"How do CTAs' Return Distribution Characteristics affect their likelihood of Survival?"

Fernando Diz[1]

Working Paper

---

[1] Fernando Diz is an Assistant Professor of Finance at Syracuse University, and President of M&E Financial Markets Research, LLC, a registered Commodity Trading Advisor.. I am grateful to Sol Waksman and Peter Nicolson from Barclay Trading Group, and to Richard Oberuc from Burlington Hall Asset Management for helpful comments and suggestions. This study was partially funded by a grant from the Foundation for Managed Derivatives Research. The opinions expressed in this study do not necessarily reflect those of Syracuse University, M&E Financial Markets Research, LLC or the Foundation for Managed Derivatives Research.

"How do CTAs' Return Distribution Characteristics affect their likelihood of Survival?"

I. Introduction

The issue of whether fund performance can be predicted has been an issue of intense academic and industry research. Most academic evidence on this issue tends to point out that performance is not predictable (for evidence on mutual funds see Malkiel, 1995 and references therein). A reflection of these findings is the statement on disclosure documents that past performance is not indicative of future performance.

One issue that has not received as much attention in the Managed Futures Industry is the question of survival. By survival we mean the ability of a program or a Commodity Trading Advisor (CTA) to stay in business. Why is this issue important? First, ignoring survivorship biases can substantially affect reported industry benchmark returns. Second, survivorship biases will tend to generate the appearance of predictability of returns. Third, the chances of survival in different industries (Managed Futures vs. Mutual Funds for example) may be markedly different rendering absolute return comparisons suspect. Fourth, programs that fail to survive may *systematically* exhibit undesirable performance characteristics. Finally, survival experience may be explained by the program's return characteristics making its study extremely important if one wants to increase the performance of a pool vehicle. If survival experience can be explained and predicted. one could potentially select programs based on their likelihood of survival and by doing so increase the future performance of the program or pool. This is in contrast to pure diversification which, in the presence survivorship problems will likely reduce returns and increase volatility.

The objective of this study is to provide some answers to these questions. The specific questions that we want to address are:

- What are the overall chances of survival in managed futures and how do they compare with the mutual fund industry?

- Do programs that survive over time have more desirable performance characteristics than those who do not?
- How do surviving programs differ from their unsuccessful counterparts in their return distribution characteristics that are associated with the program's returns mathematical expectation and its money management? We shall refer to these components as the 'edge' component and the 'money management' component respectively.
- Are common program's classifications (diversified vs. non-diversified, systematic vs. discretionary) useful in discriminating between performance and survival characteristics?
- Can a program chances of survival be reasonably predicted?

II. Data

To find answers to these questions we studied 925 programs between 1975 and June 1995. The data used was provided by the Barclay Trading Group, Ltd. and represents one of the most comprehensive managed futures databases. Of the 925 programs, 490 were still in business as of June, 1995. The 'edge' of a program was proxied by seven different and complementary measures. The first one is the annualized monthly compounded rate of return (AMCRR) over the life of the program. This measures the level of returns that the program has been able to yield over its life. To give an idea of the range of returns that one can find among different programs over their lives, the maximum AMCRR was 226.4% and the minimum AMCRR was -52.7%. The second measure of the edge was the return on the most successful month. This measure focuses on the magnitude of potential gains and gives an idea of the magnitude of the program's edge and whether the program is trading at optimal levels. The program with the most successful month had a return of 319% while the program with the smallest most successful month had a return of 0.8%. The third measure used was the Sharpe ratio. This is a measure of returns over and above the treasury bill rate per unit of standard deviation. It is useful to compare returns from programs with very different 'risk' characteristics. Suppose you have two programs A and B that have the same monthly compounded return, but program A has a Sharpe ratio that is two times as large as that of program B. This means that program A can generate the same returns as program B but with half the risk or volatility. The next two measures of the edge, skewness and kurtosis, are rather technical and measures how capable a program is to generate larger returns with a high degree of probability. The percentage of winning months provides an indication of the type of trading system that the program is using. As will be shown later in the paper low percentages are often associated with long-term trend following programs, and higher percentages are associated with very short-term arbitrage type strategies. Finally, the average monthly winning return helps us understand how profitable winning trades are. Short-term trading approaches have much lower average winning months than long term trading approaches. Table 1 summarizes the edge measures for your quick reference. Table 2 contains descriptive statistics for the edge variables for the 925 programs over the 1975-1995 period.

The quality of the program's money management was proxied by seven different measures. The first one was the standard deviation of monthly returns. The standard deviation gives an idea of the probability of large positive returns or drawdowns relative to the program's average monthly return. Take for example a program with a monthly standard deviation of 8%, and an average monthly return of 3%. Since 95.44% of all the possible monthly returns should be within plus/minus two standard deviations of the mean of the program[2], one is equally likely to see maximum drawdowns of 13% (3% - 2x8%) and large positive returns of 19% (3% + 2x8%). The second measure of money management was the maximum monthly drawdown. This measure focuses on the magnitude of potential losses and gives an idea of both risk controls and whether the program is trading at optimal levels. To give a better idea of the magnitude of drawdowns found, the largest drawdown was 81%, while the smallest drawdown was no drawdown at all. The average *monthly* drawdown for all programs studied was 16%. Although looking at any single monthly drawdown is important, perhaps even more important is the amount of time that it takes a program to come out of a loss. For example, suppose that a program's initial equity is $1,000,000 and that in the next month

[2] Of course this assumes that the distribution of a program's returns is normal. The evidence presented in this study suggests that this assumption may not be a reasonable one. Programs' returns are highly skewed, suggesting that the probability of a large positive return is higher than the probability of a large loss.

the program experiences a 25% drawdown, reducing its equity to $750,000. What is the maximum number of months that would take the program to bring the equity back to $1,000,000? This is what the third money management variable measures. It took 137 months for the worst program in this category to recover from its worst drawdown. It is not surprising that after failing to recover from its worst drawdown in more than eleven years, this program went out of business. It took an average of 21 months for the average program to recover from its worst drawdown. This measure is important because it can be theoretically shown that a program with good money management can recover from drawdowns much faster than a program with poor money management. A measure of the consistency with which programs recover from drawdowns is the standard deviation of the times to recover from any drawdown. A low number indicates greater consistency, a high number indicates poor consistency. When the time to recover from the worst drawdown is related to the time that the program has been in business a relative drawdown measure is created. Longer drawdown periods for an established program may not have the same effect on the survivability of a program as longer drawdown periods for relatively young programs. As an indirect way of measuring the level of leverage we use the reported margin to equity ratio. Finally, an important measure of money management and risk controls is the average losing monthly return. These money management' measures are summarized in Table 3 and descriptive statistics are contained in Table 4.

