Are “Market Neutral” Hedge Funds Really Market Neutral?

Andrew J. Patton
University of Oxford

Using a variety of different definitions of “neutrality,” this study presents significant evidence against the neutrality to market risk of hedge funds in a range of style categories. I generalize standard definitions of “market neutrality,” and propose five different neutrality concepts. I suggest statistical tests for each neutrality concept, and apply these tests to a database of monthly returns on 1423 hedge funds from five style categories. For the “market neutral” style, approximately one-quarter of the funds exhibit significant exposure to market risk; this proportion is statistically significantly different from zero, but less than the proportion of significant exposures for other hedge fund styles. (*JEL* G11, G23)

1. Introduction

The low correlation between hedge fund returns and market returns is an oft-cited favorable characteristic of hedge funds (Agarwal and Naik 2004; Brown, Goetzmann, and Ibbotson 1999; Fung and Hsieh 2001). Indeed, the term “hedge fund” was coined with reference to the goal of the first such funds, which was to invest in undervalued securities using the proceeds from short-sales of related securities, thereby creating a “market neutral” strategy (Caldwell 1995). The hedge fund industry is one of the fastest-growing sectors of the economy: the value of assets under the management of hedge funds has grown from \$50 billion in 1990 to around \$1.4 trillion in December 2006, and the market-neutral style of hedge funds is one of the fastest growing styles,[1] thus interest in the diversification benefits they offer is great.

Hedge funds are often classified according to their self-described investment strategies or styles, and the “equity market neutral” strategy is one of the largest of such categories, representing about 20% of funds under the management of hedge funds according to Fung and Hsieh (1999). But despite their size, what

Helpful comments and suggestions were received from the Editor, Joel Hasbrouck, an anonymous referee, and Vikas Agarwal, Carol Alexander, Ron Anderson, Andrew Ang, Giovanni Belossi, Greg Connor, Jason Hathorn, Aino Levonmaa, Oliver Linton, Andrew Lo, Sophocles Mavroeidis, Narayan Naik, Burt Porter, Joshua Rosenberg, Chester Spatt, Bas Werker, Elliot Williams, and seminar participants at the 2004 European Finance Association meeting, Inquire UK, Aarhus School of Business, University of Amsterdam, Bank of England, LSE, NYU Stern School of Business, and Tilberg University. I would like to thank IAM for providing the data used in this study, and for support through the IAM Programme in Hedge Fund Research at the London School of Economics. Send correspondence to Andrew J. Patton, Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ, UK; telephone: +44-(0)1865-281-296. E-mail: andrew.patton@economics.ox.ac.uk.

[1] Sources: *The Economist*, June 10, 2004; February 17, 2005; and March 15, 2007.

doi:10.1093/rfs/hhn113 Advance Access publication December 24, 2008

exactly is meant by the moniker market neutral can be hard to pin down. Most definitions of an equity market neutral strategy include phrases like "neutralize market risk(s) by combining long and short positions in related securities," with limited detail on how neutrality should be measured and what risks should be considered "market" risks.[2] Clarity and precision in the use of these terms would be beneficial. The case of *Weyerhaeuser Co. v. Geewax Terker & Co.*, for example, which featured several prominent academics as expert witnesses, centered directly on whether Geewax Terker had truly followed a market-neutral strategy.[3]

A very simple measure of neutrality is "dollar neutrality," whereby the dollar value of long investments in a given market/sector is offset with short positions of equal dollar value in those same markets. This measure has the great benefit of being directly verifiable: the initial value of the investments is observable, at least to the hedge fund manager. However, the extent to which the assets in the short portfolio offset the market exposure from the assets in the long portfolio is generally not observable with certainty, by the hedge fund manager or anyone else, because this depends on the unobservable risk characteristics of the long and short portfolios.

The most commonly used risk-based definition of neutrality is based on correlation or "beta": a fund may be said to be market neutral if it generates returns that are uncorrelated with the returns on some market index, or a collection of market risk factors. Several studies have observed a nonlinear relation between hedge fund returns and market returns and proposed more sophisticated methods for modeling the expected returns on hedge funds. Fung and Hsieh (2001) suggest using payoffs from "lookback straddle" options on the market to approximate the payoff structure of hedge funds. Agarwal and Naik (2004) and Mitchell and Pulvino (2001) similarly suggest using piecewise linear models for the hedge fund returns as a function of the market return. The dynamic nature of hedge funds' trading strategies and the nonlinear payoff functions generated imply that simple linear correlations and betas cannot tell the full story of the diversification benefits offered by hedge funds.

I consider the concept of neutrality more generally than that implied by the use of correlation or beta. While correlation and beta are simple to compute and easy to interpret, the dynamic strategies employed by hedge fund managers and the resulting nonlinear payoff profiles mean that these simple measures of neutrality may provide misleading information about the true diversification benefits offered by these funds. To complement these standard neutrality measures, and extend our ability to detect important forms of non-neutrality, I propose five alternative neutrality concepts: "mean neutrality," which nests the

---

2 See, for example, Hedge Fund Research's Strategy Definitions in Nicholas (2000) and Beliossi (2002).

3 Sources: *The Wall Street Journal*, March 25, 1992, and *Pension & Investments*, March 30, 1992. Geewax, Terker & Co. eventually settled the lawsuit, filed in the U.S. District Court in Seattle, by agreeing to pay Weyerhaeuser Co. $8 million, almost as much as Weyerhaeuser Co. had originally given Geewax, Terker & Co. to manage.

standard correlation- or beta-based definition of neutrality. "Variance neutrality," "value-at-risk neutrality," and "tail neutrality" all relate to the neutrality of the risk of the hedge fund returns to market returns. The final concept, "complete neutrality," corresponds to the statistical independence of fund and market returns. I suggest statistical tests for each neutrality concept, and apply the tests to a combined database of monthly hedge fund returns from the HFR and TASS hedge fund databases, across various hedge fund styles. These new measures of neutrality build on recent contributions to the literature on asymmetric correlations and other generalizations of linear correlation (e.g., Ang and Chen 2002; Ang, Chen, and Xing 2006; Erb, Harvey, and Viskanta 1994; Longin and Solnik 2001; Patton 2006).

A single series is used to represent market risk, either the S&P 500 or MSCI World index. Focusing on a single equity market index may be interpreted as testing a necessary condition for neutrality to a wider set of market variables. All of the tests proposed in this paper generalize quite straightforwardly to consider multiple sources of market risk. However, the limited length of most hedge fund return series may mean that the most obvious extensions, where the additional market risks are simply included as additional explanatory variables, will actually lead to a loss of power rather than a gain; increased estimation error may dominate the increased information from the additional risk sources. One alternative is to combine the sources of risk into a single index or portfolio, and then test for neutrality against this portfolio of market risks, using weights on each market risk in the portfolio that are guided by the preferences of the investor.

By presenting a battery of neutrality concepts and tests, I hope to aid investors' evaluation of these funds, in a similar way to the use of the "Greeks" to evaluate the exposure of an option position (e.g., Hull 2003). The concepts and tests proposed in this paper may be used as methods of analyzing the non-neutrality of hedge funds, rather than solely as strict tests of their neutrality. As the neutrality of a market neutral fund is one of its selling points, I conjecture that when comparing a collection of such funds, the risk, reward, *and* the nature of the dependence between each fund and the market are of interest to investors. Of course, if the investor's utility function was known, then funds could be directly ranked by expected utility; however, such a case is not common in practice.

The low correlation of hedge fund returns is related to the development of so-called "portable alpha" strategies. These are trading strategies that generate returns that are independent of major market risks (i.e., market neutral strategies). If true, then the returns from such a strategy can be combined with the return from another strategy that exhibits the desired amount of exposure to market risks (beta), without increasing the overall portfolio's exposure to the market risks. The methods proposed in this paper may be interpreted as tests of the "purity" of the portable alpha strategy: if some dependence on the market risk is detected, then this would need to be accounted for when incorporating the hedge fund investment into a portfolio with other exposures to market risk.

This paper makes two main contributions. First, I propose a number of different neutrality concepts relevant for hedge fund return and risk analysis, and present statistical tests of each neutrality concept. Particular attention is paid to the types of non-neutral alternatives considered in each test: I consider either a general non-neutral alternative, or only those non-neutral alternatives that are disliked by risk-averse investors with some existing exposure to market risk. For example, a risk-averse investor prefers zero correlation to positive correlation, but prefers negative correlation to zero correlation. Thus zero correlation may be tested against nonzero correlation, or only against positive correlation. Statements about the (non-)neutrality of hedge funds can be refined by using the preferences of investors.

The second contribution of the paper is a detailed study of the neutrality of a combined database of 1423 hedge funds in a variety of fund styles from the HFR and TASS hedge fund databases over the period April 1993 to April 2003. To illustrate the tests, I initially focus on the market neutral style of funds. These results are then compared with those for the "equity hedge," "equity non-hedge," "event driven," and "funds of funds" styles. I use monthly returns, net of fees, on both live and dead hedge funds to evaluate their neutrality against a market index, the S&P 500. I find that approximately one-quarter of the market neutral funds exhibit some significant non-neutrality, at the 0.05 level, while the proportions of non-neutral funds for the other fund styles vary from approximately 50% for funds of funds, to 85% for the equity non-hedge style. Thus my findings suggest that many market neutral hedge funds are in fact *not* market neutral, but overall they are, at least, more market neutral than other categories of hedge funds. A series of robustness checks verifies that the results are not overly affected by the choice of market index, the use of U.S. dollar returns, the choice of model used to deal with serial correlation in hedge fund returns, or by the last few, or first few, observations on fund returns.

The remainder of the paper is structured as follows. Section 2 describes the data used in this study. Section 3 presents definitions of different types of neutrality, tests for each definition, and the results of these tests when applied to the various hedge fund styles. Section 4 presents an array of robustness checks of the results, and Section 5 concludes. An Appendix contains details on the bootstrap methods used in the study.

2. Description of the Data

The hedge fund data set used in this study consists of monthly returns, net of all fees, on those funds in the HFR and TASS databases that categorize themselves as being in one of the following five styles: market neutral, equity hedge, equity non-hedge, event driven, or fund of hedge funds. According to

HFR's strategy definitions,[4] market neutral funds are those that seek to exploit apparent mispricings without generating any exposure to market risk. Equity hedge funds hold some exposure to the market, with the degree of exposure ranging from near 0% to over 100%, along with some hedge, either through short sales of stocks or through stock options. Equity non-hedge funds are generally long equity market risk, and engage predominantly in stock picking. These funds may also hedge their exposures, though generally not consistently. Event-driven funds seek returns from mergers, takeovers, bankruptcies, etc. These funds may or may not hedge their exposures to the market. Funds of hedge funds invest in multiple funds, which may or may not be in the same category. I employ the S&P 500 as the market index for most of this paper, and show in Section 4 that the results do not change greatly if other equity market indexes are used instead.

Summary statistics for the funds are presented in Table 1, and confirm that hedge funds, on average, performed well during the sample period, particularly when compared to common equity indexes. This table also shows that many hedge fund returns are non-normally distributed, confirming the findings of Gupta and Liang (2005) and Chan et al. (2005), with Jarque-Bera tests of normality (Jarque and Bera 1980) generating low median p-values across funds, particularly for the equity non-hedge and event driven styles. The first- and second-order hedge fund return autocorrelations reported in Table 1 indicate substantial serial correlation. The event driven style, in particular, has median first- and second-order autocorrelation coefficients of 0.19 and 0.06, contrasted with corresponding coefficients for the S&P 500 index of −0.03 and −0.03. This strong positive serial correlation in hedge fund returns has been noted by, *inter alia*, Asness, Krail, and Liew (2001) and Getmansky, Lo, and Makarov (2004). Getmansky, Lo, and Makarov (2004) consider many possible explanations for the presence of this serial correlation, and are able to rule several explanations out on the basis that they cannot generate strong enough serial correlation, or they generate serial correlation of the wrong sign. These authors conclude that the most likely causes of serial correlation in hedge fund returns are exposures to illiquid assets, which lead to positive serial correlation due to nonsynchronous trading (see Lo and MacKinlay 1990) and/or intentional performance smoothing by hedge fund managers. Getmansky, Lo, and Makarov (2004) suggest that these causes of serial correlation can be captured, though not distinguished, by the use of a moving-average model for hedge fund returns, and the analysis here follows this suggestion. Specifically, I use an MA(2) model, as in Getmansky, Lo, and Makarov (2004), to filter the hedge fund returns before subjecting them to analysis. In Section 4, I show the results obtained when MA(0) and MA(4) models are applied. The estimation error in

[4] See http://www.hedgefundresearch.com/pdf/HFR_Strategy_Definitions.pdf. It should be reiterated that these style categories are based on fund self-descriptions, and thus may be more or less appropriate for any given fund.

