Automated Finance:
The Assumptions and Behavioral Aspects of Algorithmic Trading

Andrew Kumiega
Illinois Institute of Technology

Ben Van Vliet
Illinois Institute of Technology

Abstract

Automated trading now dominates the financial markets. Yet, no philosophy of academic research into the topic exists. As the growth in automated trading suggests their greater returns and predictability, this paper examines stability and statistical control of trading process outputs as method of justifying predictions of future performance. New assumptions presented can form a foundation for positive research under this revolutionary paradigm, one almost completely ignored in the financial literature. The traditional financial literature rests on the assumption of normality of inputs, while trading systems aspire to the more rigorous engineering standard of justification. The end game is that now behavioral aspects, not of traders, but of trading system research and development projects drive market returns.

Key Words

Automated trading, quantitative finance, statistical process control.

Quantitative finance is built upon foundational assumptions. The literature derives that most notable assumption, normality, from the ability to describe past asset returns histogrammatically with that distribution. The fact that normality is widely understood to be unrealistic is inconsequential. It is a "sufficiently good approximation for the purpose [at] hand" [Friedman 1966]. The sufficiency criterion for positive research demands only that it "yield meaningful predictions…about phenomena not yet observed" [Friedman 1966]. The literature built upon the normal foundation *has* yielded significant contributions to the body of knowledge, and its application to normative finance as well as the art of finance cannot and should never be underestimated.

However, we must acknowledge also that the conversion of statistical descriptions of empirical financial data into justifications for inference occurs under a relatively loose standard: "implicitly or explicitly, it is assumed that historic results have at least some predictive ability" [Sharpe 1994]. We might contrast this with the engineering disciplines, which apply a more rigorous standard for induction: Deming [1986] states that descriptive statistics serve no useful purpose unless the underlying process is in a state of statistical control. Put another way, justification comes by stability. This begs the question then, what is a process? The American Society for Quality [ASQ 2008] defines a process is "a set of interrelated work activities characterized by a set of specific inputs and value added tasks that make up a procedure for a set of specific outputs." In engineering, a process is stable when it consists only of common-cause (or random) variation, and is absent any special-cause (or assignable) variation which originates outside the expected operating conditions of the process.

Under this definition of process, financial price data cannot be a process. There is no set of work activities that deterministically or probabilistically generates the output data. Financial data is a byproduct or epiphenomenon of human and social developments—management decisions, economic policies, and changing perceptions of value and potential future payoffs. We can, however, *model* financial data using stochastic mathematical processes—geometric Brownian motion, for example. But, of course real financial data does not answer to any such model. While a model may represent the past, the future is not obligated to fall within its rules. People don't have to go along with mathematical deductions of equilibrium or utility, or predictions of possible prices paths.

RULES

Wittgenstein [1953] in his *Philosophical Investigations* and Winch [1958] in his *The Idea of a Social Science and its Relation to Philosophy* discuss what it means for someone to follow a rule or process. The difficulty for the social sciences is that any series of human actions can be modeled with some formula or set of rules, so long as we are prepared to make that formula or those rules sufficiently complicated. And in any case, that a person's actions can be represented by a formula, is no assurance that he or she is in fact following that prescription. Winch provides a coherent example in this regard, but we will briefly develop our own.

Imagine we observe a businessman who wears brown shoes on Monday, black shoes on Tuesday, brown shoes on Wednesday, black shoes on Thursday and brown shoes on Friday. We might then derive a rule about his behavior: he wears brown shoes on Mondays, Wednesdays and Fridays, and black shoes on Tuesdays and Thursdays. On the ensuing Saturday and Sunday, we observe that he wears tennis shoes. We might augment our rule to state that on weekend days he wears tennis shoes. Based on this improved model we predict that on Monday he will wear brown shoes. However, we find that on Monday, he again wears his tennis shoes. As it turns out, he has the day off. We then change our rule again to state that on all off days he wears his tennis shoes. Armed with this new model, we confidently predict that on Tuesday he will wear either his black shoes if he goes to work or his tennis shoes if he is off. On Tuesday, however, we find that he does in fact have the day off, but that he wears his golf shoes. So, again we amend our rule to say that on off days, if he is going golfing, he wears his golf shoes. Our prediction for Wednesday is that if he goes to work, he will wear brown shoes; if he is off he will wear tennis shoes or golf shoes. On Wednesday we find that he in fact does go to work and does in fact where his brown shoes. At this point, we may be satisfied in our model of the evolution of shoe data. And since all the data has now been brought within the scope of our model (and moreover empirically validated), we even more confidently predict that on Thursday he will wear black shoes. However, on Thursday we find that he has traveled to Florida and is wearing sandals.

