Trading System Capability

Andrew Kumiega
Adjunct Faculty
Illinois Institute of Technology
565 W. Adams, Chicago, IL 60661
p. 630 484-3301 e. akumiega@iit.edu

Thaddeus Neururer
Boston University
School of Management
595 Commonwealth Ave., Boston, MA 02215
p. 617-353-4613   e. thadn@bu.edu

Ben Van Vliet, *Corresponding*
Assistant Professor
Stuart School of Business, Illinois Institute of Technology
565 W. Adams, Chicago, IL 60661
p. 312 906-6513 e. bvanvliet@stuart.iit.edu

December 12, 2012

Abstract

For firms involved in high-frequency trading, of primary concern is whether or not a particular trading system will generate sufficient profits to cover its considerable research and development, fixed and variable costs. Since these costs are allocated per accounting period, firms view market returns from a bottom line profitability perspective. The current performance metrics of finance do not provide sufficient information to assess investments in high-frequency trading systems in this context. We develop and test a capability measure that captures the performance of high-frequency trading systems in this light.

Key Words

High frequency trading, performance measurement, capability.

High-frequency trading (HFT) systems come in different shapes and sizes. They include stock and option market making systems, as well as other algorithm-driven strategies such as index arbitrage and equity long-short. What all HFT systems have in common is that technological speed is a source of competitive advantage. Latency causes slippage and, therefore, impacts the profitability of the system. The daily profits are then summed and averaged to produce the weekly profit used to pay the firm's expenses.

HFT systems generate more than 70% of the trading volume in the U.S. (Brogaard [2011]). Thus, it is no exaggeration to state that the practice of finance now focuses on their design and control. We infer from this that institutional investors have a preference for the risk-return profiles only automated systems can generate. This preference must be driven by greater returns, lower volatility of returns, and their greater predictability. But, high frequency trading is not cheap. "Years of research and development and millions of dollars go into the development of these algorithms" (Hull [2000]). For algorithmic trading firms (ATFs), of primary concern is whether or not a particular trading system will generate sufficient profits to cover its fixed and variable costs. Research and development (R&D), salaries, leases and infrastructure costs can easily run into the millions (sometimes tens of millions) of dollars. Where traders and quantitative analysts are concerned with maximizing revenue from trading systems, top management and investors in ATF are concerned with maximizing bottom-line accounting profits. Since these costs are allocated per accounting period, ATFs view profits from an earnings-before-taxes (EBT) perspective. There are most often no performance or management fees charged as HFT systems usually run by proprietary trading firms.

HFT systems execute many buy and sell transactions per minute. They never buy and hold. The daily profit on such systems is based upon thousands of trades. The large number of trades leads to the law of large numbers. The daily profit to the firm is an average of the profits across thousands of trades.

An ATF may find a profitable trading strategy, but choose to ignore it if they cannot justifiably expect to achieve some threshold level of bottom line profitability. Bottom line profitability is the total trading profits per day minus the total cost of operating the system per day. The profit over some accounting period, say two-weeks or one month, is the sum of the daily profits. Thus, ATFs must create the trading strategies that maximize their long-term business objectives. Further, they must trade-off two potential uses of capital:

- Investing in working trading systems that generate profits to cover the operating costs of the firm. We call this investment capital.
- Funding R&D of new trading strategies and their enabling technologies. We call this R&D capital.

This bifurcation of capital allocation and large number of per-period profits that are summed per day and per accounting period leads to performance measurement problems not covered by the traditional metrics of finance that ignore the business costs of operating a high technology firm.

Performance Measures in Finance

The well-known Sharpe Ratio [1994] summarizes an asset's return-to-variability. When comparing two investments, an investor should choose the one with the higher

Sharpe Ratio, defined as:

$$Sharpe = \frac{E(r) - r_f}{\sigma_r} \qquad (1)$$

Where $E(r)$ is the expected return, r_f is the risk free rate and σ_r is the standard deviation of returns. Most academic papers focus on log returns of buy and hold investments. Distributions of such returns are well-known to exhibit leptokurtosis with fat tails. Not so in high frequency trading. The output distributions of high frequency trading systems may well be non-normal, but long left tails are curtailed by trading tactics—stop losses, hedges, etc.

Aldridge [2010] estimates the Sharpe Ratio for a HFT system can run in the order of several thousands. This is suspect given that most Sharpe Ratios are below 5. Aldridge's calculation ignores operating costs which would lower that value. Aldridge's estimates can be explained by averaging of thousands of small profits and stopped-losses across dozens, hundreds or thousands of instruments. This type of process produces distributions which are approximately normal due to the central limit theorem.

The Sortino Ratio (see Sortino and van der Meer [1991]) is a modification of the Sharpe Ratio. Where Sharpe penalizes both positive and negative volatility, Sortino penalizes only for returns falling below a lower specification target for required rate of return. The Sortino Ratio is calculated as:

$$Sortino = \frac{E(r) - t}{d} \qquad (2)$$

where $E(r)$ is the expected return; t the target rate of return, and d is the downside risk. Downside risk is the target semi-deviation, as:

$$d = \left(\int_{-\infty}^{T} (t-r)^2 \cdot f(r)dr \right)^{\frac{1}{2}} \tag{3}$$

The Sortino Ratio is more appropriate for HFT since many HFT systems are designed to generate profits during jump periods. Where the Sharpe Ratio penalizes for both large positive and negative jumps, Sortino penalizes for only negative.

A similar metric, called the Information Ratio (IR), is the return of an asset relative to some benchmark (e.g. the Dow Jones Industrial Average) divided by the active risk. The IR is the main criteria used to judge trading and investment strategies.

$$IR = \frac{E(r) - r_b}{\sigma_{r^a}} \tag{4}$$

If the benchmark is the risk-free rate, then the IR is the same as the Sharpe Ratio. (This is an important point because HFT systems generally have no benchmark.) Grinold and Kahn [2001] have shown that a higher IR is the only criterion for investment selection. A higher IR is always better regardless of the risk penalty. Grinold and Kahn [2001] claim in the Fundamental Law of Active Management (FLAM) that the IR follows the following relationship:

$$IR = IC \cdot \sqrt{b} \tag{5}$$

where IC is the information coefficient, the correlation (usually Spearman) between predicted returns and their actual values. And, b is the breadth, the number of independent bets. For high-frequency trading systems, b could be in the thousands per day (i.e. hundreds of thousands per month). For HFT systems, the allocation of fixed costs on a trade-by-trade basis would be impractical. The thousands of trades per day/period result in highly skewed IC values as b increases.

