DALLASFED

# Occasional Paper

## Hedge Fund Dynamic Market Sensitivity

*Jiaqi Chen and Michael L. Tindall*

Federal Reserve Bank of Dallas
Financial Industry Studies Department

Occasional Paper 12-01
September 2012

# *Hedge fund dynamic market sensitivity*

by

Jiaqi Chen and Michael L. Tindall
*Financial Industry Studies Department*
***Federal Reserve Bank of Dallas*** [*]
September 2012

**Abstract**

***Many hedge funds attempt to achieve high returns by employing leverage. However, it is difficult to track the degree of leverage used by hedge funds over time because detailed timely information about their positions in asset markets is generally unavailable. This paper discusses how to combine shrinkage variable selection methods with dynamic regression to compute and track hedge fund leverage on a time-varying basis. We argue that our methodology measures leverage as well as hedge fund sensitivity to markets arising from other sources. Our approach employs the lasso variable selection method to select the independent variables in equations of hedge fund excess returns. With the independent variables selected by the lasso method, a state space model generates the parameter estimates dynamically. The hedge fund market sensitivity indicator is the average of the absolute values of the parameters in the excess return equations. Our indicator peaks at the time of the Long Term Capital Management meltdown in 1998 and again at a critical time in the 2008 financial crisis. In the absence of direct information from hedge fund balance sheets, our approach could serve as an important tool for monitoring market sensitivity and financial distress in the hedge fund industry.***

***Keywords: hedge fund leverage and market sensitivity, state space model, dynamic regression, lasso method***

[*] Please direct correspondence to michael.tindall@dal.frb.org. The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

## 1. Introduction

The Long Term Capital Management (LTCM) breakdown in 1998 underscored the importance of understanding and monitoring the use of leverage by hedge funds. LTCM employed leverage of extraordinary magnitudes in the various asset classes in which it operated, necessitating the intervention of Federal Reserve officials to bring about an orderly conclusion to its operation. Monitoring the degree of leverage employed by hedge funds is thus important to policymakers. Leverage, moreover, is just one way that hedge funds increase their sensitivity to the underlying markets in which they invest. Hedge funds can increase their sensitivity to markets by borrowing, by trading futures and options, and by trading assets which are intrinsically sensitive to underlying markets, and it would be useful for policy officials to have an indicator of hedge fund market sensitivity arising from all sources. Monitoring hedge fund market sensitivity with quantitative methods as we propose here falls within a broader class of relatively recent work investigating hedge funds and financial stability by Adrian (2007), Chan et al (2006), Gupta and Lang (2005), and Getmansky et al (2004).

McGuire and Tsatsaronis (2008), MT hereafter, proposed an econometric approach to calculating hedge fund leverage on a time-varying basis using rolling regressions. They estimated the average leverage for hedge funds in their database for short rolling sample periods using panel data for hedge funds' excess returns. In this paper we replace the rolling regression method used in MT with a state space model to address problems inherent with rolling regression and panel data. We propose a hybrid approach to tackle both the variable selection and leverage estimation for individual hedge funds.

First, the lasso technique is applied to select the set of independent variables for each individual hedge fund equation from a pool of 29 potential asset returns, and then a state space model is used to compute the parameters dynamically. These parameters are then used to construct the leverage indicator. We find that the results using the state space model, i.e., dynamic regression, differ markedly from those obtained with rolling regressions and that our results match anecdotal evidence from hedge fund operators concerning the use of leverage by hedge funds during periods of financial distress.

Section 2 discusses our underlying methodology. Section 3 discusses the data used in this paper. Section 4 presents our dynamic method for estimating hedge fund leverage. Section 5 presents empirical results. Section 6 presents conclusions.

**2. Methodology**

MT base their analysis on the idea that a hedge fund return stream can be modeled as a weighted average of the returns of the underlying assets of the fund (Sharpe 1992). Suppose a fund invests in risky assets and a risk-free asset. Then, the hedge fund return can be expressed as:

(1) $R_t^h = \beta_0^h + \sum_{j=1}^{k} \beta_j^h R_{j,t}^a + \beta_{k+1}^h r_{f,t} + \varepsilon_t$

where $R_t^h$ is the return of hedge fund $h$ in month $t$, $\beta_j^h$ are parameters for fund $h$, $R_{j,t}^a$ are the returns of the risky assets in which the fund invests, and $r_{f,t}$ is the risk-free rate. The term $\beta_0^h$ represents the fund manager's intrinsic ability to outperform a passive strategy. When the fund is long (short) a risky asset, the corresponding $\beta_j^h$ is expected to be positive (negative). We further assume the fund can invest at the risk-free rate and that it can also borrow to achieve leverage at the risk-free rate so that $\beta_{k+1}^h < 0$. Then, according to MT, $\sum_{j=1}^{k} |\beta_j^h|$, i.e., the sum of the absolute values of the coefficients of the excess returns of the risky assets, is an estimate of the leverage of the fund. Where the fund is fully invested, $\sum_{j=1}^{k} \beta_j^h + \beta_{k+1}^h = 1 \rightarrow \beta_{k+1}^h = 1 - \sum_{j=1}^{k} \beta_j^h$ so that (1) can be rewritten as:

(2) $R_t^h - r_{f,t} = \beta_0^h + \sum_{j=1}^{k} \beta_j^h (R_{j,t}^a - r_{f,t}) + \varepsilon_t$

eliminating the $\beta_{k+1}^h r_{f,t}$ term. Here, the dependent variable is the excess return of the hedge fund, and the independent variables are excess returns of the asset classes in which the fund invests.

