

# Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents<sup>†</sup>

Andrei A. Kirilenko and Andrew W. Lo

Over the past four decades, the remarkable growth of the semiconductor industry as embodied by Moore's Law has had enormous effects on society, influencing everything from household appliances to national defense. The implications of this growth for the financial system has been profound, as well. Computing has become faster, cheaper, and better at automating a variety of tasks, and financial institutions have been able to greatly increase the scale and sophistication of their services. At the same time, population growth combined with the economic complexity of modern society has increased the demand for financial services. After all, most individuals are born into this world without savings, income, housing, food, education, or employment; all of these necessities require financial transactions.

It should come as no surprise then that the financial system exhibits a Moore's Law of its own—from 1929 to 2009 the total market capitalization of the US stock market has doubled every decade. The total trading volume of stocks in the Dow Jones Industrial Average doubled every 7.5 years during this period, but in the most recent decade, the pace has accelerated: now the doubling occurs every 2.9 years, growing almost as fast as the semiconductor industry. But the financial industry

■ *Andrei A. Kirilenko is the Professor of the Practice of Finance, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts. During 2010–2012, Kirilenko served as the Chief Economist of the US Commodity Futures Trading Commission, Washington, DC. Andrew W. Lo is the Charles E. and Susan T. Harris Professor, Director of the Laboratory for Financial Engineering at Sloan School of Management, and a Principal Investigator at the Computer Science and Artificial Intelligence Laboratory, all at the Massachusetts Institute of Technology, Cambridge, Massachusetts. Lo is also Chairman and Chief Investment Strategist, AlphaSimplex Group, LLC, an investment management firm. Their email addresses are ak67@mit.edu and alo@mit.edu.*

<sup>†</sup>To access the disclosure statements, visit <http://dx.doi.org/10.1257/jep.27.2.51>.

differs from the semiconductor industry in at least one important respect: human behavior plays a more significant role in finance. As the great physicist Richard Feynman once said, “Imagine how much harder physics would be if electrons had feelings.” While financial technology undoubtedly benefits from Moore’s Law, it must also contend with Murphy’s Law, “whatever can go wrong will go wrong,” as well as its technology-specific corollary, “whatever can go wrong will go wrong faster and bigger when computers are involved.”

A case in point is the proliferation of high-frequency trading in financial markets, which has raised questions among regulators, investors, and the media about how this technology-powered innovation might affect market stability. Largely hidden from public view, this relatively esoteric and secretive cottage industry made headlines on May 6, 2010, with the so-called “Flash Crash,” when the prices of some of the largest and most actively traded companies in the world crashed and recovered in a matter of minutes. Since then, a number of high-profile technological malfunctions, such as the delayed Facebook initial public offering in March 2012 and an electronic trading error by Knight Capital Group in August 2012 that cost the company \$400+ million, have only added fuel to the fire. Algorithmic trading—the use of mathematical models, computers, and telecommunications networks to automate the buying and selling of financial securities—has arrived, and it has created new challenges as well as new opportunities for the financial industry and its regulators.

Algorithmic trading is part of a much broader trend in which computer-based automation has improved efficiency by lowering costs, reducing human error, and increasing productivity. Thanks to the twin forces of competition and innovation, the drive toward “faster, cheaper, and better” is as inexorable as it is profitable, and the financial industry is no stranger to such pressures. However, what has not changed nearly as much over this period is the regulatory framework that is supposed to oversee such technological and financial innovations. For example, the primary set of laws governing the operation of securities exchanges is the Securities Exchange Act of 1934, which was enacted well before the arrival of digital computers, electronic trading, and the Internet. Although this legislation has been amended on many occasions to reflect new financial technologies and institutions, it has become an increasingly cumbersome patchwork quilt of old and new rules based on increasingly outdated principles, instead of an integrated set of modern regulations designed to maintain financial stability, facilitate capital formation, and protect the interests of investors. Moreover, the process by which new regulations are put in place or existing regulations are amended is slow and subject to the vagaries of politics, intense lobbying by the industry, judicial challenges, and shifting public sentiment, all of which may be particularly problematic for an industry as quickly evolving and highly competitive as financial services.

