Alternate Histories

A Visual Comparison of Randomness

Core Insight Story™ by John Jonelis

Pre-production draft published in the Journal of Technical Analysis

Purpose

When evaluating a rule-based trading strategy, a backtest can demonstrate what *would have happened* over a slice of history, but not the many things that *could have happened* if events had unfolded differently. These are the alternate histories. They foreshadow how far actual trade results might stray from expectations. Unfortunately, this kind of analysis has been obscure and inaccessible—bogged down in the rigors of statistics or the arcane workings of the Monte Carlo Simulation. But for traders accustomed to pattern-recognition, a simple set of visual cues can transform these disciplines into intuitive tools.

A Trader's Story

To gain an appreciation of the need, let's put ourselves in the shoes of a a green trader. Consider the following hypothetical situation:

You have $10,000 set aside and want to try your hand at active trading. You attend a session put on by Tremendous Trading Seminars LLC, for which you pay $1,000, plus $500 for travel, hotel, and meals. The speaker, Mr. X, is a dynamic fellow who has been extremely successful day-trading Fish Head Futures, an important commodity in the pet food processing industry. He started in 1995 with $10 thousand in capital, just like you, and in a few years had parlayed it into $10 million. He's laid out his system test and all his trading rules in a flashy 150 page manual that comes with special software and two weeks free access to an exclusive Internet chat room—all for the one-time-only special low price of $4,000. It's called System X. Naturally, you buy it.

Your package arrives overnight express. You install the software and read the manual cover-to-cover three times. A couple weeks later, you've completed all the exercises and feel you have a handle on the system. The thing can't miss—you call your boss and quit your job.

You log onto the System X Internet chat room, register for the trial subscription, and start paper trading. It goes well and you find the interaction with other members exhilarating. After a week, it's clear you could be earning some serious cash, and you regret wasting your time trading Monopoly™ money. That profit could have been real, and after all, you won't be getting a paycheck this Friday. So you take the plunge. You start trading real money.

Monday you make $1,000 and feel like a king. Three more days like this and you will have paid for the course. You call your mom and tell her she can stop worrying—you're finally a success. You email all your friends.

Tuesday, you lose $600. It's a choppy trendless day and everybody in the chat room is down. Wednesday, you lose $200 more, but you're actually profitable except for those pesky commissions, and of course, there's that really huge trade you missed because your IP went down momentarily. Thursday, one of the really good traders is absent from the chat room (he's at the dentist) and you lose $1,600 due to whipsaw market action—but that could happen to anybody. On Friday, you make $400, ending the week on a positive note. An email arrives announcing the end of your free chat room period and you promptly subscribe for the monthly fee of $500.

That weekend, your wife asks for an accounting, so you add it all up. You are down $7,000. You have $3,000 left.

You write a nasty email to Mr. X, attaching your trades. Mr. X points out that you are really down only $1,000 in trade revenue. The rest is just normal business expense. You need to commit more money. Then he asks why you failed to take that big trade on Wednesday, which could have netted $6,000. He admonishes you to hone your skills and success will surely follow.

On Monday, you log onto the chat room, sit on your hands, and watch. Sixty members are in the room when a big trade unfolds and many post enthusiastic comments about how much money they're making. Completely frustrated, you admit your failure to the group, and ask how everybody did the previous week. Six respond. Of those, all had their orders filled on that big Wednesday trade, but each got in and out at different points. For the week, one of them reported a net total of $5,000, as he should have, and another, $5,500. But one—a very experienced and savvy trader—made $16,000. Strangely, another regular lost $2,500, which he claims was his worst week on record and nobody can understand how it happened. Another lost $3,000, and the last response moans about a $3,500 loss. So you ask yourself, WHAT'S GOING ON HERE?

Copyrighted Material

You write Mr. X again, asking for his actual trades over the past month. He's a good guy, wants you to succeed, and sends the information. You drop the data into a spreadsheet, normalize his starting equity to $10,000, then generate an equity curve—a simple graph of his account growth over time. (See Figure 1 below.) Mr. X is clearly an accomplished trader.