Table 1

Measures of the 'Edge' of a Program

1. Annualized Monthly Compounded Return	Bottom line performance of a program or trader.
2. Maximum Monthly Return	Return on the most successful month. Proxy for the program's return potential and whether it is trading at optimal levels.
3. Annual Sharpe Ratio	Measure of returns per unit of standard deviation. Allows one to compare returns of programs with different volatility.
4. Skewness of Monthly Returns	A positive number tells that the program has a tendency to generate returns higher than the mean more often than normal.
5. Kurtosis of Monthly Returns	Measures how returns are packed around the mean relative to a normal distribution. A positive number tells that there is higher probability of finding returns near the mean compared to a normal distribution.
6. Average monthly winning return %	Measures the average monthly return calculated with all the winning months.
7. Percentage of winning months	Measures the percentage of all trading months that were profitable.

Table 2

Descriptive Statistics on the 'Edge Variable'' for the 925 programs during the 1975-1995 period.

Variable	Mean	Std.	Min	Max
Average Time in Business (months)	62	43	8	246
Annual Compounded Return	16.3%	20.4%	-52.7%	226.4%
Maximum monthly return	32.7%	34.5%	0.8%	319.5%
Sharpe Ratio	0.38	0.80	-2.47	13.06
Skewness of monthly returns	0.90	1.19	-3.68	7.52
Kurtosis of monthly returns	3.90	6.26	-1.9	70.51
Percent wins	60%	12%	0%	100%
Average Win %	6.66%	5.34%	0.30%	57.64%

Table 3
Measures of the Money Management of a Program

1. Annualized Standard Deviation of Monthly Returns	Consistency with which programs achieve their returns. Gives an idea of what kind of drawdowns and large returns to expect.
2. Maximum Monthly Drawdown	Magnitude of potential loses. Gives an idea of risk controls and whether the program trades at optimal levels or not.
3. Months to Recover from Worst Drawdown	Other things equal, programs with good money management will recover quicker than programs with poor money management.
4. Standard Deviation of Times to Recover from Drawdown	Measures the consistency with which programs come out of drawdowns.
5. Time to Recover from Worst Drawdown as a percent of the Program's Life	Gives an idea of how much of a program's business life is spent recovering from the worst drawdown.
6. Reported margin to equity ratio	Measures the leverage reported by the CTA or program.
7. Average losing monthly return	Measures the average monthly loss calculated using only the losing months.

Table 4
Descriptive Statistic for Money Management Variable for the 925 program during the 1975-1995 period.

Variable	Mean	Std.	Min	Max
SD of monthly returns	29.7%	22.9%	0.3%	186.3%
Maximum monthly drawdown	-16.4%	12.5%	-81.0%	0.2%
Max months to recover from loss	21	17	0	138
SD number of months to recover	7	7	0	53
Max months to recover from loss as a percentage of program's life	0.38	0.22	-	1.00
Margin to equity ratio	20.1%	10.1%	1.0%	100.0%
Percent loss	40.2%	12.4%	-8.0%	100.0%

III. Results

III.A. **Mortality Rates: Managed Futures vs. Mutual Funds**

The results in Table 5 attempt to provide an answer to our first question. The table contains equally weighted annual returns and mortality rates in both the managed futures and mutual fund industries for the period 1982-1994.[3] The table shows that average returns in the managed futures industry have been twice as high than in the mutual fund industry. A less known fact however, is that the mortality rate in managed futures is very high. Roughly one of every two managed futures programs fails over time. The mortality rate in the mutual fund industry on the other hand, is only one in seven or 14%. These numbers are biased downward because we included the most recent data in the calculation of the average mortality rate. As more recent data is used mortality rates naturally drop. When we exclude the last two years from the calculation of the average mortality rate, the averages are 51.6% and 15.9% for managed futures and mutual funds respectively.

Table 5
Differences in Rates of Return and Mortality Rates in the Managed Futures and Mutual Fund Industries

	Managed Futures		**Mutual Funds***	
Year	**Mean Return**	**Mortality Rate**	**Mean Return**	**Mortality Rate**
1982	29.60	51.5	25.03	17.80
1983	28.90	56.0	20.23	16.10
1984	33.30	53.1	-2.08	16.20
1985	47.10	56.1	27.17	13.90
1986	40.00	56.1	13.39	16.80
1987	64.80	55.8	0.47	15.80
1988	29.90	56.1	14.44	14.60
1989	19.80	52.8	23.99	11.30
1990	33.60	47.5	-6.27	5.40
1991	15.20	45.1	N/A	N/A
1992	8.80	37.6	N/A	N/A
1993	15.20	27.4	N/A	N/A
1994	4.30	12.3	N/A	N/a
Average	28.50	46.7	12.93	14.2

*Data on mutual funds from Malkiel, B. (1995).

[3] The mutual fund results have been adapted from Malkiel (1995).

The question then becomes: If an investor chooses portfolios randomly, without consideration of survival issues, how much is he expected to lose from including programs that will not survive in his portfolio? To answer this question we need to know the expected difference in returns between portfolios of surviving and non-surviving programs. Tables 6 and 7 attempt to provide answers to this question.