The Sharpe, Information and Sortino Ratios do not correctly characterize the performance of HFT systems and their ability to operate according to specification, within tolerance or meet the firm's requirements for profitability. They are inappropriate for examining the risk-return requirements of HFT systems because they ignore R&D, deployment, operating expenses.

In three of these calculations—Sharpe, Sortino, and IC—each assumes the appropriate measure of performance is asset return and not profit. Thus, the Sharpe Ratio is an approximation to the case where fixed costs are present, but are small compared to the investing capital base. This approximation breaks down as fixed costs become large compared to the capital base as in HFT. ATFs have limited capital. Scarce resources must be allocated across a portfolio of R&D projects and systems and are reallocated periodically. Furthermore, ATF's reset trading capital periodically. Bonuses are paid. Capital may be reallocated to other trading systems (e.g. ones that are more profitable or have a portfolio effect on the ATF as a whole). Also, the trading strategy itself may not be infinitely scalable, so capital may be subtracted. The traditional metrics from finance do not account for these realities. Further, these calculations all penalize in some way for jumpy performance. Averaging profits over groups would generate a more stable metric. Most HFT systems execute tens of thousands of trades per day, where the profits per trade all occur over different holding times (measured at the microsecond level) and many result in zero profit. This forces firms to average profits from thousands of trades from (potentially) hundreds of systems to obtain a profit for the day.

In the mutual fund and hedge fund industries, the managing partner receives a fee (usually a percentage of assets under management) to cover expenses. ATFs must pay their R&D overhead out of trading profits on investment capital. A firm could go bankrupt with a Sharpe Ratio even greater than 5, since their weekly business expenses could greatly exceed the revenue from the trading system. New performance metrics that more realistically assess the performance HFT systems are needed.

Engineering Perspective

Finance is moving (and in large part *has* moved) from human to computer numerical control (CNC), a discipline well-documented in the engineering literature. ACF is increasingly similar to continuous manufacturing in this respect. However, where predictions in finance are justified under a relatively weak standard ("implicitly or explicitly, it is assumed that historic results have at least some predictive ability" (Sharpe [1994])), the engineering disciplines apply a more rigorous standard. As Deming [1986] states: "descriptive statistics serve no useful purpose unless the underlying process is in a state of statistical control."

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" (ASQ [2008]). "A process is stable if it has a constant mean and a constant variance over time." (NIST [2012])[1] A process is said to be in a state of statistical control when it is stable, when it consists only of common-cause (or random) variation, and is absent any special-cause (or assignable) variation which originates outside expected operating conditions. Without a state of statistical control, no reliable predictions can be made. If the decisions to invest in an HFT system is be based upon its ability to perform according to

[1] This definition is generally that of stationarity, rather than stability, but in the NIST definition is commonly used in the engineering disciplines.

specifications, then statistical control is a necessary pre-condition. Thus, any measure of a system's capability of meeting specifications first requires stability.

Because trading systems follow coded processes, their outputs can be made to be within statistical control. Several authors have applied statistical control to problems in trading, investment, portfolio construction and risk management, including Hassan et al. [2010], Bilson et al. [2010], Golosnoy and Schmid [2009], Hittesdorf [2009], Schmid et al. [1998, 2007, and 2008], Frisen [2003], Tomasson [2009], Bock et al. [2008], and Rowe [2003]. Our goal, however, is to examine the performance of in-control HFT systems relative to ATF specifications for profitability.

The remainder of the paper is organized as into five sections. First, we define the firm's specifications for accounting profitability. Second, we describe how to define process control limits HFT systems using backtests. Third, we describe how backtesting can prove the capability of the system to meet ROI requirements. Fourth, we provide a trading system capability study to compare Sharpe Ratio values to those our capability metric. Fifth, we summarize and conclude.

Investor Specifications

The required profit on an HFT system is net of the firm's fixed costs allocated to it. To stay in business, the firm must remain cash-flow positive. That is, rather than percentage returns, the firm is interested in dollar profits after expenses. While this seems obvious, most hedge funds nevertheless fail for operational reasons, not failure of their trading strategies (see Hamilton [2006]). Positive returns from trading activities do not guarantee positive cash flow after expenses. In Exhibit 1, assuming a trading strategy that does cover its costs, the firm's lower specification limit (LSL) is the breakeven point.[2] (An in capable trading strategy could have its expected profits below the LSL.) Thus, the LSL defines the stop-loss on the HFT system.

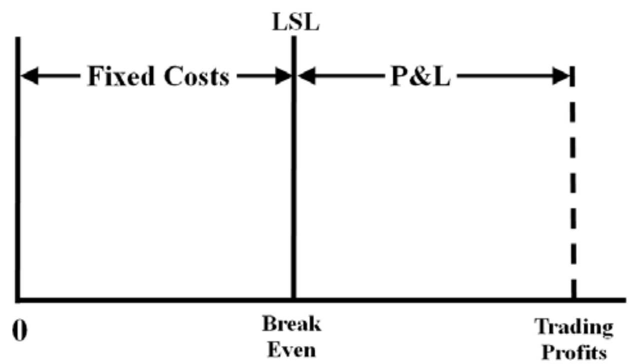

Exhibit 1: Investor Specifications

The HFT firm would like to know if a given trading system can generate enough revenue to at least breakeven (i.e. cover R&D capital expenses), and then additionally provide a return on investment capital. That is, trading revenue equals fixed costs plus ROI. Put another way, is the HFT system capable of consistently generating trading profits that exceed the LSL?

Capability indexes can relate the firm's profit specifications to the performance of an HFT system. One such index, the C_{pk} (see ISO 9000), compares the outputs of a process to specification limits. A process is capable if nearly all its outputs fall within lower and upper specification limits (USL). The C_{pk} is calculated as a ratio of the width of the specification to the width of the process outputs, where both widths are measured as six standard deviations.