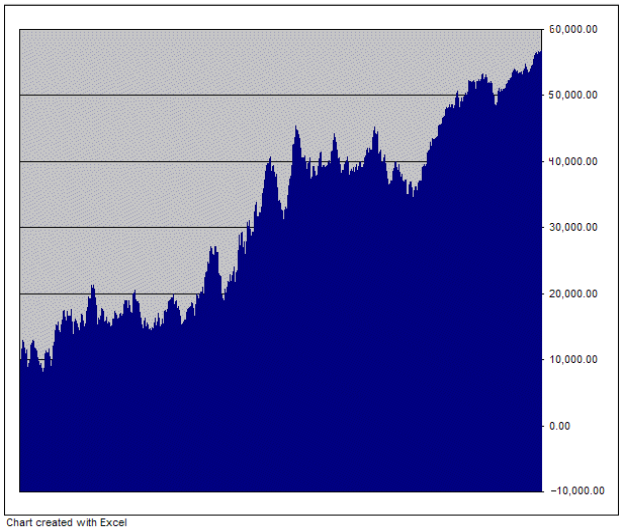

Figure 1—"*Mr. X's*" Equity Curve (1 Month)

Alternate Histories

Let's step away from the story and ask what else might have happened. What if different traders took their vacations at different times or went to the dentist on different days? What if some traders didn't get their orders filled on a few winning trades, while others missed losing trades? What if some chased entries aggressively, while others got out early or late? What if some, by instinct or design, changed position size from trade to trade? What if the commission schedule was different at various brokers, and some experienced excessive slippage? What if some trader's accounts were sufficiently capitalized while others, like yours, were not? Clearly the outcome could not be the same for each trader. In fact, the possibilities are beyond imagining. How can you even begin to collect data on so many variables? You can't. Nevertheless, such variation can be estimated, using a Monte Carlo Engine (MC). MC plots alternate histories—*the many things that might have been*. Figure 2 below shows 100 alternate equity curves for *System X* in our story. Look at the potential variety in performance:

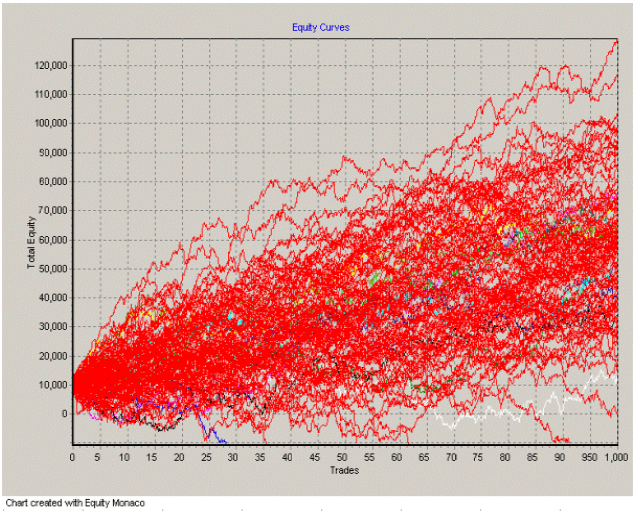

Figure 2—"*System X*" Alternate Equity Curves

To the world, the upper lines represent the wizards, the kings, the skilled ones who really know what they're doing. Note the paths that fall below zero. To the world, those are bums, goats, failures. *However, each path on the graph is generated using Mr. X's own trade results—the very same results shown in Figure 1. They are all based on exactly the same data.* Each curve is plotted in a different way by the MC engine. Taken together, they reveal more about the nature of *System X* than was evident in the simple equity curve.

Copyrighted Material

Mr. X's Dilemma

In our story, Mr. X is frustrated. He confides in you that some of his customers have gone broke. He doesn't understand why. After all, he carefully tested his strategy and then went on to build a personal fortune trading it. He truly believes in his methods. He thinks those losing traders need to practice their entry/exit technique and even suspects that some of them lack the proper psychological makeup to be successful traders.

He may be right on those points, but he is missing the main reason for the variation in performance among his customers. If he had run an MC simulation after his backtest, he would have seen a high degree of randomness inherent in his strategy. It may have given him pause. *He might never have traded his strategy and made all that money*. We now know that because of the variation built into his system, *he could just as easily have gone broke*.

There are a large variety of trading strategies in use today, some commercially available, others written for hedge funds, quant shops or by independent traders for their personal use. Some are rule-based but rely partially on a trader's skill, like *System X*. Others are more methodical. Platforms exist that identify and place trades automatically, entirely separating the human from the process. However, no matter how mechanical, each strategy has a degree of randomness. The question is not, "Is the system is random?" but rather, "How random is the system?" MC answers that question, but until recently, its use has been shrouded in mystery.

