$.

funds, for example, the average Sharpe ratio for live funds, 0.3649, is more than double the average Sharpe ratio for dead funds, 0.1686. The difference is no surprise. Anecdotal evidence suggests that hedge fund investors withdraw capital en masse following periods of poor performance. Live funds are also less volatile than the corresponding dead funds on average for all of the categories and are generally more positively skewed. The return distributions of all categories feature substantial excess kurtosis. Thick tails in the return distributions may arise from the option-like payoffs of certain trading strategies. They may also arise if the hedge fund manager switches strategies over time, even if each of the strategies used during the measurement period had normally distributed returns.[9]

The last three columns of Table I list the average AR(1) coefficient derived from a regression of fund returns on their first lag, the percentage of funds with a statistically significant and positive coefficient, and the percentage of funds with a statistically significant and negative coefficient. For the live funds, the average AR(1) coefficients for hedge funds and funds of funds are 0.1543 and 0.2091, respectively, with 30% and 37.7% of the two categories featuring significantly positive coefficients. There are at least three possible reasons why funds of funds feature higher levels of autocorrelation than individual hedge funds. First, individual funds might not always record changes in net asset value simultaneously, especially if they are invested in illiquid securities. If shocks to a particular market are not reflected simultaneously in funds exposed to it, the return series of a portfolio of those funds will exhibit positive autocorrelation.[10] Second, as noted by Getmansky, Lo, and Makarov (2004), positive autocorrelation in reported fund returns reduces measured return volatility and inflates standard estimates of the Sharpe ratio. If a fund of funds manager selects individual hedge funds at least in part based on their Sharpe ratios, he would be more likely to select funds with higher levels of autocorrelation. Third, if any of the individual funds in which a fund of funds manager invests have not finalized a monthly net asset value in time for the fund of funds manager to compute his own monthly return, the manager may use an estimate of the individual fund's return. This estimate will likely be based on last month's return and a conservative estimate of the current month's return, in which case it would contribute to serial correlation in the fund of funds return series.[11]

The AR(1) coefficients are slightly higher for live hedge funds than they are for dead funds (e.g., 0.1543 versus 0.1257). This stands to reason. Managers of dying funds are less able to "manipulate" returns. Prices used in the computation of returns are market prices resulting from forced liquidations rather than stale trade prices or model values. For commodity trading advisors (CTAs) and

[9] Fama (1965), among others, discusses how a mixture of normal distributions can explain thick tails in stock returns.

[10] This behavior is akin to the positive autocorrelation observed in stock portfolio return series when the constituent stocks trade infrequently. Fisher (1966) was the first to describe this phenomenon. The effects are modeled more formally in a number of studies including Lo and MacKinlay (1988) and Stoll and Whaley (1990),

[11] The authors thank an anonymous referee for this point.

Table II

Distribution of Length of Reported Monthly Return Series of CISDM Funds

See Appendix Table A1 for definitions of fund types. Listed are the 25^{th}, 50^{th}, and 75^{th} percentiles of the distributions of history lengths, in months, of different fund types. Data are from January 1994 through December 2005.

Type	No. of Funds	25^{th}	50^{th}	75^{th}
		Panel A: Live Funds		
HF	1,445	40	63	101
FOF	1,022	36	55	89
CTA	302	50	93	140
CPO	244	55	95	142
	3,013			
		Panel B: Dead Funds		
HF	1,622	37	53	76
FOF	373	38	56	82
CTA	513	40	53	78
CPO	637	35	49	73
	3,145			
		Panel C: All Funds		
HF	3,067	39	57	85
FOF	1,395	36	55	85
CTA	815	43	63	100
CPO	881	37	56	98
	6,158			

commodity pool operators (CPOs), there is very little evidence of serial correlation in returns of live or dead funds, consistent with their focus on relatively liquid and exchange traded futures contracts.

To provide a sense of the length of the monthly return time-series, Table II lists the 25^{th}, 50^{th}, and 75^{th} percentiles of the cross-sectional distributions of fund history lengths. Not surprisingly, the live funds have longer histories. Live hedge funds have a median history of 63 months, for example, while dead funds have a median history of 53 months. These history lengths are important because analysis of dynamic strategies is data-intensive, particularly for the case of the stochastic beta model. In their study of the market timing ability of mutual fund managers, Bollen and Busse (2001) compare the use of monthly versus daily returns and show that the use of monthly (in contrast to daily) returns tends to lack sufficient power to reject the null hypothesis of no market timing. We return to this issue and address it in detail when measuring the power of our main test statistics in the next section.

Table III contains summary statistics of the two sets of factors we use. The first set, listed in Panel A, includes the three Fama–French factors, from Kenneth French's website, and the seven asset-based style factors of Fung and Hsieh (2004). The Fama–French factors are the excess return of the market,

Table III

Summary Statistics of Factors Used to Analyze Reported Hedge Fund Returns

See Appendix Table A1 for definitions of factors. The summary statistics are: the mean monthly return, μ; the standard deviation of monthly returns, σ; the Sharpe ratio, *SR*; the skewness, *Skew*; and the excess kurtosis, *Kurt*. For *MKTXS*, *SMB*, *HML*, *SMBSQ*, *HMLSQ*, *D10YR*, and *DSPRD* the Sharpe ratio equals the average return divided by the standard deviation. Data are from January 1994 through December 2005.

Factor	μ	σ	*SR*	*Skew*	*Kurt*
		Panel A: Fung and Hsieh Factors			
MKTXS	0.0061	0.0439	0.1398	−0.7608	0.8799
SMB	−0.0030	0.0407	−0.0727	−1.5893	6.8515
HML	0.0071	0.0366	0.1936	0.7418	2.3122
SMBSQ	0.0017	0.0050	0.3329	7.5219	64.5793
HMLSQ	0.0014	0.0030	0.4666	4.0211	18.0870
D10YR	−0.0090	0.2345	−0.0385	0.3934	−0.2400
DSPRD	−0.0005	0.1244	−0.0039	0.9278	2.1367
PTFSBD	−0.0019	0.1567	−0.0320	1.5091	2.7661
PTFSFX	−0.0028	0.1898	−0.0314	1.3797	3.3402
PTFSCOM	−0.0086	0.1273	−0.0921	1.4822	4.6697
PTFSIR	−0.0016	0.1849	−0.0255	2.4967	9.4868
PTFSSTK	−0.0543	0.1309	−0.4384	1.1063	2.4438
		Panel B: Futures Contract Factors			
SP	0.0056	0.0425	0.0574	−0.6031	0.6825
ED	0.0004	0.0020	−1.3622	1.3052	4.4970
US	0.0035	0.0260	0.0150	−0.4897	1.1761
CD	0.0010	0.0179	−0.1181	−0.0219	−0.0260
JY	−0.0030	0.0353	−0.1730	1.1218	4.3408
SF	−0.0006	0.0299	−0.1238	0.3130	−0.2692
CL	0.0209	0.0906	0.1959	0.1513	0.4605
NG	0.0148	0.1630	0.0716	0.6667	0.7384
C	−0.0091	0.0646	−0.1898	−0.1296	−0.1808
GC	−0.0003	0.0368	−0.0937	0.7991	1.8903

MKTXS, and the returns of the size and value portfolios, *SMB* and *HML*. The squared returns of the size and value portfolios, *SMB*SQ and *HML*SQ, are also included. The first two Fung–Hsieh factors are *D10YR*, the change in yield of a 10-year Treasury note, and *DSPRD* (dubbed the "credit spread"), the yield on 10-year BAA corporate bonds less the yield of a 10-year Treasury note. Both series are expressed in basis points and are obtained from the U.S. Federal Reserve's website. The five remaining Fung–Hsieh variables are trend factors. In essence, they are the returns of portfolios of options on bonds, *BD*; foreign currencies, *FX*; commodities, *COM*; short-term interest rates, *IR*; and stock indexes, *STK*. The return series are obtained from David Hsieh's website.

The second set, listed in Panel B of Table III, includes monthly relative price changes of highly active futures contracts on different underlying asset classes. The trading strategy underlying each futures return series is one in which the trader holds the nearby contract until the last day *before* the contract month. The Chicago Mercantile Exchange (CME)'s S&P 500 futures, for example, has a quarterly expiration cycle (i.e., March, June, September, December), so the January and February monthly returns are computed using the March contract, the March, April, and May returns are computed using the June contract, and so on. The contracts are selected to represent the excess returns on a diverse set of asset classes.[12] The S&P 500 (*SP*) futures represents equities, and the CME's Eurodollar, *ED*, and the Chicago Board of Trade (CBT)'s long-term Treasury bond futures, *US*, contracts represent short-term and long-term bonds, respectively. Currencies are represented by the CME's Canadian dollar, *CD*, Japanese yen, *JY*, and Swiss franc, *SF*, contracts. The CME's Euro futures contract is not included in our analysis because it did not come into existence until September 17, 2001, halfway through our investigation period. To represent the petroleum complex, we choose the NYMEX's light crude oil, *CL*, and natural gas, *NG*, futures contracts. To represent commodities, we choose the CBT's corn futures contract, *C*, and the COMEX's gold futures contract, *GC*.

The results reported in Panel B show a wide range of realized returns and risks for the different factors. Unlike the factors reported in Panel A, which have differing units of measurement, the factors summarized in Panel B are all based on monthly returns and therefore can be compared to each other directly.[13] The natural gas and crude oil futures experienced the highest risk during the sample period, with monthly standard deviations of (excess) return of 16.30% and 9.06%, respectively. The corresponding mean monthly returns were 1.48% and 2.09%, respectively. Equities had lower returns and were less risky by comparison, with a mean monthly excess return on the S&P 500 futures of 0.56% and a standard deviation of 4.25%. These estimates are very close to the values reported for the *MKTXS* in Panel A, 0.61% and 4.39%, respectively. The lowest risk class was the Eurodollar futures with a volatility rate of 0.20% per month. Two factors, the Japanese yen and corn, had modest volatility rates, 3.53% and 6.46%, and had negative monthly realized returns, −0.30% and −0.91%, respectively.

In the interest of providing a better sense of the substitutability of different factors, Table IV contains a correlation matrix of the two sets of factors. In general, the first set of factors, listed in Panel A, have low correlation with each other. The three exceptions are the correlation between the Fama–French value factor *HML* and the market excess return *MKTXS*, −0.479, the correlation between the Fama–French value factor *HML* and the Fama–French size factor *SMB*, −0.501, and the correlation between the change in 10-year Treasury note yield *D10YR* and the change in the spread between BAA corporate bond yield

[12] In equilibrium, the relative price change of a futures contract equals the return on the underlying asset less the risk-free rate of interest.

[13] In Panel A, the series are constructed differently from one another: *MKTXS* represents the CRSP value-weighted stock index return in decimal format and *D10YR* represents the monthly change in the spread between a BAA corporate yield and a 10-year Treasury yield in basis points. With different scales, the parameter estimate from series to series cannot be compared. In contrast, all excess return series in Panel B are monthly returns expressed in decimal form.

Table IV

Correlation Matrix of Factors Used to Analyze Reported Hedge Fund Returns

See Appendix Table A1 for definitions of factors. Data are from January 1994 through December 2005.

Panel A: Fung and Hsieh Factors

	MKTXS	*SMB*	*HML*	*SMBSQ*	*HMLSQ*	*D10YR*	*DSPRD*	*PTFSBD*	*PTFSFX*	*PTFSCOM*	*PTFSIR*	*PTFSSTK*
MKTXS	1.000											
SMB	0.118	1.000										
HML	−0.479	−0.501	1.000									
SMBSQ	0.057	−0.571	0.341	1.000								
HMLSQ	−0.183	−0.381	0.507	0.576	1.000							
D10YR	0.023	0.188	−0.173	−0.106	−0.130	1.000						
DSPRD	−0.155	−0.329	0.095	0.149	0.118	−0.634	1.000					
PTFSBD	−0.151	−0.016	−0.063	0.046	0.038	−0.107	0.087	1.000				
PTFSFX	−0.096	0.040	0.044	0.024	−0.076	−0.173	0.142	0.145	1.000			
PTFSCOM	−0.094	−0.052	0.023	0.019	−0.046	−0.067	0.030	0.144	0.264	1.000		
PTFSIR	−0.175	−0.077	−0.037	−0.029	−0.071	−0.167	0.241	0.207	0.204	0.184	1.000	
PTFSSTK	−0.187	−0.009	−0.024	−0.062	−0.043	−0.295	0.276	0.227	0.229	0.083	0.257	1.000

Panel B: Future Contracts Factors

	SP	*ED*	*US*	*CD*	*JY*	*SF*	*CL*	*NG*	*C*	*GC*
SP	1.000									
ED	−0.079	1.000								
US	−0.066	0.426	1.000							
CD	0.407	−0.008	0.105	1.000						
JY	0.107	0.050	0.048	0.110	1.000					
SF	−0.202	0.156	0.136	0.134	0.483	1.000				
CL	−0.024	−0.032	0.026	0.188	0.092	0.064	1.000			
NG	0.016	−0.045	0.222	0.226	0.002	0.183	0.359	1.000		
C	0.155	−0.100	0.066	0.026	0.013	−0.048	0.008	0.101	1.000	
GC	−0.047	0.070	0.095	0.359	0.224	0.294	0.124	0.079	−0.020	1.000

and the 10-year Treasury yield *DSPRD*, −0.634. Comparatively speaking, the absolute levels of correlation between the different return series constructed from futures prices, shown in Panel B, are less than those reported in Panel A. Naturally, the two interest rate instruments, the short-term *ED* and long-term *US*, are positively correlated at a level of 0.426. Interestingly, with respect to the currency futures contracts, the Canadian dollar was not strongly correlated to the Japanese yen, 0.110, or the Swiss franc, 0.134. At the same time, the Japanese yen and the Swiss franc were strongly correlated, 0.483, and all three currencies were strongly correlated with gold. Crude oil and natural gas were strongly positively correlated, 0.359, as were the S&P 500 futures and the Canadian dollar futures, 0.407. While high correlation between independent variables can give rise to spurious univariate significance levels, our procedure for selecting factors is based solely on overall explanatory power and hence is unaffected by multicollinearity.