No matter how the shoe data evolves we can always derive a model to fit it, as long as we are willing to make its rules sufficiently complicated. However, our businessman is not intentionally following the model and is of course under no obligation to follow its precepts. Further, the model will be forever incapable of considering all future circumstances or states which may affect his shoe-wearing decision, and outcomes that exist outside its result set—he has recently purchased white shoes, for which the model has not yet accounted.

We could respond to these arguments by saying that there is in fact some underlying or fundamental reason why our model may yet be probabilistically true. For example, suppose that the company our businessman works for always has board meetings on Thursdays, and that the protocol at these meetings is to wear black shoes. We might find that our model predicts his Thursday choice of shoes with more certainty than other days, and that this is a contribution to the positive body of knowledge. The board meeting story provides a causal explanation for the data. We should again augment our rules to incorporate this knowledge, but then we are back where we started. We have not yet determined whether or not our businessman actually attends these meetings, regularly or otherwise. And, the firm may switch the day of its board meetings. Of course, if the man finds a new job as an ice hockey coach, we would have to throw out everything we thought we knew about shoe selection.

Furthermore, there is no statistical magic by which social behavior—as the aggregated behavior of humans—in the form of a market, can be expected to perform with any greater degree of predictability. Thus, even if financial data exhibits normality or autocorrelation or any other regularity in the past, we cannot be justified in expecting it to exhibit that behavior in the future. The changing human and social factors, emotional and otherwise, impacting business or trading practices their sets of possible outcomes cannot under this line of reasoning all be incorporated into any model, no matter how complex. What we can conclude about financial data from this exercise is that, generated as it is by human activity, it is under no obligation to perform according to the expectations set forth by any model of past experience. As such, financial data cannot meet the definition of a process.

We might, however, turn the conversation the other way around and say what if a person is aware of the process he or she is intended to follow. Wouldn't the outputs of such person then meet the definition of a process?

ALTERNATIVE

What we would like to do now is present an alternative to the traditional method of financial reasoning, one that ignores the assumption of normality, but is nevertheless systematic. As a foundation, let us assume that financial price data is not generated by any process, normal or not. Nor can past financial data be represented by any mathematical model that would yield outputs capable of meeting engineering's criteria for inferential justification—distributional stability. So, given that assumption is it possible to proceed? This is the relevant question in the age of algorithmic trading,[1] because human decision-making is being circumvented by technology, which must by its nature follow rules, which must (unlike our businessman) in the future act in accordance with the work activities it encapsulates.

Before we begin, we must admit to the presence of short-lived, but recognizable and discoverable patterns in empirical financial price data—autocorrelation of returns, sector outperformance, volatility clustering, even head and shoulders tops. And, as a necessary consequence of this, we must also admit that (theoretically at least) there exist opportunities for traders to take advantage of these patterns to generate abnormal returns[2] if only they are adept enough to recognize them. We will illustrate again with an example.

Imagine a 25 year old man standing in the middle of a sidewalk on a busy city street, with people of all ages streaming past him. If he asked each person his or her age and examined a month's worth of data, he would certainly find that no process or model could explain its evolution, and that no recognizable distribution[3]—either in the actual data or a time series sample—could encapsulate it. It's not normal or lognormal. It's just random data.

Nevertheless, inspection of that data would show that, much like financial data, short-lived but recognizable patterns do exist. Ages may be autocorrelated as a group of older folks on a shopping trip happens by, or a group of teenagers. For a time, busier middle aged folks on

---

[1] The terms automated, algorithmic and high-frequency trading have varying definitions. In this paper, we mean by these terms computerized buy-side systems that translate market data into decision to buy and sell financial instruments in order to generate profit. These systems include, for example, high-frequency equity trading systems, options market-making systems, and index arbitrage systems. The definition used in this paper can be contrasted with the sell-side definition of execution algorithms that seek volume-weighted-average-price improvement relative to a benchmark. We do not attempt to address this latter definition in this paper.