However, suppose that a hedge fund invests, without any special intrinsic ability, only in the underlying market $m$, which is the stock market. Then, equation (2) reduces to:

(3) $R_t^h - r_{f,t} = \beta_m^h (R_{m,t}^a - r_{f,t}) + \varepsilon_t$

which is an empirical form of the CAPM model, and we recognize that $\beta_m^h$ is an estimate of CAPM beta, which can take on a range of values regardless of whether the hedge fund is leveraging by borrowing. So we argue that, correctly described, equation (2) is a model of market sensitivity whether that sensitivity arises from leverage or simply from buying market-sensitive assets, and henceforth we refer to $\sum_{j=1}^{k} |\beta_j^h|$ as the market sensitivity indicator.

## 3. Data

Data for this study come from the Hedge Fund Research Inc. database of hedge fund monthly return streams. Our dataset consists of 156 hedge funds from the database, each with a continuous monthly reporting history for the period January 1998 to December 2011. In addition, we restrict the funds in our study to those with assets under management of more than $100 million as of December 2011. The assets-under-management constraint is intended to eliminate small funds with idiosyncratic strategies not representative of the hedge fund sector. We use only live hedge funds, and so we include the usual caveat that our results may be subject to the effects of survivorship bias.

The data for the $R_{j,t}^{a}$ in our study are the monthly returns of 29 asset classes. The asset classes cover U.S. and international stock markets, U.S. bond markets, commodity markets, and the volatility and variance swaps market based on the CBOE volatility index (VIX). Because of the extensive use of derivatives by hedge funds, we also included three option-like assets as suggested in Agarwal and Naik (2004). One of these is the CBOE BuyWrite Index (BXM) designed to track the performance of a hypothetical buy-write strategy of holding the S&P 500 index and writing the near-term S&P 500 index covered call option. The other two are based on a strategy of buying a 3-month maturity at-the-money synthetic call/put option of the S&P 500 index and selling it one month later. The prices of the synthetic options are calculated using the Black-Scholes formula. We use the VIX as an approximation for the implied volatility, the 3-month Treasury-bill rate as the risk-free rate, and the S&P 500 index dividend yield. The leverage of these two strategies is adjusted for derivative exposure by taking delta-adjusted values of the securities. Let $s, c$, and $p$ denote stock price, call option price, and put price. The delta-adjusted values are calculated as the estimated β multiplied by $\frac{S}{C} * \frac{\partial C}{\partial S}$ or $\frac{S}{P} * \frac{-\partial P}{\partial S}$, respectively, to take into consideration the option elasticity. The list of asset classes is presented below.

1. London PM gold price
2. Handy and Harman silver base price
3. COMEX copper spot price
4. FIBER all items industrial materials index
5. West Texas Cushing intermediate oil price
6. Federal Reserve Board nominal trade-weighted dollar index
7. Dow Jones global ex U.S. stock index
8. 90-day Treasury bill total return index
9. 10-year Treasury total return index
10. S&P 500 total return index

11. Nasdaq composite index
12. MSCI emerging markets index
13. Dow Jones equity all REIT index
14. CBOE volatility index
15. Dow Jones 10-year corporate bond index
16. S&P energy index
17. S&P materials index
18. S&P industrials index
19. S&P consumer discretionary index
20. S&P consumer staples index
21. S&P health care index
22. S&P financial index
23. S&P information technology index
24. S&P telecommunications index
25. S&P utilities index
26. Spread: Moody's seasoned BAA corporate bond yield and 90-day Treasury bill yield
27. CBOE BuyWrite index
28. S&P synthetic call option
29. S&P synthetic put option

## 4. Estimating market sensitivity with a dynamic method

MT estimated the $\beta_j^h$ using rolling regressions on panels of monthly excess returns for hedge funds in nine fund families. They used a two-stage approach. In the first stage, they applied stepwise regression to create an excess-return equation specific to the style classification of the funds in their database from a set of potential independent variables. That is, the independent variables in their excess-return equations in their study are the same for each hedge fund within a particular style classification. This approach implies that all the funds within the same investment style have the same set of risky assets in their portfolios, a condition that we view as rare in practice. Hedge funds use very idiosyncratic strategies even within the same style classification, and so in this paper we build excess return equations for individual hedge funds. In the second stage, MT used linear regression with panel data over rolling samples to calculate the $\beta_j^h$.

We employ the following two-stage approach. In our first stage, excess-return equations are built for each hedge fund individually, using the lasso method proposed by Tibshirani (1996), a shrinkage method that has gained wide acceptance in the data-mining field. Using the lasso method, a tradeoff is accepted between bias and variance in the regression. By introducing some bias in the estimation, the effects of multicollinearity can be reduced so that the variance of the model and the mean square error are reduced. Multicollinearity problems could be substantial in excess-return equations where all the independent variables are excess returns.