Table 1
Summary statistics of fund and market index returns

	Median fund						
	Market neutral	Equity hedge	Equity non-hedge	Event driven	Funds of funds	S&P 500	MSCI World
Mean	0.68	1.00	1.05	0.86	0.69	0.85	0.55
Standard deviation	2.35	4.35	5.89	2.52	1.88	4.52	4.25
Skewness	0.02	0.20	0.04	−0.17	−0.03	−0.54	−0.54
Kurtosis	3.80	3.96	4.02	4.56	4.23	3.26	3.28
Minimum	−4.71	−9.10	−15.52	−5.88	−3.91	−14.44	−13.35
Maximum	6.05	12.70	16.17	7.71	5.87	9.78	9.02
Jarque-Bera statistic	2.80	4.38	5.95	11.73	4.50	5.99	6.04
Jarque-Bera *p*-value	0.25	0.11	0.05	0.00	0.11	0.05	0.05
Autocorrel, lag 1	0.06	0.08	0.06	0.19	0.19	−0.03	−0.01
Autocorrel, lag 2	0.01	0.00	−0.07	0.06	0.03	−0.03	−0.07
Number of obs	42	44	79	58	40	121	121
Total number of funds	197	514	79	102	531	–	–

This table presents descriptive statistics on monthly fund and market index returns over the sample period, April 1993 to April 2003. The columns headed "median fund" present the medians of the statistics in the rows across the funds in each style category with at least twelve observations. The Jarque-Bera statistic refers to the Jarque-Bera (1980) test of normality; the *p*-value for this test is also reported. The penultimate row presents the number of monthly observations that are available for analysis. The final row presents the number of funds in each style category with at least twelve observations.

the MA model parameters is accounted for in the bootstrap procedure employed to conduct the tests.

Table 2 presents summary statistics on the number of observations available on each of the hedge funds in the sample. The median history across hedge fund styles varies from just forty months to seventy-nine months. Depending on the type of neutrality being considered, I impose a greater or lesser minimum number of observations for a fund to be included in the analysis. In all cases, the number of observations available on a given fund is quite constrained, and this constraint must be considered when implementing and interpreting the results of the tests of neutrality: short histories mean that parsimony in the specification is required for tests of neutrality to have reasonable power to reject the null hypothesis.

When computing measures of dependence between the market and a fund, all data from the period when the fund was in the database are employed. The database includes both live and dead funds[5] and one may question whether the behavior of some funds in the period leading up to their dropping out of the database distorts the results. I analyze this in Section 4 and find only limited differences between live and dead funds. I also check whether "backfill bias," which results from the fact that hedge funds usually enter the database with a history of 6–18 months of (good) returns, affects the conclusions, and find that it does not.

[5] As Agarwal, Daniel, and Naik (2006) point out, these funds are misnomered, since funds may drop out of the database for numerous reasons: liquidation, mergers, or simply a withdrawal from reporting to the database while continuing to operate.

Table 2
Summary statistics on the number of observations

	Market neutral	Equity hedge	Equity non-hedge	Event driven	Funds of funds	Total
Minimum	1	1	1	6	1	–
0.25 quantile	19	23	40	28	18	–
Median	42	44	79	58	40	–
Mean	49.5	52.9	74.1	61.6	48.6	–
0.75 quantile	69	77	114	90	73	–
Maximum	121	121	121	121	121	–
Number of funds with ≥ 6 obs	213	543	84	110	569	1519
Number of funds with ≥ 18 obs	171	466	77	90	463	1267
Number of funds with ≥ 24 obs	150	422	77	86	414	1149
Number of funds with ≥ 66 obs	59	182	58	47	169	515
Number of "dead" funds	42	62	11	9	23	147
Number of live funds	175	511	84	101	601	1472
Total number of funds	217	573	95	110	624	1619

This table presents descriptive statistics on the number of monthly return observations available on each of the five categories of hedge funds used in this paper. The sample period is April 1993 to April 2003 and so the maximum number of observations is 121.

3. Definitions and Tests of Versions of "Market Neutrality"

In this section, I consider refinements of the concept of market neutrality, using the preferences of a generic risk-averse investor to motivate each concept and to determine the alternative hypotheses. I will start with the simplest generalization of correlation neutrality, and proceed through to the strictest form of neutrality: that of independence between the fund return and the market return. Each test will be introduced with an application to the market neutral style of hedge funds, and I compare these results with those obtained for the other styles of hedge funds in Section 3.9. In all cases, I remove serial correlation in the hedge fund returns using the MA(2) model proposed by Getmansky, Lo, and Makarov (2004) prior to studying the neutrality of the fund,[6] and below, unless otherwise stated, I always mean the *filtered* fund return when referring to the fund return. Robustness to the particular choice of serial correlation model is studied in Section 4.

I employ a block bootstrap procedure to obtain critical values for each of the tests. This procedure incorporates the impact of all parameter estimation errors, and is designed to be robust to serial correlation, heteroscedasticity, and non-normalities of unknown form in the data. Details on this method are presented in the Appendix.

3.1 Correlation neutrality

Before moving on to consider refinements of the definition of market neutrality, I will first analyze the relationship between the market neutral funds and the

[6] Specifically, for each fund, I estimate a MA(2) model on the observed returns and use the residuals from this model as the estimate of the "true" returns on the fund. An alternative to this approach (see Asness, Krail, and Liew 2001) is to instead include lags of the market variable in the model for the observed fund return. Such an approach is less well suited to this application given the limited data available for many funds.

market index using standard linear correlation. The average correlation between the 171 market neutral hedge funds with eighteen or more observations and the market index was 0.025, and the 5th and 95th sample quantiles of the cross-sectional distribution of correlation coefficients was [−0.52, 0.53], indicating substantial cross-sectional dispersion in the degree of correlation with the market portfolio.

Using the bootstrap procedure described in the Appendix, 28.1% of the funds in this sample exhibit significant correlation with the market portfolio at the 0.05 level of significance. This statistic is surprisingly high: these funds are (self-) described as market neutral, presumably to more factors than the single market index, and yet over one-quarter of them have significant correlations with the market. If we instead focus the test only on deviations from zero correlation to *positive* correlation with the market, which is the sign of correlation a risk-averse investor seeks to avoid, we find 23.4% of funds have significant positive correlation with the market.

Under a joint null hypothesis that *all* funds in the sample are correlation neutral, we would expect 5% of funds to be rejected. To determine whether the proportion of rejections I observe is significantly different from what would be expected under the null hypothesis, I consider two methods of obtaining a critical value. The first is based on the unrealistic assumption that the test statistics for each fund are independent. In this case, the critical value can be obtained from the binomial distribution, and it is 7.74% in this sample. A bootstrap method for obtaining a critical value that allows for dependence among the test statistics is discussed in the Appendix, and it yields a critical value of 18.75%, much higher than that obtained assuming independence, but still smaller than the observed proportion of rejections. Thus, I conclude that there is significant evidence against "correlation neutrality" for this collection of funds as a whole.

Correlation neutrality is just one of many types of neutrality that may be of interest to a risk-averse investor. An investor with quadratic utility, or one facing returns that are multivariate normally distributed, will only require linear correlation as the measure of dependence, and so this standard concept of market neutrality would suffice. However, neither quadratic utility nor multivariate normality is an empirically reasonable assumption, particularly for hedge fund returns, and so I next consider alternative types of market neutrality.

3.2 Mean neutrality

The simplest neutrality concept, and the one that nests the standard "correlation neutral" concept, is that of mean neutrality. This is defined as the expected return on the fund being independent of the return on the market:

$$E[r_{it}|r_{mt}] = E[r_{it}]\,\forall\, r_{mt} \tag{1}$$

$$\text{or}\quad E[r_{it}|\mathcal{F}_{t-1}, r_{mt}] = E[r_{it}|\mathcal{F}_{t-1}]\,\forall\, r_{mt}, \tag{2}$$

and corresponds to the statement that the market return does not Granger-cause the fund return in mean. Equation (2) allows us to consider mean neutrality conditional on some other information, $\mathcal{F}_{t-1}$. I will focus primarily on Equation (1) due to data limitations.

Equation (1) nests standard correlation and beta neutrality in the sense that they both can be tested by specifying a simple linear model for the mean of the fund return, and estimating the following by ordinary least squares:

$$r_{it} = \beta_0 + \beta_1 r_{mt} + e_{it}, \tag{3}$$

and then testing the null hypothesis that $\beta_1 = 0$. Mean neutrality, as defined in Equations (1) and (2), requires not only that there is no linear relationship between the fund return and the market return, but also that there are no *nonlinear* relationships. To test mean neutrality, one could employ a number of methods. Some authors (Agarwal and Naik 2004; Asness, Krail, and Liew 2001; Fung and Hsieh 1997; Mitchell and Pulvino 2001) employ piecewise linear regressions:

$$r_{it} = \beta_0 + \beta_1^- r_{mt} 1\{r_{mt} \leq 0\} + \beta_1^+ r_{mt} 1\{r_{mt} > 0\} + e_{it} \tag{4}$$

or, more generally,

$$r_{it} = \beta_0 + \beta_1 r_{mt} 1\{r_{mt} \leq k_1\} + \beta_2 r_{mt} 1\{k_1 < r_{mt} \leq k_2\} + \cdots + e_{it}, \tag{5}$$

where $k_j,\ j = 1, 2, \ldots, K$ are estimated "kink" points in the specification. The most general test of mean neutrality would employ nonparametric regression to estimate the conditional mean function, $\mu_i(r_{mt}) = E[r_{it}|r_{mt}]$:

$$r_{it} = \mu_i(r_{mt}) + e_{it}, \tag{6}$$

and then test that μ_i is equal to a constant. Nonparametric methods are notoriously data-intensive, and, given the short data samples, I consider a simple alternative, namely to employ a Taylor series approximation[7] to the conditional mean function:

$$r_{it} = \beta_0 + \beta_1 r_{mt} + \beta_2 r_{mt}^2 + \cdots + e_{it}, \tag{7}$$

and then test:

$$\begin{aligned} &H_0 : \beta_j = 0 \text{ for all } j > 0 \\ \text{vs.}\quad &H_a : \beta_j \neq 0 \text{ for at least one } j > 0, \end{aligned} \tag{8}$$

via a standard Wald test. Under certain conditions on how the models expand as the sample size increases, both the piecewise linear and the polynomial

[7] Numerous authors have, in various contexts, proposed using a polynomial function of the market return to explain asset returns; see Bansal, Hsieh, and Viswanathan (1993); Chapman (1997); Harvey and Siddique (2000); Dittmar (2002), among others.

specifications can be considered nonparametric models for the conditional mean (Andrews 1991; Chen and Shen 1998). The use of a piecewise linear specification with estimated kink points often leads to identification problems under the null hypothesis of interest in this study of market neutrality, and so I employ the simple polynomial model in Equation (7). I estimated a third-order polynomial version of the model in Equation (7) on the 150 funds with at least twenty-four observations, and found that the null hypothesis of mean neutrality could be rejected, at the 0.05 level, for 28.0% of funds.

The definition of market neutrality given above, however, ignores the fact that there are certain types of relations between the expected return on a fund and the market return that a risk-averse investor would desire, and others that such an investor would dislike. For example, holding other things equal, a risk-averse investor would prefer a negative relation between the fund and the market when the market return is negative, and a positive relation when the market return is positive, to zero correlation in both states. Thus it may not be mean neutrality that investors truly seek, or that market neutral hedge funds truly seek to provide, but rather a restricted type of dependence between the fund and the market. Below I derive a test of mean neutrality that tests only for violations of mean neutrality that are disliked by risk-averse investors.

