NEW YORK UNIVERSITY
STERN SCHOOL OF BUSINESS
FINANCE DEPARTMENT

Working Paper Series, 1995

Rejoinder: The J-Shape of Performance Persistence Given Survivorship Bias.

Brown, Stephen J., William N. Goetzmann, Roger G. Ibbotson, and Stephen A. Ross.

FIN-95-14

Rejoinder: The J-Shape of Performance Persistence Given Survivorship Bias

Stephen J. Brown, Stern School of Business, New York University
William N. Goetzmann, Yale School of Management
Roger G. Ibbotson, Yale School of Management
Stephen A. Ross, Yale School of Management

The J-Shape of Performance Persistence Given Survivorship Bias by Hendricks, Patel and Zeckhauser (1995), hereafter HPZ, presents a useful addition to the literature on the effect of survival bias in empirical studies in financial economics.[1] Their paper takes a closer look at an effect we report in Brown, Goetzmann, Ibbotson and Ross (1992) [BGIR], and shows that survivorship may induce even more complexities in performance persistence tests that we had previously considered. In particular, HPZ show that in the context of mutual fund performance studies, survival induces a non-linear pattern of subsequent performance of performance-ranked portfolios. We have always had great respect for the subtle implications of survivorship bias and their finding may well explain results that would otherwise be difficult to reconcile with our understanding of mutual fund returns. HPZ suggest that this observation allows us to distinguish between spurious patterns of performance resulting from the effects of survival, and patterns of performance that result from what they term "true differential performance" that persists. They further propose a simple t-test based on a quadratic regression in performance ranks that appears to provide a test for spurious survival-induced performance that is reasonably powerful against a hypothesis of true persistence.

Unfortunately, though, it is over-reaching to think that simple statistical tests can resolve such complex issues. Much more likely we will find that the actual pattern observed is a mixture of survivorship bias and other effects. Numerous studies (for example, Hendricks, Patel and Zeckhauser, 1993, Brown and Goetzmann, 1995 and Malkiel, 1995) document that much of the apparent persistence in mutual fund returns arises from the persistence of poor performance, yet new money flows into mutual funds as a function of past period performance appears to be J-shaped (Goetzmann and Peles,1994 and Sirri and Tufano, 1992). One interpretation of this result is that investors are insensitive to extremely poor past period performance. An alternative point of view,

[1] See, for instance, Brown, Goetzmann, Ibbotson and Ross, 1992, Brown, Goetzmann and Ross, 1995 and Hendricks, Patel and Zeckhauser, 1993.

given the results reported by HPZ, is that the result is an artefact of survival. Note, too, that there is nothing particularly mysterious about the persistence of poor performance. It is easy to find ways to drain value from a fund and not surprising that some investors are relatively indifferent to it.

In this short rejoinder to their article, we focus upon three issues that we believe are of interest to researchers. First, we provide a simple example which shows the intuition behind the J-shape discovered by HPZ. Second, we show how the non-linearity in response is specifically related to the second moment of the distribution of performance measures. Finally, we report the results of simulations that explore the robustness of the statistical test proposed by HPZ to assumptions about survival thresholds and cross-sectional correlations.

A Simple Example

It is not immediately obvious that one should expect a J-shaped response pattern in asset returns, conditional upon surviving a performance threshold from one period to the next. However, a simple diagram provides some intuition for why this is so. Consider two funds with random returns. Fund A and fund B have equal means, and the variance of B is greater than the variance of A. Figure 1 represents the period 1 distribution of returns for the two funds. Assume that either fund will drop out of the sample if its return lies below some common threshold K. The intuition of BGIR [1995] is that higher variance funds *that survive* are more likely to be winners in both periods, and this induces a spurious performance persistence. This effect can easily be seen in the diagram by noting that conditioning upon survival has the effect of truncating the left tail of the distribution of the two types of funds by differing amounts. Rescaling the area under the density curves to one moves the center of the *conditional distribution* of the higher variance fund farther to the right.