In this paper, we provide a brief survey of algorithmic trading, review the major drivers of its emergence and popularity, and explore some of the challenges and unintended consequences associated with this brave new world. There is no doubt that algorithmic trading has become a permanent and important part of the financial landscape, yielding tremendous cost savings, operating efficiency, and scalability to every financial market it touches. At the same time, the financial system has become

much more of a *system* than ever before, with globally interconnected counterparties and privately-owned and -operated infrastructure that facilitates tremendous integration during normal market conditions, but which spreads dislocation rapidly during periods of financial distress. A more systematic and adaptive approach to regulating this system is needed, one that fosters the technological advances of the industry while protecting those who are not as technologically advanced. We conclude by proposing “Financial Regulation 2.0,” a set of design principles for regulating the financial system of the Digital Age.

## A Brief Survey of Algorithmic Trading

Three developments in the financial industry have greatly facilitated the rise of algorithmic trading over the last two decades. The first is the fact that the financial system is becoming more complex over time, not less. Greater complexity is a consequence of general economic growth and globalization in which the number of market participants, the variety of financial transactions, the levels and distribution of risks, and the sums involved have also grown. And as the financial system becomes more complex, the benefits of more highly developed financial technology become greater and greater and, ultimately, indispensable.

The second development is the set of breakthroughs in the quantitative modeling of financial markets, the “financial technology” pioneered over the past three decades by the giants of financial economics: Black, Cox, Fama, Lintner, Markowitz, Merton, Miller, Modigliani, Ross, Samuelson, Scholes, Sharpe, and others. Their contributions laid the remarkably durable foundations on which modern quantitative financial analysis is built, and algorithmic trading is only one of the many intellectual progeny that they have fathered.

The third development is an almost parallel set of breakthroughs in computer technology, including hardware, software, data collection and organization, and telecommunications, thanks to Moore’s Law. The exponential growth in computing power per dollar and the consequences for data storage, data availability, and electronic interconnectivity have irrevocably changed the way financial markets operate.

A deeper understanding of the historical roots of algorithmic trading is especially important for predicting where it is headed and formulating policy and regulatory recommendations that affect it. In this section, we describe five major developments that have fueled its growing popularity: quantitative models in finance, the emergence and proliferation of index funds, arbitrage trading activities, the push for lower costs of intermediation and execution, and the proliferation of high-frequency trading.

### Quantitative Finance

The most obvious motivation for algorithmic trading is the impressive sequence of breakthroughs in quantitative finance that began in the 1950s with portfolio optimization theory. In his pioneering PhD thesis, Harry Markowitz (1952) considered how an investor should allocate his wealth over  $n$  risky securities so as to maximize his expected utility of total wealth. Under some assumptions, he shows

that this is equivalent to maximizing the expected value of a quadratic objective function of the portfolio's return which, in turn, yields a mean–variance objective function. The solution to this well-posed optimization problem may be considered the very first algorithmic trading strategy—given an investor's risk tolerance and the means, variances, and covariances of the risky assets, the investor's optimal portfolio is completely determined. Thus, once a portfolio has been established, the algorithmic trading strategy—the number of shares of each security to be bought or sold—is given by the difference between the optimal weights and the current weights. More importantly, portfolio optimization leads to an enormous simplification for investors with mean–variance preferences: all such investors should be indifferent between investing in  $n$  risky assets and investing in one specific portfolio of these  $n$  assets, often called the “tangency portfolio” because of the geometry of mean–variance analysis.<sup>1</sup> This powerful idea is often called the “Two-Fund Separation Theorem” because it implies that a riskless bond and a single mutual fund—the tangency portfolio—are the only investment vehicles needed to satisfy the demands of all mean–variance portfolio optimizers, an enormous simplification of the investment problem.