For a linear regression $y_i = \beta_0 + \sum_{j=1}^{p} \beta_j * x_{ij} + \varepsilon_t, i = 1,2,3, \dots, n$, the lasso estimate is defined by:

(4) $$\hat{\beta}^{lasso} = \mathrm{argmin}_\beta \sum_{i=1}^{N} (y_i - \beta_0 + \sum_{j=1}^{p} \beta_j * x_{ij})^2$$

subject to $\sum_{j=1}^{p} |\beta_j| \leq T$ where $T$ is a tuning parameter. It can also be written in the equivalent Lagrangian form:

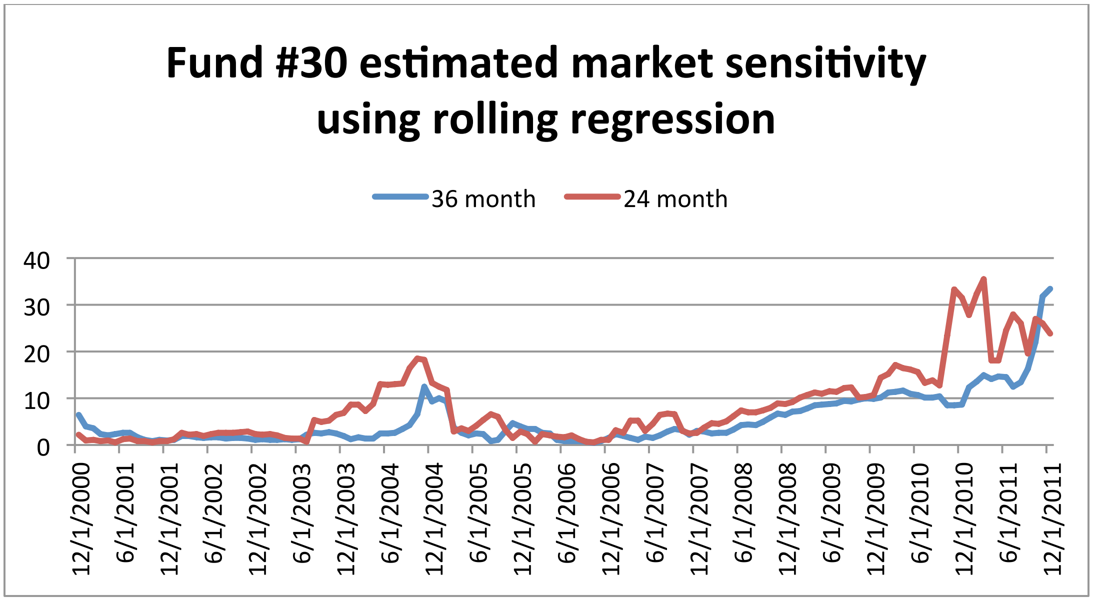
(5) $$\hat{\beta}^{lasso} = \mathrm{argmin}_\beta (\sum_{i=1}^{N} (y_i - \beta_0 + \sum_{j=1}^{p} \beta_j * x_{ij})^2 + \lambda \sum_{j=1}^{p} |\beta_j|)$$

The solutions are nonlinear in terms of the dependent variable, and there is no closed form solution for the lasso regression. By making $T$ sufficiently small, some of the coefficients are shrunk to exactly zero, which means that the lasso estimation method is also a variable selection method. By tuning $T$ continuously, the lasso in effect performs a continuous subset selection, which has less variance than the discrete process employed in stepwise selection. In our study, the choice of the tuning parameter $T$ is based on 10-fold cross-validation estimation of mean square error.

In the second stage, we estimate the parameters in equation (2) as time-varying parameters and compute market sensitivity for each hedge fund. In classical least squares used in rolling regression methods, the regression parameters are assumed to be fixed values within short rolling samples. The parameters are time-varying only in the sense that they change as the rolling sample changes. The results obtained with rolling regression are arbitrarily affected by the choice of the length of the rolling sample. Choosing a rolling sample that is too long may mean that spikes and pulses in the actual degree of hedge fund market sensitivity are smoothed away and so cannot be observed either in the estimated parameters or in the market sensitivity indicator obtained from those parameters. Yet, while rolling regressions may smooth parameter estimates so that actual spikes cannot be seen, as the sample is rolled, an outlier may be added to or deleted from the sample, causing sudden spurious changes in the estimated parameters that mimic the effects of spikes and pulses in market sensitivity. Choosing a rolling sample that is too short may mean that there are insufficient degrees of freedom for efficient estimation. Due to the small sample size in the rolling window, the regressions would have large standard errors of the estimates of the coefficients, which would lead to unstable and at times inaccurate estimates of market sensitivity. The standard errors of the estimates of the coefficients would strongly depend on the window size.

Using sums of absolute values of parameters as the market sensitivity indicator introduces another problem. Using absolute values introduces a bias in the market sensitivity estimate because of Jensen's inequality. We discuss the mathematics of this bias and calculate it and corrections for it in the appendix. The bias may be larger using rolling regression than it is using dynamic regression, and it is not constant for either approach, a topic to which we return below where we employ the bias corrections to compute results for both methods.

To provide examples of the impact of window size, we constructed excess return equations for funds #30 and #40 in our database using the lasso method and then computed the market sensitivity indicator for each fund, using both 24-month and 36-month rolling window regressions. The results are shown in figure 1, where we see that, for the same fund with the same excess-return equation, different window sizes yield very different estimates of market sensitivity.