The HPZ result can also be understood from this diagram. Notice that, at a low threshold K, the heights of the densities for the two types of funds differ considerably. If we observe a very low return by some fund in the first period, the difference in the heights of the densities near threshold K indicates that the fund was more likely of the volatile type B. If it is a more volatile fund, then its mean in the next period *conditional upon survival* is higher than the conditional mean of the less volatile fund. Thus, conditioning upon *low* first period returns can also "pick out" the more volatile funds that will have higher conditional expected returns in the following period.

Figure 1 also shows that as the threshold moves higher, say to K*, conditioning on relatively poor performance in the first period is less informative. The likelihood ratio (formed by the ratio of heights of the two densities at K*) moves closer to one. For higher thresholds, the J shape will be less pronounced, and thus less likely to be observed in sample.[2]

Expected U-Shape in Volatility

The above example suggests that it is the differential variance that drives the J-shape response in returns. If this is true, we would also expect to find a J-shaped relationship between return and variance, or more properly, between risk-adjusted returns (i.e. alphas) and residual variance. In figures 2 and 3 we show the results of simulations that address this conjecture. In these figures, we simulate risk-adjusted performance and total risk with and without a performance cut. Figure 2 has no performance threshold, and thus no bias in the subsequent performance measure. The pattern of performance conditional upon past returns is flat, however the volatility is not: the risk pattern is U shaped, since both unusually low and unusually high returns "pick-out" the high variance funds. Figure 3 shows the result when we add a performance cut after the first period. Survivorship induces a spurious covariance between the performance measure and the residual risk -- both of which are now J-shaped.

Simulating Tests About Survival

Among the important issues in the application of the test proposed by HPZ are estimates of the size, power and robustness to different conditions such as threshold levels and intercorrelation of securities. HPZ note that the J-shaped pattern that emerges as an artefact of survival "is starkly different from a monotonic increasing relation that would be predicted by a hypothesis of pure survival." They further assert that the "discovery of a monotonic pattern rather than a J-shape suggests that the performances of mutual funds are truly persistent in the short-run." HPZ would

---

[2] It is interesting to note that the magnitude of the J-shaped response bias is not invariant to the type of performance measure used. In BGIR [1992], for instance, we show that the appraisal ratio (formed by Jensen's alpha scaled by residual variance) reduces the effect of survival bias. Figure 1 shows why this is so. This scaling reduces the cross-sectional differences in the densities for extreme values. Unfortunately, the diagram also shows that using appraisal ratios as performance measures will reduce the power of any test designed to pick up survivorship bias.

therefore suggest that failure to find an interior minimum in the subsequent performance of ranked portfolios is strong evidence against the survival hypothesis.

Replicating the experiment that leads to Figure 1 in HPZ, we find that the failure to find an interior minimum in performance is a reasonably rare event. For the experimental conditions reported in HPZ, we find we fail to find an interior minimum 2.42 percent of the time[3]. HPZ note that adding intercorrelation between the residuals, while holding fixed the number of funds increases the variance of the shape across simulations[4]. Indeed, we find that this realistic intercorrelation implies that we fail to find an interior minimum in 20.38 percent of the simulations. Thus failure to find an interior minimum (as indicated in Figure 1 of HPZ) may simply be an indication of the extent of cross-sectional dependence in fund returns.

Of course, an interior minimum is a necessary but not sufficient condition for the J-shape in performance documented by HPZ. Unfortunately, the pattern of performance is almost never the striking convex pattern suggested in the discussion and illustrated in Figure 1 of HPZ. The pattern was not convex in 98.3 percent of the simulations (97.8 percent when we allowed for cross-sectional dependence in residual fund returns). With a slightly lower 5 percent performance cut which corresponds to typical mutual fund industry values we found a convex pattern in only 22 out of 5000 simulations. On the other hand, a pattern of returns that is monotone increasing up to one interchange of ranks (as found in the empirical work reported by HPZ) we find in 13.54 percent of our simulations (20.9 percent when we allow for interdependence). In short, failure to find a J-shape is relatively weak evidence against a survival artefact in the data.