Monte Carlo Demystified

In the movie *The Blues Brothers* Jake Blues, facing the muzzle of an assault rifle, blurts out a list of increasingly outrageous and contradictory excuses for jilting his fiancé. "I ran out of gas; I had a flat tire; I didn't have enough money for cab fare; my tux didn't come back from the cleaners; an old friend came in from out of town; somebody stole my car; there was an earthquake—a terrible flood; *locusts.*"

Clearly, an MC engine lacks the imagination to fabricate such scenarios. Instead, it uses a simple method: Net trade results are fed in. A random number generator scrambles the sequence of trades. The process is repeated many times, with replacement, resulting in an array of possible equity lines. Simple. In this way, MC gives a rough measure of what *could* have happened, but did not. It *estimates* the degree of randomness in a system.

Reading Monte Carlo

MC is used throughout industry—an established tool, waiting on the shelf. Until recently, it has been technically obscure and inaccessible to the bulk of the trading community. Some are unaware of it, some afraid of its complexity. Others consider it a "brute force" approach, and believe it inferior to more elegant mathematics. Some object that it treats each outcome as an independent event, assuming no serial dependency. Some rightly point out that it has been used in overly precise ways to reach highly sophisticated and spectacularly wrong conclusions.

Keeping in mind that an MC study is no more than an estimate, I'd like to steer around all the complexity and get at the basic power of the tool. I'll demonstrate a visual approach, painted with a broad brush. I'll keep calculations to a minimum. I will suggest five basic steps for reading MC, and assign descriptive names (rather than technical ones) to help the concepts stick. Each illustration represents actual trade results from well-known, mechanical trading systems. (The system names are omitted.)

Rules of the Game

For our comparisons to be meaningful, we will follow one simple rule: *Any two backtests will be compared across the same amount of time, even if each test produces a different number of trades*. To allow for individual creativity, the rest is left open.

To demonstrate the accessibility of the tool, each MC test is performed using the *Equity Monaco* package from *TickQuest*, which at time of this writing is a free download. For input, *Equity Monaco* uses net dollar result per trade. Packages are available that use percent change, random day of entry and other approaches. Each has its advantages. The same kind of simulations can be run in Excel or another spreadsheet program.

Copyrighted Material

Five Visual Steps

Slope—In Figures 3a and 3b below, the approximate Mean or Expectancy is superimposed on the MC distribution for visual clarity. The slope of the Mean for one strategy (3a) is flat to slightly negative, while the other (3b) is positive. A positive slope implies that *on average*, traders using that strategy will make money. As can be clearly seen in the MC plots, a positive slope does not indicate that any single trader will be profitable—only the average of a large group. However, it is an important measure. *(In our story, System X had a strongly positive slope. That attribute allowed Mr. X and many others to make money, even though some did not.)*