The second relevant milestone in quantitative finance was the development of the Capital Asset Pricing Model (CAPM) by Sharpe (1964), Lintner (1965), and Mossin (1966) in the 1960s, and the intense empirical and econometric investigations it launched in the following two decades. These authors took portfolio optimization as their starting point and derived a remarkably simple yet powerful result: if all investors hold the same tangency portfolio, albeit in different dollar amounts, then this tangency portfolio can only be one portfolio: the portfolio of all assets, with each asset weighted according to its market capitalization. In other words, the tangency portfolio is the total market portfolio. This more-specific form of the Two-Fund Separation Theorem was a critical milestone in both academia and industry, generating several new directions of research as well as providing the foundations for today's trillion-dollar index-fund industry (discussed in the next section).

The third milestone occurred in the 1970s and was entirely statistical and computational. To implement portfolio optimization and the Capital Asset Pricing Model, it was necessary to construct timely estimates of the expected returns and the covariance matrix of all traded equities. This seemed like an impossible task in the 1970s because of the sheer number of securities involved—almost 5,000 stocks on the New York, American, and NASDAQ Stock Exchanges—and the numerical computations involved in estimating all those parameters. For example, a 5,000-by-5,000 covariance matrix contains 12,497,500 unique parameters. Moreover, because the maximum rank of the standard covariance-matrix estimator is simply the number of time series observations used, estimates of this 5,000-by-5,000

<sup>1</sup> The set of mean-variance-optimal portfolios forms a curve when plotted in mean–variance space, and the portfolio that allows mean–variance optimizers to achieve the highest expected return per unit of risk is attained by the portfolio that is tangent to the line connecting the risk-free rate of return to the curve.

matrix will be “singular” (meaning not invertible) for all sample sizes of daily or monthly stock returns less than 5,000. Singularity is particularly problematic for employing Markowitz-type mean–variance optimization algorithms which depend on the inverse of the covariance matrix.

These challenges were met elegantly and decisively in the 1970s by Rosenberg's (1974) linear multifactor risk model in which individual stock returns were assumed to be linearly related to a smaller number  $K$  of common “factors.” The existence of such a linear relation implies that the total number of unknown covariance-matrix parameters to be estimated is now  $nK + K(K + 1)/2 + n$  instead of  $n(n - 1)/2$ , which increases linearly in  $n$  instead of as  $n^2$ . In contrast to the 12,497,500 unique parameters in the case of 5,000 stocks, a linear factor model with 50 factors requires only 256,275 parameters—a 50-fold reduction!

Rosenberg took his ideas one step further in 1975 by founding a commercial venture—Barr Rosenberg and Associates, or Barra—that provided clients with timely estimates of covariance matrices for US equities, as well as portfolio optimization software so they could implement Markowitz-style mean-variance-optimal portfolios. It is no exaggeration that Barra's software platform was largely responsible for popularizing algorithmic equity trading—particularly portfolio optimization—among institutional investors and portfolio managers throughout the world. More frequent estimation of optimal portfolios also meant that portfolio managers needed to trade more frequently. As a result, trading volumes began to rise disproportionately faster than the number of newly created securities.

The fourth milestone came in 1973 with the publication of the Black and Scholes (1973) and Merton (1973) articles on the pricing of options and other derivative securities. Although these two seminal articles contained the celebrated Black–Scholes/Merton option-pricing formula—for which Merton and Scholes shared the Nobel prize in economics in 1997—an even more influential idea to come out of this research program was Merton's (1973) insight that under certain conditions, the frequent trading of a small number of long-lived securities can create new investment opportunities that would otherwise be unavailable to investors. These conditions—now known collectively as *dynamic spanning* or *dynamically complete markets*—and the corresponding asset-pricing models on which they are based, have generated a rich literature and a multi-trillion-dollar derivatives industry. The financial services industry has subsequently written hundreds of cookbooks with thousands of recipes describing how to make complex and sometimes exotic dishes such as swaps, caps, collars, swaptions, knock-out and rainbow options, and many others out of simple ingredients—stocks and bonds—by combining them in prescribed quantities and stirring (trading) the mixture frequently to make them as appetizing as possible to investors.