The authors recognize that the visual examination of performance patterns can be misleading,

---

[3]The results reported in Figure 1 assume a 10 percent performance cut. With a slightly lower 5 percent performance cut we fail to find an interior minimum in 2.82 percent of the simulations. With intercorrelation, this number increases to 25.02 percent.

[4]HPZ provide a very useful algorithm for generating interdependence of the type observed in residual fund returns. However, aside from the minor typographical error (d should be $\sqrt{d}$ in the last term of Equation (6)), the reader should note that since the implied variance of residual returns in this factor representation is $\sigma_i^2 = f_1^2 + f_2^2 + d$, d must vary for each fund in such a way that equation (4) in HPZ is satisfied. In our implementation of the HPZ algorithm, we define the interdependence for β plus or minus two cross sectional standard deviations from the mean.

and that the shape of the resulting performance graph is a sample statistic with positive variance. They therefore propose an interesting statistical test for the hypothesis that the subsequent performance of performance ranked funds achieves a minimum that is interior to the set of rankings. This test appears to have reasonable power against an alternative hypothesis that fund returns evidence true persistence. If the true performance pattern is consistent with an interior minimum, the linear term in a quadratic regression of performance ranks on past performance ranks should be negative. In actual mutual fund data, HPZ find that the t-value associated with this statistic is +3.06, and conclude that they can "easily reject pure survivorship bias with zero persistence."

We replicate the simulation experiment reported in HPZ with similar results. With "mild pure survivorship" a simple linear regression of succeeding ranks would give some evidence of persistence in the data. In Table 1 we report that the 95 and 99 percentiles of the t-value associated with the hypothesis that the slope coefficient γ_1 is zero are 2.65 and 3.38 respectively. However, when we look at the linear term of a quadratic regression of performance ranks, we find that the 95 and 99 percentiles of the t-value are -.03 and .65 respectively. Hence we would conclude that the true performance is an interior minimum (J-shape). As HPZ indicate, this conclusion is reinforced when we examine more stringent performance cutoffs. With a ten percent cutoff, we would reject the hypothesis of no persistence more than 25 percent of the time if we were to examine a simple linear regression of sequential ranks. However, the 95 and 99 percentiles of the empirical distribution of $t(\beta_1)$ indicate that we would expect to find positive values of this statistic only about one percent of the time.

Since the observed value of $t(\beta_1)$ is +3.06 for actual mutual fund data is in excess of the 99th percentiles of the distribution of this statistic induced by pure survival, it might appear that the evidence is strongly opposed to the survival story, at least in this particular application. However, this result is not robust to the important assumption that residual returns are statistically independent in the cross-section of funds. The decisions of fund managers are not in fact independent and herd behavior is well documented (see, for example Grinblatt, Titman and Wermers (1993)). Modeling this intercorrelation not only increases the variance of the shape of fund performance, as documented by HPZ. It increases the magnitude of apparent persistence measured by a simple linear regression of successive ranks. In addition, the observed value of $t(\beta_1)$ in the HPZ study will be exceeded more than five percent of the time. The one-tail test p-value of the HPZ statistic is in fact 12.5 percent based on the simulation-based distribution of $t(\beta_1)$ with a five percent cutoff, and decreases to 8.1 percent with a ten percent cutoff.