As discussed in the following section given an HFT system that generates thousands of trades per day, these trades are averaged over the course of the trading day. Then, the daily trades per security are grouped per day and per accounting period. While

---

[2] The breakeven point is the level of revenue from a trading system that equals total fixed costs so that accounting profit equals zero. Fixed costs relate to R&D and infrastructure costs allocated to each trading period.

not guaranteed to be normal, the three levels of averaging will push the accounting period profits to normal by way of the central limit theorem. This process of averaging profits across thousands of (profits or stop-lost) trades with very small holding periods (as small as sub-second) is different from the traditional finance literature, which focuses on daily returns over long holding periods (e.g. years), market risk and fat tailed distributions.

With daily profits π_i as the output process, the daily average profit will approach $\bar{\pi} \sim N(\bar{\bar{\pi}}, s_{\bar{\pi}}^2)$. This framework can easily be expanded to a mixture of normal to handle any combination of skewness and kurtosis. (The more advanced methods of calculating SPC and capability for non-normal processes (see for example, Chen, et al. [2006] or Thissen, et al. [2005]) is not discussed in this paper to ensure the reader can follow the pedagogical example in this paper.)

The C_{pk} is the ratio of the distance from the center of the process $\bar{\bar{\pi}}$ to the nearest specification limit divided by three times the process sample standard deviation $s_{\bar{\pi}}$.

$$C_{pl} = \frac{\bar{\bar{\pi}} - LSL}{3 \cdot s_{\bar{\pi}}} \quad \text{and,} \quad C_{pu} = \frac{USL - \bar{\bar{\pi}}}{3 \cdot s_{\bar{\pi}}} \tag{6}$$

$$C_{pk} = \min\left(C_{pl}, C_{pu}\right) \tag{7}$$

Exhibit 2 shows the relationship between the C_{pu} and C_{pl}, and the sample process distribution.

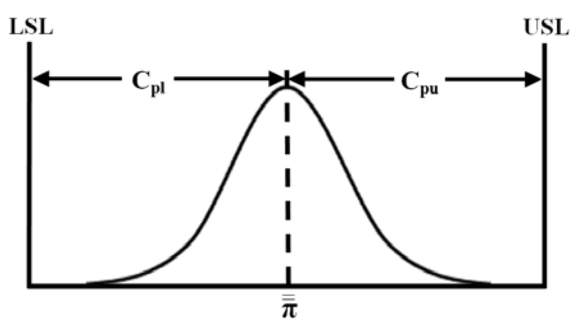

Exhibit 2: Capability and Process Distribution

An HFT system cannot make too much money. Thus, for benchmarking, the USL is set to infinity. (However, as we will show, a system that violates its upper control limit (UCL) would be out of control. An out of control system should be brought into statistical control prior to assessment of capability.) So, we will focus on only the C_{pl} to assess the capability of an HFT system in profitability rather than returns. In Exhibit 3, assuming a capable strategy, we relate the C_{pl} back to the firm's specification for profitability and we show the relationship between the C_{pl}, the firm's breakeven point, and the distribution of HFT outputs. Here, the distribution of the average trading profits is above the LSL. An incapable trading strategy would have a significant portion of its left tail below the LSL. This tail would represent the probability that the strategy would be incapable of covering its costs.

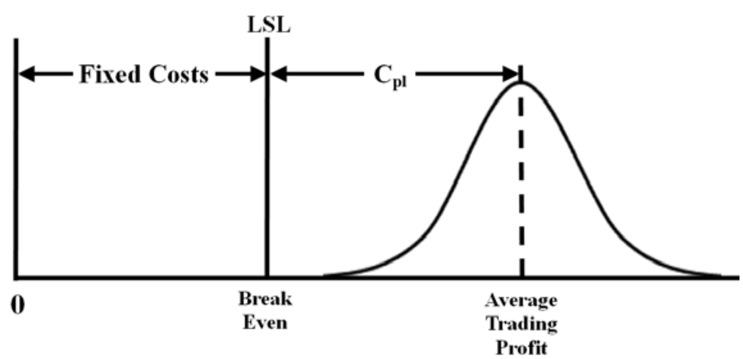

Exhibit 3: Breakeven, Profits and Average Trading Profits

C_{pl} rewards reduction of downside variability greater than an increase in profitability. This is unlike the Sharpe Ratio, where an investor should be indifferent to an increase in risk as long as he is compensated for it in increased expected return. Where the Sharpe Ratio focuses on the investor's decision on expected return per unit of risk, the ATF owner focuses on the need for positive cash flow after R&D capital expenses (rather than management fees).

To calculate the C_{pl}, sample means and standard deviations are needed, where sample size k should be at least 25, though in practice usually $k \geq 100$ (see Shewhart [1986]). To accept the process as capable, the C_{pl} should be at least 1.0, but according to Goetsch and Davis [2000] 1.33 is the preferred minimum. A C_{pl} of less than one means the process incapable of meeting specifications. In a case where the process spread exceeds the LSL, it can never be capable of meeting specifications.

For example, assume an HFT firm requires that on average the positive net cash flow from its trading systems of \$1000 per day. Furthermore, since the firm uses weekly accounting periods, each system should never generate profits below its operating costs. (The USL may be negative given capital reserves, but rolling averages should nevertheless be profitable.) Research shows that a particular trading system under consideration will produce average daily trading revenues of \$3000 with a sample standard deviation of \$250, but will cost \$1100 per day to operate. Thus, its average daily net cash flow is 3000 – 1100 = \$1900. The capability of this system to meet the firm's specification is:

$$C_{pl} = \frac{1900 - 1000}{3 \cdot 250} = .67 \qquad (8)$$

Because the C_{pl} for the HFT system is .67, we deem it incapable of meeting specifications. The ATF cannot expect to stay in business due to negative cash flows.

Assessing Stability

The C_{pl} calculation of capability is conditional upon the stability of the distribution, regardless of its shape. Stability is established by examining the distribution of the sample mean, which approaches the normal $\bar{\pi} \sim \mathrm{N}(\,\bar{\bar{\pi}}\,,\ s_{\bar{\pi}}^2\,)$ through the central limit theorem. In engineering disciplines, stability of a process is assessed according to Nelson's [1984] eight rules for statistical control (see Appendix A).

Averaging profits over accounting periods generates greater stability metrics, and proper backtesting can demonstrate this state. Assume that a particular HFT system makes several (e.g. hundreds or thousands) independent bets *b* per period, where some make money, some do not. The system holds no positions overnight. While the system demands a small amount of trading capital, it requires large amount of infrastructure investment. Placing of thousands of trades per minute requires software, servers, switches, and IT personnel.