Table 6
Annual Rates of Return for Surviving and Non-Surviving Mutual Funds

Year	Surviving	Non-Surviving	Difference
1982	26.0%	20.4%	5.6%
1983	21.7%	12.8%	8.9%
1984	-1.3%	-6.4%	5.1%
1985	28.1%	21.4%	6.7%
1986	14.4%	8.5%	5.9%
1987	0.9%	-1.9%	2.8%
1988	15.5%	8.4%	7.1%
1989	24.9%	16.7%	8.2%
1990	-6.0%	-11.1%	5.1%
Average	**13.8%**	**7.6%**	**6.2%**

Table 7
Average Annual Rates of Return for Surviving and Non-Surviving Managed Futures Programs

Year	Surviving	Non-Surviving	Difference
1980	102.0%	69.3%	32.7%
1981	32.0%	35.2%	-3.2%
1982	57.6%	36.4%	21.2%
1983	34.8%	4.3%	30.6%
1984	32.4%	40.5%	-8.1%
1985	49.4%	49.6%	-0.3%
1986	43.7%	39.1%	4.6%
1987	49.4%	50.5%	-1.1%
1988	36.6%	35.4%	1.2%
1989	22.7%	14.7%	8.1%
1990	33.3%	26.4%	6.9%
1991	17.5%	9.5%	8.0%
1992	16.4%	9.5%	6.9%
1993	19.8%	8.9%	10.9%
1994	8.7%	-7.4%	16.1%
Average	**37.1%**	**28.1%**	**9.0%**

On average, the return difference between surviving and non-surviving mutual funds has been around 6% per year for the period 1982-1990. This difference has been very uniform across that time period. In the managed futures industry the average return difference between surviving and non-surviving programs has been in the order of 9%. We can calculate the expected reduction in performance from randomly selecting programs using the following methodology.

Let's define n = the number of programs in a portfolio.

P = the probability of choosing a program that will survive.

$P_n(k) = \binom{n}{i} p^i (1-p)^{(n-i)}$ = the probability of choosing k surviving programs in a portfolio that contains n programs.

$R_n(k) = (k + (n-k).0.91)/n =$ is the relative performance of a portfolio that contains k surviving programs and n-k non-surviving programs.

$1 - \mathrm{R}_n(\mathrm{k})$ is the performance reduction relative to a portfolio of survivors.

For example if we have a portfolio composed of 2 programs (n=2) and we happened to choose programs that will survive, then $k = 2. \mathrm{R}_2(2) = (2 + (2).0.91)/2 = 1$, so that the performance reduction $(1 - \mathrm{R}_2(2))$ is 0. In other words if one happens to choose all survivors you will have full performance.

Then the expected reduction in performance in a 10 program portfolio using an average mortality rate of 46.7% would be equal to:

$$1 - \sum_{k=0}^{n} P_n(k).R_n(k) = 4.2\%$$

If we increase the mortality rate to a rate of 52.6% then the expected reduction in performance is 4.7%. In contrast the same analysis for the mutual fund industry reveals a performance reduction of only 0.88%. This fact alone highlights the importance of considering survival characteristics in the selection of managed futures programs. Ignoring survival characteristics is expected to cost investors an average of 4.2 to 4.7% per year.

III.B. **Surviving vs. Non-surviving Programs: How different are they?**

Let's now turn to the second question: Do programs that survive over time have more desirable performance characteristics than those who do not ? We have already shown that in randomly selected portfolios ignoring survival chances can have very important adverse effects on annual returns. The results of the study show that the whole return distribution characteristics of surviving and non-surviving programs is different. The results of the comparison of the 'edge' variables are presented in Table 8.

Table 8
Average Values for the 'Edge' Variables for Surviving and Non-surviving Programs

'Edge' Variable	Surviving Programs	Non-Surviving Programs
Annualized Monthly Compounded Return %	18.48%	13.91%**
Maximum Monthly Return %	30.80%	34.70%
Annual Sharpe Ratio	0.49	0.25**
Skewness of Monthly Returns	0.864	0.931
Kurtosis of Monthly Returns	4.428	3.429**
Average monthly winning return %	6.62	6.71
Percentage of winning months %	59.9	59.7

**Significantly different at the 5% level

Several patterns stand out from the results in Table 8. Programs that have survived over time have been able to yield annualized monthly compounded returns that are on average 33% higher than unsuccessful programs. This difference is statistically as well as economically significant. More importantly. not only were surviving programs able to generate larger returns for their clients but they did so with lower risk. The average annual Sharpe ratio for surviving programs was 96% larger than the one for unsuccessful programs. This difference was also statistically significant. Surviving programs have also tended to generate returns that are not as tightly packed around the mean compared to unsuccessful ones as shown by the higher average kurtosis of their return distribution. The average maximum monthly return that both types of programs have been able to generate over time has been the same. Finally, the average monthly winning return, and the percentage of winning months is statistically the same for both surviving and non-surviving programs.

Clearly, the performance of surviving programs measured by their 'edge' variables has been better than their non-surviving counterparts. These results were calculated for all surviving and non-surviving programs from 1975 to 1995. The corollary of this finding is that if one were able to select programs based on their chances of survival one could improve the 'edge' characteristics of the resulting selection. That means that if we were to select programs based on their chances of survival, we would improve not only the absolute returns of the program but its return/risk tradeoff. What most performance studies have done is to select programs with better past ‘edge’ characteristics or performance. Unfortunately, this criterion does not guarantee the selection of programs with the highest chances of survival. As an example, the annualized monthly compounded return of the 90^{th} percentile in the ***non-surviving group was 37.20%! To put it in a nutshell: survival will guarantee better performance BUT better performance does not guarantee survival!***

An important observation at this point is that the large performance differences cannot be explained by the percentage of winning months and the average monthly winning return. Why is this important? The percentage of winning months is highly associated with the type of trader and/or trading system. The average monthly win gives you an idea of how much profit traders can extract from a profitable trade. Short-term traders have lower average wins; medium to long term traders will have the largest average wins. Within each category there is considerable variation. Table 9 contains the averages for the percentage of winning months, the average monthly wins and losses for different types of traders.

Table 9
Edge and Money Management Variables in a Nutshell
1980-1994

Type of Trader	% Wins	Average Win	Average Loss	Win/Loss Ratio
Arbitrage	68%	3.46%	2.88%	1.20
Non-Classified	59%	7.10%	5.25%	1.35
Discretionary	59%	6.40%	4.45%	1.44
Systematic	57%	6.77%	5.18%	1.31
Turtles	58%	9.94%	6.10%	1.63

As expected 'Arbitrage' traders have the highest percent of winning months and the lowest average monthly wins. Systematic traders are defined by Barclay as traders where 95% or more of the decision making process is systematized and are very often 'trend following' traders. The average percentage of monthly wins is the lowest of all categories since they are typically medium to long term traders. 'Turtles' are known to be systematic (trend following) traders. The last row in Table 9 contains the average numbers for 'Turtles' that are classified as systematic traders by Barclay. Clearly, their percentage of winning months is not substantially different than the average systematic trader indicating that these Turtles may not have above average skills for picking profitable trades. The win/loss ratio suggests that the source of their performance is likely to be the dynamic use of leverage and leverage based money management.