Consider a refinement of mean neutrality, which I will call "mean neutrality on the downside." This form of neutrality imposes that the expected return on the fund is neutral or related negatively to the market return when the market return is negative. That is,

$$\frac{\partial \mu_i(r_{mt})}{\partial r_{mt}} \leq 0 \qquad \text{for } r_{mt} \leq 0, \tag{9}$$

where $\mu_i(r_{mt}) \equiv E[r_{it}|r_{mt}]$. This version of neutrality ignores the relation between the fund and the market when the market return is positive, focusing solely on the ability of the fund to provide diversification benefits when the market return is negative. If the third-order polynomial in Equation (7) is used to approximate the conditional mean function, then

$$\frac{\partial \mu_i(r_{mt})}{\partial r_{mt}} = \beta_1 + 2\beta_2 r_{mt} + 3\beta_2 r_{mt}^2. \tag{10}$$

A pointwise confidence interval for the first derivative of the conditional mean function is simple to construct using the covariance matrix of the estimated parameters from Equation (7). It can then be determined whether the first derivative of the conditional mean function is significantly greater than zero, for some $r_{mt} \leq 0$. However, using pointwise confidence intervals and searching across all values of r_{mt} leads to a size distortion in the test, and so I instead conduct a test on the average value of this derivative across values of $r_{mt} \leq 0$:

$$H_0 : E\left[\left. \frac{\partial \mu_i(r_{mt})}{\partial r_{mt}} \right| r_{mt} \leq 0 \right] \leq 0 \tag{11}$$
$$\text{vs.} \quad H_a : E\left[\left. \frac{\partial \mu_i(r_{mt})}{\partial r_{mt}} \right| r_{mt} \leq 0 \right] > 0.$$

Using the third-order polynomial model, and plugging in the parameter estimates, gives

$$E\left[\left. \frac{\partial \hat{\mu}_i(r_{mt})}{\partial r_{mt}} \right| r_{mt} \leq 0 \right] = \hat{\beta}_1 + 2\hat{\beta}_2 E\left[r_{mt} | \, r_{mt} \leq 0 \right] + 3\hat{\beta}_2 E\left[r_{mt}^2 \middle| \, r_{mt} \leq 0 \right], \tag{12}$$

and so testing that $E[\partial \mu_i(r_{mt})/\partial r_{mt} | r_{mt} \leq 0] \leq 0$ reduces to checking the pointwise confidence interval on the first derivative of the conditional mean function at the point $(\hat{E}[r_{mt}|r_{mt} \leq 0],\ \hat{E}[r_{mt}^2|r_{mt} \leq 0])$. This test was repeated, using the bootstrap distribution of the test statistic rather than the asymptotic distribution, and it was found that mean neutrality on the downside was rejected for 20.0% of the funds at the 0.05 level. Thus one-fifth of all market neutral funds exhibit conditional mean dependence on the market of a form disliked by risk-averse investors.

One concern about the above regressions and the subsequent tests on functions of the estimated parameters is potential omitted variables bias. Consider an example where a single factor, different from the market portfolio, drives a given fund's returns. If this factor is positively correlated with the market, then, with sufficient data, the above method will reject the null hypothesis that the fund is market neutral even though the factor driving returns is not the market portfolio. However, if the factor is correlated with the market then part of its risk is market risk and part is nonmarket risk, and so exposure to this factor does indeed involve some exposure to market risk. Thus the conclusion that this hypothetical fund is not market neutral seems reasonable. If one wanted to test market neutrality controlling for fund exposure to some other sources of risk, then this could be done by including the returns on the other factors in the above regression as control variables, as in Equation (2).

3.3 Variance neutrality

Another form of neutrality that one might expect from a market neutral fund is that the risk of the fund is neutral to market risk. In particular, we might expect that the risk of the fund, while not constant, does not increase at the same time as the risk of the market index. In this section, I consider risk as measured by variance, and in the next section, I consider risk as measured by value-at-risk and extreme tail probabilities. To my knowledge, this paper is the first to consider the market neutrality of the risk of a hedge fund.

In the tests below, I will control for mean non-neutrality before testing variance, VaR, or tail neutrality. The motivation for this comes from the fact that a failure to control for mean non-neutrality will generally lead to the presence of apparent variance non-neutrality, even if the (conditional) variance of the fund is truly neutral to the market return.[8] Without a control for mean non-neutrality, we should expect that the number of violations of variance neutrality are higher than when a control for mean non-neutrality is employed. Thus controlling for mean non-neutrality, combined with the limited data available, makes it unsurprising that power is limited in these tests. Nevertheless, these tests provide an alternative view on the relationship between a fund and the market and so may offer information not available in more standard mean-based tests.

The exposure of hedge funds to variance risk has been reported by Bondarenko (2004), for example, motivating this investigation of this type of non-neutrality. Further, risk-averse investors can be shown to have preferences over the dependence between the variance of the fund and the market return. Non-increasing absolute risk aversion, a property suggested by Arrow (1971) as being desirable in a utility function, leads to a preference for positive skewness in the distribution of portfolio returns. Kimball (1993) suggested further that reasonable utility functions should exhibit decreasing *absolute prudence*, which can be shown to imply an aversion to kurtosis in the distribution of portfolio returns.[9] Together these imply that risk-averse investors prefer

$$Corr\,[(r_{it} - \mu_i)^2, r_{mt} - \mu_m] \geq 0 \tag{13}$$

and

$$Corr\,[(r_{it} - \mu_i)^2, (r_{mt} - \mu_m)^2] \leq 0, \tag{14}$$

so that the skewness of a portfolio of the fund and the market is larger and the kurtosis of the portfolio is smaller. With this motivation, I define variance neutrality, controlling for mean non-neutrality, as

$$V\,[r_{it} - \mu_i(r_{mt})|r_{mt}] = V\,[r_{it} - \mu_i(r_{mt})] \tag{15}$$

or

$$V\,[r_{it} - \mu_i(r_{mt})|\mathcal{F}_{t-1}, r_{mt}] = V\,[r_{it} - \mu_i(r_{mt})|\mathcal{F}_{t-1}]. \tag{16}$$

In a manner similar to the previous section, one can obtain a test by approximating the true conditional variance function, $\sigma_i^2\,(r_{mt})$, by a Taylor series

[8] See Lumsdaine and Ng (1999) for a detailed discussion of this point in the context of testing for volatility clustering.

[9] Related papers on investor preferences over higher-order moments include Kraus and Litzenberger (1976); Harvey and Siddique (2000); and Dittmar (2002), among many others.

polynomial:

$$r_{it} = \mu_i(r_{mt}) + e_{it}, \tag{17}$$

$$e_{it} = \sigma_i(r_{mt})\varepsilon_{it}, \ \varepsilon_{it} \sim (0, 1), \tag{18}$$

$$\mu_i(r_{mt}) = \beta_0 + \beta_1 r_{mt} + \beta_2 r_{mt}^2 + \beta_3 r_{mt}^3 + e_{it},$$

$$\sigma_i^2(r_{mt}) = \alpha_0 + \alpha_1 r_{mt} + \alpha_2 r_{mt}^2,$$

or

$$\sigma_i^2 (r_{mt}, e_{it-1}) = \alpha_0 + \alpha_1 r_{mt} + \alpha_2 r_{mt}^2 + \alpha_3 e_{it-1}^2, \tag{19}$$

where the latter conditional variance specification is designed to control for an ARCH(1) effect (Engle 1982) in the fund return. To test variance neutrality, we then test:

$$\begin{aligned} H_0 &: \alpha_1 = \alpha_2 = 0 \\ \text{vs.} \quad H_a &: \alpha_1 \neq 0 \cup \alpha_2 \neq 0. \end{aligned} \tag{20}$$

I conducted this test on the 150 funds with more than twenty-four observations, with the ARCH(1) term as a control, and the null was rejected at the 0.05 level for only 4.0% of funds. Thus most of these funds appear to be variance neutral to the market portfolio.

We can also consider "variance neutrality on the downside," using the preferences of a risk-averse investor to determine that the desired sign of the first derivative of the conditional variance function is positive when the market return is negative. Again, rather than search over all values of $r_{mt} \leq 0$, I instead focus on the average value of the first derivative, and the relevant hypotheses are

$$\begin{aligned} H_0 &: E\left[\left. \frac{\partial \sigma_i^2(r_{mt})}{\partial r_{mt}} \right| r_{mt} \leq 0 \right] \geq 0 \\ \text{vs.} \quad H_a &: E\left[\left. \frac{\partial \sigma_i^2 (r_{mt})}{\partial r_{mt}} \right| r_{mt} \leq 0 \right] < 0. \end{aligned} \tag{21}$$

Following the same method as for the test of mean neutrality, a confidence interval for this first derivative can be obtained and a check conducted that it is non-negative at $\hat{E}\,[r_{mt}|r_{mt} \leq 0]$. I did this for the funds in the database and found significant violations of variance neutrality on the downside for 4.0% when an ARCH(1) term was included in the variance specification.[10]

[10] When the control for mean non-neutrality was dropped, the proportion of funds that exhibited significant variance non-neutrality increased from 4.0% to 8.1%, whereas the proportion that exhibited significant variance non-neutrality on the downside fell slightly from 4.0% to 2.0%. Thus it appears that controlling for mean non-neutrality did not substantially affect the proportion of nonvariance neutral funds.

Overall, after controlling for violations of mean neutrality, I find no statistical evidence of violations of variance neutrality. This implies that the variance risk to which many hedge funds are exposed has no significant impact on the volatility of hedge fund returns, controlling for its impact on expected fund returns. Of course, it may be that the failure to find evidence against variance neutrality is due to a lack of data, and thus limited test power.

3.4 Value-at-risk neutrality

The second risk-related neutrality concept I propose is that of value-at-risk neutrality, or "VaR neutrality." Given that the VaR of a fund is simply a quantile of its distribution of returns,[11] this could also be called "quantile neutrality." If we set the quantile to be 0.5, we would have a test of "median neutrality," though I do not pursue that here. Quantile neutrality is a special case of complete neutrality, discussed below, which implies that *all* quantiles of the fund are neutral to the market, but differs from the previous two neutrality concepts in that it focuses on quantiles rather than moments. A VaR-neutral portfolio is one with a VaR that is unaffected by the market portfolio return. That is,

$$VaR\left(r_{it}|r_{mt}\right) = VaR\left(r_{it}\right) \tag{22}$$

or

$$VaR\left(r_{it}|r_{mt}, \mathcal{F}_{t-1}\right) = VaR\left(r_{it}|\mathcal{F}_{t-1}\right). \tag{23}$$

Analogous to the impact of mean non-neutrality on tests of variance neutrality, violations of mean neutrality or variance neutrality will generally lead to violations of VaR neutrality, even if the (conditional) VaR of the fund is truly neutral to the market return.[12] This leads us to consider "conditional VaR neutrality":[13]

$$VaR\left(\left.\frac{r_{it} - \mu_i(r_{mt})}{\sigma_i(r_{mt})}\right| r_{mt}\right) = VaR\left(\frac{r_{it} - \mu_i(r_{mt})}{\sigma_i(r_{mt})}\right), \tag{24}$$

where the VaR of the standardized returns is considered and not the VaR of the returns themselves. Gupta and Liang (2005) have used VaR to examine the risk in hedge funds from a regulatory perspective. Though hedge fund returns are generally not normally or elliptically distributed (see Gupta and

[11] Formally, the value-at-risk of an asset is obtained from the following equality: $\Pr\left[r_t \leq VaR_t|\mathcal{F}_{t-1}\right] = \alpha$, where α usually equals 0.10, 0.05, or 0.01.

[12] This is because $VaR(r_{it}|r_{mt})$ can be written as $VaR(r_{it}|r_{mt}) = \mu_i(r_{mt}) + \sigma_t(r_{mt}) \cdot VaR(\varepsilon_{it}|r_{mt})$, where $\varepsilon_{it} \equiv \{r_{it} - \mu_i(r_{mt})\}/\sigma_i(r_{mt})$. Thus even if the VaR of the standardized return, ε_{it}, is neutral to the market, the VaR of the nonstandardized returns, r_{it}, will depend on the market if $\mu_i(r_{mt})$ or $\sigma_t(r_{mt})$ are not constant (i.e., if the fund is mean or variance non-neutral).

[13] Note that the term "conditional VaR" is used to refer to the quantile of a conditional distribution. Other authors have used this term to describe the expected return conditioning on the VaR being breached, that is, $E\left[r_{it}|r_{it} \leq VaR\left(r_{it}\right)\right]$, a quantity otherwise known as "expected shortfall" or "tail conditional loss."