Given the backtested performance data, we let *n* be the number of periods in a sample, and as previously described, and *k* be the number of samples in a year. Then, we let π be the time series of period-ending (e.g. second-to-second, minute-to-minute, or day-to-day[3]) *profits* on the HFT system. The daily profit is net of variable costs (i.e. commissions, etc.) and fixed costs allocated to each day (including R&D payback). Thus, π_i is the bottom-line dollar profit for time period *i.* Working at this level of control, suppose that we calculate the *n*-period (e.g. if the system must be profitable on a weekly basis, then $n = 5$ days; bi-weekly would be 10 days) sample average profits $\overline{\pi}$:

$$\overline{\pi} = \frac{1}{n}\sum_{i=1}^{n}\pi_i \tag{9}$$

From these values, we calculate the average profit of *k* (e.g. 52-weeks) samples, $\overline{\overline{\pi}}$.

$$\overline{\overline{\pi}} = \frac{1}{k}\sum_{j=1}^{k}\overline{\pi}_j \tag{10}$$

[3] For simplicity, we will use daily time periods. In practice, this would depend on the holding period for each trade. Some trading systems hold positions for milliseconds, so much shorter periods could be selected.

As mentioned, through the central limit theorem, these subgroup averages, $\overline{\pi}_j$, will tend to be normally distributed around $\overline{\overline{\pi}}$ with dispersion $s_{\overline{\pi}}$. In practice, if $n < 10$, we use the sample range to estimate $s_{\overline{\pi}}$. Range{ $\pi_1, \ldots, \pi_k$ }, the range R of profits in a sample is defined as:

$$R = \pi_{(n)} - \pi_{(1)} \tag{11}$$

And, the average range $\overline{R}$ over k samples is:

$$\overline{R} = \frac{1}{k}\sum_{j=1}^{k} R_j \tag{12}$$

The relationship between the average range and the standard deviation depends only on the sample size n (see Patnaik 1950). The process standard deviation s_π and the standard deviation of sample means $s_{\overline{\pi}}$ are estimated using d_2, a constant based on the subgroup size per Exhibit 4, as follows:

$$s_\pi = \frac{\overline{R}}{d_2} \qquad \text{and,} \qquad s_{\overline{\pi}} = \frac{s_\pi}{\sqrt{n}} \tag{13}$$

From an absolute perspective, the annualized expected profit $E(\pi)$ is equal to the sample average of backtested (or historic) profitability of the trading system $\overline{\overline{\pi}}$ times the number of periods in a year $n \cdot k$, say 52 weeks or 260 trading days[4].

$$E(\pi) = \overline{\overline{\pi}} \cdot n \cdot k \tag{14}$$

And the annualized standard deviation of sample profits $\sigma_{\overline{\pi}}$ is given by:

$$\sigma_{\overline{\pi}} = s_{\overline{\pi}} \cdot \sqrt{n \cdot k} \tag{15}$$

[4] Including holidays, the actual number of trading days per year is around 252. However, to simplify we will assume 5 trading days per week times 52 weeks is 260 trading days per year.

This enables us to define performance boundaries, or control limits, three standard deviations above and below the expected profit, as estimated by:

$$\begin{aligned} UCL_{\overline{\pi}} &= \overline{\overline{\pi}} + A_2 \cdot \overline{R} \quad \approx \overline{\overline{\pi}} + 3 \cdot s_{\overline{\pi}} \\ LCL_{\overline{\pi}} &= \overline{\overline{\pi}} - A_2 \cdot \overline{R} \quad \approx \overline{\overline{\pi}} - 3 \cdot s_{\overline{\pi}} \end{aligned} \tag{16}$$

Where A_2 is the anti-biasing constant per Exhibit 4, and:

$$A_2 = \frac{3}{d_2\sqrt{n}} \tag{17}$$

Value of *c*	Basic Factors		Factors for Averages			Factors for Ranges			
	d_2	d_3	A	A_1	A_2	D_1	D_2	D_3	D_4
2	1.128	0.853	2.121	3.760	1.880	0	3.686	0	3.267
3	1.693	0.888	1.732	2.394	1.023	0	4.358	0	2.575
4	2.059	0.880	1.500	1.880	0.729	0	4.698	0	2.282
5	2.326	0.864	1.342	1.596	0.557	0	4.918	0	2.115

Exhibit 4: Constants[5]

The true standard deviation of profits is found from the distribution of $w = R / \sigma$. The standard deviation of w is d_3 and is a function of the sample size n (see Patnaik [1950]) and is used to adjust from a sample to a population. Since $R = w \cdot \sigma$, we can estimate s_R by:

$$s_R = d_3 \cdot \frac{\overline{R}}{d_2} \tag{18}$$

As a result, the control limits for R are three standard deviations above and below the expected profit, as estimated by:

$$\begin{aligned} UCL_R &= D_4 \cdot \overline{R} \\ LCL_R &= D_3 \cdot \overline{R} \end{aligned} \tag{19}$$

[5] For example, d_2 is a proportionality factor. If X is normally distributed, then $\mu_R = \sigma_X \cdot d_2$, and being *Rbar* an estimation of μ_R, *Rbar* / d_2 becomes an estimation of σ_X. For more information on the derivation of proportionality constants used in quality control, see Woodall and Montgomery [2000], AIAG [2005], Juran [1988] and ASTM [1976].

The estimation of s_π in (7) also allows us to represent the annualized standard deviation as:

$$\sigma_\pi = \frac{\overline{R}}{d_2} \cdot \sqrt{n \cdot k} \tag{20}$$

We now have risk performance boundaries.

If in the backtest $\overline{\pi}_j$ or R_j performance occurs outside the respective limits (which occur with probability ≈ .0013), this would indicate that the process is out of statistical control, and therefore unstable.[6] Thus, the purpose of a backtest is to prove stability as justification that the HFT system can consistently cover its allocated fixed costs and provide ROI on a standard accounting cycle.

Backtesting

If backtesting proves process stability (according to the prior section), then the expectation of repeatability is justified, and capability can be assessed. Sufficient capability, using C_{pl}, justifies investment in further development. Insufficient capability may trigger additional research on process improvement until a targeted level of capability is achieved. In the case where the target cannot be achieved, specifications may be changed. Exhibit 5 shows that process improvement can enhance capability under the condition of stability. (Stability is defined as being in statistical control.)