Since on average surviving and non-surviving traders do not differ in their percentage of monthly wins and average monthly winning return, a large proportion of their differences could potentially be explained by their money management. Table 10 contains the results of this comparison. The comparison of money management variables between surviving and non-surviving programs reveals important differences that are both statistically and economically significant. Surviving programs average monthly returns are 11% less variable than those of their unsuccessful counterparts. Moreover, surviving programs have been able to experience lower average maximum drawdowns, and they have recovered from drawdowns faster and more consistently than their non-surviving counterparts.

Table 10

Average Values for Money Management Variables for Surviving and Non-surviving Programs

Money Management Var.	Surviving Programs	Non-Surviving Programs
Standard Deviation of Returns	28.00%	31.53%**
Maximum Monthly Drawdown	15.30%	17.80%**
Months Recovering from worst Drawdown	19	22**
Variability in Drawdown Recovery in months	6	7**
Time Recovering from worst Drawdown as % of program life	0.31	0.41**
Margin to equity ratio (Leverage)	19.83%	20.58%
Average losing monthly return	-4.60%	-5.33%

**Significantly different at the 5% level.

No significant differences in the use of (reported) leverage was detected between both types of CTAs or programs[4]. Unsuccessful programs have spent a larger proportion of their business life recovering from their worst drawdown. A very important difference between surviving and non-surviving programs is in their average losing monthly returns. The average losing monthly return for the surviving programs was -4.60%. Programs that did not survive had an average losing monthly return of -5.33%. This difference is not only statistically significant but economically very large. In fact, this difference can account for a large proportion of the differences between surviving and non-surviving programs. Why? Suppose for the moment that the monthly return process of any arbitrary fund or CTA can be approximately modeled by the following binomial process:

$$E_t(R_{t+1}) = p(\overline{w}) + (1-p)(\bar{l}) \qquad (1)$$

[4] Reported leverage in terms of margin to equity is a very poor proxy for actual leverage. Two programs that report the same margin to equity could have very different leverage levels.

$$Var(R_{t+1}) = p(1-p)(\overline{w}-\overline{l})^2 \qquad (2)$$

where:

$E_t(R_{t+1})$: is the expected return on the fund in the next month. R_{t+1} = ln($Vami_{t+1}/Vami_t$).

P: is the probability that the fund will have a winning month.

$\overline{w}$: is the average return on winning months.

$\overline{l}$: is the average return on losing months.

$Var(R_{t+1})$: is the variance of monthly returns.

If this is a reasonable description of any fund or CTA's return generating process, then the annual expected return and the variance of annual returns can be calculated using the following formulae:

$$E_t(R_{t+1}) = (p(\overline{w}) + (1-p)(\overline{l})) * 12 \qquad (3)$$

$$Var(R_{t+1}) = p(1-p)(\overline{w}-\overline{l})^2 * 12 \qquad (4)$$

We assume that reasonable estimates of $p, \overline{w}, \overline{l}$ can be obtained with the historical data for any fund or CTA. In comparing surviving and non-surviving programs we shall use the estimates in Tables 8 and 10. Since equation (1) estimates the expected monthly return, it can also be used to estimate the expected amount of time that it would take a program or CTA to recover from a given drawdown. The expected time to recover (ttr) from a given drawdown DD would be:

$$ttr = \frac{E_t(R_{t+1})}{DD} \qquad (5)$$

Table 11 contains the estimates of $p, \overline{w}, \bar{l}$ for both surviving and non-surviving programs, as well as the expected annual returns, and the expected annual standard deviation of returns for both types of programs.

Table 11

Estimated Performance Variables based on Binomial Process for CTAs' Returns. Parameter Values Estimated with 1975-1995 Data.

Statistic	*Surviving*	*Non-Surviving*
P	0.599	0.597
$\overline{w}$	6.62%	6.71%
$\bar{l}$	-4.60%	-5.33%
$E_t(R_{t+1})$	25.44%	22.29
$\sqrt{Var(R_{t+1})}$	19.05	20.46
Sharpe Ratio	1.07	0.85
ttr	7	9.6

The results in Table 11 demonstrate that the assumption of a binomial process for CTAs' returns is consistent with the actual relative performance between survivors and non-survivors:

1. Annual returns on surviving programs are expected to be 14% greater than returns on their non-surviving counterparts.
2. Surviving programs are expected to be 7% less volatile than their non-surviving counterparts.
3. The Sharpe ratio of surviving programs is expected to be 26% larger than for non-surviving programs.
4. Surviving programs are expected to recover from drawdowns in 73% of the time that it takes non-surviving programs.

These differences are very consistent with the actual differences between surviving and non-surviving programs and suggest that the binomial assumption may be a reasonable one. Further, these results suggest that the binomial model could be used to investigate other relative performance issues like: the effect of leverage on a program's performance results, specific differences between discretionary and systematic traders, sources of performance among different trading systems, etc.

III.C. **Using the Binomial Model to understand the Source of Performance Differences**

The binomial model can also be used to understand the sources of performance between different type of traders. Table 9 contains the average percent wins and loses, and the average monthly win and loss for five types of traders or systems. The numbers for the 'Turtles' were calculated only for those Turtles who are classified by Barclay as systematic traders. To illustrate the use of the binomial model, we shall focus on systematic traders and compare the average systematic trader to the average systematic turtle. The typical systematic trader is not very different than the systematic turtle in that their percentage of monthly profitable trades is around 57-58%. The major difference between them lies in the profitability of profitable trades and their control of losses. How would this translate in expected measures of performance? Table 12 illustrates the output of the binomial model. Annual returns for the systematic Turtle are expected to be around double the returns of a systematic trader. These returns are expected to be achieved with only 34% more volatility. The expected Sharpe ratio of a systematic Turtle is almost double that of a typical systematic trader, and they are expected to recover from a 20% drawdown in half the time it takes an average systematic trader.