### Index Funds

One of the most enduring legacies of Markowitz, Sharpe, Lintner, Tobin, and Mossin is the idea of “passive” investing through index funds. The recipe for an index fund is now well-known: define a collection of securities by some set of easily

observable attributes, construct a portfolio of such securities weighted by their market capitalizations, and add and subtract securities from this collection from time to time to ensure that the portfolio continues to accurately reflect the desired attributes.

The original motivation behind fixing the set of securities and value-weighting them was to reduce the amount of trading needed to replicate the index in a cash portfolio. Apart from the occasional index addition and deletion, a value-weighted portfolio need never be rebalanced since the weights automatically adjust proportionally as market valuations fluctuate. These “buy-and-hold” portfolios are attractive not only because they keep trading costs to a minimum, but also because they are simpler to implement from an operational perspective. It is easy to forget the formidable challenges posed by the back-office, accounting, and trade reconciliation processes for even moderate-sized portfolios in the days before personal computers, automated order-generating engines, and electronic trading platforms. A case in point is the precursor to the very first index mutual fund, a \$6 million equal-weighted portfolio of 100 New York Stock Exchange (NYSE) equities managed by Wells Fargo Bank for Samsonite’s pension fund starting in 1969. An equal-weighted portfolio—a portfolio in which equal dollar amounts are invested in each security—does not stay equally weighted as prices fluctuate, and the process of rebalancing a portfolio of 100 stocks back to equal weighting at the end of each month was such an operational nightmare back then that the strategy was eventually abandoned in favor of a value-weighted portfolio (Bogle 1997). Since then, most investors and managers equate “passive” investing with low-cost, static, value-weighted portfolios (portfolios in which the dollar amount invested in each security is proportional to the total market capitalization of the company issuing that security).

However, with the many technological innovations that have transformed the financial landscape over the last three decades, the meaning of passive investing has changed. A functional definition of passive investing is considerably more general: an investment process is “passive” if it does not require any discretionary human intervention—that is, if it is based on a well-defined and transparent algorithm. Such a definition decouples active investing from active trading; today, a passive investor may be an active trader to minimize transaction costs, manage risks more adroitly, participate in new investment opportunities such as initial public offerings, or respond more quickly to changing objectives and market conditions. Moreover, new investment products such as target-date funds, exchange-traded funds, and strategy indexes such as 130/30, currency carry-trade, hedge-fund replication, and trend-following futures strategies are growing in popularity and acceptance among passive investors despite the active nature of their trading, thanks to the automation facilitated by algorithms. At the same time, the much more active participation of investors has created new technological challenges for the issuers of new financial instruments. We provide an example of this later in this paper when discussing the Facebook and BATS initial public offerings.

### **Arbitrage Trading**

Arbitrage strategies are among the most highly visible applications of algorithmic trading over the past three decades. These strategies are routinely implemented by

broker-dealers, hedge funds, and institutional investors with the sole objective of generating profits with lower risk than traditional investments. Arbitrage trading is as old as financial markets, but using algorithms to identify and exploit arbitrage-trading opportunities is a thoroughly modern invention, facilitated by the use of computers, applications of probability and statistics, advances in telecommunications, and the development of electronic markets.

The most common form of algorithmic arbitrage trading is a transaction that attempts to exploit situations where two securities that offer identical cashflows have different market prices. The law of one price implies that such opportunities cannot persist, because traders will quickly construct arbitrage portfolios in which the lower-priced asset is purchased and the higher-priced asset is sold (or shorted) yielding a positive and riskless profit by assumption (because the underlying cashflows of the two securities are assumed to be identical). More generally, an arbitrage strategy involves constructing a portfolio of multiple securities such that the combined cashflows are riskless, and if the cost of constructing such a portfolio is nonzero for reasons other than trading costs, then there exists a version of the arbitrage strategy that generates positive riskless profits, which is a definition of an arbitrage opportunity.