To benchmark the propensity and magnitude of switches in factor loadings of the individual hedge funds in our sample, we estimate corresponding measures for a sample of individual equity mutual funds. The mutual fund return data are drawn from the Center for Research in Security Prices (CRSP) Survivor-Bias Free U.S. Mutual Fund Database. To be included, a mutual fund must have at least 24 observations of monthly return data during the period January 1994 through December 2005. The mutual fund sample includes 6,840 equity funds in the Aggressive Growth, *AG*, Growth and Income, *GI*, Long-term Growth, *LG*, Balanced, *BL*, and Total Return, *TR*, categories.

IV. Model Performance

In this section, we evaluate the performance of the two models of time-varying fund exposures—the changepoint regression and stochastic beta model—using simulated monthly return data. First, we calibrate the simulations by turning to the actual monthly return data and finding the "optimal" constant parameter factor model for each live fund during the sample period January 1994 through December 2005. We identify the optimal factor model by choosing the subset of factors that minimizes the Bayesian Information Criterion (BIC) using a maximum of three factors.[14] Second, given the magnitudes of the factor loadings, we choose parameter values for a changepoint regression and a stochastic beta model and generate random returns. We then compare the ability of the changepoint regression and stochastic beta model to reject the null hypothesis that risk exposures are constant. We estimate both models using data generated from the stochastic beta model and the changepoint regression. This procedure allows us to measure the power of each model when it is both correctly specified and misspecified—in practice, we do not know the underlying data generating process for a given fund. We find that the changepoint regression is superior to

[14] Though we can estimate constant parameter models with more then three factors, we limit the number of factors to three to allow comparison with the time-varying models, which have more parameters per factor and limit the number of factors we can allow.

the stochastic beta model. Finally, we perform a robustness check to determine whether fund fees could generate false rejections of the null in the changepoint regression.

A. Calibrations

Table V contains a summary of the results of the OLS regression procedure applied to 3,013 live CISDM fund return series. In this analysis, we use the Fung and Hsieh factors. To simplify comparison of hedge fund exposures to the different factors, we scale each Fung and Hsieh factor to have a standard deviation equal to that of the market excess return. Panel A contains summary regression statistics for the optimal factor model. The reported adjusted-R^2 levels are averages across all funds in each category. For hedge funds, for example, the average adjusted-R^2 is 28.8%, whereas the average for funds of funds is 37.6%. The averages for CTAs and CPOs are 19.2% and 22.6%, respectively. The average abnormal performance on a monthly basis ranges between 0.25% for funds of funds and 0.84% for CTAs.[15] The average number of factors ranges from 1.8 for CTAs to 2.6 for funds of funds. The larger number of factors required to model funds of funds returns is consistent with their diversification across strategies.

Panel B shows the percentage of funds for which each factor is used in the optimal factor model, and Panel C shows the average parameter estimate across funds for which the factor is included in the optimal subset. For hedge funds, 53.5% of the funds have the excess return of the CRSP value-weighted index, *MKTXS*, as a risk factor. Of those that do, the average factor exposure is 0.4841. For hedge funds, the two other Fama–French factors, *SMB* and *HML*, enter the optimal model in 24.8% and 23.9% of the funds, respectively, with average coefficients of 0.3783 and 0.2638. For funds of funds, 75% of the funds have the excess return of the CRSP value-weighted index, *MKTXS*, as a risk factor, with an average factor exposure of 0.2364. For hedge funds and funds of funds, stock market exposure plays an important role in explaining fund returns. For CTAs and CPOs, the trading strategies are more commodities-based, and hence stock market exposure is less relevant. The remaining variables/columns can be interpreted in a similar fashion. Based on the results of the table, we decided to generate random data using two sets of factors: (1) *MKTXS* and *HML*, and (2) *D10YR* and *PTFSFX*. The results are similar for the two sets. For the sake of brevity, we report below results from the first set only.

As an aside, Table VI reports regression statistics using the futures contract factors, again limiting the number of factors per fund to at most three. In this analysis, we require 36 return observations (as opposed to 24 observations for the other factors analyzed above) because we estimate two parameters of a quadratic function for each futures contract factor. Consequently, the sample size falls from the original 3,013 to 2,481. In comparing the results of

[15] For comparison, Fung and Hsieh (2004) report abnormal returns of 0.7% to 0.9% for four hedge fund indexes over the period 1994 to 2002.

Table V

Summary Statistics of Factor Models Estimated Using Reported Monthly Returns of 3,013 Live CISDM Funds

See Appendix Table A1 for definitions of factors and fund types. Listed are summary statistics of factor models estimated by OLS using the Fung and Hsieh factors listed in Appendix Table A1. For each fund, an optimal subset of factors is selected using the Bayesian Information Criterion. Panel A contains the number of funds of each type; the average adjusted-R^2; alpha, α; total volatility, σ; residual volatility, σ_ε; and the number of factors used in each regression. Panels B and C contain the percentage of funds for which a factor is included in the optimal subset and the exposure to each factor averaged across funds for which the factor is included in the optimal subset, respectively. Data are from January 1994 through December 2005.

Panel A: Regression Statistics

Statistic	All	HF	FOF	CTA	CPO
No. of funds	3,013	1,445	1,022	302	244
Adjusted-R^2	30.3%	28.8%	37.6%	19.2%	22.6%
α	0.53%	0.64%	0.25%	0.84%	0.67%
σ	3.26%	3.58%	1.68%	5.50%	5.20%
σ_ε	2.66%	2.86%	1.27%	4.87%	4.50%
No. of factors	2.3	2.1	2.6	1.8	2.1

Panel B: Percent of Funds with Factor Exposure

Factor	All	HF	FOF	CTA	CPO
MKTXS	54.8%	53.5%	75.0%	19.5%	20.9%
SMB	19.9%	24.8%	21.4%	4.3%	4.1%
HML	21.6%	23.9%	27.9%	4.3%	2.5%
SMBSQ	16.2%	21.0%	14.3%	8.3%	5.3%
HMLSQ	9.9%	12.2%	9.1%	5.0%	5.7%
D10YR	23.0%	16.5%	33.7%	13.9%	28.3%
DSPRD	26.0%	23.0%	40.8%	7.3%	5.3%
PTFSBD	12.0%	8.9%	7.2%	30.1%	28.3%
PTFSFX	18.1%	8.5%	15.1%	42.7%	57.0%
PTFSCOM	9.8%	7.6%	5.9%	22.2%	23.8%
PTFSIR	4.5%	5.5%	2.1%	8.3%	4.5%
PTFSSTK	11.4%	9.0%	9.2%	18.2%	26.2%

Panel C: Average Factor Exposure

Factor	All	HF	FOF	CTA	CPO
MKTXS	0.3555	0.4841	0.2364	0.2697	0.2967
SMB	0.2921	0.3783	0.1724	0.0379	0.1488
HML	0.2286	0.2638	0.1867	0.3171	0.0024
SMBSQ	0.0607	−0.0107	0.1366	0.7774	−0.5059
HMLSQ	0.0848	0.0246	0.0938	0.0993	0.7690
D10YR	−0.1816	−0.1825	−0.1434	−0.2563	−0.3241
DSPRD	−0.1710	−0.1984	−0.1574	−0.1235	0.0117
PTFSBD	0.1341	−0.0394	0.0019	0.3413	0.3267
PTFSFX	0.2978	0.1504	0.1142	0.4908	0.4526
PTFSCOM	0.2301	0.1842	0.1268	0.3076	0.3346
PTFSIR	−0.0524	−0.0991	−0.0595	0.0453	0.0746
PTFSSTK	0.1373	0.0797	0.1356	0.1295	0.2637

Table VI

Summary Statistics of Factor Models Estimated Using Reported Monthly Returns of 2,481 Live CISDM Funds

See Appendix Table A1 for definitions of fund types. Listed are summary statistics of factor models estimated by OLS using the futures contract factors listed in Appendix Table A1. For each fund, an optimal subset of factors is selected using the Bayesian Information Criterion. Listed are the number of funds of each type; the average adjusted-R^2; alpha, α; total volatility, σ; residual volatility, σ_ε; and the number of factors used in each regression. Data are from January 1994 through December 2005.

Statistic	All	HF	FOF	CTA	CPO
No. of funds	2,481	1,198	797	271	215
Adjusted-R^2	20.8%	20.7%	23.0%	16.8%	18.1%
α	0.54%	0.93%	0.49%	−0.18%	−0.49%
σ	3.41%	3.74%	1.67%	5.52%	5.39%
σ_ε	2.97%	3.22%	1.42%	4.95%	4.82%
No. of factors	1.3	1.3	1.4	1.4	1.5

Tables V and VI, we see that the level of the adjusted-R^2 falls across all categories. Averaged over all funds, it drops from 30.3% to 20.8%.[16] For CTAs and CPOs, the differences in explanatory power are smallest. This is not surprising since the primary trading vehicles for CTAs and CPOs are futures contracts. Note that the average alpha drops from 0.84% to −0.18% for CTAs, and from 0.67% to −0.49% for CPOs. This suggests that performance rankings can be quite sensitive to the choice of factors, a point we return to in Section V.

B. Critical Values and Power Comparisons

In this subsection, we first construct bootstrapped critical values for the changepoint regression and stochastic beta model, and then we compare the power of the two models to reject the null hypothesis when data are generated under the two alternatives.

We begin by generating data under the null hypothesis of constant factor exposures. For each simulated hedge fund return series of length T, we draw T factor returns with replacement, scale the factor returns by hypothetical factor loadings, and add randomly generated residuals. Two types of residuals are generated: independently and identically distributed normal variates as well as conditionally normal variates with variance given by a GARCH(1,1) process. We then estimate parameters and compute test statistics for both the changepoint regression and stochastic beta model using each simulated hedge fund return series. The test statistics are the changepoint regression's F-statistic in (9) and a standard likelihood ratio test for the stochastic beta model in (10). The null hypothesis imposes five restrictions—the two AR(1) coefficients and the three

[16] The comparisons of the values in Tables V and VI are not strictly correct since they are based on different, albeit, strongly overlapping samples.

Table VII

Critical Values

Listed are bootstrapped critical values for a changepoint regression and a stochastic beta model. Monthly data are generated under the null of constant factor exposures by drawing observations of the market excess return and the value factor with replacement:

$$R_t = 0.50MKTXS_t + 0.30HML_t + \varepsilon_t.$$

Residuals are either i.i.d. normal with volatility of 3% or simulated from the following GARCH(1,1) process:

$$\sigma_t^2 = \phi + 0.8\sigma_{t-1}^2 + 0.1\varepsilon_{t-1}^2,$$

where ϕ is calibrated so that the unconditional volatility is also 3% monthly. 10,000 series are generated at each of three history lengths (T) of 36, 60, or 120 observations. Panel A lists simulated critical values of a changepoint regression's F-statistic when the regression is run on the randomly generated data. Panel B lists simulated critical values of a stochastic beta model's likelihood ratio statistic when the model is estimated on the same set of randomly generated data.

	Normal Residuals			GARCH(1,1) Residuals		
T	10%	5%	1%	10%	5%	1%
Panel A: Changepoint Regression Model						
36	2.04	2.53	3.68	2.05	2.55	3.72
60	1.83	2.20	3.11	1.85	2.24	3.20
120	1.69	1.97	2.67	1.71	2.01	2.89
Panel B: Stochastic Beta Model						
36	3.81	5.48	9.47	3.96	5.73	9.68
60	3.69	5.30	8.84	3.91	5.62	9.80
120	3.56	5.10	8.64	3.88	5.71	10.09

parameters of the variance–covariance matrix for the innovations are jointly equal to zero.

For each history length T, we repeat the procedure 10,000 times, and then sort the values of the test statistics to determine the critical values; for example, the 500$^{\text{th}}$ largest test statistic out of 10,000 gives the critical value at a 5% significance level. We use history lengths of 36, 60, and 120 observations to assess the impact of the lengths of the series on model performance.

Table VII lists the results. In Panel A, critical values for the changepoint regression decline as the history length increases. At the 5% significance level, for example, the critical F-statistic is 2.53 using 36 observations and 1.97 using 120 observations. The 5% critical values are slightly higher when using GARCH(1,1) residuals (i.e., 2.55 and 2.01, respectively), so we use these in the ensuing power tests to be conservative. In Panel B, critical values for the stochastic beta model with normal residuals decline with history length, but they do not with GARCH residuals. Apparently, time variation in residual volatility can lead to rejections of the null, especially for longer time-series. Here, too, critical values are higher for the GARCH residuals and again we use these in the power tests.