[2] Whether they do or not is immaterial, but theoretically the possibility is there.

[3] Of course, ages are bounded by zero and say one hundred, so the analogy doesn't quite fit. But let's assume that for my intents and purposes, the distribution is unbounded.

their way to work may pass retired window shoppers. We may also see strollers driven by grandparents, a sort of volatility clustering. The point is that there can be causal explanations for short-term patterns in the data otherwise uncaptured by descriptive statistics of the entire data set.

Now, let's assume that our young man has a particular interest in meeting women between the ages of 23 and 27, and that he is remarkably adept at selecting them from the crowd. Of course, we do not know his selection process; only can we see its outputs. Now, it is entirely possible that his selection process would produce output age data that is normally distributed, with say a mean of 25 and a standard deviation of one year. (If serial correlation exists in the crowd of data, all the more opportunity for his strategy to select eligible women.) He takes random input data, runs it through his process, and produces an output distribution that is not only normal, but stable. He can produce this distribution of ages every day. Based upon this stability, we would be justified (according to the higher standard of engineering) in inferring that he will be able to produce this same output distribution tomorrow. Based upon this stability, we now also have substantive conditions for knowing when this inference can no longer be justified. If his eye wanders, and the data violates the predefined conditions for stability, we might ask him what has changed and why.

What we can conclude about financial data from this exercise is that given random input data and known decision rules, humans can produce outputs which may potentially meet the definition of a process. Sticking to a process, however, is not something humans are particularly good at. Overconfidence, cognitive dissonance, regret and other cognitive biases are at odds with rule-based financial thinking [see Ricciardi and Simon 2000].

Significantly, these biases in practice introduce variation into what otherwise might in theory be a stable trading process. Consider as well that our young man may choose at any time to violate his own rules. He may forego certain opportunities if he deems them undesirable though within the bounds of his process. He may interrupt his process to go on lunch or take vacation. All these introduce inconsistencies in his output data which invalidate statistical analysis under the engineering standard. Indiscernible then is whether deviation from stability is due to the failure of the strategy or human factors affecting its implementation.

AUTOMATION

In the financial markets, decision rules are based upon estimates of probability (often under the assumption of normality). In the financial literature, when combined with a level of risk aversion, these estimates form a utility function unique to each investor. But, people are incapable of forming estimates of probability without bias. They have a preference for known risks versus unknown risks. This rejection of risk based upon its measure of certainty is unaccounted for in expected utility theory. Recent research by Easley and O'Hara [2010], Epstein and Schneider [2008], and Lo and Mueller [2010] has examined this ambiguity aversion in financial decision making. Yet, despite the devastating effect of contradictory empirical evidence (Ellsberg's paradox being the most notable in this regard), expected utility theory persists. Since information regarding probabilities is unavailable to traders, they must subjectively quantify the ambiguity. They pay for their subjectivity by not maximizing their utility.

Computers, on the other hand, face no such subjectivity. They *can* be expected to follow rules, and furthermore, they *can* form objective, unbiased estimates of risk. This is why automated systems represent a revolution in both the art and theory of finance. No longer do we need to rely on descriptive statistics of human outputs as a sort of loose justification of inference.

This would explain the recent growth in automated trading, the growth[4] of which, despite a dearth of empirical data[5], we must attribute to greater returns and their greater predictability. This conclusion is especially relevant when considering the considerable research, development and deployment costs required to launch and automated trading system. Through their objectivity and consistency automated trading systems fully capture the utility left by human traders. We call this the trading system's ambiguity alpha. Trading systems capture this ambiguity alpha in three ways:

1.) They dispassionately perform complex data-driven calculations of probability beyond the ability of humans.
2.) They perform calculations faster and more consistently than humans.
3.) They can watch many, many, hundreds or even thousands of inputs—both structured and unstructured—concurrently, beyond the ability of humans.

What we can conclude is that a three-fold set of foundational assumptions upon which to build a systematic study of automated trading exists.