Violations of the law of one price have been routinely exploited in virtually every type of financial market ranging from highly liquid securities such as foreign currencies and exchange-traded futures to highly illiquid assets such as real estate and emerging-market debt. However, in most practical settings, pure arbitrages do not exist because there are subtle differences in securities that cause their prices to differ despite seemingly identical cashflows, like differences in transactions costs, liquidity, or credit risk. The fact that hedge funds like Long-Term Capital Management have suffered severe losses from arbitrage strategies implies that such strategies are not, in fact, pure arbitrages or completely riskless profit opportunities.

However, if the statistical properties of the arbitrage portfolios can be quantified and managed, the risk/reward profiles of these strategies might be very attractive to investors with the appropriate tolerance for risk. These considerations led to the development of a new type of proprietary trading strategy in the 1980s, so-called “statistical arbitrage strategies” in which large portfolios of equities were constructed to maximize expected returns while minimizing volatility. The risks embedded in statistical arbitrage strategies are inherently different from market risk because arbitrage portfolios are, by construction, long and short, and hence they can be profitable during market downturns. This property provides attractive diversification benefits to institutional investors, many of whom have the majority of their assets in traditional long-only portfolios of stocks and bonds. The details of statistical arbitrage strategies are largely unknown because proprietary traders cannot patent such strategies, and thus they employ trade secrecy to protect their intellectual property. However, simple versions of such strategies have been proposed and studied by Lehmann (1990), Lo and MacKinlay (1990), and Khandani and Lo (2007, 2011), and we provide a more detailed exposition of them in the sections that follow.

Apart from the attractive risk/reward profile they offer to investors and portfolio managers, arbitrage strategies play two other critical roles in the financial system: liquidity provision and price discovery. The presence of arbitrageurs almost always increases the amount of trading activity, and larger volume is often interpreted as greater liquidity, meaning that investors often can buy or sell securities more quickly, in larger quantities, and with lower price impact. Moreover, because arbitrage trading exploits temporary mispricings, it tends to improve the informational efficiency of market prices (assuming that the mispricings are genuine). However, if arbitrageurs become too dominant in any given market, they can create systemic instabilities. We provide an example of this in our later discussion of the so-called “Quant Meltdown” in August 2007.

### **Automated Execution and Market Making**

Algorithmic trading is also central to the automation of large buy and sell orders of publicly traded securities such as exchange-traded equities. Because even the most actively traded stocks have downward-sloping demand curves over a short period of time, executing a large “parent” order in a single transaction is typically more costly than breaking up the order into a sequence of smaller “child” orders. The particular method for determining the timing and sizes of these smaller orders is called an “execution strategy,” and optimal execution strategies can be derived by specifying an objective function and a statistical model for stock-price dynamics.

For example, Bertsimas and Lo (1998) consider the problem of minimizing the expected cost of acquiring  $S_0$  shares of a given stock over  $T$  discrete trades. If  $S_0$  is a small number, like a “round lot” of 100 shares, then the entire block can be executed in a single trade. However, institutional investors must often trade hundreds of thousands of shares as they rebalance multi-billion-dollar portfolios. By modeling the short-run demand curve for each security to be traded—also known as the “price-impact function”—as well as other state variables driving price dynamics, Bertsimas and Lo (1998) are able to derive the expected-cost-minimizing sequence of trades as a function of those state variables using stochastic dynamic programming. These automated execution algorithms can be computationally quite complex for large portfolios of diverse securities, and are ideally suited for automation because of the accuracy and significant cost savings that they offer, especially when compared to human traders attempting to do this manually. However, under certain market conditions, automated execution of large orders can create significant feedback-loop effects that cascade into systemic events as in the case of the so-called “Flash Crash” of May 6, 2010, which we discuss in the next section.

A closely related activity to automated execution is market making, when an intermediary participates in buying and selling securities to smooth out temporary imbalances in supply and demand because buyers and sellers do not always arrive at the same time. A participant of a trading venue, typically a broker-dealer, can voluntarily apply to register as a designated market maker on a security-by-security basis. To qualify, a potential market maker must satisfy certain net capital requirements and be



willing to provide continuous two-sided quotes during trading hours, which means being willing to purchase securities when the public wishes to sell, and to sell securities when the public wishes to buy. Registration does not guarantee profits or customer order flow; it only provides lower trading fees and a designation that can help attract orders from potential customers. Note that participants need not register to function as market makers. Market making is a risky activity because of price fluctuations and adverse selection—prices may suddenly move against market makers and force them to unwind their proprietary positions at a loss. To protect themselves against possible losses, market makers demand compensation, typically in the form of a spread that they charge buyers over sellers known as the “bid–offer spread.”