With the critical values in hand, we can now generate data under the alternative hypotheses, estimate parameters of the two models, and compute the percentage of simulations for which the null hypothesis is rejected. For the changepoint regression, we randomly generate 1,000 return series for each history length as follows:

$$\begin{aligned}
&\text{for}\quad t = 1, \ldots, T_1 : \\
&R_t = 0.00 + 0.30 MKTXS_t + 0.20 HML_t + \varepsilon_t \\
&\text{for}\quad t = T_1 + 1, \ldots, T : \\
&R_t = 0.02 + 0.60 MKTXS_t + 0.40 HML_t + \varepsilon_t,
\end{aligned} \tag{11}$$

where *MKTXS* and *HML* are drawn with replacement.[17] Six versions of the data are generated. Three use constant volatility normal residuals and three use residuals with variance generated by a GARCH(1,1) process. Within each category of residual, the changepoint is either fixed at the middle of the series (labeled "Center" in Table VIII), selected at random from the 13th to the $T-11$th observations ("Uniform-A"), each equally likely, or drawn at random from the 2nd to the penultimate observations ("Uniform-B").

We also generate 1,000 sets of data at each history length under the alternative hypothesis of stochastic betas using

$$\begin{aligned}
R_t &= 0.01 + \beta_{1,t} MKTXS_t + \beta_{2,t} HML_t + \varepsilon_t \\
\beta_{1,t} &= 0.0225 + 0.95 \beta_{1,t-1} + \sqrt{0.10}\, \nu_{1,t} \\
\beta_{2,t} &= 0.0100 + 0.95 \beta_{2,t-1} + \sqrt{0.05}\, \nu_{2,t}.
\end{aligned} \tag{12}$$

We assume that innovations to the two betas, ν_1 and ν_2, are uncorrelated. These parameters generate unconditional factor loadings equal to those of the discrete structural change model.[18] We choose a relatively parsimonious stochastic beta model with a diagonal transition matrix to generate data. Further, when estimating the model, we only estimate the diagonal elements of the transition matrix. If anything, then, the power of the stochastic beta model to identify

[17] In unreported analysis, we compare the small sample properties of the model in equation (11) to a model with constant alpha and switching volatility. When data are generated with switches in alpha, a model with constant alpha has poor power. When data are generated with constant alpha, a model with switching alpha does not over reject the null. Similar to the GARCH results, using data with switches in residual volatility does not affect inference. We conclude that the characterization in equation (11) is a reasonable choice for the changepoint regression.

[18] To see this, manipulate the expectation of β in the stochastic beta model as follows:

$$\begin{aligned}
E[\beta] &= \mu + TE[\beta] \\
\mu &= (I - T)E[\beta] \\
E[\beta] &= (I - T)^{-1}\mu,
\end{aligned}$$

where I is the identity matrix. Plugging in the values from equation (12) in the RHS generates values for $E[\beta]$ equal to their averages in equation (11).

changes in risk exposure is likely to be lower in practice than measured in this exercise. As with the changepoint data, we generate returns using both constant volatility and GARCH(1,1) residuals.

As an aside, we test the ability of the Kalman filter to infer the latent beta process by simulating a single series of 120 betas from the autoregressive process in (12), generating 10,000 return series by randomly drawing factors with replacement, scaling by the betas, and generating residuals as described above. We estimate the stochastic beta model on each series, and record the inferred time-series of betas. Figure 1 shows the confidence bands as measured by percentiles of the inferred betas across simulations at each observation. When residual volatility is 1% per month, the confidence bands are relatively tight around the true beta process. As residual volatility increases, the fit loosens, as expected, but the model still captures the general time variation in beta.

Returning to the comparison of the changepoint regression and the stochastic beta model, we estimate parameters of each using the simulated data for both models. Table VIII lists the percentage of the simulations for which the null hypothesis is rejected at the 5% level. Panel A shows results when data are analyzed using the changepoint regression. When the switch occurs in the center of each series, and residuals have constant volatility, the changepoint regression rejects the null 44.1% of the time using 36 observations, 67.2% of the time using 60 observations, and 94.7% of the time using 120 observations. Power is almost identical with GARCH residuals. Power deteriorates when the changepoint is drawn at random, though at 120 observations it is still 77.9% when the changepoint occurs between the 13^{th} and $T - 11^{\text{th}}$ observation and 66.5% when the changepoint occurs between the 2^{nd} and the penultimate. When the data are generated by the stochastic beta model, the power of the changepoint regression to reject the null hypothesis falls to 47.7% at 120 observations.

Panel B shows results when the stochastic beta model is estimated using the same data. The power is much lower when the data are generated by the changepoint regression, which is, perhaps, no surprise since the stochastic beta model is misspecified. When the data are generated by the stochastic beta model, the correctly specified stochastic beta model has power about the same as the misspecified changepoint regression at 36 and 60 observations and slightly higher at 120 observations, 55.4% versus 47.7%. Presumably, with a sufficiently long time-series, a correctly specified model should outperform a misspecified model. In the context of hedge fund risk exposures, however, we do not know which process is generating the data, and we often have fewer than 60 observations, as shown in Table II. Moreover, even if we knew the process generating the data, we still may not be able to detect it. As discussed earlier, transition from one risk factor to another is likely to occur over days not months. Using monthly data to estimate the AR(1) coefficient in such a case may be simply impossible. The data are too crude. For these reasons, we focus on the generally more powerful changepoint regression in the empirical analysis of actual funds described in the next section.

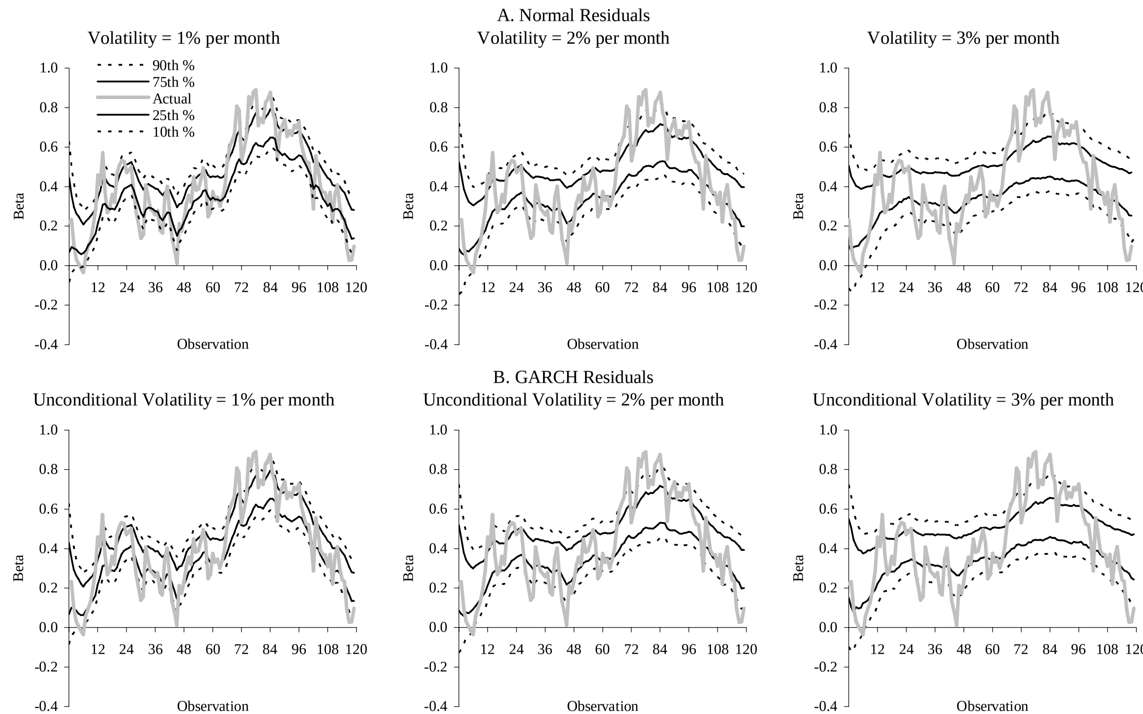

Figure 1. Stochastic beta estimation. The figures illustrate the precision with which stochastic betas are estimated using the Kalman filter. Random hedge fund returns are generated 1,000 times in two steps. First, a single series of 120 monthly betas is drawn from $\beta_t = \mu + T\beta_{t-1} + \nu_t$, where β is a 2 × 1 vector of factor loadings on market excess returns and the Fama–French *HML* series, T is a 2 × 2 transition matrix, and ν is a 2 × 1 vector of random normal variables. Parameters are selected so that the market's factor loading, depicted as a solid gray line in the figures, has an unconditional mean of 0.45. Second, 1,000 sets of hedge fund returns are constructed using the single series of betas by drawing factors with replacement and adding randomly generated residuals ε as follows: $R_t = \alpha + \beta_t^{\mathrm{T}} F_t + \varepsilon_t$. Residuals in Panel A are i.i.d. normal with volatility 1%, 2%, or 3% monthly. Residuals in Panel B are normally distributed with variance generated from the following GARCH (1, 1) process: $\sigma_t^2 = \phi + 0.8\sigma_{t-1}^2 + 0.1\varepsilon_{t-1}^2$, where ϕ is calibrated so that the unconditional volatility is also 1%, 2%, or 3% monthly. Black lines are percentiles of the estimated betas from a Kalman Filter.

Table VIII
Power Tests

Panel A lists the percentage of 1,000 simulations for which the null of constant factor exposures is rejected using the changepoint regression at each of three history lengths (T) of 36, 60, or 120 observations. Data are generated four ways. Center denotes a changepoint regression with switch date set to the middle observation of each sample. Uniform-A denotes a changepoint regression with switch date selected at random between the 13[th] and the $T - 11^{\text{th}}$ observation of each series. Uniform-B denotes a changepoint regression with switch date selected at random between the 2[nd] and penultimate observation of each series. Stochastic β denotes a stochastic beta model. Hedge fund returns are constructed using the market excess return, the value factor, and either i.i.d. normal residuals or residuals simulated from a GARCH(1,1) process. For each set of simulated data, the null is tested using an F-statistic. Panel B lists the percentage of the simulations rejecting the null using a stochastic beta model and its likelihood-ratio statistic.

	Normal Residuals			GARCH Residuals		
T	36	60	120	36	60	120
			A. Changepoint Model			
Center	44.1%	67.2%	94.7%	46.6%	68.7%	94.7%
Uniform-A	38.5%	48.8%	77.9%	41.7%	52.5%	79.3%
Uniform-B	22.7%	35.8%	66.5%	24.7%	38.4%	67.5%
Stochastic β	17.7%	28.5%	47.7%	18.8%	28.6%	47.7%
			B. Stochastic Beta Model			
Center	10.8%	15.8%	35.5%	12.4%	19.0%	37.5%
Uniform-A	10.6%	16.1%	28.7%	11.4%	18.1%	29.7%
Uniform-B	8.8%	14.3%	26.2%	9.6%	16.0%	26.6%
Stochastic β	16.0%	27.8%	55.4%	17.7%	30.4%	57.1%

C. Performance Fees and Spurious Rejections

Before using the changepoint regression to assess the importance of dynamic factor exposures in actual funds, we investigate whether hedge fund performance fees could cause spurious rejections of the null hypothesis that exposures are constant. A hedge fund's α can be interpreted as the average return of the manager's idiosyncratic trading strategies minus the fund's fees. Management fees are a constant percentage of fund assets, so these have a constant effect on α unless a fund manager decided to change the management fee. Performance fees, however, could impose a discrete shift on α. To see this, note that performance fees are a percentage of a fund's profits, but are accrued only when a fund's net asset value (NAV) exceeds the fund's high water mark (H). If a fund's NAV spends an extended period of time below H before rising above it, the fund's α would undergo a discrete reduction. This could result in a rejection of the null hypothesis even though the fund's exposures are unchanged.

We test whether this alternative explanation can account for the rejections we document by randomly generating 10,000 hedge fund return series under the null hypothesis exactly as in Subsection B and modifying each to reflect the impact of management fees and performance fees. Without loss of generality, we initialize NAV and H to \$1. We assume that performance fees accrue monthly if the NAV is above H and that performance fees are paid quarterly, at which point H is reset. We assume a "1 and 20" fee structure: (a) management fees are 1% per year and accrue monthly, and (b) performance fees are 20% of profits. Within each quarterly cycle we compute after-fee returns first by updating the NAV each month to reflect pre-fee returns R_t^{pre} and the management fee:

$$NAV_t = NAV_{t-1}\left(1 + R_t^{\text{pre}}\right)(1 - .01/12). \tag{13}$$

Then, if $NAV_t > H$, the performance fee accrues as follows:

$$NAV_t = NAV_t - .20(NAV_t - H). \tag{14}$$

After-fee returns are computed as the percentage change in NAV from $t-1$ to t. Every 3 months, H is reset to the prevailing NAV if it exceeds H.

We estimate a changepoint regression on the after-fee returns and compute the test statistic in (9) for each of the 10,000 sets of returns. For each of the 10,000 simulations, we use history lengths ranging from 24 to 120 observations. We then assess statistical significance using the bootstrapped critical values corresponding to each history length. The performance fees do result in additional spurious rejections of the null, but only in a handful of cases. At a history length of 60 observations and 10%, 5%, and 1% significance levels, for example, the after-fee returns reject the null 11.9%, 6.3%, and 1.2% of the time. These results suggest that performance fees will not distort our results.