[6] Other signals with similar probability exist. See Nelson [1984] and Bilson, et al. [2010].

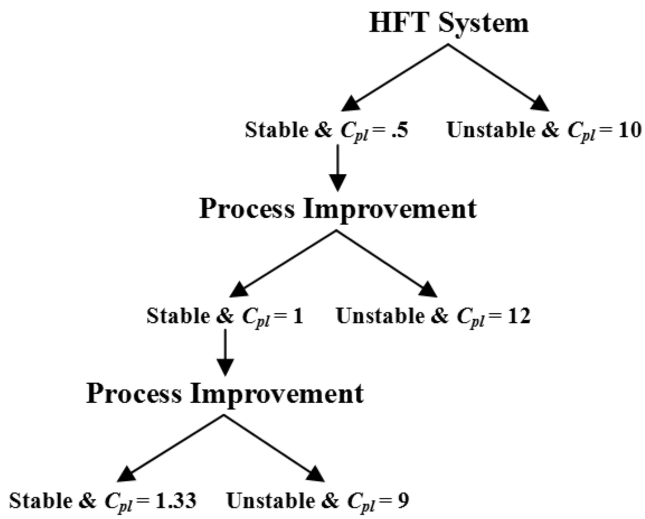

Exhibit 5: Backtesting with Process Improvement to Achieve Target C_{pl}

Continued research into quantitative and technological methods will produce differing versions of the same HFT system. Unstable versions of a trading strategy are discarded regardless of capability, while stable versions are improved until the targeted level of capability (e.g. C_{pl} = 1.33) is met. Exhibit 6 shows how control limits defined in the backtest fit within the C_{pl} framework.

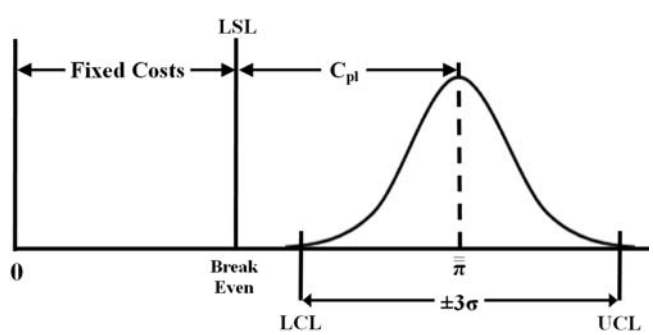

Exhibit 6: C_{pl} and Process Control Limits

Simulation Example

For example, let's assume that the total annual fixed costs for an ATF firm—including payroll, rent, R&D amortization, technology costs—are (for round numbers) $1 M. Over a 252 day trading year, operating 6.5 hours per trading day, this works out to per second costs of $.17. Positive returns from trading activities do not guarantee positive cash flow after expenses. The firm must develop an HFT system that at a minimum covers this cost.

To illustrate the use of the C_{pl}, we generate simulated data as a proxy for (potentially non-normal) empirical HFT system returns which are highly secret. While many families of distributions exist to handle various combinations of skewness and/or kurtosis—Pearson, Johnson, beta—Cooper and Van Vliet [2012] have used the generalized lambda distribution (GLD) for fitting non-normal financial data in part because of its ease of use in simulation. The four parameter GLD was developed by Ramberg and Schmeiser (RS) [1974]. The RS generalization is most often defined by its quantile function *Q(p)*.

$$Q(p) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2} \tag{21}$$

Where λ_1 is a location parameter, λ_2 the scale parameter, and λ_3 and λ_4 determine the shape. p is the probability $0 < p < 1$. To generate the data, we set the parameters shown in Exhibit 7.

Parameters	Values
λ_1	.75
λ_2	.50
λ_3	.30
λ_4	.05

Exhibit 7. Lambda parameters for simulated HFT return data

These parameters generated non-normal 2000 data points with the descriptive statistics shown in Exhibit 8.

Statistic	Value
Mean	0.37633
Standard Deviation	0.43183

Skewness	-0.54204
Kurtosis	-0.16027

Exhibit 8. Descriptive statistics of the simulated HFT data

The histogram of the data is shown in Exhibit 9. Clearly, the distribution is non-normal, having a large left tail.

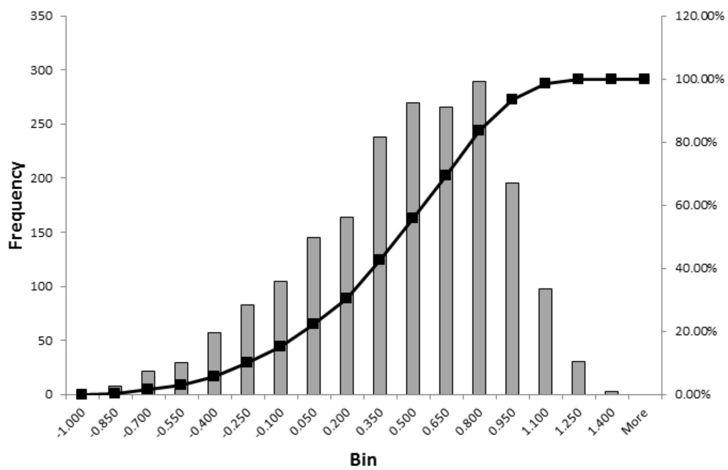

Exhibit 9. Histogram of simulated HFT returns

Using $n = 5$, SPC generated the statistics on the resulting 400 samples shown in Exhibit 10:

Statistic	Value
$\overline{\overline{\pi}}$	0.37633
$\overline{R}$	1.00102
Estimate of $3 \cdot s_{\overline{\pi}}$	0.55757

Exhibit 10. Sample statistics

First, we look at the Sharpe Ratio of this system (which ignores stability). Assuming it requires \$1 M of margin to operate. The return on this investment is:

$$r_i = \ln\left(\frac{\text{Margin} + \text{Cash Flow}}{\text{Margin}}\right) \tag{22}$$

Thus, the average return on this system per second is expected to be:

$$\bar{r} = \ln\left(\frac{1M + .37633}{1M}\right) = .000037633\%\ \text{per second} \tag{23}$$