1.) The inputs into automated financial systems (i.e. financial price data) are driven by human and societal decisions, and therefore are random and cannot be stable. Furthermore, human traders cannot implement a trading strategy with the consistency necessary for their outputs to fit the definition of process.
2.) Regularities in financial data do exist, but only for short periods of time. A window of opportunity may open, and then at some future time it will close. While the window is open, an automated trading system that exploits the opportunity can earn

[4] While empirical data is scant, industry research firm TABB Group estimates that 70% of trading volume is driven by automated systems. See Sussman et al. [2009], and Gomber et al. [2009].
[5] There is a tremendous amount of secrecy around automated trading algorithms due to their portability.

excess returns. Therefore, perceiving in a timely fashion the openings and closings of windows of opportunity is a key driver of competitive advantage.

3.) While a particular exploitable regularity exists, the outputs of an automated trading strategy can generate stable process outputs. That is, the output data will the engineering standard of justification: that it is under statistical control. This provides a more sound basis for belief in the repeatability of those outputs. At some future time, then such outputs will go out of statistical control, belief in the output process is no longer justified.

These assumptions do not depend upon the normality assumption that is the basis of much of quantitative finance.

BEHAVIORAL ASPECTS

Automated trading does, however, contain behavioral aspects. Automated strategies, which make use of models, and their encapsulating technologies are intellectual constructs, and as such require research and development to prove stability and establish control limits for deployment (see Bilson et al. [2010] and Hassan et al. [2010]). In this respect, they are projects. But, projects themselves have behavioral components. Statman and Caldwell [1987] show that behavioral factors—namely, regret aversion—in project budgeting negatively impacts firm value. They state that "managers tend to become entrapped in losing projects and throw good money after bad as they attempt to rescue them." Staw's [1976] experiment showed a high level of personal responsibility increases resistance to project termination. (We suspect that in a trading environment, where financial incentives are significant, this phenomenon is pushed to its extreme.) The behavioral aspects of trading systems are not in emotional reactions to market moves, but rather in the emotional reactions to various facets of trading system project management. As with much of the project management literature, rigorous methodology can reduce financial risk through objective criteria for project evaluation and termination.

Trading system development projects must trade-off *project cost, time, scope,* and *quality,* as well as *market risk, expected return,* and *diversification effects*. The answers are not a matter of one or the other, but rather of how much of one and how much of the other. As de Mast and Bisgaard [2007] point out: “if problems are not quantified, the trade-off nature is obscured and people tend to treat them as either/or problems and frequently politicize them.” The condition of sustainable market returns, then, is that reliable cognitive processes in this respect trump behavioral decisions.

Furthermore, humans monitoring automated trading systems may adjust trading decision parameters in real-time or make on/off decisions. The engineering literature has examined behavioral aspects of systems monitoring and adjustment. Much of this research has found that humans cannot keep up with and modify processes in real time. Their tendency is to over-correct and hence add to rather than subtract from process variation. Adaptive controllers have proven to better manage high-speed machines. The research of Suh and Cheon [2002], Williams and Davies [1986] is notable in this regard. This is an area for additional research in finance.

SUMMARY

While the assumption of normality in financial data is not incorrect according to the looser standards of the social science of finance, we believe that automated systems’ abilities to capture ambiguity alpha by producing outputs which meet the more rigorous engineering standard is a primary driver of the growth of automated trading, despite its considerable research and development costs.

Since a majority of trading volume is driven by computers, the practice of finance revolves largely around the operation of automated trading systems. Yet, the academic financial literature is only just beginning to address the topic. Where little work has been done on the philosophy of finance (Frankfurter and McGoun [1996]), none has yet been attempted in automated finance.

In this paper, we described the impossibility of stable inputs into financial models. Such data cannot meet the definition of a process. Next, we described how the outputs of human traders cannot be sufficiently consistent to meet the engineering criteria for stability. We then argued that through automation, trading strategy outputs can meet such criteria. We developed a method of justification that explains the possible existence of additional descriptive power and excess returns through stability and statistical control, what we call ambiguity alpha. We believe this is the philosophical core of the revolution in the practice of finance.