A typical market-making algorithm submits, modifies, and cancels limit orders to buy and sell a security with the objective of regularly capturing the bid–offer spread and liquidity rebates (payments made to participants who provide liquidity to the market), if any, while also continuously managing risky inventory, keeping track of the demand–supply imbalance across multiple trading venues, and calculating the costs of doing business, including trading and access fees, margin requirements, and the cost of capital. As a result, automation of the trading process means that the rewards from market making activities accrue not necessarily to those who register with the exchanges as their designated market makers, but to those with the best connectivity, best algorithms, and best access to customer order flow.

The central issue with respect to algorithmic market making is whether this activity has improved overall market quality, thus allowing investors to raise capital and manage risks more efficiently. To analyze this issue, Hendershott, Jones, and Menkveld (2011) study the introduction of “autoquoting”—the automated transmission of improved terms of trade for larger trade sizes—that was introduced in 2003 on the New York Stock Exchange. Autoquoting did favor algorithmic traders because they could receive valuable information about changes in the order book faster than humans, but did not otherwise alter the advantages and obligations of the NYSE-designated specialists. The authors show that the introduction of autoquoting increased the informativeness of quoted prices, narrowed bid–offer spreads, and reduced the degree of adverse selection associated with trading. At the same time, automation makes technological glitches in the ultracompetitive business of market making extremely costly. We illustrate this point later in the paper with an example of an algorithmic market maker whose fate was sealed minutes after it launched a new trading algorithm.

### **High-Frequency Trading**

A relatively recent innovation in automated financial markets is a blend of technology and hyperactive trading activity known as “high-frequency trading”—a form of automated trading that takes advantage of innovations in computing and telecommunication to consummate millions upon millions of trades per day. High-frequency trading is now estimated to account for 40 to 60 percent of all trading activity across the universe of financial markets, including stocks, derivatives, and liquid foreign currencies (Tabb 2012). However, the number of entities that engage in high-frequency trading is reportedly quite small and

what is known about them is not particularly illuminating. Baron, Brogaard, and Kirilenko (2012) examine high-frequency trading in the “E-mini S&P 500 futures contract,” an extremely popular futures contract on the Standard & Poor’s 500 index that owes its name to the fact that it is electronically traded and in smaller denominations than the traditional S&P 500 index futures contract. Their study finds that high-frequency traders (as designated by their trading activity) earn large, persistent profits while taking very little risk. In contrast to a number of public claims, high-frequency traders do not as a rule engage in the provision of liquidity like traditional market makers. In fact, those that do not provide liquidity are the most profitable and their profits increase with the degree of “aggressive,” liquidity-taking activity.

High-frequency trading is a recent innovation in financial intermediation that does not fit neatly into a standard liquidity-provision framework. While the net contribution of high-frequency trading to market dynamics is still not fully understood, their mere presence has already shaken the confidence of traditional market participants in the stability and fairness of the financial market system as a whole. Recent revelations of manipulative trading activity, discussed later in this paper, have only added fuel to the debate about the usefulness of high-frequency trading.

## Ghosts in the Machine

As in every other industry that has reduced its costs via automation, the financial services industry has also been transformed by technology. In the modern trading environment, an investor’s trading strategy—whether to liquidate a large position, to make markets, or to take advantage of arbitrage opportunities—is typically executed by an automated trading system. Such systems are responsible for the initiation of trading instructions, communication with one or more trading platforms, the processing of market data, and the confirmation of trades. But technology that supersedes human abilities often brings unintended consequences, and algorithmic trading is no exception. A chainsaw allows us to clear brush much faster than a hand saw, but chainsaw accidents are much more severe than handsaw accidents. Similarly, automated trading systems provide enormous economies of scale and scope in managing large portfolios, but trading errors can now accumulate losses at the speed of light before they’re discovered and corrected by human oversight. Indeed, the enhanced efficiency, precision, and scalability of algorithms may diminish the effectiveness of those risk controls and systems safeguards that rely on experienced human judgment and are applied at human speeds. While technology has advanced tremendously over the last century, human cognitive abilities have been largely unchanged over the last several millennia. Thus, due to the very success of algorithmic trading, humans have been pushed to the periphery of a much faster, larger, and more complex trading environment.