V. Frequency and Magnitude of Switches in Factor Loadings

The focus now turns to measuring the frequency and magnitude of switches in factor loadings using the set of 3,013 live funds and 3,145 dead funds in the CISDM database with at least 24 observations between 1994 and 2005. For each fund, a subset of factors is selected under the null hypothesis of constant betas. Each fund is then tested for switches in factor loadings using the changepoint regression. If the F-statistic in (9) rejects the null hypothesis at a 10% probability level, the fund is earmarked as having undergone a significant switch in factor loadings. Critical values are computed using the bootstrap procedure described in Section IV. For each fund with history length T, we use the critical values established with T observations per simulation.

Table IX summarizes the frequency of significant switches in factor loadings for the live funds using the Fung and Hsieh factors.[19] For a fund with T monthly observations, valid switching dates are months 13 though $T-11$, so that at least 12 observations are used to estimate parameters of each regime. Results are separated by type of fund (arranged vertically in the table) and the number of monthly return observations (arranged horizontally). Panel A lists the number of funds in each subcategory. Panel B reports the frequency with which funds in each subcategory reject the null hypothesis of constant betas. Overall,

[19] For the sake of brevity, we do not report similar results for dead funds.

Table IX

Frequency of Significant Parameter Changes in Factor Models Estimated Using Reported Monthly Returns of 3,013 CISDM Live Funds

See Appendix Table A1 for definitions of fund types. Panel A shows the number of active funds categorized by fund type and history length in months. Panel B shows the percentage of funds for which a constant-beta model can be rejected in favor of a switching-beta model at the 10% probability level using the Fung and Hsieh factors listed in Appendix Table A1. Panels C and D compare the average adjusted-R^2 of funds with significant switches in factor loadings when loadings are restricted to be constant (Panel C) and when loadings are allowed to vary (Panel D). Data are from January 1994 through December 2005.

	History Length			
Type	All	$n < 36$	$36 \leq n < 60$	$n \geq 60$
Panel A: Number of Funds				
All	3,013	532	860	1,621
HF	1,445	247	422	776
FOF	1,022	225	335	462
CTA	302	31	65	206
CPO	244	29	38	177
Panel B: Percent of Funds with Significant Switches				
All	41.2%	19.4%	38.8%	48.6%
HF	41.3%	23.1%	35.6%	49.1%
FOF	49.9%	15.1%	48.1%	66.0%
CTA	28.8%	22.6%	23.1%	31.6%
CPO	20.1%	17.2%	21.1%	20.3%
Panel C: Adjusted-R^2 Constant Beta				
All	29.9%	34.9%	31.6%	28.5%
HF	27.5%	28.7%	26.0%	27.7%
FOF	35.4%	47.4%	38.2%	32.7%
CTA	18.9%	26.9%	18.9%	18.1%
CPO	22.7%	31.3%	26.2%	20.8%
Panel D: Adjusted-R^2 Switching Beta				
All	45.4%	53.5%	47.7%	43.3%
HF	43.8%	50.6%	44.1%	42.6%
FOF	50.5%	61.3%	53.0%	48.1%
CTA	32.1%	45.2%	33.1%	30.4%
CPO	35.5%	45.8%	38.2%	33.4%

41.2% of the funds display significant changes in factor loadings. The frequency increases dramatically with observation history, from 19.4% for funds with less than 36 observations, to 38.8% for funds between 36 and 59 observations and 48.6% for funds with at least 60 observations. The longer the time-series, the greater the likelihood of having a large number of return observations before and after the changepoint and, hence, the greater likelihood of being able to identify its location. Recall that, in Table VIII, we showed that our ability to reject the null hypothesis is lower for funds with short histories.

Panel B of Table IX also shows that, of the four different categories of funds, the greatest propensity to switch factor loadings occurs in funds of funds with more than 60 observations, with significant switches occurring in 66% of the 462 funds. One possible explanation for this result is that it might be easier for a fund of funds manager to switch strategies than a hedge fund manager: While the former can simply shift the allocation to a different type of hedge fund, the latter must actually switch investing styles in potentially illiquid markets. Another is that, because funds of funds are more diversified (i.e., have lower idiosyncratic risk), changes in factor loadings are easier to identify.

For the subsample of funds experiencing a statistically significant change in factor loadings, Panels C and D of Table IX show the improvement in adjusted-R^2 that is achieved when factors are allowed to change. Panel C lists the average adjusted-R^2 when parameters are held constant. Overall, the constant parameter model explains 29.9% of the variability of hedge fund excess returns. Interestingly, the adjusted-R^2 is generally decreasing in history length, presumably because the restriction on factor loadings becomes more binding for funds with longer histories. Panel D shows the average adjusted-R^2 for the same set of funds using the changepoint regression. The average adjusted-R^2 increases to 45.4%. In general, all subcategories experience large increases in explanatory power. The largest relative increase in adjusted-R^2 is for CTAs, rising from 18.9% to 32.1%. Overall, the results show that allowing factor loadings to change is important for about 40% of the funds in our sample and that, for these funds, the increase in explanatory power is substantial.

Table X contains the same analysis as is reported in Table IX except that the TASS database is used. Recall that the TASS sample is largely independent in the sense that only it includes only 23% of the funds contained in CISDM and therefore provides us with an important robustness check. Interestingly, the TASS results reported in Table X are very similar to the results in Table IX. In particular, 41.2% of all funds experienced a significant switch in the CISDM sample, and 38.9% of all TASS funds did the same. Again, the length of the return history is important in being able to identify a significant switch. With less than 36 observations, 20.8% of funds experienced a significant switch, 34.4% of funds with between 36 and 59 observations, and 47.9% of funds with at least 60 observations. For the subsample of funds experiencing a statistically significant change in factor loadings, Panels C and D of Table X are also remarkably similar to their counterparts in Table IX. The constant parameter model explains 29.9% of the variability of TASS fund excess returns, and this figure rises to 45.6% using the changepoint regression. Recall these numbers are 29.9% and 45.4%, respectively, for the CISDM funds.

Table XI reports the magnitude of the changes in factor loadings for funds with statistically significant changes. Listed for each factor is (a) the number of funds for which the optimal factor model includes the factor, (b) the average factor loading prior to the switch, and (c) the interquartile range of switches in factor loading. The market excess return *MKTXS* is the most commonly

Table X

Frequency of Significant Parameter Changes in Factor Models Estimated Using Reported Monthly Returns of 2,751 Live TASS Funds

See Appendix Table A1 for definitions of fund types. Panel A shows the number of active funds categorized by fund type and history length in months. Panel B shows the percentage of funds for which a constant-beta model can be rejected in favor of a switching-beta model at the 10% probability level using the Fung and Hsieh factors listed in Appendix Table A1. Panels C and D compare the average adjusted-R^2 of funds with significant switches in factor loadings when loadings are restricted to be constant (Panel C) and when loadings are allowed to vary (Panel D). Data are from January 1994 through December 2005.

	History Length			
Type	All	$n < 36$	$36 \leq n < 60$	$n \geq 60$
Panel A: Number of Funds				
All	2,751	572	800	1,379
HF	1,652	327	477	848
FOF	796	206	276	314
CTA/MF	303	39	47	217
Panel B: Percent of Funds with Significant Switches				
All	38.9%	20.8%	34.4%	47.9%
HF	39.1%	22.0%	32.1%	48.5%
FOF	44.9%	20.9%	41.7%	61.2%
CTA/MF	22.4%	10.3%	14.9%	26.3%
Panel C: Adjusted-R^2 Constant Beta				
All	29.9%	31.9%	32.6%	28.6%
HF	27.8%	29.0%	26.6%	28.3%
FOF	35.5%	37.2%	41.4%	31.9%
CTA/MF	20.2%	25.5%	19.9%	19.8%
Panel D: Adjusted-R^2 Switching Beta				
All	45.6%	50.4%	48.8%	43.6%
HF	44.2%	47.7%	44.9%	43.5%
FOF	50.6%	55.5%	54.7%	47.4%
CTA/MF	32.9%	44.5%	35.0%	31.8%

included factor. For the live fund results reported in Panel A, the average factor loading on *MKTXS* prior to the switch is 0.3357. The 25^{th} and 75^{th} percentiles for the change in the factor loadings are −0.2530 and 0.3051, respectively. This means that while the average factor loading is 0.3352 before the switch, it is less than 0.0827 or greater than 0.6408 after the switch in 50% of the cases. Changes in other factor loadings are similar in magnitude. These results show that switches in hedge fund risk exposure are often quite large, and hence economically important, as should be expected considering the improvement in adjusted-R^2 reported in Table IX. Panel B of Table XI shows the changes in factor loadings for dead funds. In general, the interquartile range is substantially wider than that of live funds. One interpretation of this result is that failing

Table XI

Significant Parameter Changes in Factor Models Estimated Using Reported Monthly Returns of 1,242 Live CISDM Funds and 1,207 Dead CISDM Funds

See Appendix Table A1 for definitions of factors. Listed are summary statistics of factor exposures of funds for which a constant-beta model can be rejected in favor of the following switching-beta model at the 10% probability level:

$$R_t = \alpha_0 + \beta_0^{\mathrm{T}} F_t + \varepsilon_t \qquad \text{for } t = 1, \ldots, T\pi$$

$$R_t = \alpha_0 + \alpha_1 + \left(\beta_0^{\mathrm{T}} + \beta_1^{\mathrm{T}}\right) F_t + \varepsilon_t \quad \text{for } t = T\pi + 1, \ldots, T,$$

where $T\pi$ is the switch date. Listed for each factor are the number of funds for which the factor is selected, the average factor loading prior to the switch in factor loadings, and the 25th, 50th, and 75th percentiles of the distributions of switch magnitudes. Data are from January 1994 through December 2005.

			β_1		
Factor	No. of Funds	β_0	25th	50th	75th
		Panel A: Live Funds			
MKTXS	764	0.3357	−0.2530	0.0558	0.3051
SMB	286	0.3640	−0.3420	−0.1273	0.0391
HML	292	0.1838	−0.0869	0.1252	0.3064
SMBSQ	245	0.3902	−0.7679	−0.3567	−0.0437
HMLSQ	164	0.0945	−0.4368	−0.1039	0.3456
D10YR	225	−0.0868	−0.1376	−0.0645	0.0373
DSPRD	369	−0.2367	−0.1133	0.0139	0.2623
PTFSBD	143	0.0627	−0.3409	0.0209	0.2188
PTFSFX	164	0.2767	−0.2261	−0.0228	0.1625
PTFSCOM	95	0.4678	−0.7184	−0.1159	0.0961
PTFSIR	51	0.0425	−0.2591	−0.0339	0.2037
PTFSSTK	104	−0.0189	−0.0501	0.0540	0.2616
		Panel B: Dead Funds			
MKTXS	596	0.4970	−0.3933	−0.1055	0.3135
SMB	222	0.3872	−0.3600	−0.0877	0.1532
HML	227	−0.5324	−0.1508	0.1108	0.4862
SMBSQ	237	0.2487	−0.7171	−0.2932	0.2025
HMLSQ	155	0.2820	−0.8385	−0.2151	0.3305
D10YR	109	−0.1432	−0.3064	0.0281	0.2471
DSPRD	241	−0.2001	−0.3880	−0.0267	0.3374
PTFSBD	221	0.0118	−0.4140	0.0012	0.3047
PTFSFX	147	0.3319	−0.3062	0.0439	0.4653
PTFSCOM	144	0.4766	−0.6976	−0.2990	0.2829
PTFSIR	73	−0.0220	−0.5260	−0.0403	0.1955
PTFSSTK	155	0.1209	−0.3598	−0.0488	0.2119

funds tend to make more radical changes in strategy in an attempt to survive. For funds with exposure to *D10YR* (i.e., interest rate changes), for example, the 25th and 75th percentiles for the change in factor loadings are −0.1376 and 0.0373 for live funds versus −0.3064 and 0.2471 for dead funds.

Thus far, we have documented that about 40% of the funds experience significant shifts in factor loadings. For these funds, the increase in adjusted-R^2 and the magnitude of change in factor loadings are quite dramatic. A remaining issue is whether the changepoints are unique to individual funds or are shared across funds. Fung and Hsieh (2004), for example, document a structural break in hedge fund indices in September 1998, around the collapse of LTCM. In a similar vein, Kosowski, Naik, and Teo (2007) document a structural break in December 2000, which was the height of the U.S. bull market. If the changes in factor loadings we identify with the changepoint regression generally occur near these shared dates, it might be the case that switches in factor loadings are generated not by any action of managers, but rather by exogenous shifts in the relative importance of different factors in the economy. If changes in factor loadings occur throughout our sample period, however, it must be the case that individual fund managers are choosing when to shift allocations across strategies.

To investigate the possibility that our results are being driven by such macroeconomic events, we divide the number of funds that experience a structural break in a given month by the total number of funds that could have experienced a structural break during that month (i.e., all funds with at least 12 months of data prior to the month and at least 11 months of data following the month). Note that this computation controls for the dramatic change in the size of the fund industry over time. The monthly percentages for live funds together with the monthly levels of the S&P 500 index are displayed in Figure 2. As the figure shows, while there are concentrations of switches near the macro event dates noted in the prior literature, there remain large masses away from these dates—masses that indicate fund-specific changes in strategy. Naturally, of the different fund categories, the largest masses on the two macro-event dates are for funds of funds. Funds of funds diversify away the idiosyncratic risk of individual funds, the very risk that allows us to identify fund-specific changes. An analysis of dead funds produces qualitatively similar results. Thus, it appears that individual managers are actively changing factor loadings, as opposed to factor loadings shifting as a consequence of breaks in the time-series of underlying strategy returns.