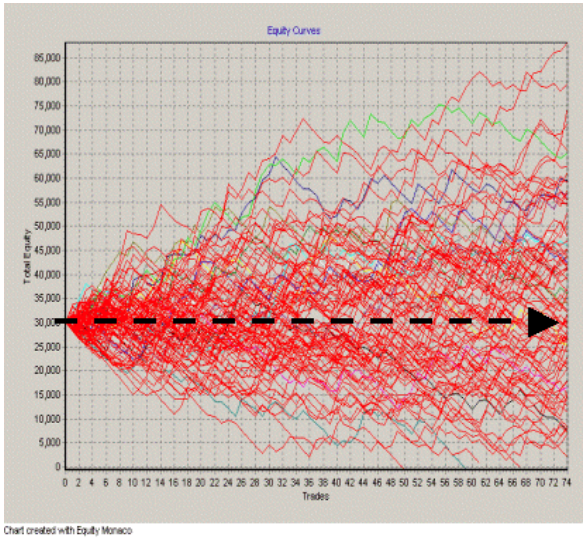

Figure 3a

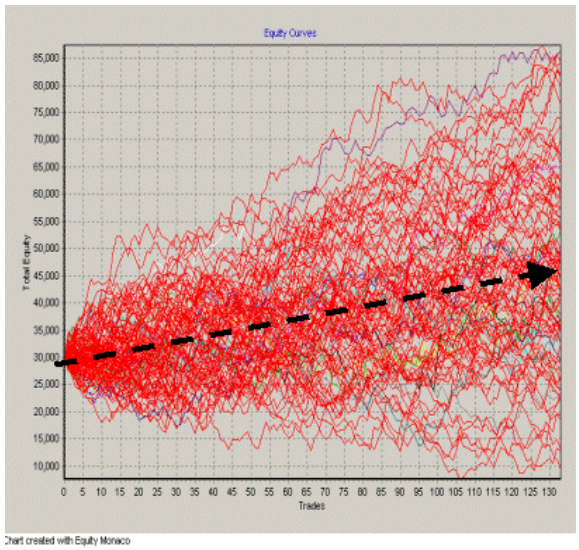

Figure 3b

A numerical value for Slope can be derived by using the original data to calculate the rise over the run. Divide the equity gain **(Δe)** *by calendar days* **(t)** *and compare. If* $(\Delta e/t)_a < (\Delta e/t)_b$ *then "***b***" is better. Because our "rules of the game" dictate that we run both tests across the same time period, the variable "***t***" cancels out of the equation, leaving* $(\Delta e)_a < (\Delta e)_b$ *which can be interpreted directly from the MC plot.*

Spread—Figure 4a and 4b visually compare the random nature of one strategy to that of another. This is quite literally, a picture of risk. A wide spread implies widely diverse outcomes, while the tightly clustered distribution is less random. *(System X's spread is wide, which explains the wild variation in results.)*

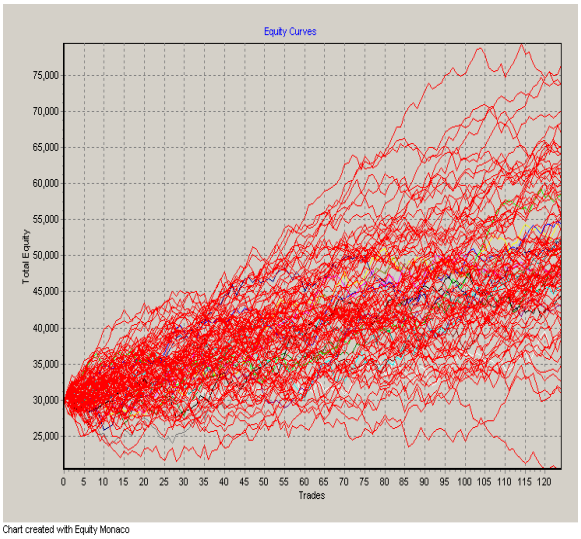

Figure 4a

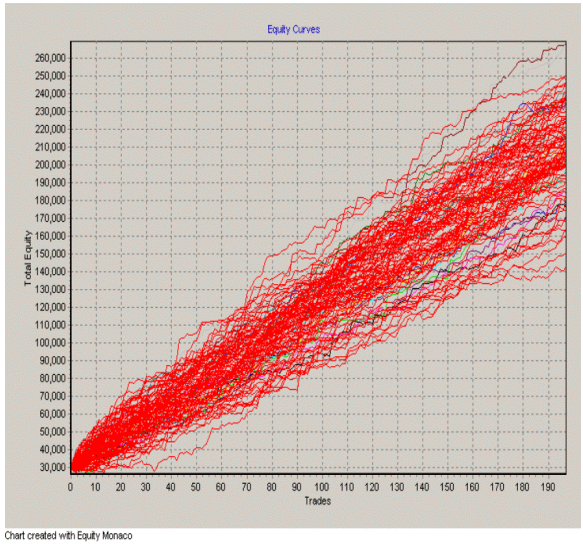

Figure 4b

Copyrighted Material

Spread is also used to determine the minimum initial capital considered safe to trade a given system, by finding an entry level at which no paths lead to ruin. This is visually apparent on the MC display. If a path has led to ruin, re-run the test with a larger initial equity to determine maximum expected drawdown.