The standard deviation of one second returns turns about to be:

$$\sigma_r = 0.000043183\% \tag{24}$$

To annualize these numbers, we multiply by 5,896,800 (assuming 252 trading days of 6.5 hours), the number of seconds in a year, and the square root of that number respectively, so that:

$$\begin{aligned} r_{annual} &= 221.9\% \\ \sigma_{annual} &= .10486\% \end{aligned} \tag{25}$$

Thus, the Sharpe ratio for this system would be:

$$Sharpe = \frac{r_{annual}}{\sigma_{annual}} = \frac{221.9\%}{.10486\%} = 2116.20 \tag{26}$$

While the Sharpe Ratio has been calculated correctly, the value of 2116.20 is effectively stating that the system is all profit (221.9% annually) with extremely low risk (.105% annually). Clearly, any fund manager would place all their assets into such an investment. If, on the other hand, we view the time series of profits from a statistical control and capability perspective, we generate different values relative to the Sharpe Ratio, and therefore possibly different decisions as to which HFT systems to invest R&D capital in.

Next, we assess the stability of the underlying distribution by examining the distributions of the sample mean and range. SPC is unconcerned with the shape of the underlying distribution, only its stability. Exhibit 11 shows the distribution of the sample means $\bar{\pi}$. Even at n = 5, we can see that the distribution of the mean approaches normal. Larger values of n increase the tendency toward normality.

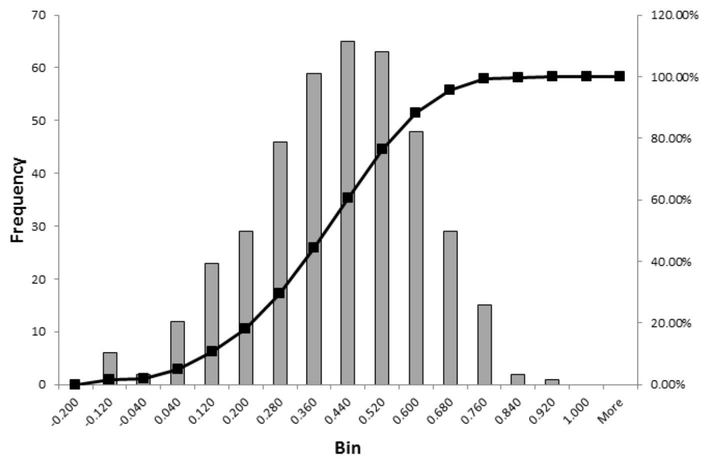

Exhibit 11. Histogram of sample mean returns

As can be seen from the X-bar chart in Exhibit 12, the sample means are in statistical control according to Nelson's [1984] rules.

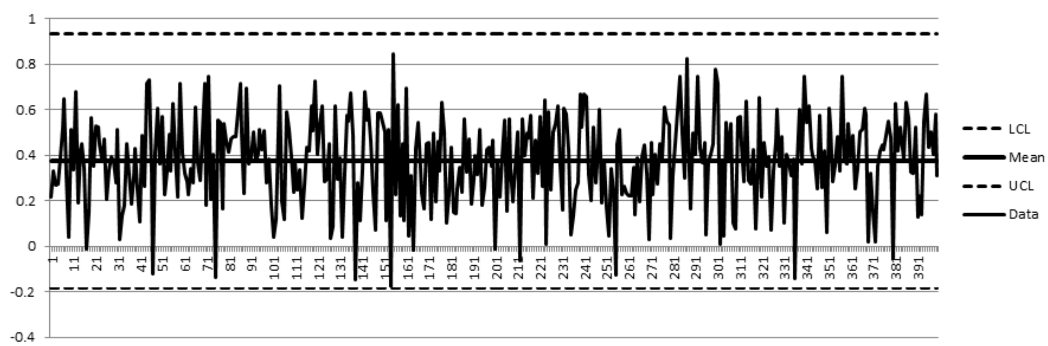

Exhibit 12. X-bar chart of simulated means

The X-bar chart control limits over the data set are:

$$\begin{aligned} LCL_{\bar{\pi}} &= \bar{\bar{\pi}} - A_2 \cdot \bar{R} = .37633 - .557 \cdot 1.00102 = -.18124 \\ UCL_{\bar{\pi}} &= \bar{\bar{\pi}} + A_2 \cdot \bar{R} = .37633 + .557 \cdot 1.00102 = .93389 \end{aligned} \tag{27}$$

The R chart in Exhibit 13 also shows stability.

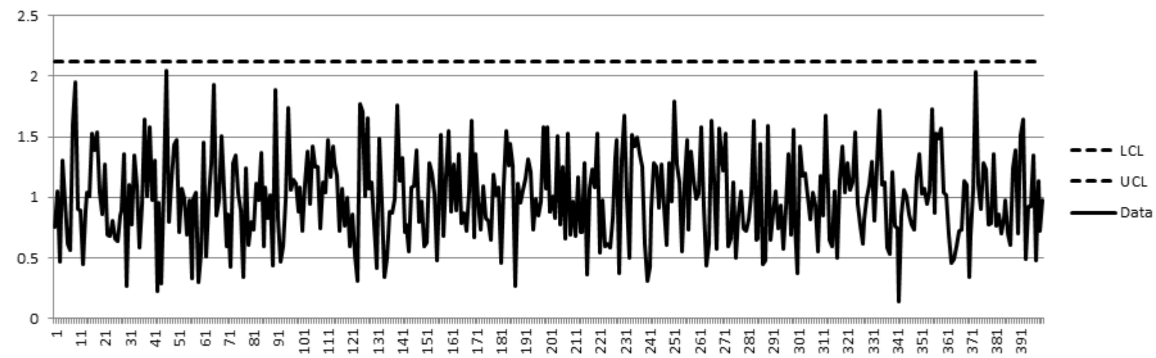

Exhibit 13. R chart of simulated HFT returns

The R chart control limits over the data set are:

$$\begin{aligned} LCL_R &= D_3 \cdot \overline{R} = 0 \cdot 1.00102 = 0 \\ UCL_R &= D_4 \cdot \overline{R} = 2.115 \cdot 1.00102 = 2.11715 \end{aligned} \tag{28}$$