Table 12
Expected Measures of Performance for the Average Systematic Trader and the Average Systematic Turtle

Performance Measure	*Avg. Systematic*	*2X Avg. Systematic*	*1.6X Avg. Systematic*	*Systematic Turtle*
E. Annual Return	19.58%	39.10%	37.76%	38.44%
E. Standard Deviation	20.50%	41%	30.65%	27.40%
E. Sharpe Ratio	0.71	0.83	1.07%	1.22
E. Months to recover from 20% Drawdown	12.25	6.14	6.35	6.25

It is clear from the numbers in Table 9 that the average systematic Turtle does not have above average skills in picking profitable trades. The size of the average winning trade however is substantially larger than the one of an average systematic trader. There are only two possible complementary explanations for this fact: leverage and the ability to participate in every single profitable market. If leverage is one of the factors, one question that can be answered using the binomial model is: how much more leverage does a typical systematic trader need to look more like a systematic Turtle? The results of such an exercise are contained in Table 12 under the column '2x Avg. Systematic'. Using the binomial model, this column was generated by simply doubling the leverage of the typical systematic trader. Two important observations can be made. Doubling the leverage of a typical systematic trader gets us returns that are comparable to those of the typical systematic Turtle. Notice that since we have not made any changes in the loss controls of the typical systematic trader, these returns are expected to be achieved with volatility levels that are 50% higher than the volatility of a systematic Turtle. Suppose that we force the typical systematic trader to reduce their average losses to a level comparable to what a typical discretionary trader will experience (-4.40% instead of -5.18%). This reduction will likely be achieved less with money stops and more through de-leveraging during drawdown periods. This dynamic use of leverage will likely reduce the overall leverage level and achieve the loss control result. The results of performing this exercise using the binomial model (replacing -5.18% with -4.40% and lowering the leverage amount to 1.6x) are shown in Table 12 under the heading '1.6x Avg. Systematic'. The results in Table 12 confirm the hypothesis that the right combination of leverage and loss control can transform the typical systematic trader into a systematic Turtle.

III.D. **Performance and Survival of Diversified vs. Non-Diversified Programs**

In this section of the paper we examine the issue of performance and survival as it is related to the type of program or CTA. Specifically we compare performance and survival experience between *diversified* and *non-diversified* (specialists) programs. The objective of the comparison is to provide some factual information on the relationship (if any) between reported diversification, performance and survival experience. Table 13 contains the results of a comparison between diversified and non-diversified programs. The results contained in Table 13 suggest that, on average, there is no significant performance advantage for diversified programs over non-diversified ones. What is most striking is that the results suggest diversified programs have substantially higher volatility than non-diversified programs. This difference persists even when we compare the median annual standard deviation instead of the average. The median standard deviation for diversified programs is 29.13%, while it is only 21.75% for non-diversified programs. More interestingly, it appears that a program that calls itself ‘diversified’ does not have any better chances of surviving than a program that does not. These findings are important in view of the popular belief that a diversified program is more likely to have lower volatility and higher chances of long-term survival. The results presented so far indicate that neither one of these two popular beliefs can be substantiated by the data. The classification of a program as ‘diversified’ or ‘non-diversified’ does not appear to provide any meaningful information regarding the potential future volatility of the program or its long-term survival prospects. One potential criticism of these results is that their calculation may include programs with very short track records. With this criticism in mind, we constructed a data-set that includes programs and/or CTAs that have track records greater than three years. The results of the same analysis on this subset of data is contained in Table 14.

Table 13

Average Values for 'Edge' and Money Management Variables for Diversified and Non-diversified Programs, 1975-1995.

Variables	*Diversified - N=340*	*Non-Diversified- N=585*
Edge Variables		
Annualized Monthly Comp. Return	16.58%	16.19%
Average Maximum Monthly Return	38.74%	29.13%**
Average Annual Sharpe Ratio	0.32	0.42
Average monthly winning return	7.66%	6.08%**
Percentage of Winning months	58%	61%**
Money Management Variables		
Average Standard Deviation of Returns	33.90%	27.21%**
Average Maximum Monthly Drawdown	19.60%	14.60%**
Months Recovering from worst drawdown	8	6**
Average Losing monthly return	-5.67%	-4.53%**
Mortality Rate %	46.7%	47.2%

**Significantly different at the 5% level.

Table 14

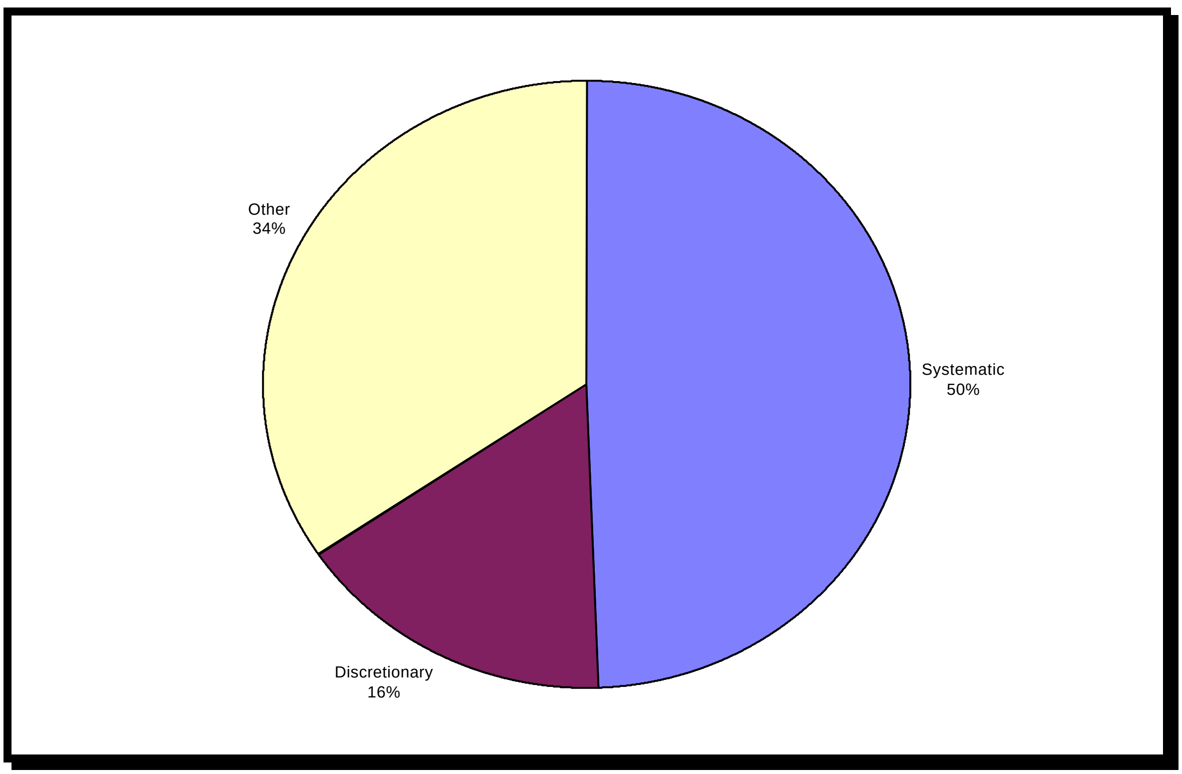
Average Values for 'Edge' and Money Management Variables for Diversified and Non-diversified Programs with records greater than three years, 1975-1995.