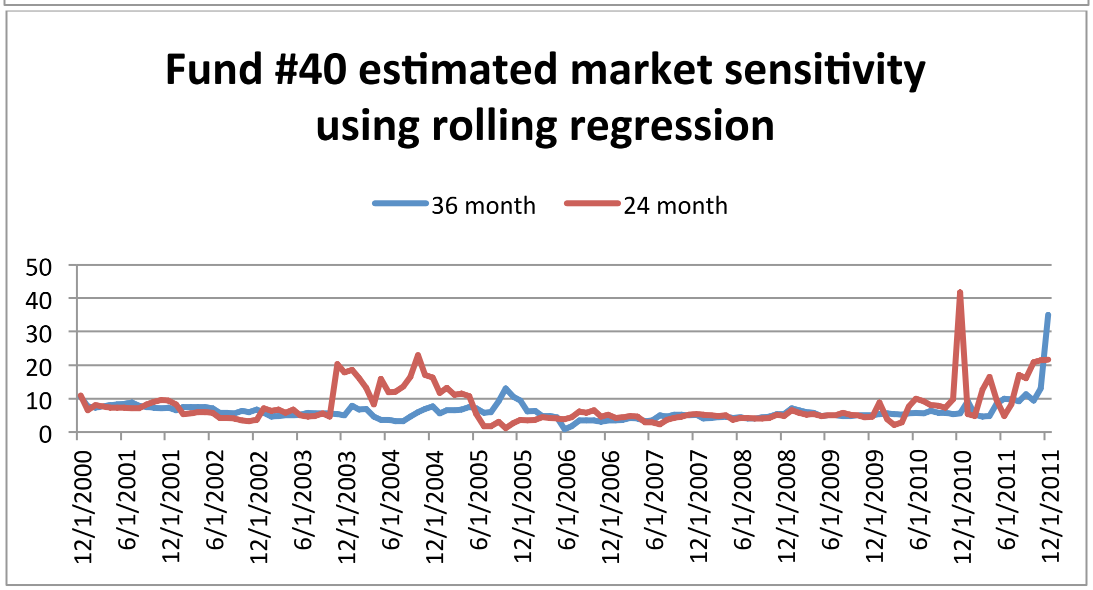




Figure 1. 24-month and 36-month rolling regression estimation of market sensitivity for funds #30 and #40.

To solve the problems inherent in rolling regression, we propose using a state space model to estimate the regression parameters dynamically. The state space model, or dynamic linear model, has an order one vector autoregression as its state equation.

(6) $$\boldsymbol{x}_t = \boldsymbol{\phi}\boldsymbol{x}_{t-1} + \boldsymbol{w}_t$$

The p*1 state vector $\boldsymbol{x}_t$ is generated from the previous state vector $\boldsymbol{x}_{t-1}$ for $t = 1,2,3,\dots,n$. The vector $\boldsymbol{w}_t$ is a p*1 independent and identically distributed zero-mean normal vector having a covariance matrix of $Q$. The starting state vector $\boldsymbol{x}_0$ is assumed to have mean $\boldsymbol{\mu}_0$ and a covariance matrix of $\Pi_0$. The state process is a Markov chain, but we do not have direct observation of it. We can only observe a q-dimensional linear transformation of $\boldsymbol{x}_t$ with added noise. Thus, we have the following observation equation:

(7) $\boldsymbol{y}_t = A_t\boldsymbol{x}_t + \boldsymbol{v}_t$

where $A_t$ is a q*p observation matrix. The observed data are in the q*1 vector $\boldsymbol{y}_t$, and the additive observation noise $\boldsymbol{v}_t$ is white Gaussian noise with a q*q covariance matrix $R$. We also assume $\boldsymbol{v}_t$ and $\boldsymbol{w}_t$ are uncorrelated. Our primary interest is to produce the estimator for the underlying unobserved $\boldsymbol{x}_t$ given the data $\boldsymbol{Y}_s = \{\boldsymbol{y}_1, \dots, \boldsymbol{y}_s\}$. In the literature, when s<t, this is referred to as forecasting; when s=t, this is Kalman filtering; and when s>t, this is Kalman smoothing.

With the definitions $\boldsymbol{x}_t^s = E(\boldsymbol{x}_t|\boldsymbol{Y}_s)$, $\boldsymbol{P}_t^s = E[(\boldsymbol{x}_t - \boldsymbol{x}_t^s)(\boldsymbol{x}_t - \boldsymbol{x}_t^s)]$ and the initial conditions $\boldsymbol{x}_0^0 = \boldsymbol{x}_0 = \boldsymbol{\mu}_0$ and $\boldsymbol{P}_0^0 = \Pi_0$, we have the recursive Kalman filter equations as follows:

(8) $\boldsymbol{x}_t^{t-1} = \boldsymbol{\phi}\boldsymbol{x}_{t-1}^{t-1}$

(9) $\boldsymbol{P}_t^{t-1} = \boldsymbol{\phi}\boldsymbol{P}_{t-1}^{t-1}\boldsymbol{\phi}' + Q$

for $t = 1,2,3, \dots, n$ with

(10) $\boldsymbol{x}_t^t = \boldsymbol{x}_t^{t-1} + K_t(\boldsymbol{y}_t - A_t\boldsymbol{x}_t^{t-1})$

(11) $\boldsymbol{P}_t^t = [I - K_tA_t]\boldsymbol{P}_t^{t-1}$

(12) $K_t = \boldsymbol{P}_t^{t-1}A_t'[A_t\boldsymbol{P}_t^{t-1}A_t' + \boldsymbol{R}]^{-\mathbf{1}}$

The Kalman smoother is set up using $\boldsymbol{x}_n^n$ and $\boldsymbol{P}_n^n$ calculated from the Kalman filter as the initial conditions as follows:

(13) $\boldsymbol{x}_{t-1}^n = \boldsymbol{x}_{t-1}^{t-1} + J_{t-1}(\boldsymbol{x}_t^n - \boldsymbol{x}_t^{t-1})$

(14) $\boldsymbol{P}_{t-1}^n = \boldsymbol{P}_{t-1}^{t-1} + J_{t-1}[\boldsymbol{P}_t^n - \boldsymbol{P}_t^{t-1}]J_{t-1}'$

(15) $J_{t-1} = \boldsymbol{P}_{t-1}^{t-1}\boldsymbol{\phi}'[\boldsymbol{P}_t^{t-1}]^{-1}$

for $t = n, n-1, n-2, \dots, 1$.