Turning to another robustness issue, we noted earlier that we interpret alpha as the mean excess return generated by a fund manager beyond that attributable to investment in the chosen set of strategy-mimicking factors. This implies, of course, that the set of factors used in the regression model are tradable instruments. While the set of factors used in Tables IX through XI are commonly used in studies of fund performance, it is not clear whether any of the factors other than *MKTXS* are easily mimicked. For this reason, we repeat our analysis using tradable instruments, specifically, highly liquid futures contracts. Table XII reports the frequency of significant changes in the futures contract factor loadings. Only funds with at least 36 observations are studied. For a fund with T monthly observations, valid switch dates are months 19 though $T-17$, so that at least 18 observations are used to estimate parameters of each regime. The results are, again, arranged by type of fund (vertically)

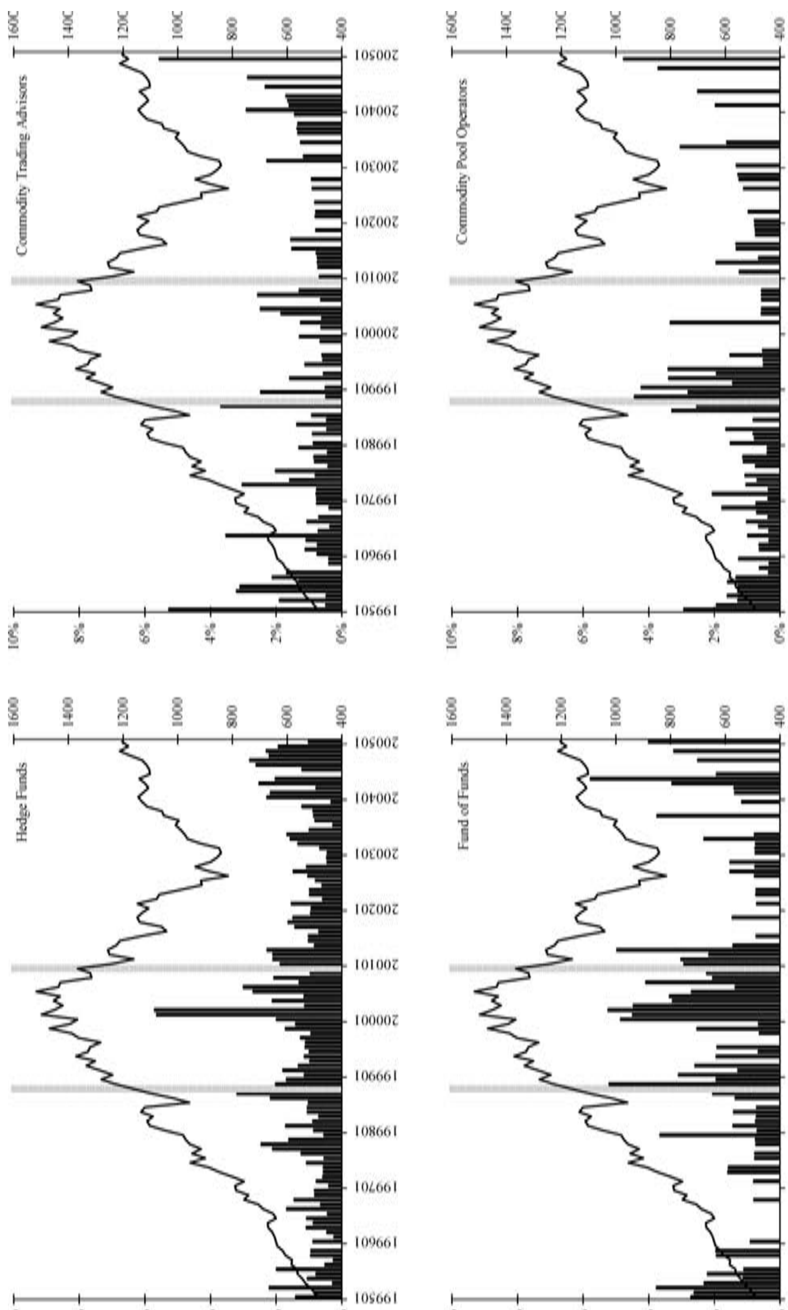

Figure 2. Switching frequency of live funds. Dark bars are the percentage of funds that featured a statistically significant change in parameters in a given month. The percentage is the number of funds featuring a change in parameters divided by the number of funds that had a valid switching opportunity. Only active funds as of December 2005 in the CISDM database are included. Light bars indicate September 1998 and December 2000. The line is the level of the S&P 500 index (right axis).

Table XII

Frequency of Significant Parameter Changes in Factor Models Estimated Using Reported Monthly Returns of 3,013 CISDM Live Funds

See Appendix Table A1 for definitions of fund types. Panel A shows the number of active funds categorized by fund type and history length in months. Panel B shows the percentage of funds for which a constant-beta model can be rejected in favor of a switching-beta model at the 10% probability level using the futures contract factors listed in Appendix Table A1. Panels C and D compare the average adjusted-R^2 of funds with significant switches in factor loadings when loadings are restricted to be constant (Panel C) and when loadings are allowed to vary (Panel D). Data are from January 1994 through December 2005.

	History Length		
Type	All	$36 \leq n < 60$	$n \geq 60$
Panel A: Number of Funds			
All	2,481	860	1,621
HF	1,198	422	776
FOF	797	335	462
CTA	271	65	206
CPO	215	38	177
Panel B: Percent of Funds with Significant Switches			
All	39.7%	30.9%	44.3%
HF	42.1%	35.8%	45.5%
FOF	40.4%	25.7%	51.1%
CTA	32.5%	29.2%	33.5%
CPO	32.6%	26.3%	33.9%
Panel C: Adjusted-R^2 Constant Beta			
All	20.7%	21.2%	20.6%
HF	20.4%	19.2%	20.9%
FOF	23.4%	26.3%	22.4%
CTA	14.7%	16.6%	14.2%
CPO	18.4%	15.9%	18.8%
Panel D: Adjusted-R^2 Switching Beta			
All	35.0%	37.4%	34.1%
HF	35.3%	36.1%	35.0%
FOF	37.0%	40.9%	35.6%
CTA	28.6%	35.3%	26.8%
CPO	30.9%	30.3%	31.0%

and the number of monthly return observations (horizontally). Panel A lists the number of funds in each subcategory, and Panel B reports the frequency with which funds in each subcategory reject the null hypothesis of constant betas. As the table shows, about 39.7% of all funds display significant changes in factor loadings. The frequency of shifts increases dramatically with observation history, from 30.9% for funds with between 36 and 59 observations to 44.3% for funds with at least 60 observations. Similar to the results reported in Table IX, the longer the time-series, the greater the likelihood of having a large number of return observations before and after the changepoint and, hence, the greater the likelihood of being able to identify its location. Also similar to Table IX, the greatest propensity to switch factor loadings occurs in funds of funds with more than 60 observations.

Panels C and D of Table XII document the increase in adjusted-R^2 when factors are allowed to change. For the funds with statistically significant changes in factor loadings, Panel C shows the average adjusted-R^2 when factor loadings are restricted to be constant. Overall, the adjusted-R^2 is 20.7%. Panel D shows the average adjusted-R^2 for the same set of funds when factor loadings are allowed to change. Like in Table IX, the improvement is substantial, with the average adjusted-R^2 increasing from 20.7% to 35%. Clearly, allowing futures factor loadings to change is important in explaining the overall variability of hedge fund returns.

The success of the futures factor loadings in explaining the variability of hedge fund returns is not specific to the CISDM database. Table XIII contains the same analysis as in Table XII except that the TASS database of hedge fund returns is used. Again, the results from the two data sources are remarkably similar even though the overlap in sample composition is small. With the CISDM sample, 39.7% of all funds experienced a significant switch, while 39.6% of all TASS funds did the same. And, where the constant parameter model explains 20.9% of the variability of TASS fund excess returns, it rises to 35.2% using the changepoint regression. Recall these numbers are 20.7% and 35%, respectively, for the CISDM funds.

The adjusted-R^2s when using the futures factors in the changepoint regression (Tables XII and XIII) are considerably lower than those reported for the Fung and Hsieh factors (Tables IX and X). While these results are not directly comparable since slightly different samples and estimation periods are used, part of the difference in explanatory power is attributable to variables such as the change in credit spread, *DSPRD*, and the Fama–French size and value factors, *SMB* and *HML*. These factors are not easily mimicked using liquidly traded instruments, but appear to have significant explanatory power. The chosen set of factors has important practical implications. Consider Figure 3. On the vertical axis is the performance rank of a fund using the alpha from a constant parameter regression using the Fung and Hsieh factors, and on the horizontal axis is the rank of a fund from a constant parameter regression using futures contract factors. If all funds were similarly ranked using both methods, the coordinates of all fund rankings would appear on a diagonal line emanating upward from the origin. Clearly, such is not the case. While the darkness of the diagonal indicates that the two methods often produce similar rankings, there remains considerable mass in the off-diagonal areas. Indeed, there are many instances in which the two rankings procedures produce diametrically opposite results. Since past research uses the Fung and Hsieh factors in assessing abnormal performance, we will adopt their use in the abnormal performance tests that follow, and leave the tradability issue for future research.

Table XIII

Frequency of Significant Parameter Changes in Factor Models Estimated Using Reported Monthly Returns of 2,751 Live TASS Funds

See Appendix Table A1 for definitions of fund types. Panel A shows the number of active funds categorized by fund type and history length in months. Panel B shows the percentage of funds for which a constant-beta model can be rejected in favor of a switching-beta model at the 10% probability level using the futures contract factors listed in Appendix Table A1. Panels C and D compare the average adjusted-R^2 of funds with significant switches in factor loadings when loadings are restricted to be constant (Panel C) and when loadings are allowed to vary (Panel D). Data are from January 1994 through December 2005.

	History Length		
Type	All	$36 \leq n < 60$	$n \geq 60$
Panel A: Number of Funds			
All	2,179	800	1,379
HF	1,325	477	848
FOF	590	276	314
CTA/MF	264	47	217
Panel B: Percent of Funds with Significant Switches			
All	39.6%	30.5%	44.8%
HF	42.0%	34.6%	46.1%
FOF	36.6%	23.2%	48.4%
CTA/MF	34.1%	31.9%	34.6%
Panel C: Adjusted-R^2 Constant Beta			
All	20.9%	22.4%	20.3%
HF	20.3%	19.4%	20.6%
FOF	23.9%	29.5%	21.5%
CTA/MF	17.6%	24.6%	16.1%
Panel D: Adjusted-R^2 Switching Beta			
All	35.2%	39.0%	33.7%
HF	35.5%	37.3%	34.7%
FOF	36.8%	42.7%	34.3%
CTA/MF	30.0%	41.5%	27.7%

Finally, to see how the changepoint regression works in more traditional environments, we repeat the analysis using a set of open-ended equity mutual funds. In addition to the excess return of the market, *MKTXS*, we use the standard Fama–French size and value factors, *SMB* and *HML*, as well as the Carhart (1997) momentum factor, *UMD*. Since equity mutual fund managers use relatively simple strategies in a well-defined asset class, we do not have the factor selection problem that complicates the hedge fund analysis. As before, we choose a maximum of three factors to minimize the Bayesian Information Criterion. Table XIV contains a summary of the results. In the table, the subcategories of 6,840 mutual funds are Aggressive Growth, AG; Growth and

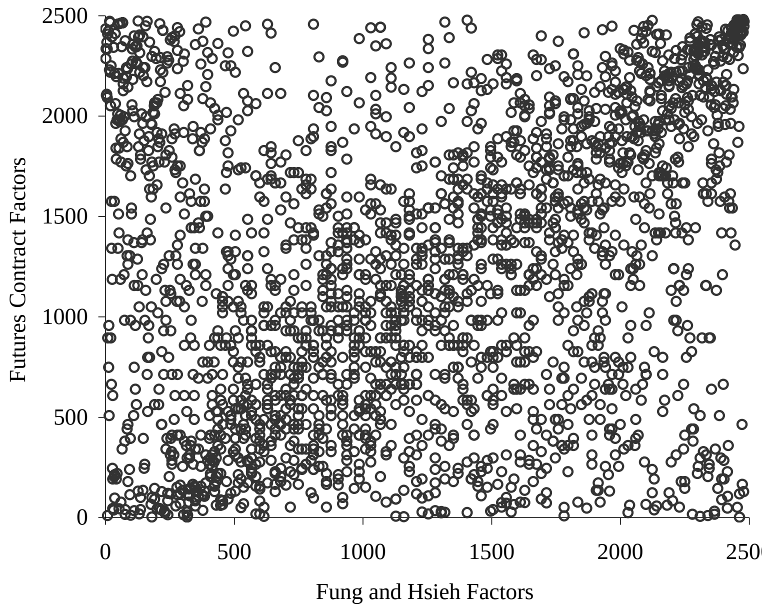

Figure 3. Impact of factor selection on fund ranking. Each of the 2,481 live CISDM funds with at least 36 monthly observations is ranked using two metrics: alpha from a constant parameter regression using the Fung and Hsieh factors and alpha from a constant parameter regression using futures contract factors. The horizontal axis presents rankings based on Fung and Hsieh factors and the vertical axis presents rankings based on futures contract factors.