Implications For Future Research

The discovery of a survival-induced J-shaped pattern in asset returns is a useful contribution to the growing literature about the issue of survival biases in empirical finance. It may help to explain puzzling results reported in the mutual fund literature, and may provide a guide for future experimental design. Our investigation of the HPZ results led us to a more complete understanding of how differential volatility affects survival-conditioned returns. Our simulations of the test statistic proposed by HPZ suggest that power of the test is dependent upon the absolute level of the threshold, as well as upon the magnitude of the cross-sectional differences in variance. While it would be useful to have a reliable test of the conjecture that survivorship is *not* driving an observed empirical result, we are only beginning to understand the kind of empirical regularities that survival may induce. Perhaps most significantly, Neither BGIR [1992] nor HPZ [1995] consider any truncation rule more complex than a common lower bound. The actual rules for fund survival are likely to be much more complex. These may well include "up and out" bounds as well as lower bounds, and may also depend upon the sum or the average return over the two-period interval. For a richer set of rules governing survival it is not clear what kind of conditional response to expect, however this question is a potentially fruitful one for future research. HPZ suggest that failure to find a J-shape serves to distinguish spurious survival-induced performance from "true performance persistence" which they claim has been observed with mutual funds. Even taking into account the complicating factors of sampling error and intercorrelation of fund returns which reduce power, the HPZ procedure may be useful for testing the null about *one specific implication* of survivorship. It is important to stress that their procedure is not a general test of the proposition that an empirical result is unaffected by survivorship concerns. Indeed, as we noted before, persistence is really only mysterious at the upper end of the scale -- probably above the "J" -- since there is nothing very troublesome about persistent poor performance. Research in this field will not produce black and white answers but rather should focus on better dissecting the differential impact of survivorship bias.