In our studies $\boldsymbol{y}_t$ is the observed excess return of a hedge fund at time $t$, and $A_t$ is the 1*p row vector representing the excess returns of the selected assets for that particular hedge fund. The unobserved state $\boldsymbol{x}_t$ is the vector of time-varying parameter estimates of the $\beta_j^h$ in equation (2). As suggested in Lai and Xing (2008), in the case of the dynamic capital asset pricing model, a popular choice is to make $\boldsymbol{\phi}$ the identity matrix and $Q$ diagonal. Then the $Q$ and $R$ matrices are computed using maximum likelihood estimation. In our paper we employ the Kalman smoother to compute market sensitivity since the method can utilize all the available data and so provide more accurate estimates of the time-varying regression parameters than that obtained with rolling regression, which uses only subsets of the data. In the discussion below, we use the terms Kalman smoother and dynamic regression interchangeably.

## 5. Empirical results

In the first stage we use lasso to select the variables for each hedge fund. The frequency of occurrence of each asset class in the 156 equations is shown below.

| | Asset | Occurrences |
|---|---|---|
| 1 | London PM gold price | 11 |
| 2 | Handy and Harman silver base price | 45 |
| 3 | COMEX copper spot price | 23 |
| 4 | FIBER all items industrial materials index | 44 |
| 5 | West Texas Cushing intermediate oil price | 20 |
| 6 | Federal Reserve Board nominal trade-weighted dollar index | 28 |
| 7 | Dow Jones global ex U.S. stock index | 55 |
| 8 | 90-day Treasury bill total return index | 36 |
| 9 | 10-year Treasury total return index | 39 |
| 10 | S&P 500 total return index | 12 |
| 11 | Nasdaq composite index | 45 |
| 12 | MSCI emerging markets index | 81 |
| 13 | Dow Jones equity all REIT index | 21 |
| 14 | CBOE volatility index | 33 |
| 15 | Dow Jones 10-year corporate bond index | 26 |
| 16 | S&P energy index | 30 |
| 17 | S&P materials index | 33 |
| 18 | S&P industrials index | 27 |
| 19 | S&P consumer discretionary index | 19 |
| 20 | S&P consumer staples index | 24 |
| 21 | S&P health care index | 16 |
| 22 | S&P financial index | 23 |
| 23 | S&P information technology index | 30 |
| 24 | S&P telecommunications index | 17 |

| | | |
|---|---|---|
| 25 | S&P utilities index | 15 |
| 26 | Spread: Moody's seasoned BAA corporate bond yield and 90-day Treasury bill yield | 18 |
| 27 | CBOE BuyWrite index | 30 |
| 28 | S&P synthetic call option | 14 |
| 29 | S&P synthetic put option | 65 |

We see that all the proposed asset classes enter into the excess return equations. There are about 5-6 excess returns picked up by each hedge fund equation on average. The three option-like strategies appear frequently, which agrees with the results in Agarwal and Naik. The MSCI emerging markets index appears very frequently in the equations. In addition, it is interesting that the S&P sub-indexes each appear more frequently than the S&P 500 total return index. This may be due to the fact that hedge funds tend to specialize in some sectors of the stock market but not in the overall market. The sub-indexes and S&P index are highly correlated, and so this demonstrates the power of the lasso method to handle multicollinearity.

With the variables selected for each hedge fund equation by the lasso method, we estimate equation (2) using rolling regressions for purpose of comparison as in MT and equation (13) using the Kalman smoother. We use a 36-month window for the rolling regressions. We also employ the bias correction procedure presented in the appendix to compute average market sensitivity. The bias correction is employed for both the rolling regressions and the Kalman smoother. In our work with rolling regression, the standard errors are calculated using the Newey-West estimator, which is robust in the presence of heteroskedasticity and autocorrelation. The sample period for the dynamic regression is the 168 months from January 1998 to December 2011. We estimate the parameters for each hedge fund using both methods and then compute the market sensitivity indicator $(\sum_{h=1}^{156}\sum_{j=1}^{k}\left|\beta_{j,t}^{h}\right|)/156\,, t = 1,2,3,\dots n$, i.e., the average of the market sensitivities for the 156 funds in our database in each period t. The chart below shows the results.