Liang 2005 or Table 1 of this paper for an example), it is interesting to note that if the market and fund returns were jointly elliptically distributed, then the portfolio VaR would be an affine function of the portfolio variance, and VaR neutrality would then follow directly from mean and variance neutrality. Under normality, conditional VaR neutrality would always hold, even if mean and variance neutrality did not, but, for other elliptical distributions, this need not be the case. See Embrechts, McNeil, and Straumann (2001) for further discussion on VaR for portfolios, and Artzner et al. (1999) for a criticism of VaR as a measure of risk.

There are a number of ways that one might test the null hypothesis:

$$\begin{aligned} H_0 &: VaR\,(r_{it}|r_{mt}) = VaR\,(r_{it})\,\forall\, r_{mt} \\ \text{vs. } H_a &: VaR\,(r_{it}|r_{mt}) \neq VaR\,(r_{it}) \text{ for some } r_{mt} \end{aligned} \tag{25}$$

or

$$\begin{aligned} H_0 &: VaR\left(\frac{r_{it}-\mu_i(r_{mt})}{\sigma_i(r_{mt})}\middle| r_{mt}\right) = VaR\left(\frac{r_{it}-\mu_i(r_{mt})}{\sigma_i(r_{mt})}\right)\forall\, r_{mt} \\ \text{vs. } H_a &: VaR\left(\frac{r_{it}-\mu_i\,(r_{mt})}{\sigma_i(r_{mt})}\middle| r_{mt}\right) \neq VaR\left(\frac{r_{it}-\mu_i(r_{mt})}{\sigma_i\,(r_{mt})}\right) \text{ for some } r_{mt}. \end{aligned} \tag{26}$$

With sufficient data, one could use quantile regression (see Koenker and Bassett 1978) to test for the influence of the market return on a quantile of the fund return distribution in a way similar to the tests for mean and variance neutrality. However, hedge fund return histories are generally short and the quantiles of interest are in the tail, so it is likely that data shortages will be a problem.

A simple alternative way of testing a necessary condition for VaR neutrality is via a test of Christoffersen (1998). This test examines whether the probability of one variable exceeding its VaR is affected by another variable exceeding or not exceeding its VaR. Specifically,

$$\begin{aligned} H_0 &: \Pr\left[\varepsilon_{it} \leq \widehat{VaR}(\varepsilon_{it})\middle|\varepsilon_{mt} \leq \widehat{VaR}(\varepsilon_{mt})\right] = \Pr\left[\varepsilon_{it} \leq \widehat{VaR}(\varepsilon_{it})\right] \\ \text{vs. } H_a &: \Pr\left[\varepsilon_{it} \leq \widehat{VaR}(\varepsilon_{it})\middle|\varepsilon_{mt} \leq \widehat{VaR}(\varepsilon_{mt})\right] \neq \Pr\left[\varepsilon_{it} \leq \widehat{VaR}(\varepsilon_{it})\right], \end{aligned} \tag{27}$$

where $\varepsilon_{it} \equiv (r_{it} - \mu_i(r_{mt}))/\sigma_i(r_{mt})$ and $\varepsilon_{mt} \equiv (r_{mt} - \mu_{mt})/\sigma_{mt}$. For the fund, I again use a third-order polynomial for the conditional mean, and a second-order polynomial for the conditional variance with an ARCH(1) term. For the market, I use a simple AR(1)-ARCH(1) model. $\widehat{VaR}(\varepsilon_{it})$ and $\widehat{VaR}(\varepsilon_{mt})$ are estimated by the empirical quantiles of ε_{it} and ε_{mt}.

Due to the data-intensive nature of studies of VaR, only funds that had at least sixty-six months of observations available were considered, which left fifty-nine market neutral funds, and I tested the 10% VaR rather than the more common 1% or 5% VaR. I conducted the conditional VaR neutrality test on these funds and found evidence against VaR neutrality for none of the funds at the 0.05 level. Thus, there is no evidence against VaR neutrality for these funds, having controlled for mean and variance non-neutrality.

We can also consider a "downside" version of this test, which focuses specifically on testing whether the probability of the fund breaching its VaR is greater given that the market return has breached its VaR:

$$H_0 : \Pr\left[\varepsilon_{it} \leq \widehat{VaR}(\varepsilon_{it})\middle|\varepsilon_{mt} \leq \widehat{VaR}(\varepsilon_{mt})\right] \leq \Pr\left[\varepsilon_{it} \leq \widehat{VaR}(\varepsilon_{it})\right] \quad (28)$$

$$\text{vs.} \;\; H_a : \Pr\left[\varepsilon_{it} \leq \widehat{VaR}(\varepsilon_{it})\middle|\varepsilon_{mt} \leq \widehat{VaR}(\varepsilon_{mt})\right] > \Pr\left[\varepsilon_{it} \leq \widehat{VaR}(\varepsilon_{it})\right].$$

This version of VaR neutrality uses the fact that a risk-averse investor would be averse to a fund that has a higher probability of a VaR exceedance given that the market has exceeded its VaR, and would have a preference for the opposite. Conducting this test on the funds revealed that none of these funds are rejected at the 0.05 level, and thus there is no evidence against the downside VaR neutrality of these funds.[14]

3.5 Tail neutrality

The concept of neutrality during extreme events, or tail neutrality, can be thought of as an extension of VaR neutrality to the extreme tail: a market neutral fund should have a probability of extreme events that is unaffected by the market return. The formal definition of tail neutrality that I will use is

$$\tau^L \equiv \lim_{q \to 0} \Pr\left[F_i\left(r_i\right) < q | F_m\left(r_m\right) < q\right] = \lim_{q \to 0} \Pr\left[F_i\left(r_i\right) < q\right] = 0, \quad (29)$$

where $r_i|\mathcal{F}_{t-1} \sim F_i$ and $r_m|\mathcal{F}_{t-1} \sim F_m$. In other words, this definition implies that the probability of an extremely low return on the fund is not affected by conditioning on the fact that an extremely low return on the market is observed. The variable τ^L is known as the coefficient of lower tail dependence (e.g., Joe 1997). If the fund return and the market return have zero lower tail dependence, then the probability of an extreme negative return on the fund is unaffected by an extreme negative return on the market portfolio, and limits to zero as more and more extreme returns are considered. The alternative to tail neutrality is tail dependence, when $\tau^L > 0$. If the tail dependence coefficient is positive, then there is a nonzero chance that both the fund and the market will simultaneously experience an extremely low return. It is intuitively clear that risk-averse investors would prefer tail neutrality to positive tail dependence: a higher probability of a joint crash increases the probability of a large negative

[14] When the control for mean and variance non-neutrality was dropped, the proportion of funds that exhibited significant VaR non-neutrality increased from 0% to 13.8%, both for general VaR neutrality and for VaR neutrality on the downside. This is not very surprising given the proportions of mean non-neutral funds reported in Section 3.2. By controlling for mean and variance non-neutrality in the test of VaR neutrality, I seek evidence of non-neutrality beyond that already detected in the tests of mean and variance neutrality.

return on a portfolio of these two assets. That is, positive lower tail dependence will generally lead to a fatter left tail of the portfolio distribution. The null and alternative hypotheses for this neutrality concept then are

$$\begin{aligned} H_0 &: \tau^L = 0 \\ \text{vs. } H_a &: \tau^L > 0, \end{aligned} \tag{30}$$

which makes clear that this test is already a "downside" test, as the only possible alternative to tail neutrality is dependence of a form disliked by a risk-averse investor.

A number of recent studies have proposed methods for detecting dependence in the tails of joint distributions. Longin and Solnik (2001) propose specifying a specific copula for the joint tails, Gumbel's copula in Nelsen (1999), and then testing that the parameter of this copula is such that no tail dependence is present. Bae, Karolyi, and Stulz (2003) model the probability of the joint occurrence of large returns across assets using parametric multinomial logistic regression, while Brown and Spitzer (2006) test for an excess probability, relative to a normal distribution or Student's t-distribution, of large joint negative returns via a test of the odds ratio. I employ the method of Quintos (2003), who proposes a nonparametric approach using extreme value theory, to derive a statistic to test for tail dependence.[15] Due to the heavy data requirements of tail analyses, I restricted the sample to the twenty-eight funds in the database with at least 100 observations. Of these, fifteen had enough observations in the joint tail to complete the test, and only one of these fifteen funds rejected the null of no tail dependence at the 0.05 level.[16] Thus, I conclude that no evidence of violations of tail neutrality is present for the funds in the database. This conclusion, however, may be overturned in the future when more data become available and estimates of tail behavior become more precise.[17]

It should be noted that the heavy data requirements of the VaR neutrality and tail neutrality tests introduce the possibility that survivorship bias affects the results. It may be that the funds that survive for a minimum of 66 or 100 months are those that live up to the name market neutral. This may be because surviving funds are those that have maintained a "good" return regardless of the market (which is a definition of market neutrality) or because investors desire market neutral funds that are truly market neutral and so these

[15] In unreported work, I also implemented the method of Longin and Solnik (2001) on these hedge funds. While this method is quite different in implementation from the method of Quintos (2003), the same conclusion emerges.

[16] I used the asymptotic theory provided by Quintos (2003) rather than the bootstrap for this test.

[17] Brown and Spitzer (2006), on the other hand, find strong evidence of dependence between hedge fund returns and S&P 500 index returns in their lowest decile. Possible explanations for the differences between our findings include: Quintos's test focuses on the (asymptotic) tail dependence coefficient, τ^L, rather than the 10% quantile of returns, with the latter test using more data than the former; Quintos's test accounts for serial correlation and volatility clustering in the fund and index returns, while Brown and Spitzer assume these returns to be *iid*; and Brown and Spitzer rely on a parametric copula assumption for the 10% joint tail (independent, normal, or Student's t_3) rather than remaining nonparametric as in Quintos's test.

funds remain alive. In either of these scenarios, the surviving funds would be more likely to pass VaR and tail neutrality tests, and thus the low proportion of rejections of VaR neutrality and tail neutrality would not be representative of the VaR and tail neutrality of market neutral funds with shorter histories. I investigate fund longevity and market neutrality further in Section 4.

3.6 Complete neutrality

Complete neutrality is the strictest form of neutrality, and requires that the distribution of fund returns is completely independent of the market return. The formal definition is

$$r_i | r_m =^d r_i, \tag{31}$$

where "$=^d$" indicates equality in distribution. If $r_{it} \sim F_i$ and $r_{mt} \sim F_m$, and $(r_{it}, r_{mt}) \sim F$, this implies that

$$f(r_{it}, r_{mt}) = f_i(r_{it}) \cdot f_m(r_{mt}), \tag{32}$$

whereas in general the joint distribution of the fund return and the market return is written as

$$f(r_{it}, r_{mt}) = f_i(r_{it}) \cdot f_m(r_{mt}) \cdot c(F_i(r_{it}), F_m(r_{mt})), \tag{33}$$

where c is the "copula density" [18] or "dependence function" of the fund and the market returns. Under complete neutrality, the fund's copula with the market is the "independence copula," denoted by c_I, which takes the value 1 everywhere. The preferences of a risk-averse investor can be used to derive a ranking of copulas between the fund and the market, using a result of Epstein and Tanny (1980).

A general alternative to complete neutrality is a dependence function, C^*, that differs from C_I by a "correlation-increasing transformation" (CIT) of Epstein and Tanny (1980). A CIT is the dependence equivalent of the better-known "mean-preserving spread" of Rothschild and Stiglitz (1970). A CIT shifts some probability mass toward realizations where both variables are "large" or "small," and away from realizations where one is "large" and the other is "small," in such a way that the marginal distributions of the variables are preserved. From Epstein and Tanny (1980), it is known that

$$C_I(u, v) \leq C^*(u, v) \;\forall\; (u, v) \in [0, 1] \times [0, 1] \tag{34}$$

and we say that C^* is "more concordant"[19] than C_I, or simply that $C_I \leq C^*$. This ordering is a multivariate first-order stochastic dominance ordering.

[18] The copula *cdf* is denoted with an uppercase C while the copula density is denoted with a lowercase c. See Nelsen (1999) for an introduction to copulas.

[19] Epstein and Tanny (1980) interpret the condition in Equation (34) by saying that C^* exhibits "greater correlation" than C_I but the term "correlation" will not be used here unless referring directly to Pearson's linear correlation or Spearman's rank correlation.