Income, GI; Long-term Growth, LG; Balanced, BL; and Total Return, TR. Long-term Growth funds are the single largest subcategory, followed by Aggressive Growth.

The results reported in Table XIV show that mutual funds produce more significant changes in factor loadings than do hedge funds. For funds with more than 60 observations, for example, 73.9% of the mutual funds experience significant parameter changes, compared to the 49.1% of hedge funds and 66% of funds of funds reported in Table IX. Time variation in mutual fund factor loadings might seem surprising given their restrictions on short selling, leverage, and derivatives trading, as well as the well-defined strategies provided in many mutual fund prospectuses. The strategy shifts we document in mutual funds, however, are consistent with the vast market timing literature, which describes how mutual fund managers shift equity holdings in an effort to increase factor exposures prior to high factor returns. Even if a fund manager is constrained to a fixed allocation to equities, she can change factor exposures by holding individual stocks that possess the desired level of correlation with the factors. Jiang, Yao, and Yu (2007), for example, use changes in mutual fund portfolio holdings derived from the Thomson Financial data set and find that the median fund has significant timing ability.

Note also that the percentage of funds that switch factor loadings does not reveal information regarding the magnitude of the changes. Minor switches

Table XIV

Frequency of Significant Parameter Changes in Factor Models Estimated Using Reported Monthly Returns of 6,840 CRSP Mutual Funds

See Appendix Table A1 for definitions of fund types. Panel A shows the number of funds categorized by fund type and history length in months. Panel B shows the percentage of funds for which a constant-beta model can be rejected in favor of a switching-beta model at the 10% probability level. Panels C and D compare the average adjusted-R^2 of funds with significant switches in factor loadings when loadings are restricted to be constant (Panel C) and when loadings are allowed to vary (Panel D). Data are from January 1994 through December 2005.

	History Length			
Type	All	$n < 36$	$36 \le n < 60$	$n \ge 60$
Panel A: Number of Funds				
All	6,840	578	1,559	4,703
AG	2,076	142	448	1,486
GI	1,527	116	323	1,088
LG	3,008	314	761	1,933
BL/TR	229	6	27	196
Panel B: Percent of Funds with Significant Switches				
All	62.6%	29.1%	40.7%	73.9%
AG	66.3%	24.7%	36.2%	79.4%
GI	68.7%	31.9%	42.4%	80.4%
LG	56.6%	29.9%	42.6%	66.4%
BL/TR	66.4%	33.3%	44.4%	70.4%
Panel C: Adjusted-R^2 Constant Beta				
All	81.9%	82.5%	86.3%	81.1%
AG	77.0%	78.3%	82.0%	76.2%
GI	85.1%	83.1%	88.7%	84.6%
LG	84.4%	84.3%	87.8%	83.5%
BL/TR	76.7%	60.2%	75.3%	77.1%
Panel D: Adjusted-R^2 Switching Beta				
All	87.3%	89.0%	91.0%	86.5%
AG	83.0%	85.4%	87.1%	82.3%
GI	90.9%	91.4%	93.0%	90.5%
LG	88.9%	89.7%	92.2%	88.0%
BL/TR	83.3%	74.4%	89.3%	83.0%

in factor loadings can likely be detected in mutual funds due to the relatively low level of residual volatility in mutual fund factor models. In Panel C of Table XIV, the average adjusted-R^2 of the mutual funds is 81.9% when betas are restricted to be constant, indicative of the low residual volatility. The adjusted-R^2 increases only slightly to 87.3% when betas are allowed to change.

To assess further the degree to which factor loadings change in mutual funds using our approach, we summarize the magnitude of the parameter changes in Table XV. Clearly the most important factor exposure for the mutual funds

Table XV

Magnitude of Significant Parameter Changes in Factor Models Estimated Using Reported Monthly Returns of 6,840 CRSP Mutual Funds

See Appendix Table A1 for definitions of factors. Listed are summary statistics of factor exposures of funds for which a constant-beta model can be rejected in favor of the following switching-beta model at the 10% probability level:

$$R_t = \alpha_0 + \beta_0^{\mathrm{T}} F_t + \varepsilon_t \qquad \text{for } t = 1, \ldots, T\pi$$

$$R_t = \alpha_0 + \alpha_1 + \left(\beta_0^{\mathrm{T}} + \beta_1^{\mathrm{T}}\right) F_t + \varepsilon_t \quad \text{for } t = T\pi + 1, \ldots, T,$$

where $T\pi$ is the switch date. Listed for each factor are the number of funds for which the factor is selected, the average factor loading prior to the change in factor loadings, and the 25th, 50th, and 75th percentiles of the distributions of switch magnitudes. Data are from January 1994 through December 2005.

			β_1		
Factor	No. of Funds	β_0	25th	50th	75th
MKTXS	4,272	1.0126	−0.1303	0.0195	0.1920
SMB	2,148	0.2992	−0.2729	−0.0074	0.2126
HML	2,404	0.0813	−0.2582	0.0731	0.3576
UMD	2,756	0.0850	−0.3110	0.0717	0.2994

is the excess return of the market, *MKTXS*, selected in 4,272 of the 4,282 mutual funds experiencing a significant change in factor loading. In the period before the switch, the average factor loading on *MKTXS* is 1.0126. The 25th and 75th percentiles for the change in the *MKTXS* factor loadings are −0.1303 and 0.1920, respectively. In other words, where the average factor loading on the excess return of the market is 1.0126 before the switch, it is between 0.8823 and 1.2046 after the switch in 50% of the cases. Thus, for the factor that constitutes the bulk of mutual fund risk, the changes in exposure are generally quite small.[20] These results are comparable to those reported by Mamaysky et al. (2008), who use a stochastic beta model to find an unconditional exposure to the market of 0.93 and a monthly standard deviation of 0.14 for mutual funds with moderate turnover.

VI. Causes and Effects of Changing Risk Exposures

The evidence reported thus far shows that over 40% of the live hedge funds and almost 50% of the live funds of funds in our sample experience a statistically significant shift in risk exposures. We have also shown that allowing risk

[20] Changes in exposure to the other three factors can be substantial relative to the initial exposures. Unlike the case of hedge funds, though, where changes in risk exposure can indicate a fundamental switch in strategy, changes in exposure to the *SMB*, *HML*, or *UMD* factors simply represent a reallocation of portfolio weights from one type of stock to another.

exposures to change increases substantially the explanatory power of factor models. Taken together, these results suggest that many fund managers make fundamental changes to strategy over time. In this section, we study these changes in greater detail. More specifically, we examine (a) the length of time or duration that a particular strategy is held in place, (b) fund performance before and after switches in strategy, and (c) the time-series intensity of strategy switches—all in an attempt to better understand when, during a fund's life, a change in strategy is likely to occur. After these examinations, we gauge the effects that changing risk exposures have on fund performance measurement. If risk parameters are restricted to be constant when, in fact, they change, alpha will be measured incorrectly.

A. When Funds Switch

In this subsection, we examine the conditions under which funds switch strategies. First, we document when strategy changes occur in relation to a fund's life cycle. Next, we examine the performance of funds before and after switches. Finally, we document the intensity of switches in calendar time and by strategy to illustrate the extent to which common events can lead to common switches.

This study characterizes a fund strategy as a unique set of factor loadings. At the beginning of a fund's life, the fund manager chooses a particular strategy. As time passes, the manager periodically evaluates whether to maintain his original strategy or switch to a new strategy characterized by a new and different set of factor loadings. We call the time until the factor loadings change the *duration* of the fund's original strategy. For each month t during the sample period January 1994 through December 2005, we compute the ratio of the number of funds that switched strategies in the month t to the number of funds that had a history at least as long as month t but did not switch prior to month t. As shown in the Appendix, this ratio can be interpreted as the hazard rate for switching strategies, that is, the probability of a switch in strategy conditional on not yet having switched.

Figure 4 shows the hazard rates of the live and dead CISDM funds in our sample. No switches occur before month 13 or after month 132 since we require at least 12 observations in each regime. At least two features of Figure 4 are noteworthy. First, the spike in month 13 suggests that many funds switch very early in their lives, as least as reflected by our sample. These are all clustered in month 13. Possible reasons for early switches include initially focusing on a specific strategy and then branching out to other strategies as the fund ages and reducing exposures to particular risk factors after attracting a target level of capital. Second, other than at month 13, there are pronounced spikes. One such spike occurs in month 75 for both live and dead funds. For those funds that begin reporting in January 1994, this corresponds to March 2000. For dead funds, there is another spike at month 96. For those funds that begin reporting in January 1994, this corresponds to December 2001. We study each of these events in greater detail below.

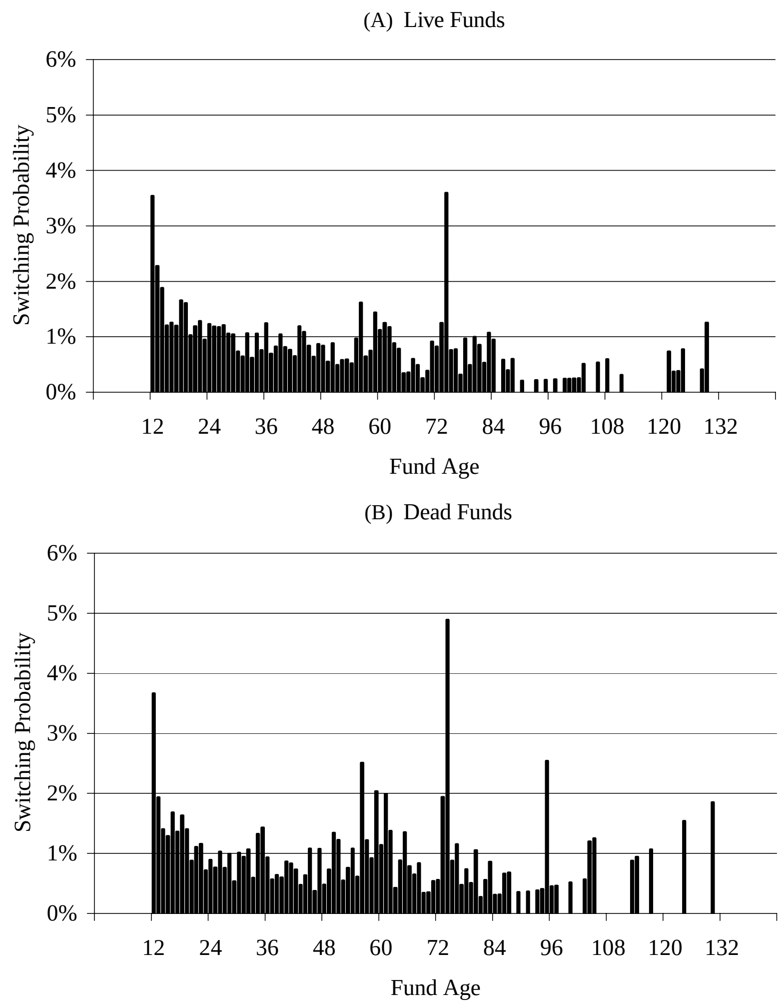

Figure 4. Duration of strategies. The figures show the probability with which funds switch strategies as a function of the age of the fund. For funds with a significant switch in factor exposures, we record the age at which the switch occurs. The horizontal axis is fund age, in months. The height of the vertical bars shows the number of switches at each age as a percentage of the number of funds that have a history at least as long as the age and that did not switch prior to the age. Figures 4A and 4B show the results for live and dead CISDM funds, respectively.

Figure 4 also suggests that the preponderance of funds that switch strategies switch early in their lives. What motivates the decision to switch? One possibility is fund performance. Table XVI reports summary statistics for the returns and performance of live and dead CISDM funds during the sample period January 1994 through December 2005, with each group split into two parts based on whether the fund switched using the changepoint regression. The results are striking in two ways. First, the average Sharpe ratio for live

Table XVI

Comparison of Funds That Switch Exposures to Those That Do Not

See Appendix Table A1 for definitions of fund types. Listed are summary statistics of 3,013 live funds and 3,145 dead funds in the CISDM database. Funds are split into Switchers, for which a constant-beta model can be rejected in favor of a switching-beta model at the 10% probability level, and Non-switchers. The summary statistics are the number of funds and the equally-weighted averages of the mean monthly return, μ; the standard deviation of monthly returns, σ; the Sharpe ratio, *SR*; the skewness, *Skew*; and the excess kurtosis, *Kurt*. Data are from January 1994 through December 2005.