Given sufficient iterations, the lines at the extremes will approximate the third standard deviation of final equity from the original data (**3σ**) or ultimately, the largest gain (or loss) multiplied by the total number of trades. These are unlikely but possible events, and perhaps an analogy is in order. *It is not safe to venture across the ocean in a rowboat. On the other hand, an unexpected iceberg can sink even the Titanic.*

> *Note—If the Y-axis for each plot has a similar range, the comparison of Spread can be performed visually. However, there are exceptions. As shown in Figures 4a and 4b, a computer program will auto-scale the Y-axis, and the scale of one chart may be significantly different from the other, making the comparison misleading. Double check using simple arithmetic. Take the difference between the terminal mean value "***e***" and the worst terminal value "***w***" and divide the result by "***e***" as in* **(e-w)/e.** ***A****pproximate numbers from the charts will do. This gives downside risk as a ratio. For "***a***" it's about* **0.7**, *while for "***b***" it's* **0.3**. *The smaller number is better, in this case, "***b***," confirming the visual pattern in Figures 4a and 4b.*
>
> *Statisticians take a different slant on what I have termed Spread, calculating the Relative Efficiency using the variance (***σ^2***) of each set of original data. For example, given the two systems shown (Figures 4a and 4b), calculate the ratio (***σ_b^2/σ_a^2***). If* **eff(a,b) < 1**, *then "***b***" is better, confirming the visual interpretation. This calculation also makes a number of other statistical comparisons possible, but if the Y-axis is similar for each, a visual comparison is more intuitive and compelling.*

Scatter—the figures below display one plot (Figure 5a) with little or no discernable grouping of paths, and another (5b) more weighted to the center. A scattered distribution implies that one outcome is nearly as likely as another, (as in *System X*). There is also a higher probability of outliers beyond the plotted range—a possibility not frequently considered. A clustered grouping, on the other hand, implies a statistical distribution, however skewed, with some outcomes much more likely than others. It is less random. Statisticians refer to this phenomenon as *Kurtosis*.

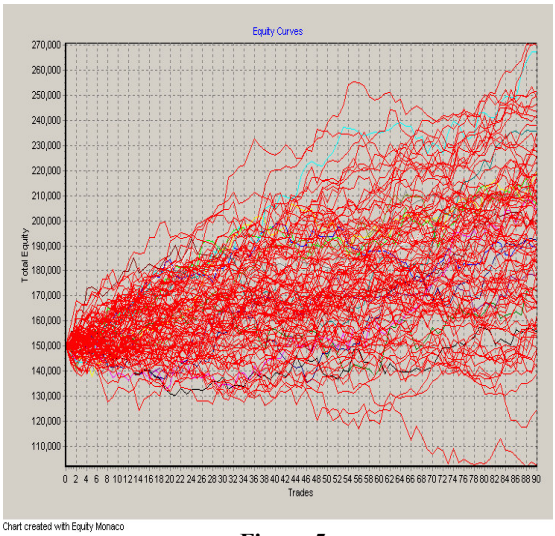

Figure 5a

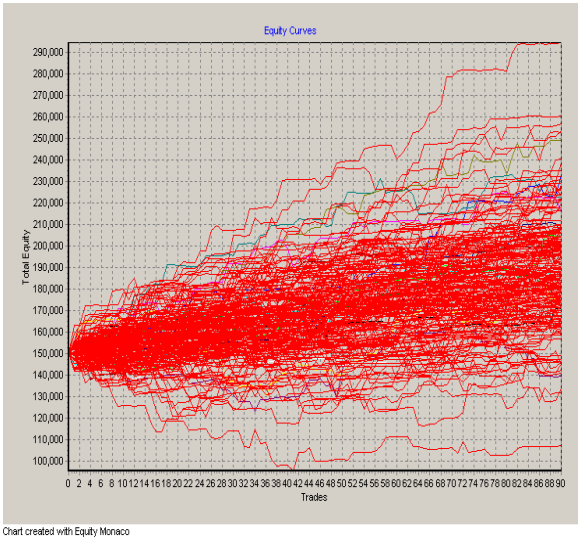

Figure 5b

Copyrighted Material

Streak—A backtest package will report the largest streak of consecutive losing trades *encountered during the test*. MC can go further, estimating the largest streak that *might have occurred*, and then plot the outcomes by percentile. How this information is used depends upon an individual's tolerance for risk. As a rule of thumb, consider the number of consecutive losses posted at the 95th percentile. The strategy in Figure 6a below returned 14 consecutive losses while 6b only 5. Obviously, the fewer consecutive losses, the better.