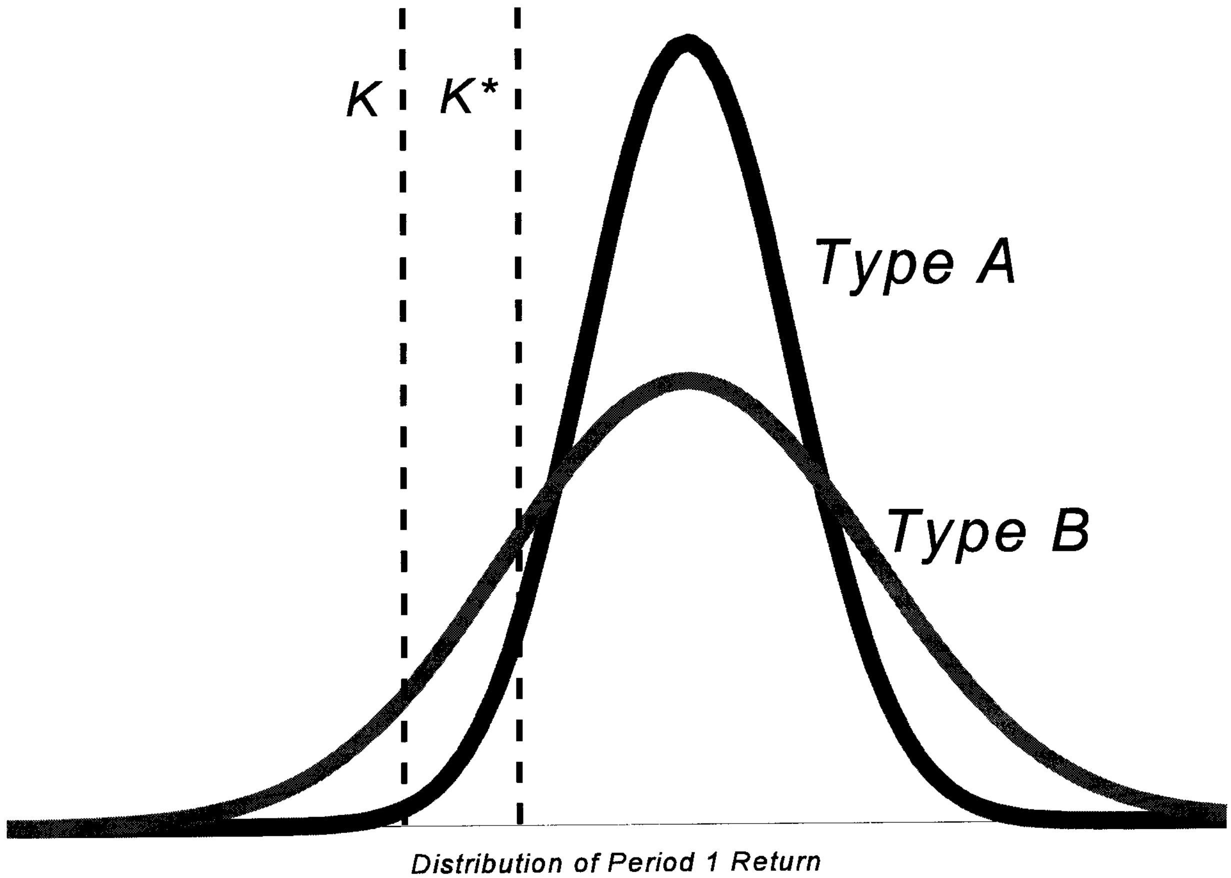
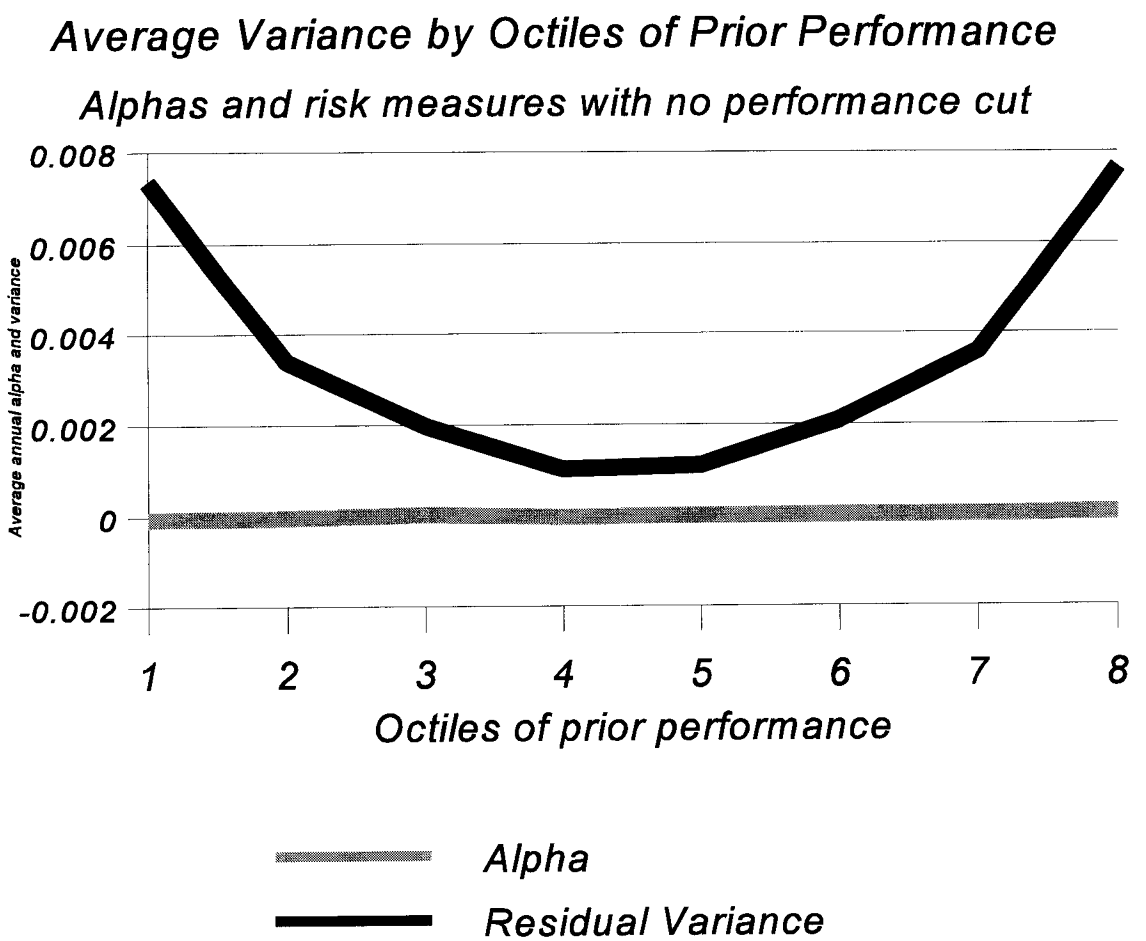