Given that the process is in control, capability can be assessed. Using C_{pl} the LSL is set to the fixed cost allocated per period plus a maximum allowable dollar drawdown times some number of periods:

$$LSL = \text{Expense per Period} - \text{Maximum Drawdown} \cdot \text{Number of Periods} \tag{29}$$

The expense per period is \$.17 and management will carry a losing system for three times the maximum sample loss over the backtest. Thus,

$$LSL = .17 - .17678 \cdot 3 = -.36035 \tag{30}$$

As graphically depicted in Exhibit 14, the C_{pl} for the system is:

$$C_{pl} = \frac{\overline{\overline{\pi}} - LSL}{3 \cdot s_{\overline{\pi}}} = \frac{.37633 - (-.36035)}{.55757} = \frac{.73668}{.55757} = 1.32 \tag{31}$$

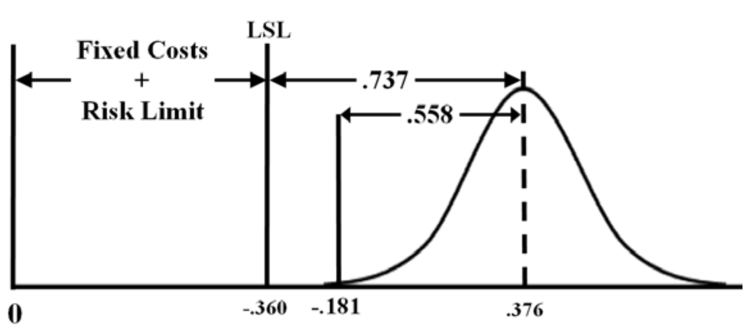

Exhibit 14. C_{pl} for Hypothetical HFT System

As the capability dividing line is 1.33, this system would be barely capable of long term operation due to its dubious ability to consistently pay for its fixed operating costs. C_{pl} leads to the opposite decision relative to the Sharpe Ratio of 2116. The C_{pl} metric gives a better indication about the future profitability of the HFT system.

Empirical Comparison of Sharpe and Capability Metrics

To illustrate the difference between the Sharpe Ratio and the C_{pl}, we use as a proxy for HFT system returns, daily Hedge Fund Research Index[7] returns over the period 1/5/2009 to 9/2/2010. In lieu of actual returns from HFT systems (which are highly secret), we chose the most liquid, non-fee-based return data available. We also considered that the strategies these indexes represent are in some cases similar to HFT strategies (e.g. long-short equity). Lastly, these returns imply alpha driven, as opposed to beta driven, returns.

To assess capability of these return streams using C_{pl}, we calculated cash flows net of a weekly fixed cost allocation of .10. We set the LSL to the maximum allowable dollar drawdown of capital invested:

$$LSL = -Initial\ Capital \cdot Drawdown \cdot \frac{Number\ of\ Weeks}{52}$$

We set initial capital to 100, the drawdown of 2.50, and number of weeks to 4. Thus, the stop-loss on the trading system is -.1923. Furthermore, we reset the amount of trading capital quarterly to 100 by subtracting the accumulated trading profits. Table 2 shows the corresponding values this scenario generates for the Sharpe and C_{pl} ratios.

Index Description	Sharpe	C_{pl}

[7] HFRX® and Hedge Fund Research™ are registered trademarks of Hedge Fund Research, Inc.

HFRX GL: Global Hedge Fund Index	.159	.460
HFRX NA: North America Index	.038	.451
HFRX EMN: Equity Market Neutral Index	-.064	.522
HFRX M: Macro Index	-.070	.326
HFRX RVA: Relative Value Multi-strategy Index	.372	.160
HFRX EH: Equity Hedge Index	.061	.272
HFRX MREG: Multi-Region Index	-.033	.306
HFRX SDV: Systematic Diversified Index	-.021	.238
HFRX SS: Special Situations Index	.168	.304

Exhibit 15: Comparing Sharpe and C_{pl}

Given the data in Exhibit 15, if we view the time series of profits from the perspective of an ATF, and apply C_{pl} we generate different values relative to the Sharpe Ratio, and therefore possibly different decisions as to which HFT systems to invest R&D capital in. This shows that a strategy that makes for a good hedge fund does not necessarily make for a good HFT system at an ATF, or vice versa. None the indexes achieve the 1.33 C_{pl} threshold. This may explain the large number of hedge fund failures since 2008 relative to the few number of ATF failures.

Capability Studies

From preferences for risk and return, we can infer specification limits against which processes potentially capable of satisfying those preferences can be assessed. Then, financial engineers can R&D automated strategies capable of generating return processes that meet those specifications. Thus, research and development of a trading system is in part a capability study, as Bothe [2001] defines it "a formal procedure for undertaking a systematic analytical investigation to reliably assess a process's ability to consistently meet a specific requirement." Since financial innovation is often a euphemism for new ways to take advantage of customers, capability demands that new products perform to customers specifications, not in conflict with them.

Without capability, a trading system may be profitable, even in control, but yet fail to be profitable leading to operational failure of the ATF or hedge fund. A process capability study should be performed whenever:

- The capability of a trading system to meet investor specifications needs to be determined. That is, when a new trading strategy is backtested.
- Specified tolerances are assessed against the observed variability of the trading system output process.
- Changes and/or improvements to an existing trading system need to be evaluated. As Steiner et al. [1997] note, "the effect of a process change can be assessed by comparing capability indices calculated before and after the change."

Conclusion

The current metrics of finance do not correctly quantify the ability of an HFT system to meet an ATF's specifications for ROI. The Sharpe Ratio quantifies expected return per unit of risk. Because the Sharpe Ratio is not path dependent, it does not take into consideration drawdown limits or capital resetting. The C_{pl} is a better formula for ATFs to assess the profitability of HFTs per accounting period.

Given statistical control and capability as decision framework for research and development of HFT systems, we can match the firm's specifications with systems capable of generating sufficient profit to cover fixed costs. Without a capability study, ATFs cannot know if an HFT system will cover its own R&D costs. If the C_{pl} for a given system is too low, the firm can choose to reduce costs, shift specifications, or reduce variation through process improvement. With this methodology, ATFs can evaluate portfolios of HFT systems to determine the best mix of investments. This problem is similar to the job shop framework, which has been used for decades in manufacturing.

This framework can be expanded similar to manufacturing where alternative distributions can be used to create process capability measurements to insure a system can meet a predefined reliability.