To relate the above statistical ordering of dependence functions to some economic ordering, Epstein and Tanny (1980) introduce the concept of "correlation aversion." An individual with a utility function involving two random variables is "correlation averse" if expected utility is reduced by a CIT. These authors show that this can be checked directly for utility functions that are differentiable twice by checking whether

$$\frac{\partial^2 u(x, y)}{\partial x \partial y} < 0. \tag{35}$$

In this paper, I consider the case that

$$u(r_{it}, r_{mt}) = \mathcal{U}(w_i r_{it} + w_m r_{mt}),$$

where $\mathcal{U}$ is some utility function, so

$$\begin{aligned} \frac{\partial^2 u(r_{it}, r_{mt})}{\partial r_{it} \partial r_{mt}} &= \mathcal{U}''(w_i r_{it} + w_m r_{mt}) w_i w_m \\ &\leq 0 \text{ if } w_i, w_m \geq 0 \end{aligned} \tag{36}$$

for any concave utility function $\mathcal{U}$. If short selling is ruled out (or not relevant) for both the hedge fund and the market portfolio, both reasonable restrictions for most investors, then $w_i, w_m \geq 0$. The weak inequality in Equation (36) holds strictly if $w_i w_m > 0$; of course, if either portfolio weight is zero then the investor is correlation neutral. Thus, any risk-averse investor subject to short-selling constraints will be (weakly) correlation averse.

The concept of correlation aversion can be used to derive an economic ranking of dependence functions for risk-averse investors: if F_1 and F_2 are two possible joint distribution functions for (r_{it}, r_{mt}) with common marginal distributions, and if

$$E_{F1}[u(r_{it}, r_{mt})] \leq E_{F2}[u(r_{it}, r_{mt})] \tag{37}$$

for all correlation-averse utility functions u, then Epstein and Tanny (1980) write $F_2 \preccurlyeq_u F_1$ and say that F_1 exhibits greater correlation than F_2. The main theorem in Epstein and Tanny (1980) shows that the ranking obtained from the expected utility of risk-averse investors is equivalent to the purely statistical concordance ordering discussed above. That is,

$$F_2 \preccurlyeq_u F_1 \Leftrightarrow F_2 \leq F_1. \tag{38}$$

In terms of dependence functions and neutrality, this implies that

$$C_I \preccurlyeq_u C^* \Leftrightarrow C_I \leq C^*, \tag{39}$$

and so any dependence function that is a CIT away from complete neutrality will be less preferred by risk-averse investors. Epstein and Tanny thus show theoretically that general risk-averse investors care about the dependence (not just the correlation) between hedge fund returns and market returns.

The above results could be used to motivate tests for a concordance ordering of hedge funds, using tests for multivariate first-order stochastic dominance. Instead, I propose the more modest task of examining the ordering of a scalar measure of dependence, namely Spearman's rank correlation. Nelsen (1999) shows that Spearman's rank correlation, denoted by ρ_S, will reflect the concordance ordering of two dependence functions. That is,

$$C^{**} \leq C^{*} \Rightarrow \rho_S(C^{**}) \leq \rho_S(C^{*}) \text{ for any copulas } C^{**},\ C^{*}, \tag{40}$$

and so we have

$$C^{**} \preccurlyeq_u C_I \preccurlyeq_u C^{*} \Leftrightarrow C^{**} \leq C_I \leq C^{*} \Rightarrow \rho_S(C^{**}) \leq 0 \leq \rho_S(C^{*}). \tag{41}$$

Thus an approximate ordering of the funds for a general risk-averse investor may be obtained by categorizing the funds as having significant negative rank correlation, non-significant rank correlation, or significant positive rank correlation with the market index. Rank correlation can detect monotonic nonlinear relationships, in addition to the linear relationships that the usual correlation coefficient may be used to detect.

Average rank correlation across the 171 funds with at least eighteen observations was 0.018, a figure similar to that obtained using linear correlation. From tests for nonzero rank correlation, I found 28.1% of funds had significant rank correlation at the 0.05 level, and 22.8% of funds had significantly positive rank correlation at the 0.05 level. Of course, complete neutrality implies neutrality of any other type, and so all other tests in this paper may also be thought of as tests of necessary conditions for complete neutrality.

Above rank correlation was used to obtain an approximate ordering of the copula of the fund return and the market return, which captures all types of dependence between the fund and the market. One could also consider using rank correlation to obtain an approximate ordering of the copula of the fund return and the market return, *controlling for* mean or mean and variance non-neutrality. In such cases, rank correlation would be used to obtain an approximate ordering of the copula of the fund's residual (or standardized residual) and the market return. In these cases we would expect the proportion of funds exhibiting significant *conditional* rank correlation with the market to be lower than those exhibiting significant unconditional rank correlation, and this is indeed what I find: when controlling for mean non-neutrality, the proportion of funds with significant (significantly positive) rank correlation fell from 28.1% (22.8%) to 8.6% (5.2%), and when controlling for mean *and* variance non-neutrality, the proportions fell to 5.2% (6.9%). Thus it seems that mean non-neutrality is the primary source of violations of complete neutrality, with some additional violations being attributable to variance non-neutrality.

3.7 Summary: are market-neutral hedge funds really market neutral?

In this section, I combine the results of the tests introduced above to draw an overall conclusion about the market neutrality of a hedge fund. Given that so few of the funds in the sample had sufficient data for the test of tail neutrality to be applied, I will not consider this test when drawing overall conclusions.

Declaring a fund to be non-neutral if it fails at least one test for market neutrality leads to a size distortion. For example, the probability that a truly market neutral fund fails at least one of five independent tests of market neutrality, each with size 0.05, is 0.23. Further, we must take into account the fact that the test statistics for each of the five tests considered here (correlation neutrality, mean neutrality, variance neutrality, VaR neutrality, and complete neutrality) are not likely to be independent. I deal with these two problems by looking at the number of tests failed by a set of bootstrapped data series, generated imposing the independence of the fund and the market returns. If the actual number of tests failed is greater than the 95th percentile of the number of tests failed by the bootstrapped data sets, then I conclude that the fund fails a joint test of market neutrality. Further details are discussed in the Appendix.

At the 0.05 level, I found that 29.2% of the funds failed a joint test of market neutrality against general non-neutral alternatives, while 20.5% of the funds failed a joint test of market neutrality against alternatives that are disliked by risk-averse investors. The 95% critical value for the proportion of funds failing a test of neutrality is 18.75%, and so both of these proportions represent significant evidence against market neutrality for these funds as a whole.

Thus these results, using the joint test, are roughly in line with the results above: about 30% of market neutral funds fail a test of general neutrality, and about 20% of market neutral funds fail a test of neutrality that focuses specifically on the preferences of a risk-averse investor. The sample sizes here are not extremely large (the median sample size is just forty-two observations), which means that the power of the tests employed may be low, suggesting that the true proportion of non-neutral funds may be even higher. These results suggest that a careful analysis of individual fund returns is required to reap the widely cited diversification benefits of hedge funds.

Finally, each of the tests of neutrality were conducted on the HFR Equity Market Neutral index, to see whether a portfolio of market neutral funds is more or less neutral than the individual funds.[20] The results of this analysis could have gone in either direction: if the (generally small) market exposures of individual funds offset each other in the cross section, then the index would be market neutral even though a substantial fraction of the individual funds failed a joint test of market neutrality. Conversely, if the market exposures of the individual funds are relatively constant, while the non-market risk exposures of the funds are offsetting, then the index may appear more exposed to the market than any individual fund. The results of the tests of the index's neutrality favor

[20] I thank a referee for suggesting this analysis.

the first scenario: the p-values for each of the individual tests of neutrality ranged from .14 to .75, and the p-values on the two joint tests of neutrality (against general alternatives, and against alternatives disliked by a risk-averse investor) were both .14. Thus although approximately 25% of the funds failed a test of market neutrality, a simple portfolio of these funds passed all tests of neutrality, suggesting that the market exposures of these funds, if present, are offsetting. Thus, it appears that combining "market neutral" funds to form a market neutral (no quotation marks) fund is indeed possible.

3.8 How do neutral and non-neutral funds differ?

A natural question to ask is how neutral and significantly non-neutral funds differ along various dimensions. In Table 3, I compare some simple summary statistics on two portfolios of these funds: the neutral (non-neutral) portfolio is constructed as an equally weighted average of the returns on the funds that passed (failed) the joint test for neutrality on the downside. Table 3 reveals that the average age of non-neutral funds was substantially greater than neutral funds, with non-neutral funds having an average age of around six years, while neutral funds have an average age of 4.5 years. The fact that non-neutral funds are, on average, 33% older than neutral funds may either reflect the fact that more observations lead to greater power to reject the null hypothesis of neutrality if it is false, or it may genuinely reflect the fact that older funds tend to stray from market neutrality more than younger funds. Non-neutral funds also tend to have more funds under management than neutral funds: \$79.3 million compared with \$57.5 million.

Table 3 also shows that the non-neutral portfolio yielded a significantly greater average return than the neutral portfolio: 1.19% per month compared with 0.79% per month. Non-neutral funds also had significantly greater standard deviation: 1.78% per month versus 0.74% per month. The correlation with the market was much greater for the non-neutral portfolio (a predictable outcome, given the way the two portfolios were constructed), and this is particularly true in months following a six-month period of below-average market returns: the "neutral" fund portfolio had a correlation with the market of −0.30 in such months, compared with the "non-neutral" fund portfolio, which had a correlation of 0.82. Interestingly, the correlation of the "neutral" fund with the market is lower when I condition on the market return being negative or on the six-month market return being below average, whereas the correlation of the non-neutral fund with the market is *higher* in these states. Of course, the tests of neutrality "on the downside" are designed to detect this type of dependence structure and so this is not overly surprising.

In the third panel of Table 3, I investigate whether the "non-neutral" portfolio behaved more like one of the other fund styles, by computing its correlation with an equally weighted portfolio of the funds in each of the four other styles in the data set (I also report the correlation with the market neutral index for comparison). These correlations vary from 0.58 (with the funds of funds index)

Table 3
Comparison of neutral and non-neutral hedge funds

	Neutral fund portfolio	Non-neutral fund portfolio	Difference
Average size	\$57.5m	\$79.3m	−\$21.8m
Average age	55.4 months	75.7 months	–20.3 months
Mean	0.79	1.19	−0.40*
Standard deviation	0.74	1.78	−1.04*
Skewness	0.20	−0.34	0.43
Kurtosis	3.19	3.59	0.40
Correl with market	−0.08	0.78	−0.86*
Correl with market, given $r_{mt} > 0$	−0.20	0.61	−0.81*
Correl with market, given $r_{mt} \leq 0$	−0.21	0.70	−0.91*
Correl with market, given $\bar{r}_{m,t,t-6} > \bar{r}_m$	0.10	0.72	−0.62*
Correl with market, given $\bar{r}_{m,t,t-6} \leq \bar{r}_m$	−0.30	0.82	−1.12*
Correl with market neutral index	0.78	0.77	0.01
Correl with equity hedge index	0.14	0.74	−0.60*
Correl with equity non-hedge index	0.00	0.77	−0.77*
Correl with event driven index	0.05	0.72	−0.66*
Correl with funds of funds index	0.16	0.58	−0.42*

This table presents descriptive statistics on two portfolios constructed using the results of the joint test of neutrality on the downside, reported in Section 3.7. The first (second) portfolio is constructed as an equally weighted average of the neutral (non-neutral) funds in the "market neutral" category, for each month in the sample period, April 1993 to April 2003. Table 4 reports that 20.5% of the 171 funds studied failed the joint test of neutrality on the downside, and so the "neutral" portfolio contains up to 136 funds, while the "non-neutral" portfolio contains up to thirty-five funds. The fact that some funds dropped out of the database and others joined sometime after April 1993 implies that the two portfolios contained varying numbers of funds. The first panel reports the average age and size (in millions of U.S. dollars, as at the last available observation for each fund) of the funds that comprise the two portfolios. The second panel reports standard summary statistics for these two portfolios. The third panel reports correlations of the two portfolios with the market return: the first row is the sample linear correlation; the next two rows report the correlation conditional on the market return in that month being positive or negative; the final two rows report the correlation conditional on the market return in the preceding six months (denoted $\bar{r}_{m,t,t-6}$) being greater than or less than the average market return (denoted $\bar{r}_m$). The final panel reports the correlation of the two portfolios with an equally weighed index of all funds in a given style. In all cases the third column reports the difference between the "neutral" and "non-neutral" portfolios, and an asterisk denotes a difference that is significantly different from zero at the 0.05 level.

to 0.77 (with the equity non-hedge index), compared with a correlation of 0.78 with the S&P 500 index. These indexes, however, are correlated with the market return,[21] and so a more useful analysis is to consider these correlations after controlling for market risk exposure. In unreported regression results, I found that the equity hedge index explained the greatest proportion of variation in the non-neutral portfolio after controlling for exposure to market risk, among all four style indexes[22] (for obvious reasons, I did not consider the market neutral index as a regressor). This suggests that these "non-neutral" funds are perhaps better described as equity hedge than market neutral in their investment style.