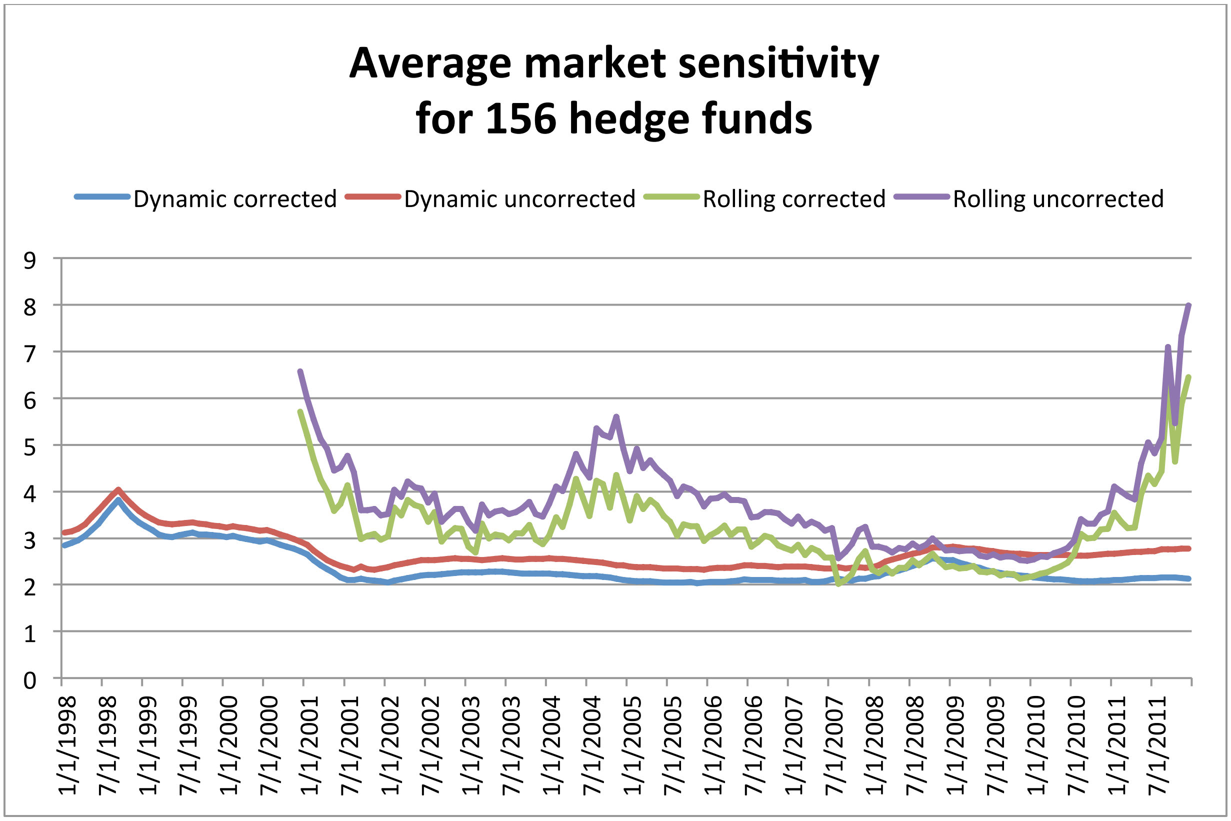



Figure 2. Average market sensitivity for 156 funds
using rolling and dynamic regression

At a glance, the results using rolling regression and dynamic regression are very different. The correction factors are much larger for the rolling regression than for dynamic regression. This is understandable since the Kalman smoother uses all the information so that the degrees of freedom are greater and there is, thus, less estimation error. Under dynamic regression, the Kalman smoother has its largest bias correction at the end of the sample. This is because near the end of sample there is less information about the future. So the standard errors are larger. This means that our proposed bias correction is critical to the analysis since the end of sample is most important for the purpose of monitoring hedge fund market sensitivity. Second, the 36-month rolling regression shows a spike in 2004 that we cannot link to any significant event. In contrast, the market sensitivity estimate using the Kalman smoother is less noisy, and it seems to capture important changes in market sensitivity. The chart below shows the bias-corrected dynamic measure in isolation.