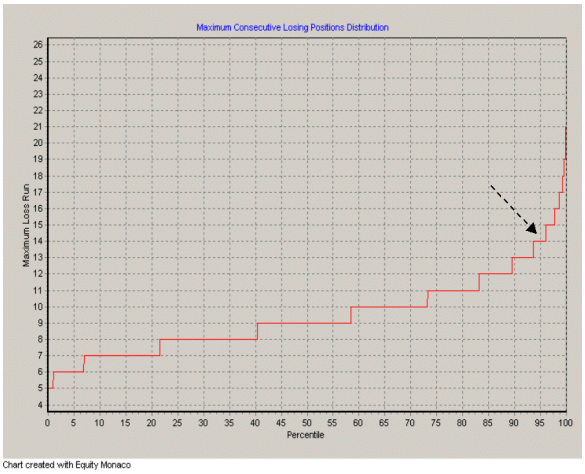

Figure 6a

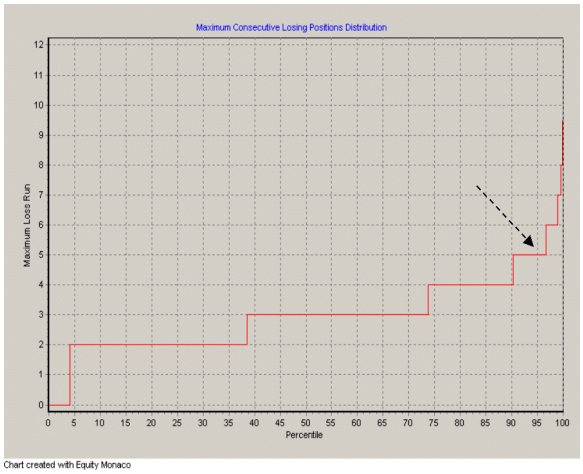

Figure 6b

For a fair comparison, verify that the number of trades on the X-axis are somewhat similar, as is the case for Figures 6a and 6b. Some strategies intentionally rely on many small losses, offset by a few large gains. To assess such strategies, multiply the number of consecutive losses by the average size of a loss, and weigh that in light of your personal tolerance for risk.

Shape—Figure 7a and 7b compare *position-detail charts*, providing a visual overview or *shape* of each strategy by plotting the number of occurrences of every trade outcome. Consider the area under the curves in relation to the black dotted zero line. 7a is skewed negative. 7b is skewed positive. The more results to the right of zero, the better.

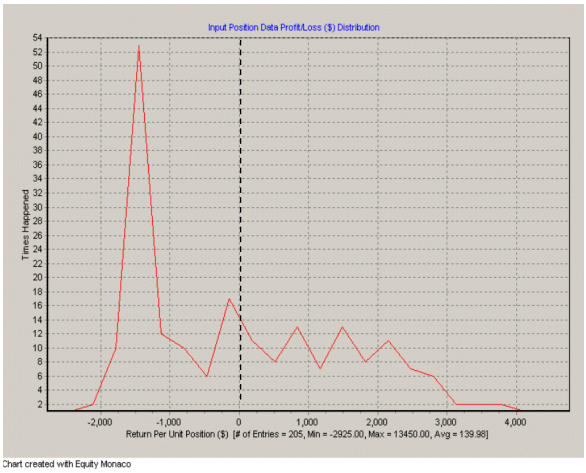

Figure 7a

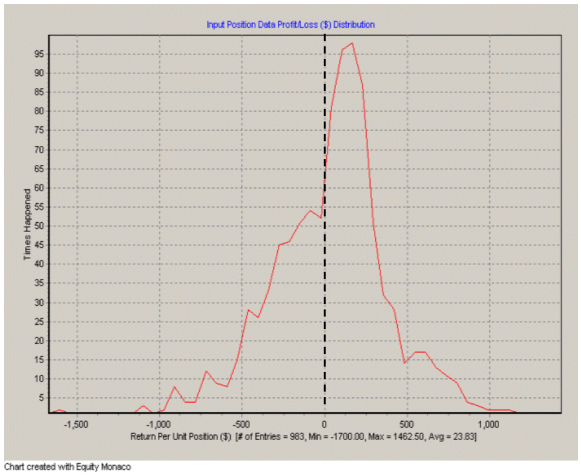

Figure 7b

Another system's shape may tightly hug the zero line, demonstrating that it depends on small gains and is easily overwhelmed by commissions and slippage. On the other hand, a shape leaning heavily to the right of zero implies a positive average reward/risk ratio and resistance to small price swings and noise. Always check the Y-axis of each test to verify that the scales are similar.

Copyrighted Material

An Application—*System Y* and *System Z*

Slope, Spread, and Scatter, Streak, and Shape— five visual cues. Let's create two systems and see what we can learn about them, using this simple toolbox.