[21] The correlations for the Market Neutral, Equity Hedge, Equity Non-hedge, Event Driven, and Funds of Funds indexes with the S&P 500 index over the sample period were 0.42, 0.68, 0.81, 0.68, 0.49, respectively.

[22] Specifically, I regressed the "non-neutral" portfolio returns on a constant, the first three powers of the market index returns, and on the returns on each of the fund style indexes separately. The robust t-statistic on the equity hedge index was 5.15, which was the highest of all four styles. Furthermore, when I regressed the "non-neutral" portfolio returns on a constant, the first three powers of the market index returns, and on the returns on all four of the fund indexes jointly, the equity hedge index was the only one that was significant, with a robust t-statistic of 2.16. Detailed results are available upon request.

Table 4
Proportion of funds that fail tests of neutrality, by hedge fund style

	Market neutral	Equity hedge	Equity non-hedge	Event driven	Funds of funds
Panel I: Proportion rejecting null of neutrality					
Correlation neutrality	0.281*	0.543*	0.857*	0.656*	0.482*
Mean neutrality	0.280*	0.481*	0.805*	0.558*	0.459*
Variance neutrality	0.040	0.031	0.065	0.035	0.070
VaR neutrality	0.000	0.022	0.035	0.021	0.024
Complete neutrality	0.281*	0.569*	0.857*	0.589*	0.516*
Joint test	0.292*	0.539*	0.857*	0.611*	0.497*
Panel II: Proportion rejecting null of neutrality on the downside					
Correlation neutrality	0.234*	0.560*	0.870*	0.700*	0.510*
Mean neutrality	0.200*	0.543*	0.779*	0.709*	0.611*
Variance neutrality	0.040	0.019	0.000	0.012	0.012
VaR neutrality	0.000	0.022	0.035	0.021	0.024
Tail neutrality	0.067	0.106	0.250*	0.067	0.021
Complete neutrality	0.228*	0.571*	0.870*	0.667*	0.523*
Joint test	0.205*	0.562*	0.844*	0.700*	0.538*

This table presents the results of tests of various types of market neutrality presented in the paper, applied to five different hedge fund styles. The tests are conducted on monthly hedge fund returns over the period April 1993 to April 2003. The S&P 500 is used as the market index. Panel I reports the proportion of funds for which the null hypothesis of market neutrality could be rejected at the 0.05 level in favor of a general alternative. Panel II reports the proportion of funds for which the null hypothesis of market neutrality could be rejected at the 0.05 level in favor of a dependence structure of a type that is disliked by U.S.-based risk-averse investors with a preexisting exposure to the S&P 500 (a test of "neutrality on the downside"). All tests were conducted using a bootstrap procedure described in the Appendix. An asterisk indicates that the proportion of rejected funds is significantly different from zero and thus that collection of funds taken as a whole are significantly non-neutral.

3.9 Are market-neutral hedge funds more market neutral than other funds?

In this section, I present the results of applying all of the above tests to the full set of 1423 hedge funds in all five style categories. Table 4 represents the main empirical contribution of this paper, summarizing the results of these tests as applied to the full set of funds, categorized by style.

Table 4 reveals that although there is significant evidence against neutrality for the collection of market neutral hedge funds, this style does represent the most market neutral of the five styles considered. This table shows that a far higher proportion of funds in the remaining four categories exhibit significant exposure to equity market risk. The proportion of funds that fail the joint test of neutrality, by style, are market neutral: 29.2%; funds of funds: 49.7%; equity hedge: 53.9%; event driven: 61.1%; and equity non-hedge: 85.7%. Thus the market neutral style category is the most neutral, and the equity non-hedge category is easily the least neutral.

As only approximately one-quarter of market-neutral funds exhibited significant violations of market neutrality, and the average correlation coefficient across funds was 0.025, I draw the conclusion that while not all market neutral funds are truly market neutral, they are, as a category, substantially more market neutral than other hedge fund categories.

Table 4 also shows that, in general, allowing for a more flexible specification of the conditional mean of fund returns leads to a slight reduction in power. The third-order polynomial used for the conditional mean model nests the linear correlation case, and so, in the absence of estimation error, should detect at least as many cases of non-neutrality. Of course, the presence of estimation error means that the added flexibility may be outweighed by lower precision, which appears to be the case here.

All the tests of neutrality of hedge fund risk indicate limited power to reject the null hypothesis. The rejection frequency of these tests is generally approximately equal to the size of the tests, indicating no evidence against the neutrality of these risk measures to market risks. This may indeed be the case, or it may be the more prosaic explanation that short sample sizes have reduced our ability to detect violations of neutrality in risk.

4. Robustness Checks

In this section, I conduct range robustness checks of the results reported above. First, I consider an alternative index for the "market" portfolio, the MSCI World index. I next consider changing the outlook of our hypothetical investor from one who cares about U.S. dollar returns to one who cares about British pound returns. I then analyze whether my results change when I drop the last six months, or the first twelve months, of available data on firms. Further, I examine alternative model specifications for capturing the serial correlation in hedge fund returns, and I look at the relation between the number of observations available on a fund and its dependence characteristics. I next analyze whether any significant differences arise from results based on live versus dead funds. Finally, I conduct the analysis on two subsamples of the full sample, to see whether neutrality varies with market conditions. In the interest of parsimony, I only present the robustness checks for the market neutral hedge fund style; the robustness of the other styles is very similar to that reported below.

4.1 Choice of market portfolio

Obviously, the choice of market index is an important input to tests of market neutrality. Column 1 of Table 5 contains a summary of the results from Table 4, which used the S&P 500 index as the market index. Corresponding results when the MSCI World index is instead used are presented in column 2 of Table 5. A comparison of these two columns shows that the results are relatively robust to this choice. The proportion of rejections of the null of neutrality falls slightly, while the proportion of rejections of the null of neutrality on the downside increases slightly. I also obtained results (not reported) when the MSCI Europe index was employed, and again no substantial differences were found.

4.2 Choice of currency

To consider the impact of the choice to examine the neutrality of these hedge funds from the perspective of a U.S. investor, I recomputed all results from the perspective of a U.K. investor. The results for the U.K. investor using the

Table 5
Results from robustness checks

Home country	U.S.	U.S.	U.K.	U.S.	U.S.	U.S.	U.S.
Market index	S&P 500	MSCI World	MSCI World	S&P 500	S&P 500	S&P 500	S&P 500
Dropped obs	None	None	None	Last 6	First 12	None	None
MA order	2	2	2	2	2	0	4
Panel I: Proportion rejecting null of neutrality							
Correlation neutrality	0.281*	0.269*	0.222*	0.247*	0.267*	0.316*	0.269*
Mean neutrality	0.280*	0.220*	0.287*	0.235*	0.169	0.213*	0.260*
Variance neutrality	0.040	0.040	0.120	0.027	0.022	0.040	0.020
VaR neutrality	0.000	0.017	0.000	0.000	0.000	0.000	0.000
Complete neutrality	0.281*	0.298*	0.193*	0.247*	0.260*	0.275*	0.257*
Joint test	0.292*	0.269*	0.216*	0.230*	0.247*	0.281*	0.275*
Panel II: Proportion rejecting null of neutrality on the downside							
Correlation neutrality	0.234*	0.240*	0.257*	0.212*	0.227*	0.287*	0.199*
Mean neutrality	0.200*	0.247*	0.247*	0.175	0.154	0.207*	0.193*
Variance neutrality	0.040	0.040	0.020	0.027	0.052	0.053	0.047
VaR neutrality	0.000	0.017	0.000	0.000	0.000	0.000	0.000
Tail neutrality	0.067	0.125	0.000	0.071	0.118	0.067	0.067
Complete neutrality	0.228*	0.257*	0.240*	0.247*	0.247*	0.257*	0.193*
Joint test	0.205*	0.240*	0.246*	0.212*	0.220*	0.257*	0.175

This table presents the results of tests of various types of market neutrality presented in the paper, applied to variations on the original "market neutral" hedge fund data set to check the robustness of the conclusions obtained. The tests are conducted on monthly hedge fund returns over the period April 1993 to April 2003. Panel I reports the proportion of funds for which the null hypothesis of market neutrality could be rejected at the 0.05 level in favor of a general alternative. Panel II reports the proportion of funds for which the null hypothesis of market neutrality could be rejected at the 0.05 level in favor of a dependence structure of a type that is specifically disliked by U.S.-based risk-averse investors with a preexisting exposure to the S&P 500 (which I call a test of "neutrality on the downside"). All tests were conducted using a bootstrap procedure described in the Appendix. An asterisk indicates that the proportion of rejected funds is significantly different from zero and thus that the collection of funds taken as a whole are significantly non-neutral. The first column of results are for the benchmark case, and are the same as those reported in the first column of Table 4. The second column considers the impact of switching the market index from the S&P 500 to the MSCI World index. The third column considers the MSCI World index as the market index, from the perspective of a U.K.-based investor, and so all U.S. dollar returns are converted to British pound returns. The fourth and fifth columns check the robustness of the results to "end-game" behavior and to "backfill bias," by using all but the last six months of data, or all but the first twelve months of data. The sixth and seventh columns examine the sensitivity of the results to the order of the moving average (MA) model employed to capture autocorrelation in the hedge fund returns.

MSCI World index are presented in column 3 of Table 5. The differences in the results are small, and are similar to those obtained for a U.S. investor using the MSCI World index.

4.3 End-game behavior

In the months leading up to a fund dropping out of the HFR or TASS databases, it is conceivable that the behavior of a fund's returns changes. If a fund is doing poorly and is about to be liquidated, then the investment decisions of the hedge fund manager may place greater emphasis on objectives other than maintaining the market neutrality of the fund. For this reason, I recomputed all the results for the U.S.-based investor using the S&P 500 index as the market index, dropping the last six observations on each fund. The results are presented in column 4 of Table 5, and are not substantially different from the results for the benchmark case in column 1.

4.4 Backfill bias

Hedge funds usually enter databases with a history of returns, usually ranging from six to eighteen months in length. It might be reasonable to think that the emphasis placed on maintaining market neutrality during the first year of a fund's life, relative to simply achieving positive returns, is lower than that for the rest of a market-neutral fund's life. Thus a rejection of market neutrality may come from the fund's first year of life, and not be representative of its neutrality following that first year. To allow for this, I recomputed all the results for the U.S.-based investor using the S&P 500 index as the market index, dropping the first twelve observations on each fund. The results are presented in the fifth column of Table 5 and are not substantially different from the results for the benchmark case in column 1.

4.5 Choice of return smoothing model

The work of Asness, Krail, and Liew (2001) and Getmansky, Lo, and Makarov (2004), which shows the prevalence of serial correlation in hedge fund returns, impels us to treat the problem of serial correlation very seriously. All results in this paper are reported using a second-order moving average, MA(2), model for the fund returns, as suggested by Getmansky, Lo, and Makarov (2004). In columns 6 and 7 of Table 5, I present the results obtained when using no filter, denoted MA(0), and when using an MA(4) filter. The sensitivity of the results to the choice of MA(q) model is relatively low. For the market neutral style of hedge funds, this is not surprising, as these funds exhibit relatively little serial correlation (Table 1). The results for the other styles of hedges were also relatively insensitive to moving from an MA(2) to an MA(0) or an MA(4) model. This is possibly because the increased estimation error in the MA filter partially offsets the better specification of the reported hedge fund returns.

4.6 Live versus dead funds

To further study the impact of survivorship bias, I present the results of the neutrality tests on live and dead funds separately. Table 6 reveals that the set of live funds had a slightly higher proportion of significantly non-neutral funds than the set of dead funds, though this may be influenced by the fact that dead funds on average had a shorter history of data than live funds (46.8 months versus 50.1 months). The only test that revealed a significant difference across fund status was the test of mean neutrality on the downside: all but one of the funds that fail this test are "alive." In general, the proportion of funds that failed a given test of neutrality was not substantially different from the proportion obtained when restricting attention solely to live funds (the differences ranged from 0% to 4.0%, with an average absolute difference of 1.0%); thus I conclude that my results are not overly affected by the presence (or absence) of "dead" funds in the data set.