Variables	*Diversified - N=257*	*Non-Diversified- N=352*
Edge Variables		
Annualized Monthly Comp. Return	17.68%	17.24%
Average Maximum Monthly Return	45.42%	36.23%**
Average Annual Sharpe Ratio	0.32	0.48**
Average monthly winning return	8.28%	6.42%**
Percentage of Winning months	57%	61%**
Money Management Variables		
Average Standard Deviation of Returns	37.00%	29.60%**
Average Maximum Monthly Drawdown	22.10%	16.83%**
Months Recovering from worst drawdown	9	7**
Average Losing monthly return	-5.93%	4.63%**
Mortality Rate %	47.5%	44.3%

**Significantly different at the 5% level.

Interestingly, the conclusions drawn from these results are identical to those drawn from the results in Table 13. In both samples, diversified programs have substantially higher volatility than non-diversified programs. This higher volatility is largely attributable to the lower percentage of monthly wins and higher average win/loss spread for these programs. The relatively low average percentage of winning months of 57% suggests that the proportion of systematic traders in this category may be high. Figure 1 contains the breakdown of the diversified sample into systematic, discretionary and other programs.

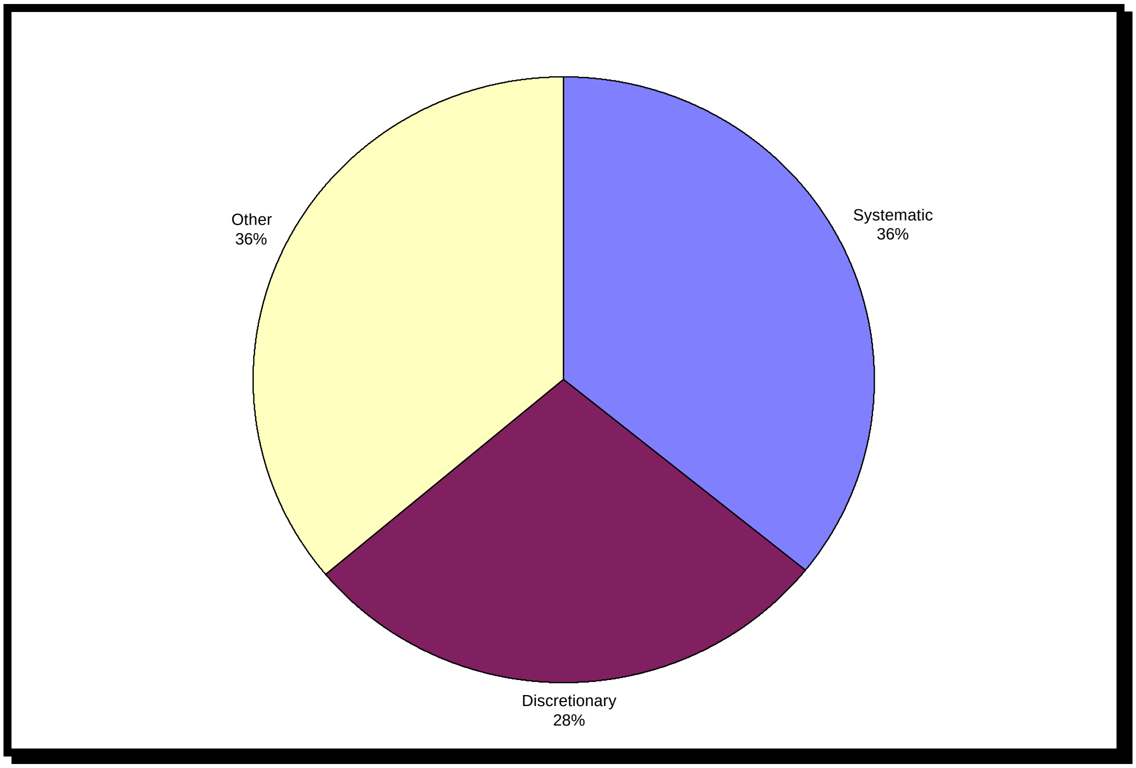
Figure 1
Composition of the Diversified Sample



As expected almost fifty percent of the sample consists of Systematic programs, and only sixteen percent is represented by Discretionary traders. Thirty four percent of the sample consists of programs that are neither Systematic nor Discretionary. Figure 2 contains the breakdown of the Non-Diversified sample into the same categories. Unlike the breakdown of the Diversified sample, the Non-Diversified sample is roughly evenly split between Systematic, Discretionary and Other.

Figure 2

Composition of the Non-Diversified Sample



III.E. **Performance and Survival of Systematic vs. Discretionary Programs**

The results presented in the previous section suggest that it may be the type of trading system and not the degree of diversification what may discriminate among programs' performance and survival characteristics. With this idea in mind we classified programs or CTAs into systematic and discretionary and we compared the average 'edge' and money management variables as well as the mortality rate in each category. The results of this comparison are contained in Table 15.

Table 15

Average Values for 'Edge' and Money Management Variables for Discretionary and Systematic Programs, 1975-1995.

Variables	*Discretionary - N=209*	*Systematic- N=364*
Edge Variables		
Annualized Monthly Comp. Return	16.19%	14.78%
Average Maximum Monthly Return	32.95%	31.83%
Average Annual Sharpe Ratio	0.33	0.31
Average monthly winning return	6.40%	6.77%
Percentage of Winning months	60.3%	57.4%**
Money Management Variables		
Average Standard Deviation of Returns	28.60%	29.73%
Average Maximum Monthly Drawdown	15.42%	17.07%
Months Recovering from worst drawdown	7.5	7.5
Average Losing monthly return	-4.46%	-5.18%**
Mortality Rate %	45.5%	36.8%**

**Significantly different at the 5% level.

Several distinctive results emerge from the results in Table 15. First, Discretionary programs have had higher mortality rates than their systematic counterparts. The mortality rate for discretionary programs is 45.5% while it is only 36.8% for Systematic programs. Second, the variability in the performance of Discretionary programs has been much larger than that found for Systematic traders even though the average program standard deviation[5] has been comparable to that found in Systematic programs. The standard deviation of the annual compounded return for Diversified traders was 20% as opposed to only 15% for systematic traders.[6] This difference accounts for the fact that, statistically speaking, discretionary programs have not yielded larger compounded returns than systematic programs. The previous two results suggest that in practice it may be more difficult to select discretionary traders than systematic traders. Systematic traders or programs appear to be more homogeneous than discretionary ones.