		Switchers					Non-switchers					
Type	No. of Funds	μ	σ	*SR*	*Skew*	*Kurt*	No. of Funds	μ	σ	*SR*	*Skew*	*Kurt*
Panel A: Live Funds												
HF	596	0.0117	0.0348	0.4261	0.1416	5.1561	849	0.0113	0.0365	0.3219	0.1615	2.2699
FOF	510	0.0070	0.0169	0.3530	−0.2838	3.5995	512	0.0071	0.0168	0.3710	−0.2040	1.3677
CTA	87	0.0115	0.0524	0.1756	0.6098	3.2476	215	0.0114	0.0562	0.1623	0.3401	1.4741
CPO	49	0.0093	0.0493	0.1550	0.6113	4.1756	195	0.0082	0.0527	0.1247	0.3472	1.2119
	1,242						1,771					
Panel B: Dead Funds												
HF	670	0.0096	0.0545	0.1757	−0.1133	4.9137	952	0.0091	0.0536	0.1636	0.0391	2.9104
FOF	176	0.0063	0.0276	0.1895	−0.4956	5.4286	197	0.0051	0.0282	0.1431	−0.1445	2.6349
CTA	169	0.0092	0.0585	0.0607	0.5695	4.6683	344	0.0083	0.0662	0.0523	0.3450	2.1665
CPO	192	0.0033	0.0530	0.0083	0.0323	4.1746	445	0.0044	0.0523	0.0154	0.2693	1.8480
	1,207						1,938					

CISDM funds that switch is 0.4261, while the average Sharpe ratio for nonswitchers is 0.3219. This suggests that switching funds are associated with superior performance. This is consistent with the story that funds with strong performance quickly attract mimickers. The resulting flood of capital into the existing strategy dampens (and eventually eliminates) its profitability, forcing funds to search for "virgin terrain."

Another interesting feature of Table XVI is that both switching live funds and switching dead funds have high levels of excess kurtosis relative to nonswitchers. One possible explanation for this phenomenon is that, for switchers, the return distribution over the entire time-series is a mixture of distributions before and after the switch date. To test for this possibility, we compute summary statistics for the return distributions for switching funds before and after the switch dates as defined by the changepoint regression. Table XVII contains the results. Note that, for both the pre-switch and post-switch periods, the excess kurtosis of switchers is now comparable to the non-switchers in Table XVI. Driving this result is, among other things, a dramatic shift in the means of the return distributions before and after the switch date. For live funds, the mean monthly return falls from 1.81% a month to 0.88%, and, for dead funds, the mean monthly return falls from 2.07% to −0.21%. For live funds, we see substantial risk reduction—the standard deviation of monthly returns falls from 4.34% to 2.84%. Thus, it seems that we have found the explanation for the excess kurtosis reported in Table XVI.

Table XVII also shows that while live hedge funds have higher Sharpe ratios before the switch, they continue to have high Sharpe ratios after the switch—more than double the Sharpe ratio of the excess market return, *MKTXS*, reported in Table III. In contrast, dead hedge funds see their Sharpe ratios collapse after the switch. These results suggest that shifts in risk exposure are fundamentally related to the life cycle of hedge funds.

In Section V and earlier in this section, we note that switch dates identified by the changepoint regression tend to cluster. Some of the clustering is due to macro events such as the LTCM debacle in September 1998. Other clustering, however, is due to more narrowly defined events. To illustrate, we measure the intensity of switches by fund strategy. In this context, the term "strategy" is defined as a statistically significant exposure to a given risk factor. For each month in which a switch can occur (i.e., January 1995 through January 2005 to allow at least 12 return observations in each regime), we count the number of funds with a significant exposure to a given risk factor. Of these funds, we then count the number that experienced a switch in that month. The ratio of switches to total in each month is our measure of switching intensity for that strategy.

Figure 5 documents the switching intensities by strategy where the Fung and Hsieh factors are used in the changepoint regressions. The five factors we use based on the Fama–French model (*MKTXS*, *SMB*, *HML*, *SMBSQ*, and *HMLSQ*) all have pronounced peaks on March 2000. The switching frequency on this date is not as pronounced in Figure 2 because the figure aggregates across all funds. Note that the *HML* factor also has peaks in September 1998 and February 1999.

Table XVII

Summary Statistics of Switchers Before and After Switch

See Appendix Table A1 for definitions of fund types. Listed are summary statistics of 1,242 live funds and 1,207 dead funds in the CISDM database for which a constant-beta model can be rejected in favor of a switching-beta model at the 10% probability level. Statistics for each fund are computed separately before and after the switch in exposures occurs. The summary statistics are the number of funds and the equally-weighted averages of the mean monthly return, μ; the standard deviation of monthly returns, σ; the Sharpe ratio, *SR*; the skewness, *Skew*; and the excess kurtosis, *Kurt*. Data are from January 1994 through December 2005.

		Pre-Switch					Post-Switch				
Type	No. of Funds	μ	σ	*SR*	*Skew*	*Kurt*	μ	σ	*SR*	*Skew*	*Kurt*
					Panel A: Switching Live Funds						
HF	596	0.0181	0.0434	0.5496	−0.0092	2.0575	0.0088	0.0284	0.3148	−0.0067	1.8244
FOF	510	0.0082	0.0194	0.3892	−0.2648	1.8713	0.0063	0.0150	0.3142	−0.3988	0.9350
CTA	87	0.0214	0.0678	0.3188	0.4601	1.4371	0.0086	0.0448	0.0771	0.0662	1.2608
CPO	49	0.0146	0.0661	0.2063	0.5092	1.5216	0.0063	0.0414	0.0943	0.1963	1.6591
	1,242										
					Panel B: Switching Dead Funds						
HF	670	0.0207	0.0547	0.4041	0.0442	2.1409	−0.0021	0.0524	−0.0733	−0.1895	1.9274
FOF	176	0.0087	0.0288	0.3960	−0.2478	2.2882	0.0021	0.0250	0.0271	−0.3707	2.1929
CTA	169	0.0222	0.0656	0.2331	0.3995	1.5799	0.0005	0.0540	−0.8328	0.2011	2.2324
CPO	192	0.0073	0.0534	0.0836	0.1189	2.0490	−0.0036	0.0529	−0.0804	0.0162	1.6451
	1,207										

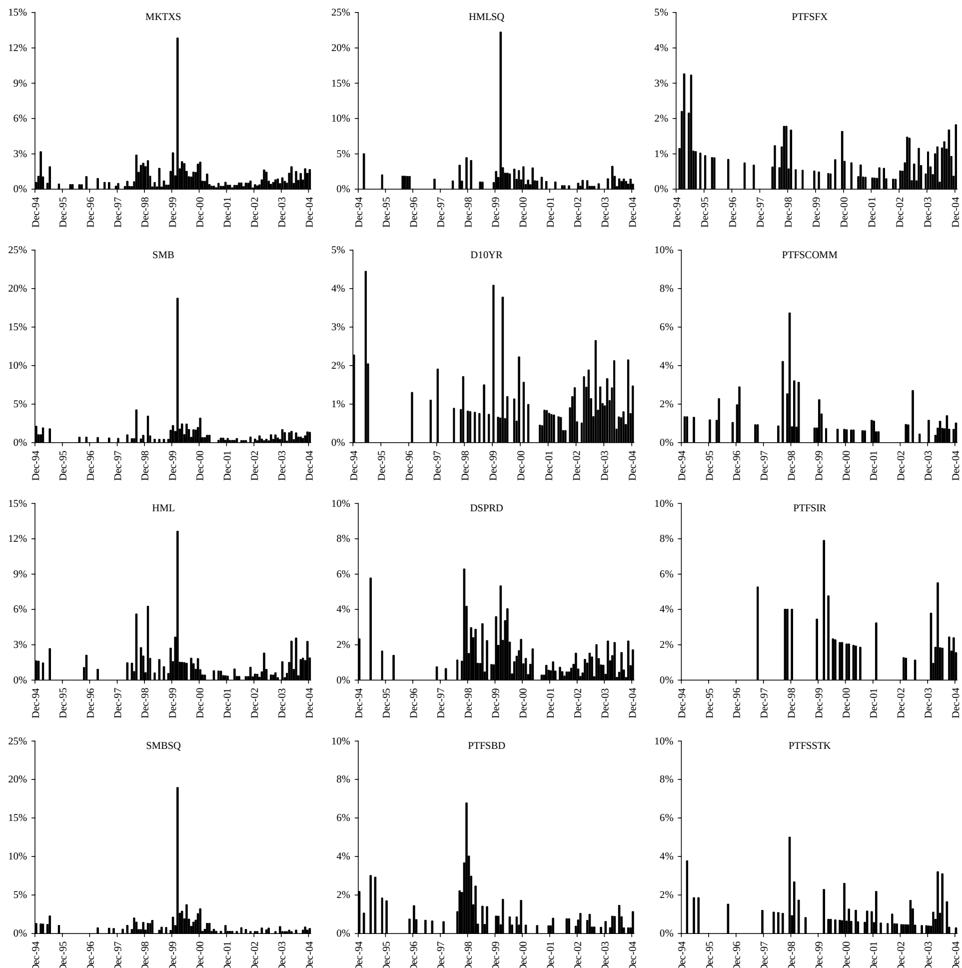

Figure 5. Switching frequency of live funds by strategy. See Appendix Table A1 for definitions of factors. Dark bars are the percentage of funds by strategy that featured a statistically significant change in parameters in a given month. The percentage is the number of funds with exposure to the strategy that feature a change in parameters divided by the number of funds with exposure to the strategy that had a valid switching opportunity in a given month. Only active funds as of December 2005 in the CISDM database are included.

This is no great surprise. Value stocks posted particularly poor performance during 1998. Strategy switches in funds exposed to the change in the 10-year yield, *D10YR*, have spikes in May and June of 1995. This follows 2 months of significant declines in the yield amid fears in the market of a recession and a subsequent inversion of the yield curve. The Federal Open Market Committee dropped the federal funds rate on July 6, 1995 for the first time in almost 3 years as a result. Switches involving the trend-following strategy *PTFSFX* also peaked in the first part of 1995, reflecting the linkage between the profitability of interest rate and exchange rate strategies.

In sum, while Figures 2 and 5 both show clustering of switches at macro, market-specific, and sector-specific levels, considerable mass remains away from the spikes. These are idiosyncratic, fund-specific changes—changes that motivate the fund-by-fund use of the changepoint regression in assessing fund performance.

B. Performance Appraisal

Our focus now turns to examining whether measures of ex-post abnormal performance are affected by allowing risk exposures to change. We begin by annually estimating the parameters of three models: (a) a constant parameter model using all available observations, (b) a constant parameter model using the most recent 24 observations, and (c) the optimal changepoint regression using all funds with at least 24 return observations. Funds are then ranked twice: once using alpha from a constant parameter model and once using alpha from the second regime of a changepoint regression.

Figure 6A compares the ex-post abnormal performance of the changepoint regression with that of the constant parameter model. On the vertical axis is the rank of the fund using the alpha from the constant parameter model, and on the horizontal axis is the rank of the fund using the alpha from the changepoint regression. If the two methods provided similar rankings, the scatter plot would be a diagonal line emanating upward from the origin. The dispersion off the diagonal indicates the two rankings are different—substantially different if located in the upper left or bottom right corners of the figure. To quantify this dispersion, we regress changepoint rank on the constant parameter rank. The slope coefficient is highly significant but has an estimate of just 0.53, and the adjusted-R^2 of the regression is just 28%. Clearly ex-post abnormal performance is affected by risk measurement, as the changepoint regressions allow for shifts in asset classes, strategies, and leverage.

A quick fix alternative to the changepoint regression (and one that is frequently used in practice) is to use a constant parameter regression but limit the number of observations to, say, the most recent 24 months. In effect, adopting this practice is like fixing a changepoint to a rolling arbitrary date 24 months in the past and estimating parameters using only the most recent data. To test the effectiveness of this approach, we repeat the ranking experiment used to generate Figure 6A and compare performance metrics from the second regime of the optimal changepoint regression to those obtained from a constant parameter model using only the most recent 24 observations. Figure 6B shows the relation between the rankings of the two methods. Not surprisingly, the results are more peaked along the diagonal since a constant parameter model using more recent data is closer to the second regime of a changepoint regression than is a constant parameter model using all available data. Nonetheless, a significant mass remains in the off-diagonal area, indicating that the ex-post abnormal performance measures can lead to significantly different results.

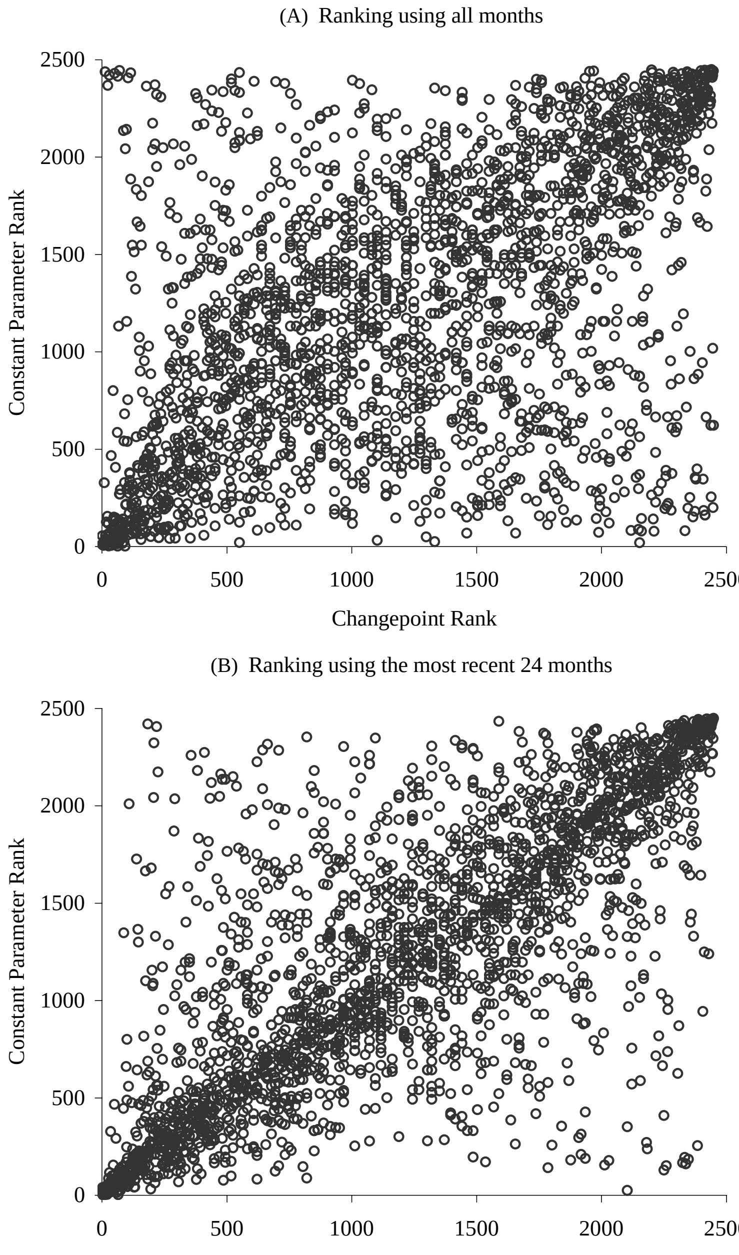

Figure 6. Impact of switching parameters on fund ranking. Parameters of a changepoint regression are estimated using all available data for each of the 2,481 live CISDM funds with a statistically significant switch in parameters. The two sets of rankings are: (1) the post-switch alpha from the changepoint regression and (2) the alpha from a constant parameter model. Figure 6A shows rankings when the constant parameter model uses all available data. Figure 6B shows rankings when the constant parameter model uses the most recent 24 months of data. In both figures, the horizontal axis presents rankings based on the changepoint regression and the vertical axis presents rankings based on the constant parameter model.