Table 6
Proportion of funds that fail tests of neutrality, by fund status

	All	Live	Dead
Panel I: Proportion rejecting null of neutrality			
Correlation neutrality	0.281*	0.300*	0.194*
Mean neutrality	0.280*	0.303*	0.179*
Variance neutrality	0.040	0.041	0.036
VaR neutrality	0.000	0.000	0.000
Complete neutrality	0.281*	0.279*	0.290*
Joint test	0.292*	0.307*	0.226*
Panel II: Proportion rejecting null of neutrality on the downside			
Correlation neutrality	0.234*	0.243*	0.194*
Mean neutrality	0.200*	0.238*	0.036
Variance neutrality	0.040	0.041	0.036
VaR neutrality	0.000	0.000	0.000
Tail neutrality	0.067	0.077	0.000
Complete neutrality	0.228*	0.229*	0.226*
Joint test	0.205*	0.214*	0.161*

This table presents the results of tests of various types of market neutrality presented in the paper, applied to the "market neutral" hedge fund data, stratified into funds that were still present in the database at the end of the sample period ("live" funds) and funds that were no longer in the database at the end of the sample period ("dead" funds). There were 175 live funds and forty-two dead funds in the sample. The tests are conducted on monthly hedge fund returns over the period April 1993 to April 2003. The S&P 500 is used as the market index. Panel I reports the proportion of funds for which the null hypothesis of market neutrality could be rejected at the 0.05 level in favor of a general alternative. Panel II reports the proportion of funds for which the null hypothesis of market neutrality could be rejected at the 0.05 level in favor of a dependence structure of a type that is disliked by U.S.-based risk-averse investors with a preexisting exposure to the S&P 500 (a test of "neutrality on the downside"). The first column of results are for the benchmark case, and are the same as those reported in the first column of Table 4. All tests were conducted using a bootstrap procedure described in the Appendix. An asterisk indicates that the proportion of rejected funds is significantly different from zero and thus that collection of funds taken as a whole are significantly non-neutral.

4.7 Age of the fund and its market neutrality

Above I reported proportions of rejections of market neutrality concepts, averaging across all funds with sufficient data to conduct the test. But an interesting, and possibly important, question is whether the older funds have market dependence properties different from those of newer funds. Getmansky (2004) studies how the properties of hedge fund returns change over time, though she studies performance and size rather than dependence with the market portfolio. As an example, in Figure 1, I plot the linear correlation between a fund and the S&P 500 market index against the number of observations available on that fund. This plot indicates a significant positive relation between the correlation coefficient and the age of the fund.[23] The robust t-statistics associated with each of these correlation coefficients also have a positive relation with the number of observations available. A similar picture emerged when comparing the average age of funds that passed and failed the joint test of market neutrality on the downside, presented in Table 3.

These findings suggest that market neutral hedge funds that survive for a relatively long time are more positively dependent on the market return

[23] The reader may notice two observations in the upper right-hand corner of this plot, representing two funds that have the maximum number of observations (121) and correlation coefficients of over 0.9. The regression was rerun without these observations and the relation was still significantly positive.

than younger funds. This may be related to the fact that many hedge fund prospectuses commit the fund to a market neutral strategy only for a fixed period of time (three years seems to be a common choice), if at all. "Material" changes of investment strategy within that time often require the approval of shareholders, but changes after that time are not discussed. It is thus possible that hedge funds exhibit "style drift," and that funds that were once correctly classified as market neutral may not follow such strategies anymore. This observation suggests that a classification method based on relating fund returns to various risk factors, as in Brown and Goetzmann (1997) and Fung and Hsieh (1997) for example, may be more reliable than one based on the self-reported styles of funds. However, it is also possible that my finding is due to the fact that older funds are those that existed during the "dot com" bubble of the late 1990s.[24] I investigate this in the following robustness check.

4.8 Subsample analysis of market neutrality

It is possible that the relationship between the age of a fund and its degree of (non-)neutrality is purely due to the fact that older funds were ones that existed during the mid- to late 1990s, which was a period of rapidly rising stock market values. To investigate this, and also shed some light on how market neutrality may be affected by general market conditions, I reran all of the results in this paper over two subsamples, April 1993 to December 1998 (sixty-nine months) and January 1999 to April 2003 (fifty-two months). These subsamples roughly correspond to the tranquil and rising stock market (first subsample), and the peak and then decline in the market (second subsample). Given the dramatic increase in the rate at which new hedge funds entered the database over the sample period, only those funds that existed for the entire full-sample period are included in this analysis. This substantially increases the comparability of the results across the two subsamples.

Table 7 presents the results of joint tests of market neutrality, and market neutrality on the downside, across the five styles of hedge funds.[25] For the purpose of comparison, the first row of each panel is for the benchmark case, using all funds for the entire sample. The second row presents the results for the entire sample, using just those funds with a complete history of returns. Comparing these two rows reveals that studying only those funds with a complete history of returns leads to more rejections of neutrality than also including funds with an incomplete history of returns. This may be due to older funds being more significantly non-neutral than younger funds, or may simply be due to the tests having greater power when more data are available.

The third and fourth rows in each panel of Table 7 present the results of the joint test of market neutrality for each subsample. The results are quite

[24] I thank a referee for making this observation.

[25] The results for the full set of neutrality tests for this subsample analysis are available upon request.

Table 7
Proportion of funds that fail a joint test of neutrality, by style and subsample

	Market neutral	Equity hedge	Equity non-hedge	Event driven	Funds of funds
	Panel I: Proportion rejecting null of neutrality				
All funds, full sample	0.292*	0.539*	0.857*	0.611*	0.497*
Full sample	0.375*	0.784*	0.900*	1.000*	0.755*
Apr 1993–Dec 1998	0.375*	0.706*	0.850*	0.857*	0.755*
Jan 1999–Apr 2003	0.375*	0.549*	0.900*	0.500*	0.469*
	Panel II: Proportion rejecting null of neutrality on the downside				
All funds, full sample	0.205*	0.562*	0.844*	0.700*	0.538*
Full sample	0.375*	0.745*	0.950*	1.000*	0.816*
Apr 1993–Dec 1998	0.375*	0.686*	0.850*	0.929*	0.735*
Jan 1999–Apr 2003	0.188	0.529*	0.950*	0.643*	0.490*

This table presents the results of a joint test of market neutrality presented in the paper, applied to five different hedge fund styles. The tests are conducted on monthly hedge fund returns over the period April 1993 to April 2003. The S&P 500 is used as the market index. The sample is split into two parts: April 1993 to December 1998 and January 1999 to April 2003. The funds studied are those that had a complete set of historical returns. Panel I reports the proportion of funds for which the null hypothesis of market neutrality could be rejected at the 0.05 level in favor of a general alternative. Panel II reports the proportion of funds for which the null hypothesis of market neutrality could be rejected at the 0.05 level in favor of a dependence structure of a type that is disliked by U.S.-based risk-averse investors with a preexisting exposure to the S&P 500 (a test of "neutrality on the downside"). All tests were conducted using a bootstrap procedure described in the Appendix. An asterisk indicates that the proportion of rejected funds is significantly different from zero and thus that collection of funds taken as a whole are significantly non-neutral. The first row of results in each panel is for the benchmark case, as reported in the last rows of each panel in Table 4, and are repeated here to make comparing easier.

striking: for four out of five hedge fund strategies, I find that the proportion of significantly non-neutral funds was higher in the first subsample than in the second subsample. The only hedge fund style for which this is not true is the equity non-hedge style, which is the most non-neutral of styles. These results suggest that funds that existed in the early period of a booming stock market were more exposed to market risk than funds that existed in the second period when the stock market peaked and then fell. This is true even when I restrict the analysis to funds that existed in both subperiods, which is presumably a more homogeneous set of funds. Overall this subsample analysis suggests that market neutrality may vary with aggregate market conditions (which is not surprising, given the dynamic nature of most hedge fund strategies), and that it is not yet possible to disentangle the effects on the degree of market neutrality of the age of a fund and aggregate market conditions.

5. Conclusion

One of the most widely cited benefits of hedge fund investing is that hedge funds in general, and so-called market neutral hedge funds in particular, offer returns that have low correlation with the usual market indexes. This paper provided an in-depth study of the dependence between hedge fund returns and the S&P 500 index. I proposed generalizing the concept of market neutrality to consider the "completeness" of a fund's neutrality to market risks. This relates directly to the diversification benefits offered by hedge funds, a problem that has received less attention in the academic literature to date. I proposed

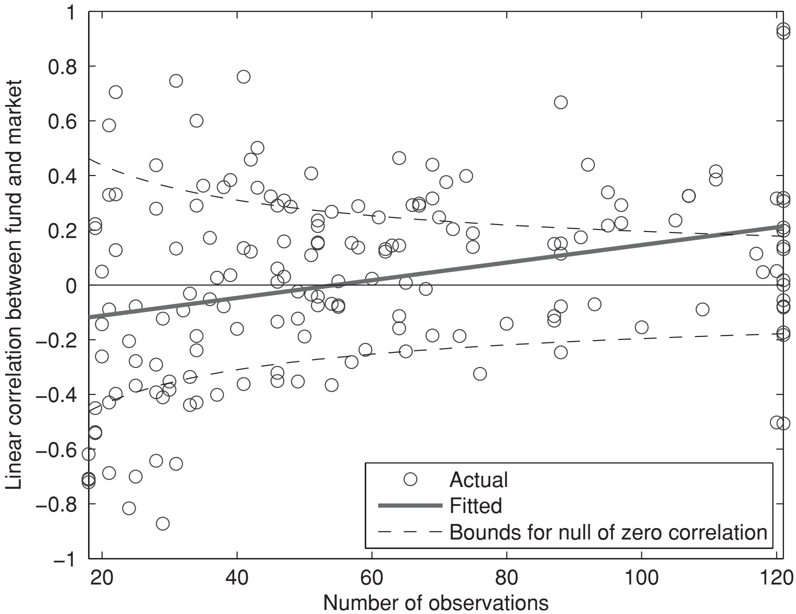

Figure 1
Relationship between survival and correlation with the market
The relation between linear correlation and the number of available observations on a hedge fund. The circles represent the number of observations/linear correlation with the market return observations for each fund in the sample with at least eighteen observations. The solid line is the result of a linear regression of correlation on a constant and the number of observations available. The dashed lines are simple Bartlett 95% confidence bounds for testing a null of zero correlation, and equal $\pm 1.96/\sqrt{n}$. For formal tests, I use serial correlation and heteroscedasticity robust confidence intervals based on a block bootstrap.

five new neutrality concepts for hedge funds: mean neutrality, which nests the standard correlation-based definition of neutrality; variance neutrality, value-at-risk neutrality, and tail neutrality, which examine the neutrality of the risk of a fund to market risk; and complete neutrality, which corresponds to the independence of the fund and the market returns.

I suggested statistical tests of each of these neutrality concepts. These tests take neutrality as the null hypothesis and compare it with either a general non-neutral alternative hypothesis, or a non-neutral alternative hypothesis that focuses solely on deviations from neutrality that would be disliked by a risk-averse investor. I applied the tests to a combined database of monthly returns on 1423 hedge funds from the HFR and TASS hedge fund databases over the period April 1993 to April 2003, using a block bootstrap method to deal with serial correlation, volatility clustering, and non-normality of the returns, and employing the S&P 500 index as the market index.

I found that about one-quarter of funds in the market neutral category are significantly non-neutral at the 0.05 level. In a series of robustness checks, I verified that my results are not overly affected by the choice of market index, the use of U.S. dollar returns, the choice of model to deal with hedge fund return autocorrelation, or by the last few observations, or first year of observations, on fund returns. I compared the results for the market neutral style with those obtained by looking at equity hedge, equity non-hedge, event driven, and "funds of funds" hedge funds, and found strong evidence that market neutral funds are more neutral to market risks than these funds.

Overall, my results suggest that the dependence between hedge fund returns and market returns is often significant and positive, even for market neutral funds. The widely cited diversification benefits from investing in hedge funds thus may not be as great as first thought. Some analysis of a fund's co-movements with the market is required to determine whether the fund is offering the degree and type of market neutrality desired by the investor. The neutrality concepts and tests proposed in this paper may assist investors in coming to a fuller understanding of the relationship between a given hedge fund and the market index.