Figure 1



Average Variance by Octiles of Prior Performance
Alphas and risk measures with no performance cut
0.008
0.006
0.004
0.002
0
-0.002
Average annual alpha and variance
1
2
3
4
5
6
7
8
Octiles of prior performance
Alpha
Residual Variance

Figure 2

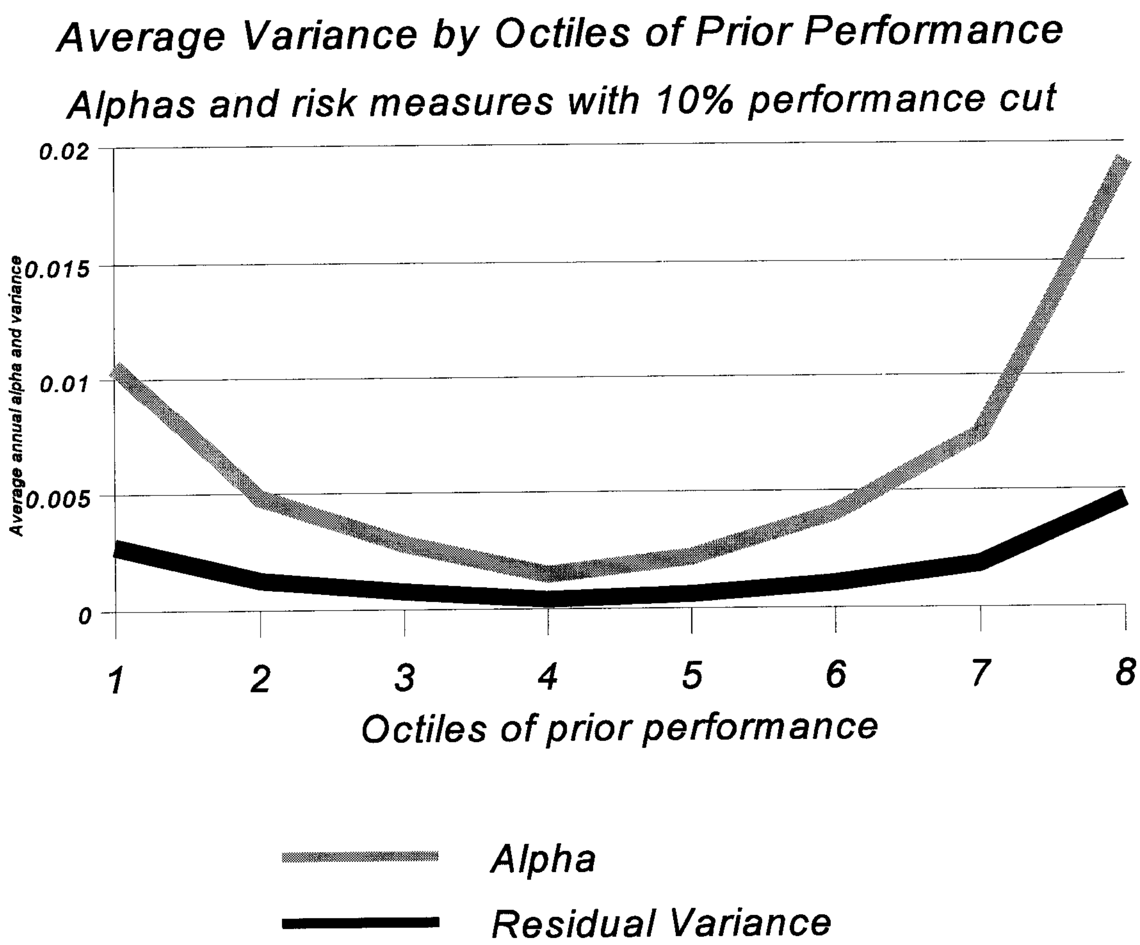

Average Variance by Octiles of Prior Performance
Alphas and risk measures with 10% performance cut
Average annual alpha and variance
0.02
0.015
0.01
0.005
0
1
2
3
4
5
6
7
8
Octiles of prior performance
Alpha
Residual Variance