The first strategy—*System Y*—tests the idea of buying on the first trading day of each month and selling on the last of each month, essentially asking the question, "How safe is it to go Long at any given time?" Any number of technical indicators could be chosen to time our entries and exits, but for this comparison, we will use our monthly rule as a proxy for all indicators. The strategy is tested on daily data from the Dow Jones Industrial Average, running from January 1915 to January 2004, a period of almost 90 years. For this test, all positions are Long 100 shares and executed at the average daily price. Currently available commissions and nominal slippage are included. Figure 8a below displays the equity curve from the backtest of *System Y*. Figure 8b shows its MC distribution of alternate equity lines.

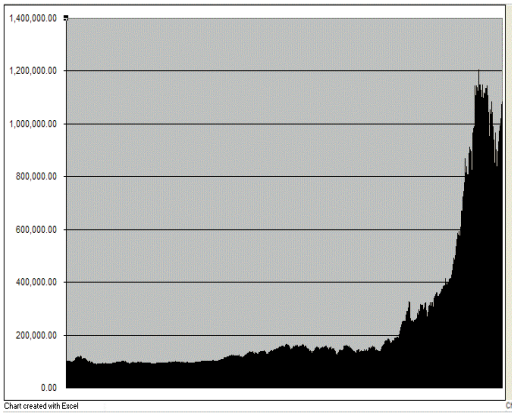

Figure 8a—*System Y* Equity

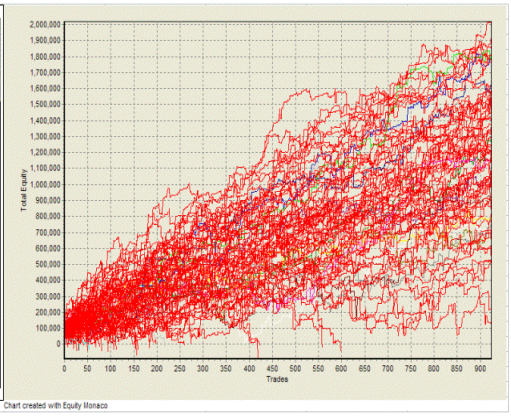

Figure 8b—*System Y* Alternates

The next strategy—*System Z*—is identical to *Y* in every respect except that a *Stop Loss* has been added. If equity drops 8% below the purchase price for any given month, the trade is terminated, then re-entered at the start of the following month. Figure 9a below displays the equity curve from the backtest of *System Z*. Figure 9b displays its MC distribution of alternate equity lines.

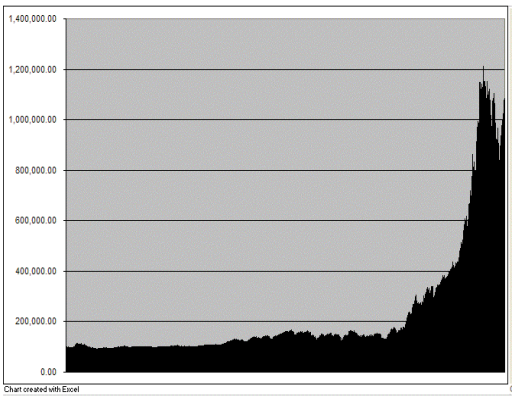

Figure 9a—*System Z* Equity

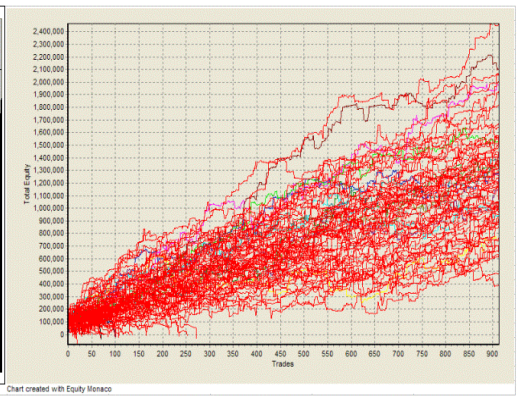

Figure 9b—*System Z* Alternates

Comparison—Both strategies survived the Great Depression, the malaise of the 70's, and the Internet bubble. Both had an average annualized return slightly better than 11%, with significant gains occurring after 1982. Both had reward/risk ratios better than 1.0 and both were on the right side of the trade more than 50% of the time. The *Shape* for each (not shown) was virtually identical. The *Streak* for both (not shown) was 10 consecutive losses.