[5] These are the numbers in Tables 13-16.

[6] These statistics should not be confused with the 'Average Standard Deviation of Returns' in Tables 13-16. The numbers in the Tables measure the average volatility of a program, while the standard deviation of the annual compounded return measures the variability of yields for a particular group of programs.

We conducted the same analysis on programs that had track records greater than three years to avoid the potential criticism that programs with short track records might bias the analysis. The results of such analysis are contained in Table 16.

Table 16

Average Values for 'Edge' and Money Management Variables for Discretionary and Systematic Programs with records greater than three years, 1975-1995.

Variables	*Discretionary - N=131*	*Systematic- N=247*
Edge Variables		
Annualized Monthly Comp. Return	19.06%	16.00%
Average Maximum Monthly Return	40.36%	39.57%
Average Annual Sharpe Ratio	0.48	0.31**
Average monthly winning return	6.76%	7.68%
Percentage of Winning months	60.5%	56.6%**
Money Management Variables		
Average Standard Deviation of Returns	30.98%	33.81%
Average Maximum Monthly Drawdown	17.92%	20.11%
Months Recovering from worst drawdown	8	8
Average Losing monthly return	-4.51%	-5.67%**
Mortality Rate %	41.2%	36.4%**

**Significantly different at the 5% level.

The conclusions from the previous analysis are not changed by reducing the sample to only those programs that have track records of more than three years.

III.F. **Summary of Findings on Program Classification**

The results presented so far suggest that:

a. Programs classified as diversified did not have any less volatility than those that were classified as non-diversified. In fact, diversified programs had on average greater volatility than those classified as non-diversified. This difference was statistically significant and economically meaningful.
b. The classification of programs into 'diversified' and 'non-diversified' did not have any power discriminating between program with different chances

of survival. On average, the mortality rate of Diversified programs was the same as the mortality rate of Non-Diversified programs.

c. The classification of programs into ‘Systematic’ and ‘Discretionary’ did not have any power discriminating between programs with different annual compounded rates of return and/or standard deviation of returns. Discretionary programs are a more heterogeneous population than their systematic counterparts as evidenced by the larger dispersion of the population performance.
d. The classification of programs into ‘Systematic’ and ‘Discretionary’ had power discriminating between programs with different chances of survival. On average, Systematic programs had mortality rates around 37% while Discretionary programs had mortality rates around 46%.
e. The conclusions that emerge from the comparisons between Diversified vs. Non-Diversified, and Systematic vs. Discretionary do not change when shorter track records (less than three years) are removed from the analysis.
f. The percentage of winning months seems to be very useful in discriminating between Systematic and discretionary programs. Systematic traders are characterized by a percentage of winning months around 57% while Discretionary traders are well characterized by a percentage around 61%.

III.G. **Predicting Survival**

The results presented so far clearly show that there are important differences between surviving and non-surviving programs. They show that if one were able to discriminate between programs with good chances of survival and programs with poor chances of survival, the overall performance of a pool could be improved. In contrast, the usual classification of traders and/or programs into Diversified vs. Non-diversified, or Systematic vs. Discretionary did not provide with a powerful means of discriminating between good performers or bad performers. How can we tell which programs have better chances of surviving in the future? In order to do this it is necessary to know how particular combinations of return characteristics (edge and money management) can affect the program's survival chances. For example, program A can have a higher monthly return than program B but may recover from drawdowns in a less consistent manner than B. What are the comparative chances of survival of A and B? The information in Tables 8 and 10 or Tables 13 to 16 cannot help answer these kind of questions. To help answer these questions we constructed a model of survival behavior that incorporates 'edge' and money management variables as well as variables related to the type of program/trader (i.e. discretionary, systematic, etc.). The variables included in the survival model are listed in Table 17. The construction of a survival model had several objectives. First, we wanted to identify which variables affect the probability of survival. Second, we wanted to know how well survival experience could be explained with a survival model. Finally, we wanted to know whether survival is a predictable characteristic. Since the estimation of a reliable model of survival behavior requires a substantial amount of data we divided the entire Barclay database into two sub-samples. The first sub-sample was used for estimation and the second for validation and testing of predictability. The model was estimated on data for programs that traded between January 1975 and December 1992. This sub-sample contains data on programs that started trading at different times up to December 1991; thus only programs with one year or more of trading experience were included.[7] This sub-sample contained 665 programs of which only 519 were in business at the end of 1992. Table 17 also contains information about whether each variable had a statistical relationship with survival experience. The annualized monthly compounded return did not have a significant relationship with survival experience. This finding is consistent with the general rule that says that survival is related to performance but performance may not be related to survival. The maximum monthly return achieved by a program or trader, its Sharpe ratio, and the average monthly winning return are found to have positive impact

[7] We included programs with less of a three years track record based on the results in Section 5 that suggest that this population is very unlikely to bias the results.