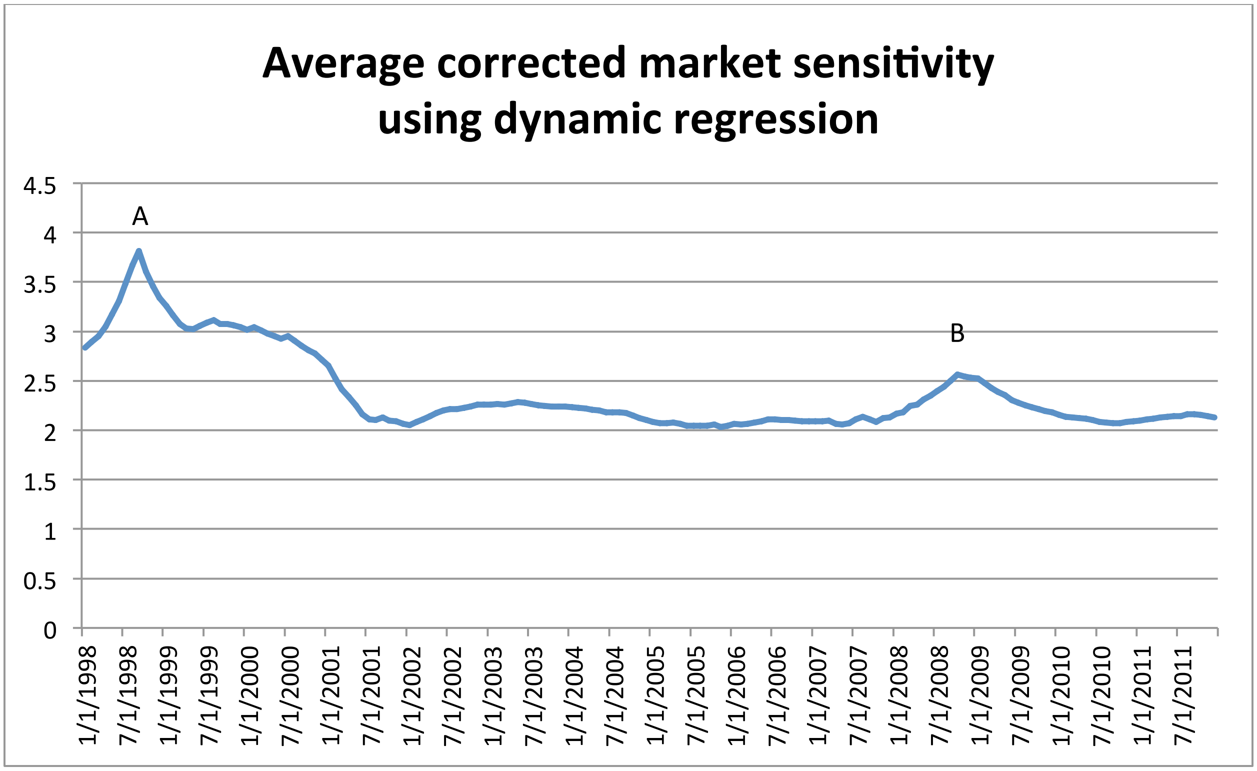



Figure 3: Average corrected market sensitivity

The indicator spiked in September 1998 (point A). This was the month the LTCM meltdown became known to markets, causing investors to withdraw huge sums from other hedge funds. The market sensitivity indicator reached another peak at about the time of the Lehman bankruptcy filing (point B). Thus, in both of these cases the market sensitivity indicator computed from dynamic regression seems to serve as an indicator of financial distress. As for the dot-com bubble burst in 2001, we do not see any substantial rise in market sensitivity using our dynamic approach. This is in line with the findings in Brunnermeier and Nagel (2004). They found that, although hedge funds were heavily invested in technology stocks in the run-up to the bubble, they reduced their positions before prices collapsed and avoided much of the downturn.

The chart below shows the market sensitivity indicator by major style classification.

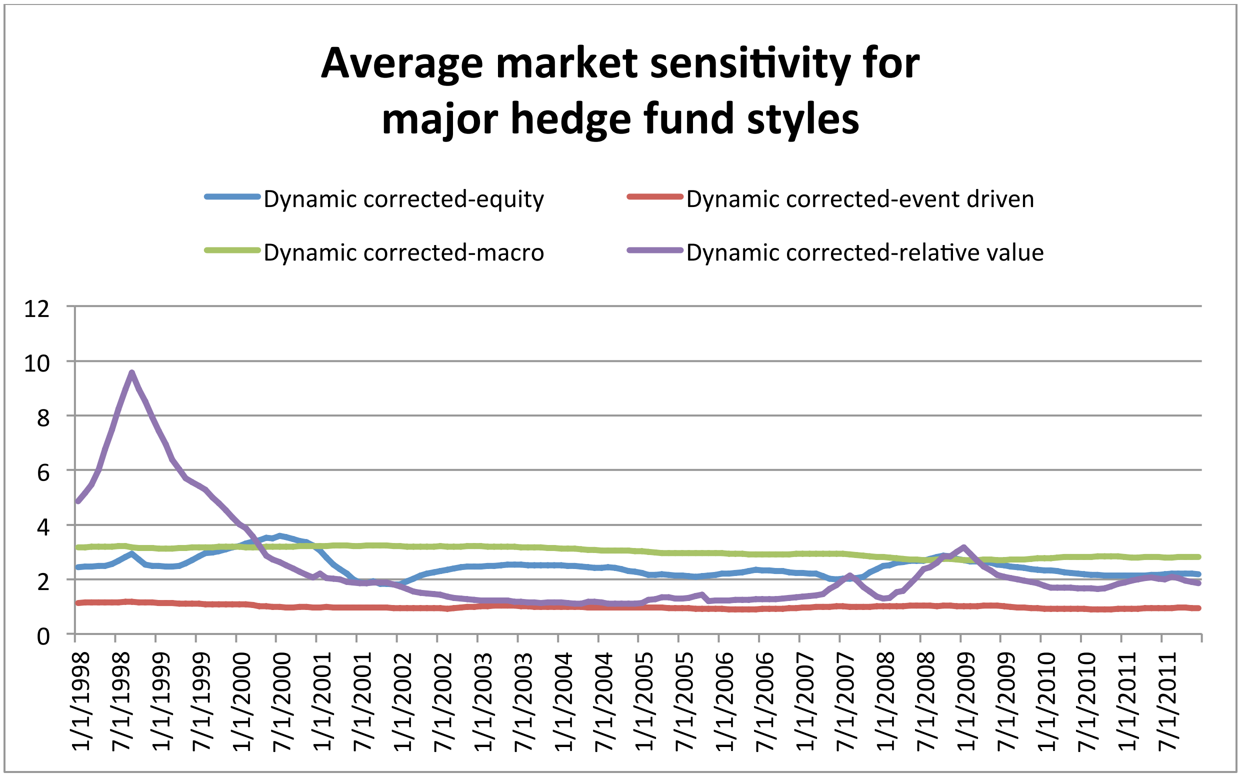



Figure 4. Market sensitivity by major style classification

Here we see that relative value funds are highly market sensitive at times while event-driven funds are relatively insensitive to market fluctuations.

**6. Conclusion**

In this paper we have introduced a hybrid two-stage approach, employing the lasso variable selection method and dynamic regression, to analyze hedge fund market sensitivity with a market sensitivity indicator computed from our two-stage approach. We compared the results to those generated by rolling regression. We find that rolling regression has several shortcomings, one of which is the arbitrary choice of rolling window length. Rolling regression methods may produce spurious spikes in estimated parameters and the market sensitivity indicator that are not rooted in the actual data, and rolling regression may introduce large bias into the market sensitivity indicator.

Dynamic regression produces estimates of the parameters that are not subject to the particular problems inherent in rolling regressions. Our methodology produces smoother results, and the peaks in our market sensitivity indicator are more consistent with anecdotal information about historical events in financial markets. Thus, the market sensitivity indicator described in this paper may provide analysts and policymakers with more accurate information about the degree of market sensitivity in hedge fund space. The sharp increases and peaks of market sensitivity around crisis events revealed by our approach could be used as an indicator of stress in financial markets.