Now compare Figure 8a to Figure 9a. Note that the backtested equity curves for Systems Y and Z are virtually indistinguishable. But Figures 8b and 9b demonstrate that the MC distributions are quite different. *Slope* and *Scatter* are about the same. However, *System Z* displays a significantly narrower *Spread*. *Z*'s distribution is actually quite tight and fewer paths lead to ruin. The stop loss made it more predictable, less random, and less risky. So, which system would one choose to take a Long position on a Dow derivative? Given a choice between *System Y* and *System Z*, one would clearly choose *Z*.

Copyrighted Material

Analysis of System X

Now let us return to *System X*. Putting aside the viability of the mythical Fish Head Futures market, let's ask whether the system has any technical merit. Referring to Figures 1 and 2, the *Slope* is strongly positive, demonstrating that the system can make a lot of money. However, its *Spread* is wide, with a number of paths leading to ruin. In terms of *Scatter,* there is no discernible cluster about the mean, making it relatively unpredictable. Its *Streak* of consecutive losses is eleven, but its *Shape* (not shown) is strongly positive. The system is high-risk/high-reward, with a significant probability of loss and a much greater probability of profit. Some may say that trading such a strategy is gambling, but consider that the odds are slanted toward the player.

System X may have merit if the risk can be managed. The common approach is total commitment, hoping for a few initial good years. However, a glance at the MC distribution warns that this is a financial form of Russian Roulette. A low-risk money management strategy requires sufficient capital to implement.

1.) Commit a small portion of the portfolio to *System X*, making sure that starting capital is above the predicted threshold of ruin.
2.) Trade *System X* as one element in a group of dissimilar strategies.
3.) Trade across a diversified basket of non-correlated markets.
4.) Examine the MC distribution of several aggregate combinations.

The optimum solution will be found in a balance of diversification techniques.

Conclusion

MC is a readily available tool that provides an important measure—the degree of randomness in a trading strategy. However, its utility has been hidden behind a veneer of technical obscurity. No matter how precise and sophisticated the mathematics, it is still no more than an estimate of the likely range of events. Therefore, I have proposed a simple visual method of analysis, using pattern-recognition, and have painted it with a broad brush. In this way, MC is intuitive and accessible, adding a wealth of information not evident in a system backtest. MC points to the range of outcomes that *might have been*, and by extension, the range of *what might occur* in actual trading.

* * *

About the Author—I translate ideas and technology into written, visual and strategic content that is readily understood. I've been a writer, artist, inventor and trader. I've served as vice president and on the boards of two corporations, patented 7 inventions and published in professional journals. My paintings have hung in museums, galleries and private collections and I've written my first novel. I hold a Bachelor of Fine Arts degree and Kellogg MBA. Contact me at john.jonelis@gmail.com

Suggested Reading—

STATISTICS WITHOUT TEARS, by Derek Rowntree; Scribners 1981, ISBN 0-02-404090-8
FOOLED BY RANDOMNESS, by Nassim Nicholas Taleb; Texere 2001, ISBN 1-58799-071-7.

Limitations*—In system testing, it is important to have an intuitive grounding in statistics. One needs a general feel for probability distributions, an ability to distinguish trend from noise, a way to judge the statistical significance of a price move, and a method for avoiding the phenomenon known as curve fitting. These subjects are outside the scope of this article.*

The names and organizations portrayed are fictitious and any similarity to reality is entirely random. All that can be known with any degree of certainty is what actually happened, and even that information is suspect. All analysis is based on assumptions, there are no guarantees, and all traders should be prepared for the unexpected. This material is not intended as investment advice. This author has been known to make errors and reach erroneous conclusions. No doubt, some of that has crept into this paper. That said, these ideas are put forward with the intention of proposing a simple but meaningful method for using an otherwise obscure tool. It is hoped that other writers will add to what has been presented.

Acknowledgements—This article has grown out of a collaboration with my research team, Project 304. Thanks also go to Erik Olson as well as the members of the Barrington Writer's Workshop.
This article is a pre-production draft published in the JOURNAL OF TECHNICAL ANALYSIS—
Issue 63, Market Technicians Association, 2005, © 2004 John Jonelis, revised 2011.
Equity Monaco is a product of *TickQuest*, http://www.tickquest.com/
Monopoly is a trademark of *Parker Brothers Inc*.

Copyrighted Material