Table 17

List of Variables used in the construction of a Survival Model

VARIABLE #	DESCRIPTION	SIGNIFICANT?	P-VALUE*
1	Annualized Monthly Compounded Return	NO	0.7203
2	Maximum Monthly Return	YES	0.0001
3	Sharpe Ratio	YES	0.0662
4	Average Monthly Winning Return	YES	0.0001
5	Percentage of Winning Months	NO	0.4778
6	Management Fee	YES	0.0157
7	Incentive Fee	YES	0.0008
8	Excess Skewness of Monthly Returns	NO	0.6480
9	Excess Kurtosis of Monthly Returns	NO	0.3439
10	Standard Deviation of Monthly Returns	YES	0.0001
11	Maximum Monthly Drawdown	YES	0.0001
12	Maximum time to Recover from Drawdown as a percent of business life	YES	0.0001
13	Average Losing Monthly Return	NO	0.2645
14	Maximum Number of Months to Recover from a drawdown	YES	0.0001
15	Standard Deviation of the times to recover from a drawdown	NO	0.1535
16	Arbitrage Dummy (for Arbitrage programs)* *	NO	0.7712
17	Systematic Dummy (for Systematic programs)**	YES	0.0403
18	Other Dummy (for non classified programs)**	YES	0.0535
19	Diversified Dummy ***	NO	0.3760

* P-value is that probability of drawing the obtained coefficient from a distribution with zero expected value.
**The coefficient on these dummies is relative to discretionary programs.
*** The coefficient on this dummy is relative to non-diversified programs.

on the program's probability of survival. Management fees are negatively related to survival, but incentive fees are positively related to survival experience. The standard deviation of monthly returns, the maximum monthly drawdown, and the maximum time to recover from a drawdown as a percentage of a program's life are all negatively related to a program's chances of survival. The type of trader or program has an effect on survival as evidenced by significant coefficients for systematic traders and Other traders. Systematic traders are found to have greater probability of survival than discretionary traders. 'Other'[8] traders are found to have lower probability of survival than discretionary traders. Arbitrage traders have comparable survival probability to discretionary traders. Finally, as expected from the previous section's descriptive analysis, Diversified traders were found to have not significantly different survival experience than non-diversified ones.

The results of the model estimation were used to construct a ***survival index.*** This index was used to rank programs by their chances of survival into three categories: low, medium, and high chances of future survival. To gain an appreciation of how well the model explained 'in sample' survival experience, we ranked the non-survivors based on their survival index. The results of this classification are contained in Table 18. Seventy percent of the non-survivors came from the Low survival rank. When we add the number

Table 18
Programs that failed during the 1975-1992 period ranked by their Survival Index

Survival Index Rank (Chances of Survival)	Number of programs that failed	Percent of Total Non-Survivors
Low	102	70%
Medium	34	23%
High	10	7%

[8] Other traders are traders or programs that were not classified as either Systematic or Discretionary in the Barclay database.

of programs in the Low and Medium survival ranks we were able to account for ninety three percent of the non surviving programs. These results clearly demonstrate that the model was quite successful in discriminating between the survivors and non-survivors.

A more relevant question from the practical point of view is whether we can rely on survival rankings to forecast *future* survival experience. With this question in mind we ranked all programs that were trading on December 1992 based on their survival indices. We then followed these programs through December 1995 and kept track of the number of programs that did not survive together with the rank they came from. If the rankings had no predictive power we would expect that on average, only thirty three percent of the non survivors would come from each of the survival ranks. If on the other hand, the rankings had predicting power, we would expect the largest percentage of non-survivors coming from the Low rank and the lowest percentage of non-survivors from the High rank. The results in Table 19 demonstrate that the survival rankings have predictive power. Eighty three percent of the programs that did not survive came from the Low and Medium survival ranks. Only seventeen percent of the non-survivors came from the High survival rank.

Table 19
Programs that were trading on December 1992 but failed during the January 1993-December 1995 period ranked by their Survival Index

Survival Index Rank (Chances of Survival)	Number of programs that failed	Percent of Total Non-Survivors
Low	66	45%
Medium	56	38%
High	25	17%

One potential criticism of these rankings is that they could be biased if the Low and Medium categories contain a larger number of programs relative to the High category. To illustrate this, suppose that the model does not predict better than a random 33% for each category. Suppose also that of the 519 programs 220 were classified as low, 190 as medium and 109 as high. Under this scenario, you would expect to find 73 (220x0.33) non-surviving programs coming from the low group, 63 (190x0.33) from the medium group and 36 (109x0.33) from the high group. The percentages of non-survivors coming from each group would then be: 42.4% for Low, 36.6% from medium and 21% from high. The results in Table 20 show that the results of this study do not contain this bias. If anything, the bias would be towards finding ***more*** non-survivors in the High category since there are slightly more programs in this category relative to the other two. Table 20 also contains the actual mortality rates in each of the groups from Jan-93 to Dec-95. If the survival rankings had predictive power, one would expect to find the highest mortality rates in the low group and the smallest mortality rates in the high group. The results in Table 20 further demonstrate that the survival rankings had predictive power. The highest mortality rate was experienced in the Low group (38%) followed by the medium group (35%) and the high group (14%). Notice that the high group experienced mortality rates

Table 20
Mortality Rates and Number of Programs in each Survival Category for Jan-93 to Dec-95

Survival Index Rank	Total Number of Programs as of Jan-93	Programs that survived through Dec-95	Programs that did not survive through Dec-95	Mortality Rate
Low	175	109	66	38%
Medium	160	104	56	35%
High	184	159	25	14%
Total	519	372	147	28%

that are less than half the size of the other two groups and comparable to the mortality rates experienced in the Mutual Fund industry as show in Table 5.

IV. **Summary**

This article demonstrates the importance of considering survival issues when selecting managed futures programs. Mortality in managed futures has been historically very high relative to mortality in the mutual fund industry. Ignoring survival issues in the selection of managed futures programs is likely to translate into reduced performance of the final

program. The empirical evidence suggests that survival is associated with performance but that performance is not necessarily associated with survival.
We develop a simple model to understand the sources of performance in managed futures. Results from the binomial model are consistent with observed results, suggesting that this model could be used to compare the potential performance of different programs as well as to understand the sources of performance of a program or trader.

We find that the classification of programs into ‘Diversified’ and ‘Non-Diversified’ does not provide any useful information regarding the likelihood of survival. Surprisingly, ‘Diversified’ programs are significantly more volatile than ‘Non-Diversified’ without any corresponding larger returns. The classification of programs and/or traders into ‘Systematic’ and ‘Discretionary’ is found to be associated with their likelihood of survival. Systematic programs/traders are found to have higher chances of survival that Discretionary ones.

We show that a model of survival behavior can explain a substantial proportion of past survival experience as well as ***future survival experience.*** Return distribution characteristics associated with the ‘edge’ of a program and its money management can have important effects on future survival and performance. The finding that survival experience can be reasonably explained and forecasted has important practical implications for the formation of managed futures portfolios.

References

Malkiel, Burton G., 1995, ‘Returns from Investing in Equity Mutual Funds 1971 to 1991’, *The Journal of Finance, Vol 50, No. 2, June 1995.*