Indeed, a regression of changepoint rank on the constant parameter rank has an adjusted-R^2 of only 59%. This result indicates substantial disagreement across the two ranks. We believe the changepoint regression provides more accurate performance measurement because it allows the data to dictate if and when a change in strategy actually occurred.

VII. Conclusions

Hedge fund managers are free to change asset classes, strategies, and leverage in response to changing market conditions and arbitrage opportunities. Typical measures of hedge fund performance, however, fail to recognize these dynamics. The standard approach of measuring exposure to underlying sources of risk is to regress investment returns on risk factors that proxy for different trading strategies. Assuming constant coefficients in an environment where they are time-varying implies performance evaluation can be unreliable.

To remedy the problem, we study two econometric techniques that accommodate changes in risk exposures. The optimal changepoint regression searches for a discrete number of dates on which factor loadings can shift. We allow for a single shift in parameters for each fund. The stochastic beta model specifies an autoregressive process for risk exposures. Through simulation, we demonstrate that in the hedge fund context, the changepoint regression is generally more powerful. Next, we apply the changepoint regression to a sample of live and dead funds during the period January 1994 through December 2005. We find significant changes in the risk factor parameters in about 40% of our sample of hedge funds. We also find that, for live funds, switches tend to occur early in the fund's life and that switchers tend to have higher performance on average than non-switchers.

With significant changes in risk factor parameters, the alphas from a constant parameter regression will be misleading measures of abnormal performance. We investigate this issue through historical performance appraisal. We show that for the subset of funds with significant shifts in risk exposure, a substantial number are ranked incorrectly when constant parameter models are used for evaluation. This underscores the importance of having the correct structural model in assessing performance.

Appendix

A. *Estimation of the Stochastic Beta Model Using a Kalman Filter*

Our task is to estimate parameters of

$$\begin{aligned} R_t &= \alpha + \beta_t^{\mathrm{T}} F_t + \varepsilon_t \\ \beta_t &= \mu + T\beta_{t-1} + \nu_t. \end{aligned} \tag{A1}$$

We assume that the disturbance terms are homoskedastic, uncorrelated both contemporaneously and at all lags, and serially uncorrelated. To construct a likelihood function, we also assume normality, so that

$$\varepsilon_t \sim N(0, S), \quad \nu_t \sim MVN(0, Q). \tag{A2}$$

Following the derivation in Harvey (1989), let $b_{t-1|t-1.}$ denote the optimal estimator of β_{t-1} given return observations of R up to and including time $t-1$, and let $P_{t-1|t-1.}$ denote the variance–covariance matrix of the estimator. Step forward one time increment, and consider the optimal estimator for β_t without an additional observation. By exploiting the structure imposed by (A1), we can define the resulting optimal estimator and its associated variance–covariance matrix as

$$\begin{aligned} b_{t|t-1} &= \mu + Tb_{t-1|t-1} \\ P_{t|t-1} &= TP_{t-1|t-1}T^{\mathrm{T}} + Q. \end{aligned} \tag{A3}$$

Now, using a property of multivariate normal distributions, we can construct the optimal estimator for β_t given R_t, F_t, and the associated variance–covariance matrix as

$$\begin{aligned} b_{t|t} &= b_{t|t-1} + P_{t|t+1}F_t^{\mathrm{T}}\left(F_t P_{t|t+1}F_t^{\mathrm{T}} + S\right)^{-1}\left(R_t - \alpha - b_{t|t-1}^{\mathrm{T}}F_t\right) \\ P_{t|t} &= P_{t|t+1} - P_{t|t+1}F_t^{\mathrm{T}}\left(F_t P_{t|t+1}F_t^{\mathrm{T}} + S\right)^{-1}F_t P_{t|t+1}. \end{aligned} \tag{A4}$$

To construct the likelihood function, rewrite the measurement equation as

$$R_t = \alpha + b_{t|t-1}^{\mathrm{T}}F_t + \left(\beta_t^{\mathrm{T}} - b_{t|t-1}^{\mathrm{T}}\right)F_t + \varepsilon_t. \tag{A5}$$

For each observation of fund returns, we can now compute the optimal forecast and associated variance:

$$\begin{aligned} E[R_t] &= \alpha + b_{t|t-1}^{\mathrm{T}}F_t \\ \sigma^2[R_t] &= F_t P_{t|t-1}F_t^{\mathrm{T}} + S. \end{aligned} \tag{A6}$$

The log-likelihood is then given by

$$L = -\frac{T}{2}\ln(2\pi) - \sum_t \ln(\sigma[R_t]) - \frac{1}{2}\sum_t\left(\frac{R_t - E[R_t]}{\sigma[R_t]}\right)^2. \tag{A7}$$

Maximizing (A7) produces MLE estimates of the parameters α, S, μ, T, and Q.

B. Hazard Rate of Switching Strategies

Let T be the random time until a hedge fund manager switches strategies. Denote the distribution of durations $f(t)$, so that the unconditional probability of switching at or before time t, evaluated at time 0, is the cumulative density function:

$$F(t) = \int_0^t f(s)\,ds. \tag{A8}$$

Table A1
Definitions of Fund Types and Factors

Panel A: Fund Types	
HF	Hedge fund
FOF	Fund of fund
CTA	Commodity trading advisor
CPO	Commodity pool operator
MF	Managed futures
AG	Aggressive growth
GI	Growth and income
LG	Long-term growth
BL	Balanced
TR	Total return
Panel B: Fung and Hsieh Factors	
MKTXS	Excess return of the CRSP value-weighted index
SMB	Fama–French size factor
HML	Fama–French value factor
SMBSQ	Fama–French size factor squared
HMLSQ	Fama–French value factor squared
D10YR	Change in the 10-year treasury yield
DSPRD	Change in the spread between BAA yield and 10-year treasury yield
PTFSBD	Primitive trend follower strategy bond
PTFSFX	Primitive trend follower strategy currency
PTFSCOM	Primitive trend follower strategy commodity
PTFSIR	Primitive trend follower strategy interest rate
PTFSSTK	Primitive trend follower strategy stock
Panel C: Futures Contract Factors	
SP	S&P 500
ED	Eurodollar
US	30-year U.S. Treasury
CD	Canadian Dollar
JY	Japanese Yen
SF	Swiss Franc
CL	Crude Oil
NG	Natural Gas
C	Corn
GC	Gold
Panel D: Mutual Fund Factors	
MKTXS	Excess return of the CRSP value-weighted index
SMB	Fama-French size factor
HML	Fama-French value factor
UMD	Carhart momentum factor

The unconditional probability of survival up to at least time t, evaluated at time 0, is

$$S(t) = 1 - F(t). \tag{A9}$$

The "hazard rate" at time t is defined as

$$\lambda(t) = \frac{f(t)}{S(t)} \tag{A10}$$

and can be interpreted as the probability of switching per time increment conditioned on not switching up until the beginning of the time increment.

REFERENCES

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik, 2007, Role of managerial incentives and discretion in hedge fund performance, Working Paper, Georgia State University.

Agarwal, Vikas, William F. Fung, Yee Cheng Loon, and Narayan Y. Naik, 2006, Risk and return in convertible arbitrage: Evidence from the convertible bond market, Working Paper, Georgia State University.

Agarwal, Vikas, and Narayan Naik, 2004, Risks and portfolio decisions involving hedge funds, *Review of Financial Studies* 17, 63–98.

Andrews, Donald W.K., Inpyo Lee, and Werner Ploberger, 1996, Optimal changepoint tests for normal linear regression, *Journal of Econometrics* 70, 9–38.

Asness, Clifford S., Robert Krail, and John M. Liew, 2001, Do hedge funds hedge? *Journal of Portfolio Management* 28, 6–19.

Bollen, Nicolas P.B., and Jeffrey A. Busse, 2001, On the timing ability of mutual fund managers, *Journal of Finance* 56, 1075–1094.

Bollen, Nicolas P.B., and Veronika Krepely Pool, 2008, Conditional return smoothing in the hedge fund industry, *Journal of Financial and Quantitative Analysis* 43, 267–298.

Brown, R.L., J. Durbin, and J.M. Evans, 1975, Techniques for testing the constancy of regression relationships over time, *Journal of the Royal Statistical Society* B37, 149–192.

Busse, Jeffrey A., and Paul J. Irvine, 2006, Bayesian alphas and mutual fund persistence, *Journal of Finance* 61, 2251–2288.

Carhart, Mark, 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Fama, Eugene, 1965, The behavior of stock-market prices, *Journal of Business* 38, 34–105.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.

Fisher, Lawrence, 1966, Some new stock market indexes, *Journal of Business* 39, 191–225.

Fung, William, and David Hsieh, 1997, Empirical characteristics of dynamic trading strategies: The case of hedge funds, *Review of Financial Studies* 10, 275–302.

Fung, William, and David Hsieh, 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

Fung, William, and David Hsieh, 2004, Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal* 60, 65–80.

Fung, William, David Hsieh, Narayan Y. Naik, and Tarun Ramadorai, 2008, Hedge funds: Performance, risk and capital formation, *Journal of Finance* 63, 1777–1803.

Getmansky, Mila, Andrew W. Lo, and Igor Makarov, 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* 74, 529–609.

Ghysels, Eric, 1998, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance* 53, 549–573.

Harvey, Andrew C., 1989, *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge University Press).

Henriksson, Roy D., and Robert C. Merton, 1981, On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills, *Journal of Business* 54, 513–533.

Jagannathan, Ravi, and Robert A. Korajczyk, 1986, Assessing the market timing performance of managed portfolios, *Journal of Business* 59, 217–235.

Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov, 2006, Do hot hands persist among hedge fund managers? An empirical evaluation, Working Paper, Northwestern University.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 389–416.

Jiang, George, Tong Yao, and Tong Yu, 2007, Do mutual funds time the market? Evidence from portfolio holdings, *Journal of Financial Economics* 86, 724–758.

Jostova, Gergana, and Alexander Philipov, 2005, Bayesian analysis of stochastic betas, *Journal of Financial and Quantitative Analysis* 40, 747–778.

Kosowski, Robert, Narayan Y. Naik, and Melvyn Teo, 2007, Do hedge funds deliver alpha? A Bayesian and bootstrap analysis, *Journal of Financial Economics* 84, 229–264.

Lin, Chien-Fu, and Timo Terasvirta, 1994, Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics* 62, 211–228.

Lo, Andrew W., 2002, The statistics of Sharpe ratios, *Financial Analysts Journal* 58, 36–50.

Lo, Andrew W., 2004, The adaptive markets hypothesis, *Journal of Portfolio Management* 30, 15–29.

Lo, Andrew W., and A. Craig MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 137–158.

Lowenstein, Roger, 2000, *When Genius Failed* (Random House, NY).

Mamaysky, Harry, Matthew Spiegel, and Hong Zhang, 2008, Estimating the dynamics of mutual fund alphas and betas, *Review of Financial Studies* 21, 233–264.

Mitchell, Mark, and Todd Pulvino, 2001, Characteristics of risk and return in risk arbitrage, *Journal of Finance* 56, 2135–2175.

Ohlson, James, and Barr Rosenberg, 1982, Systematic risk of the CRSP equal-weighted common stock index: A history estimated by stochastic-parameter regression, *Journal of Business* 55, 121–145.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Sharpe, William F., 1992, Asset allocation: Management style and performance measurement, *Journal of Portfolio Management*, 7–19.

Stoll, Hans R., and Robert E. Whaley, 1990, The dynamics of stock index and stock index futures returns, *Journal of Financial and Quantitative Analysis* 25, 441–468.

Treynor, Jack L., and Kay Mazuy, 1966, Can mutual funds outguess the market? *Harvard Business Review* 44, 131–136.