References

AIAG Automotive Industry Action Group. 2005. *Statistical Process Control*. 2nd edition.

Aldridge, Irene. 2010. "How Profitable Are High-Frequency Trading Strategies?" Accessed on Sept. 21, 2010 at: www.irenealdridge.com/how-profitable-are-high-frequency-trading-strategies/

American Society for Quality (ASQ). 2008. See http://www.asq.org/glossary/p.html.

American Society for Testing and Materials (ASTM). 1976. *ASTM Manual on Presentation of Data and Control Chart Analysis.* ASTM Publication STP15D, 134-135.

Bilson, John, Andrew Kumiega and Ben Van Vliet. 2010. "Trading Model Uncertainty and Statistical Process Control." *Journal of Trading*. 5, (3).

Bock, D. 2008. "Aspects on the control of false alarms in statistical surveillance and the impact on the return of financial decision systems." *Journal of Applied Statistics.* 35, (2), 213-227.

Bothe, D.R. 2001. *Measuring Process Capability*. Cedarburg: Landmark Publishing, Inc.

Brogaard, Jonathan A. 2010. "High frequency trading and its impact on market quality." Northwestern University. Available at SSRN: http://ssrn.com/abstract=1641387 or http://dx.doi.org/10.2139/ssrn.1641387

Chen, Tao, Julian Morris, and Elaine Martin. 2006. "Probability density estimation via an infinite Gaussian mixture model: application to statistical process monitoring." *Journal of the Royal Statistical Society*. 55, (5), 699–715.

Cooper, Rick and Ben Van Vliet. 2012. "Whole Distribution Statistical Process Control for High Frequency Trading." *Journal of Trading*. 7(2), pp. 57-68.

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

Frisen, M. 2003. "Statistical surveillance: Optimality and methods." *International Statistical Review, 71*, (2), 403-434.

Goetsch, David L. and Stanley B. Davis. 2000. *Quality Management*. Prentice-Hall. p. 590.

Golosnoy, Vasyl and Wolfgang Schmid. 2009. "Statistical Process Control in Asset Management." Chapter in *Applied Quantitative Finance*, 2nd ed. Eds. Wolfgang Härdle, Nikolaus Hautsch, and Luger Overbeck. Springer. Berlin, Germany. pp. 399-416.

Hamilton, Dane. 2006. Quoting Paul Roth, partner in Schulte Roth & Zabel in "S&P may form hedge fund operational risk service." Reuters. Dec. 20.

Hassan, Zia, Andrew Kumiega, and Ben Van Vliet. 2010. "Trading Machines: Using SPC to Assess Performance of Financial Trading Systems." *Quality Management Journal*. 17, (2).

Hittesdorf, Mick. 2009. "Get Real." *Automated Trader*. Q4. pp. 50-56.

Hull, Blair. 2000. "The Future of Trading." *Futures Industry Magazine*. December/January.

International Standards Organization (ISO). 2011. See www.iso.org/iso/home.htm.

Juran, J.M. (ed.) 1988. *Juran's Quality Control Handbook, 4th edition*. Table A.

National Institute of Standards and Technology (NIST). 2012.
See www.itl.nist.gov/div898/handbook/ppc/section4/ppc45.htm

Nelson, Lloyd. 1984. "Technical Aids." *Journal of Quality Technology.* 16 (4), 238-239.

Patnaik, P. B. 1950. "The use of mean range as an estimator of variance in statistical tests." *Biometrika*. 37, 78-87.

Ramberg, J. and Schmeiser, B. 1974. "An approximate method for generating asymmetric random variables." *Communications of the ACM*. 17 (2), 78–82.

Rowe, David. 2003. "Statistical process control." Accessed at www.risk.net. OpRisk & Compliance.

Schmid, Wolfgang, and Taras Bodnar. 2008. "Estimation of optimal portfolio compositions for Gaussian returns." *Statistics & Decisions*. 26, 179-201.

Schmid, Wolfgang, and V. Golosnoy. 2007. "EWMA control charts for monitoring optimal portfolio weights." *Sequential Analysis*. 26, 195-224.

Schmid, Wolfgang, and Y. Okhrin. 2008. "Estimation of optimal portfolio weights." *International Journal of Theoretical and Applied Finance*. 11, 249-276.

Schmid, Wolfgang, and T. Severin. 1998. "Statistical process control and its application in finance." *Contributions to Economics: Risk Measurement, Econometrics, and Neural Networks*, Physica-Verlag, Hei. pp. 83-104.

Sharpe, William. 1994. "The Sharpe Ratio." *Journal of Portfolio Management*. 21, (1), 49-58.

Shewhart, Walter A. 1986. *Statistical Method from the Viewpoint of Quality Control*. Dover Publications, New York. p. 37.

Sortino, F. and R. van der Meer. 1991. "Downside Risk: Capturing What's at Stake." *Journal of Portfolio Management*. 17, 27-31.

Steiner, S., B. Abraham, and J. MacKay. 1997. "Understanding Process Capability Indices." Institute for Improvement in Quality and Productivity research report. University of Waterloo, Ontario, Canada.

Thissen, U., H. Swierenga, . P. de Weijer, R. Wehrens1, W. J. Melssen1, and L. M. C. Buydens. 2005. "Multivariate statistical process control using mixture modeling." *Journal of Chemometrics*, 19, (1), 23–31.

Tomasson, H. 2009. "Statistical models in finance." In M. Frisen (Ed.), *Financial Surveillance.* New York: John Wiley & Sons.

Woodall, William and Montgomery, Douglas. 2000. "Using Ranges to Estimate Variability. " *Quality Engineering*. 13, (2).

Appendix A

Nelson [1984] defines eight signals in total, each with probability approximately equal to that of 3 standard deviations (.135%):

1. Any single measurement above or below the 3 standard deviation UCL or LCL.
2. 9 points in a row on one side of the mean.
3. 6 points in a row increasing or decreasing.
4. 14 points in a row toggling back and forth between increasing and decreasing.
5. 2 out of 3 points in a row more than 2 standard deviations from the mean in the same direction.
6. 4 out of 5 points in a row more than 1 standard deviation from the mean in the same direction.
7. 15 points in a row all within plus or minus 1 standard deviation.
8. 8 points in a row all outside 1 standard deviation in either direction.