Figure 3

References

Brown, Stephen J, William N. Goetzmann, Roger G. Ibbotson and Stephen A. Ross, 1992, Survivorship Bias in Performance Studies, *Review of Financial Studies* 5, 553-580.

Brown, Stephen J, William N. Goetzmann and Stephen A. Ross, 1995, "Survival," *Journal of Finance*, 50 (3) 853-874.

Goetzmann, William and Nadav Peles, 1994, "Cognitive Dissonance and Mutual Fund Investors," *Working Paper, Yale School of Management.*

Grinblatt, Mark, Sheridan Titman and Russell Wermers, 1994, Momentum Strategies, Portfolio Performance and Herding, Working Paper, Johnson School of Management, UCLA.

Hendricks, Darryl, Jayendu Patel and Richard Zeckhauser, 1993, Hot hands in Mutual Funds: the Persistence of Performance, 1974-88, *Journal of Finance* 48, 93-130.

Daryl Hendricks, Jayendu Patel and Richard Zeckhauser, 1995,"The J-Shape of Performance Persistence Given Survivorship Bias," *Review of Economics and Statistics*, Forthcoming.

Malkiel, Burton G., 1995, "Returns Form Investing in Equity Mutual Funds 1971-1991," *Journal of Finance* 50(2) 549-572.

Erik Sirri and Peter Tufano, 1992, "The Demand For Mutual Fund Services by Individual Investors," Working Paper, Harvard Business School.

Table 1: Simulation-based distribution of rank-based cross-section test statistics subject to survival biases

	Cross sectional Independence				**Contemporaneous Residual Correlation**			
	5 percent cut		*10 percent cut*		*5 percent cut*		*10 percent cut*	
Fractile	$t(\beta_1)$	$t(\gamma_1)$	$t(\beta_1)$	$t(\gamma_1)$	$t(\beta_1)$	$t(\gamma_1)$	$t(\beta_1)$	$t(\gamma_1)$
0.5	-1.947	0.781	-2.966	1.467	-1.508	0.724	-2.064	1.055
0.75	-1.146	1.582	-2.028	2.294	1.103	3.856	0.435	3.956
0.95	-0.032	2.656	-0.669	3.502	5.152	9.042	4.105	10.250
0.975	0.322	3.021	-0.214	3.852	6.365	10.792	5.521	12.996
0.99	0.655	3.389	0.232	4.299	7.967	12.885	6.987	16.427
0.995	0.933	3.666	0.613	4.665	8.845	14.669	7.708	18.753

In this table, we report the upper fractiles of the distribution of test statistics based on the cross-section quadratic regression of the rank of second two-year performance against the rank of first two year performance:

$$\text{Rank}[r_{i2}] = \beta_0 + \beta_1 \text{Rank}[r_{i1}] + \beta_2 \text{Rank}[r_{i1}]^2 + \nu_i$$

and the linear regression:

$$\text{Rank}[r_{i2}] = \gamma_0 + \gamma_1 \text{Rank}[r_{i1}] + \eta_i$$

where the subscripts 1 and 2 denote the two different periods in which the performances are measured, and Rank[.] denotes the relative ranking of the performance within the period. The Cross sectional independence results correspond to the simulations reported in HPZ, with a five percent performance cut leading to a final sample of 500 managers, and 5000 simulations. We also report results with a ten percent cut. The contemporaneous residual correlation results correspond to simulations where the Cross sectional correlation of residual returns correspond to the data for mutual funds. This pattern of correlations is described in HPZ.