# References

Adrian, T., 2007, "Measuring risk in the hedge fund sector," Federal Reserve Bank of New York, *Current Issues in Economics and Finance*, 13 (3).

Agarwal, V., and Naik, N., 2004, "Risks and portfolio decisions involving hedge funds," *Review of Financial Studies*, 17, 63-98.

Brunnermeier, M., and Nagel, S., 2004, "Hedge funds and the technology bubble," *Journal of Finance*, 59 (5), 2013-2040.

Chan, N., M. Getmansky, S. Haas, and A. Lo, 2006, "Do hedge funds increase systemic risk?" Federal Reserve Bank of Atlanta, *Economic Review*, fourth quarter.

Getmansky, M., A.W. Lo, and I. Makarov, 2004, "An econometric model of serial correlation and illiquidity in hedge fund returns," *Journal of Financial Economics*, 74, 529-610.

Gupta, A., and B. Lang, 2005, "Do hedge funds have enough capital? A value-at-risk approach," *Journal of Financial Economics*, 77, 219-253.

Lai, T., and Xing, H., 2008, *Statistical models and methods for financial markets*, Springer.

McGuire, P., and K. Tsatsaronis, 2008, "Estimating hedge fund leverage," *BIS Quarterly Review*, no. 260.

Sharpe, W., 1992, "Asset allocation: management style and performance measurement," *The Journal of Portfolio Management*, Winter, 7-19.

Tibshirani, R., 1996, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B*, 58 (1), 267-288.

### *Appendix: Correction for market sensitivity estimation bias*

Suppose $x \sim N(u, \sigma^2)$. Then:

$$E[|x|] = \frac{1}{\sqrt{2\pi}\sigma}\int_0^{+\infty} x * exp(\frac{-(x-u)^2}{2\sigma^2})dx + \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{0}(-x) * exp(\frac{-(x-u)^2}{2\sigma^2})dx$$

Let $y = x - u$ and let $\Phi$ denote the cumulative distribution function of a standard normal distribution. Then, the first term can be rewritten as:

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-u}^{+\infty}(y+u) * exp(\frac{-y^2}{2\sigma^2})dy$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-u}^{+\infty} y * exp(\frac{-y^2}{2\sigma^2})dy + \frac{1}{\sqrt{2\pi}\sigma}\int_{-u}^{+\infty} u * exp(\frac{-y^2}{2\sigma^2})dy$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{-u}^{+\infty} exp(\frac{-y^2}{2\sigma^2})d\frac{y^2}{2} + u\left[1 - \Phi\left(\frac{-u}{\sigma}\right)\right]$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{\frac{u^2}{2}}^{+\infty} exp(\frac{-z}{\sigma^2})dz + u\left[1 - \Phi\left(\frac{-u}{\sigma}\right)\right]$$

$$= \frac{\sigma}{\sqrt{2\pi}} exp(\frac{-u^2}{2\sigma^2}) + u\left[1 - \Phi\left(\frac{-u}{\sigma}\right)\right]$$

Similarly, the second term becomes:

$$\frac{-1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{-u}(y+u) * exp(\frac{-y^2}{2\sigma^2})dy$$

$$= \frac{-1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{-u} y * exp(\frac{-y^2}{2\sigma^2})dy + \frac{-1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{-u} u * exp(\frac{-y^2}{2\sigma^2})dy$$

$$= \frac{-1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{-u} exp(\frac{-y^2}{2\sigma^2})d\frac{y^2}{2} + u\left[-\Phi\left(\frac{-u}{\sigma}\right)\right]$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int_{\frac{u^2}{2}}^{+\infty} exp(\frac{-z}{\sigma^2})dz + u\left[-\Phi\left(\frac{-u}{\sigma}\right)\right]$$

$$= \frac{\sigma}{\sqrt{2\pi}} exp(\frac{-u^2}{2\sigma^2}) + u\left[-\Phi\left(\frac{-u}{\sigma}\right)\right]$$

Then:

$$E[|x|] = \frac{2\sigma}{\sqrt{2\pi}} exp(\frac{-u^2}{2\sigma^2}) + u\left[1 - 2\emptyset\left(\frac{-u}{\sigma}\right)\right]$$

which is biased from $|u|$ by $\frac{2\sigma}{\sqrt{2\pi}} exp\left(\frac{-u^2}{2\sigma^2}\right) - 2|u|\emptyset\left(\frac{-|u|}{\sigma}\right)$. It is clear that with a small $\frac{|u|}{\sigma}$ and a large $\sigma$ we could have a large positive bias. The ideal cases, where bias is zero, are where $u > 0\ and\ \frac{u}{\sigma} \gg 1$ so that:

$$\frac{2}{\sqrt{2\pi}\sigma} exp\left(\frac{-u^2}{2\sigma^2}\right) \sim 0, \emptyset\left(\frac{-u}{\sigma}\right) \sim 0 \text{ and } E[|x|] \sim u$$

Where $u < 0\ and\ \frac{|u|}{\sigma} \gg 1$ we have:

$$\frac{2}{\sqrt{2\pi}\sigma} exp\left(\frac{-u^2}{2\sigma^2}\right) \sim 0, \emptyset\left(\frac{-u}{\sigma}\right) \sim 1 \text{ and } E[|x|] \sim -u.$$

In this paper we propose to modify the estimate by subtracting $\frac{2\sigma}{\sqrt{2\pi}} exp\left(\frac{-u^2}{2\sigma^2}\right) - 2|u|\emptyset\left(\frac{-|u|}{\sigma}\right)$. And if the modified market sensitivity is negative, which only happens occasionally, we force it to be 0.