Finally, we established a set of foundational assumptions that explicitly avoids the assumption of normality upon which so much of quantitative finance is based. This perspective on finance leads to the conclusion that financial returns are driven by the behavioral aspects of trading system research and development project management rather than the behavioral aspects of market participants, which automated systems seek to circumvent.

REFERENCES

American Society for Quality (ASQ). 2008. Accessed at: http://www.asq.org/glossary/p.html, on December 16, 2008.

Bilson, John F.O., Andrew Kumiega, and Ben Van Vliet. 2010. "Trading Model Uncertainty and Statistical Process Control." *Journal of Trading.* Forthcoming.

De Mast, Jeroen and Soren Bisgaard. 2007. "The Science of Six Sigma." *Quality Progress.* January, pp. 25 ff.

Deming, W. Edwards. 1986. *Out of the Crisis.* MIT Press. Cambridge, MA.

Easley, David, and Maureen O'Hara. 2010. "Microstructure and Ambiguity." *Journal of Finance*. Forthcoming.

Epstein, Larry G., and Martin Schneider. 2008. "Amibiguity, Information Quality, and Asset Pricing." *Journal of Finance*. Vol. 62, No. 1, pp. 197-228.

Gomber, Peter, Markus Gsell, Gregor Pujol, and Adrian Wranik 2009. **"**Regulation and Technology in Equity Trading **–** The Impact of Regulatory and Technological Changes on Order Execution and the Trading Venue Landscape in Europe" Chapter in *Facing New Regulatory Frameworks in Securities Trading in Europe*. Engelen, P.-J. and K. Lannoo, K. Editors. Intersentia and Centre for European Policy Studies (CEPS), pp. 31-54.

Frankfurter, George, and Elton McGoun. 1996. *Toward Finance With Meaning: The Methodology of Finance : What It Is and What It Can Be*. JAI Press.

Friedman, Milton. 1966. "The Methodology of Positive Economics." In *Essays in Positive Economics.* Univ. of Chicago Press, pp. 3-16, 30-43.

Hassan, M. Zia, Andrew Kumiega, and Ben Van Vliet. 2010. "Trading Machines: Using SPC to Assess Performance of Financial Trading Systems." *Quality Management Journal*, Vol. 17, No. 2, pp. 42-53.

Lo, Andrew, and Mark Mueller. 2010. "WARNING: Physics Envy May Be Hazardous To Your Wealth!" *Journal of Investment Management.* Forthcoming.

Marshall, Alfred. 1885. "The Present Position of Economics" reprinted in *Memorials of Alfred Marshall*, ed. A. C. Pigou (London: Macmillan & Co., 1925), p. 164.

Ricciardi, Victor and Helen Simon. 2000. "What is Behavioral Finance?" *Business, Education and Technology Journal.* Fall.

Schmid, Wolfgang and T. Bodnar. 2008. "Estimation of Optimal Portfolio Compositions for Gaussian Returns." *Statistics and Decisions.* 26, pp. 179-201.

Sharpe, William. 1994. "The Sharpe Ratio." *The Journal of Portfolio Management.* Fall.

Statman, Mier and David Caldwell. 1987. "Applying Behavioral Finance to Capital Budgeting: Project Terminations" *Financial Management.* Winter, pp. 7-15.

Staw, B.M. 1976. "Knee-deep in the Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action." *Organizational Behavior and Human Performance.* June, pp. 27-44.

Suh, S.H., and S.U. Cheon. 2002. "A Framework for an Intelligent CNC and Data Model." *The International Journal of Advance Manufacturing Technology*. Vol. 19, No. 10, June, pp. 727-735.

Sussman, Adam, Larry Tabb, and Robert Iati. 2009. "US Equity High-Frequency Trading: Strategies, Sizing and Market Structure." *TABB Group*.

Williams, J.H., and A. Davies. 1986. "Controller base manufacturing system management-How close to reality?" *The International Journal of Advance Manufacturing Technology*. Vol. 1, No. 5, November, pp. 61-74.

Winch, Peter. 1958. *The Idea of a Social Science and its Relation to Philosophy.* Quoting Ludwig Wittgenstein in *Philosophical Investigations* (1953). Blackwill Publishing.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Blackwell.