The work in this paper leaves unanswered several interesting questions. One important extension of the methods in this paper is to allow for time variation in the exposure of hedge funds to market risks. It is well known that hedge fund managers employ dynamic trading strategies, and by studying unconditional correlations, for example, we average across these dynamics. The subsample analysis in Section 4 found some evidence that the market neutrality of hedge funds is time-varying. A second important extension of this paper is to consider tests of neutrality to a collection of market risks. If the collection of market risks to which an investor is exposed can be summarized in a single index, for example if the investor's existing portfolio is known, then the methods of this paper apply directly. If such an index is not available, then each risk factor would have to enter the model separately. With large sample sizes, simple generalizations of the methods presented in this paper could be employed, but, with the short sample sizes currently available, some care must be taken in specifying such flexible models.

Appendix: Details of the Bootstrap Tests

Concerns about short sample sizes, serial correlation, volatility clustering, and non-normality of hedge fund returns prompted the use of a block bootstrap to obtain critical values for the various tests proposed in this paper. I used the stationary bootstrap of Politis and Romano (1994), and the algorithm of Politis and White (2004) to determine the optimal average block size for each asset. Specifically, for the fund and the market, I applied the Politis and White algorithm to the return, squared return, and the product of the fund return and the market return. I then selected the largest of these three lengths to use as the block length for that fund. The block lengths selected ranged from 1 to 10, and averaged 2.4. I used 1000 bootstrap replications in all cases.

To obtain the distribution of each test statistic under the null hypothesis of neutrality, I resampled the fund and market returns separately, rather than resampling the vector of fund and market returns. By using the stationary bootstrap and imposing independence between the bootstrapped fund and market data, I ensure that the null hypothesis in each of the tests is satisfied, while not changing the univariate distributions of the fund and market returns, at least asymptotically. For all but the test of complete neutrality, independence is a sufficient but not necessary condition for the null hypothesis to hold.

On each block-bootstrapped hedge fund return series, prior to the running of any tests, I estimated a moving average (MA) model of order q. For the main results, I followed Getmansky, Lo, and Makarov (2004) and set $q = 2$, and then, in robustness checks, I also considered $q = 0$ and $q = 4$. The block bootstrap maintains the serial dependence structure of the original hedge fund returns, and by estimating the MA model on each bootstrap sample, I incorporate the impact of estimation error in the MA model on the bootstrap distribution of the test statistics.

Obtaining a joint test

Using five individual tests (correlation neutrality, mean neutrality, variance neutrality, VaR neutrality, and complete neutrality) and obtaining a "joint test" by simply checking whether at least one test failed clearly leads to a size distortion. Instead, I employed a method related to that of Westfall and Young (1993): on each bootstrap sample, I conduct the five tests, and then count the number of tests that lead to a rejection of a null hypothesis. If the test statistics were independent, then we would expect $0.05 \times 5 = 0.25$ tests to fail for each sample; however, these test statistics are almost certainly not independent, and this bootstrap procedure captures this dependence. I then compute the 95th percentile of the distribution of the number of tests failed. Across all funds, this quantile ranged from 0 to 2, with a median of 1 and a mean of 0.8, though for each fund the quantile is of course an integer between 0 and 5. (If the tests were independent then the 95th percentile would be 1: the *cdf* of a binomial (5, 0.05) evaluated at 1 is 0.98.) I conclude that a fund failed the *joint* test of market neutrality if it failed more tests than the 95th percentile of the distribution of the number of tests failed by the bootstrap data. This test controls the size of the joint test, and enables me to draw an overall conclusion about the neutrality of a fund.

Obtaining critical values on the proportion of funds failing a test

Under the null hypothesis that all funds in a given style are neutral, the expected proportion of funds to fail a given test is 0.05, but in order to conclude that the proportion of funds that fail a given test is significantly more than would be expected under the null, the 95th percentile of the distribution of the proportion of funds that fail a given test under the null hypothesis must be computed. I consider two approaches. The first approach relies on the assumption that the test statistics are independent across funds, and so the number of funds that fail a given test is a binomial $(K, 0.05)$ random variable, where K is the number of funds that are tested. The 95th percentile of this distribution, divided by K, ranged from 0.077 to 0.102 depending on the test (which affected the value for K).

The assumption that the test statistics are independent across funds is not realistic, and I considered a second approach again using the bootstrap to estimate the joint distribution of these test statistics under the null. Ideally, a vector bootstrap using all K funds would be used to obtain this distribution. However, because some of the funds have missing observations, either because they entered the sample late or left the sample early, I cannot use this method: it would generate bootstrap samples that have missing observations in the middle of the sample, which would mean that I could not employ all of our tests (the estimation of a GARCH model, for example, requires a contiguous sample). I employed an alternative approach, using a vector bootstrap only for the K^* funds that are present for the full sample.[26] For each bootstrap sample, I compute the proportion of these funds that failed each test, and then use the 95th percentile of the bootstrap distribution of these proportions as the critical value. This method relies on the assumption that the proportion of the K^* funds that fail each test has the same distribution as the proportion of

[26] The number of funds that were present for the full sample was 16, 51, 20, 14, and 49 for the five hedge fund styles considered here: market neutral, equity hedge, equity non-hedge, event driven, and funds of funds.

all K that fail each test, which may or may not be a reasonable assumption, but is likely a more reasonable assumption than that of independence between all test statistics. For market neutral hedge funds, this method yielded a critical proportion of 0.125 for the VaR-neutrality test, and 0.188 for the remaining tests.

References

Andrews, D. W. K. 1991. Asymptotic Normality of Series Estimators for Nonparametric and Semi-Parametric Regression Models. *Econometrica* 59:307–46.

Ang, A., and J. Chen. 2002. Asymmetric Correlations of Equity Portfolios. *Journal of Financial Economics* 63:443–94.

Ang, A., J. Chen, and Y. Xing. 2006. Downside Risk. *Review of Financial Studies* 19:1191–239.

Agarwal, V., and N. Y. Naik. 2004. Risks and Portfolio Decisions Involving Hedge Funds. *Review of Financial Studies* 17:63–98.

Agarwal, V., N. D. Daniel, and N. Y. Naik. 2006. Role of Managerial Incentives and Discretion in Hedge Fund Performance. Working Paper, London Business School.

Arrow, K. J. 1971. *Essays in the Theory of Risk Bearing*. Chicago: Markham Publishing Co.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath. 1999. Coherent Measures of Risk. *Mathematical Finance* 9:203–28.

Asness, C., R. Krail, and J. Liew. 2001. Do Hedge Funds Hedge? *Journal of Portfolio Management* 28:6–19.

Bae, K.-H., G. A. Karolyi, and R. M. Stulz. 2003. A New Approach to Measuring Financial Contagion. *Review of Financial Studies* 16:717–63.

Bansal, R., D. A. Hsieh, and S. Viswanathan. 1993. A New Approach to International Arbitrage Pricing. *Journal of Finance* 48:1719–47.

Beliossi, G. 2002. Market Neutral Strategies. *Journal of Alternative Investments* 5:1–4.

Bondarenko, O. 2004. Market Price of Variance Risk and Performance of Hedge Funds. Working Paper, Department of Finance, University of Illinois at Chicago.

Brown, S. J., and W. N. Goetzmann. 1997. Mutual Fund Styles. *Journal of Financial Economics* 43:373–99.

Brown, S. J., W. N. Goetzmann, and R. G. Ibbotson. 1999. Offshore Hedge Funds: Survival and Performance 1989–1995. *Journal of Business* 72:91–117.

Brown, S. J., and J. F. Spitzer. 2006. Caught by the Tail: Tail Risk Neutrality and Hedge Fund Returns. Working Paper, NYU Stern School of Business.

Caldwell, E. 1995. Introduction: The Model for Superior Performance, in J. Lederman and R. A. Klein (eds.), *Hedge Funds*. New York: Irwin.

Chan, N., M. Getmansky, S. M. Haas, and A. W. Lo. 2005. Systemic Risk and Hedge Funds, in M. Carey and R. Stulz (eds.), *The Risks of Financial Institutions*. Chicago: University of Chicago Press.

Chapman, D. 1997. Approximating the Asset Pricing Kernel. *Journal of Finance* 52:1383–410.

Chen, X., and X. Shen. 1998. Sieve Extremum Estimates for Weakly Dependent Data. *Econometrica* 66:289–314.

Christoffersen, P. F. 1998. Evaluating Interval Forecasts. *International Economic Review* 39:841–62.

Dittmar, R. F. 2002. Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns. *Journal of Finance* 57:369–403.

Embrechts, P., A. McNeil, and D. Straumann. 2001. Correlation and Dependence Properties in Risk Management: Properties and Pitfalls, in M. Dempster (ed.), *Risk Management: Value at Risk and Beyond*. Cambridge: Cambridge University Press.

Engle, R. F. 1982. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica* 50:987–1008.

Epstein, L. G., and S. M. Tanny. 1980. Increasing Generalized Correlation: A Definition and Some Economic Consequences. *Canadian Journal of Economics* 13:16–34.

Erb, C. B., C. R. Harvey, and T. E. Viskanta. 1994. Forecasting International Equity Correlations. *Financial Analysts Journal* 50:32–45.

Fung, W., and D. A. Hsieh. 1997. Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds. *Review of Financial Studies* 10:275–302.

Fung, W., and D. A. Hsieh. 1999. A Primer on Hedge Funds. *Journal of Empirical Finance* 6:309–31.

Fung, W., and D. A. Hsieh. 2001. The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers. *Review of Financial Studies* 14:313–41.

Getmansky, M. 2004. The Life Cycle of Hedge Funds: Fund Flows, Size and Performance. Working Paper, MIT Sloan School of Management.

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. Liang. 2005. Do Hedge Funds Have Enough Capital? A Value-at-Risk Approach. *Journal of Financial Economics* 77:219–53.

Harvey, C. R., and A. Siddique. 2000. Conditional Skewness in Asset Pricing Tests. *Journal of Finance* 55:1263–95.

Hull, J. C. 2003. *Options, Futures, and Other Derivatives*. Upper Saddle River, NJ: Prentice Hall.

Jarque, C. M., and A. K. Bera. 1980. Efficient Tests for Normality, Heteroskedasticity, and Serial Independence of Regression Residuals. *Economics Letters* 6:255–59.

Joe, H. 1997. *Multivariate Models and Dependence Concepts*. Monographs on Statistics and Applied Probability 73. London: Chapman and Hall.

Kimball, M. 1993. Standard Risk Aversion. *Econometrica* 61:589–611.

Koenker, R., and G. Bassett. 1978. Regression Quantiles. *Econometrica* 46:33–50.

Kraus, A., and R. H. Litzenberger. 1976. Skewness Preference and the Valuation of Risk Assets. *Journal of Finance* 31:1085–100.

Lo, A. W., and A. C. MacKinlay. 1990. An Econometric Analysis of Nonsynchronous Trading. *Journal of Econometrics* 45:181–212.

Longin, F., and B. Solnik. 2001. Extreme Correlation of International Equity Markets. *Journal of Finance* 56:649–76.

Lumsdaine, R. L., and S. Ng. 1999. Testing for ARCH in the Presence of a Possibly Misspecified Conditional Mean. *Journal of Econometrics* 93:257–79.

Nelsen, R. B. 1999. *An Introduction to Copulas*. New York: Springer.

Nicholas, J. G. 2000. *Market-Neutral Investing*. Princeton, NJ: Bloomberg Press.

Mitchell, M., and T. Pulvino. 2001. Characteristics of Risk and Return in Risk Arbitrage. *Journal of Finance* 56:2135–75.

Patton, A. J. 2006. Modelling Asymmetric Exchange Rate Dependence. *International Economic Review* 47:527–56.

Politis, D. N., and J. P. Romano. 1994. The Stationary Bootstrap. *Journal of the American Statistical Association* 89:1303–13.

Politis, D. N., and H. White. 2004. Automatic Block-Length Selection for the Dependent Bootstrap. *Econometric Reviews* 23:53–70.

Quintos, C. 2003. Extremal Correlation for GARCH Data. Working Paper, Simon Graduate School of Business, University of Rochester.

Rothschild, M., and J. E. Stiglitz. 1970. Increasing Risk: I. A Definition. *Journal of Economic Theory* 2:225–43.

Westfall, P. H., and S. S. Young. 1993. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. New York: John Wiley.