Moreover, in a competitive trading environment, increased speed of order initiation, communication, and execution become a source of profit opportunities for the fastest market participants. Given these profit opportunities, some market

participants, who either trade on their own account or provide execution services to their customers, may choose to engage in a “race to the bottom,” forgoing certain risk controls that may slow down order entry and execution. This vicious cycle can lead to a growing misalignment of incentives as greater profits accrue to the fastest market participants with less-comprehensive safeguards, and may become a significant source of risk to the stability and resilience of the entire financial system.

In this section, we review five specific incidents that highlight these new vulnerabilities created or facilitated by algorithmic trading. We consider them in approximate chronological order to underscore the progression of technology and the changing nature of the challenges that financial innovation can bring.

### **August 2007: Arbitrage Gone Wild**

Beginning on Monday, August 6, 2007, and continuing through Thursday, August 9, some of the most successful hedge funds in the industry suffered record losses. The *Wall Street Journal* reported on August 10, 2007: “After the close of trading, Renaissance Technologies Corp., a hedge-fund company with one of the best records in recent years, told investors that a key fund has lost 8.7% so far in August and is down 7.4% in 2007. Another big fund company, Highbridge Capital Management, told investors its Highbridge Statistical Opportunities Fund was down 18% as of the 8th of the month, and was down 16% for the year. The \$1.8 billion publicly traded Highbridge Statistical Market Neutral Fund was down 5.2% for the month as of Wednesday . . . Tykhe Capital, LLC—a New York-based quantitative, or computer-driven, hedge-fund firm that manages about \$1.8 billion—has suffered losses of about 20% in its largest hedge fund so far this month . . .” (Zuckerman, Hagerty, and Gauthier-Villars 2007). On August 14, the *Wall Street Journal* reported that the Goldman Sachs Global Equity Opportunities Fund “lost more than 30% of its value last week . . .” (Sender, Kelly, and Zuckerman 2007). What made these losses even more extraordinary was the fact that they seemed to be concentrated among quantitatively managed equity market-neutral or “statistical arbitrage” hedge funds, giving rise to the monikers “Quant Meltdown” and “Quant Quake” of 2007.

Because of the secretive nature of hedge funds and proprietary trading firms, no institution suffering such losses was willing to comment publicly on this extraordinary event at the time. To address this lack of transparency, Khandani and Lo (2007) analyzed the Quant Meltdown of August 2007 by simulating the returns of the contrarian trading strategy of Lehmann (1990) and Lo and MacKinlay (1990), and proposed the “Unwind Hypothesis” to explain the empirical facts (see also Goldman Sachs Asset Management 2007; Rothman 2007a, b, c). This hypothesis suggests that the initial losses during the second week of August 2007 were due to the forced liquidation of one or more large equity market-neutral portfolios, primarily to raise cash or reduce leverage, and the subsequent price impact of this massive and sudden unwinding caused other similarly constructed portfolios to experience losses. These losses, in turn, caused other funds to deleverage their portfolios, yielding additional price impact that led to further losses, more deleveraging, and so on. As with Long-Term Capital Management and other fixed-income arbitrage funds in August 1998, the deadly feedback loop of coordinated forced liquidations

leading to the deterioration of collateral value took hold during the second week of August 2007, ultimately resulting in the collapse of a number of quantitative equity market-neutral managers, and double-digit losses for many others.

This Unwind Hypothesis underscores the apparent commonality among quantitative equity market-neutral hedge funds and the importance of liquidity in determining market dynamics. In a follow-on study, Khandani and Lo (2011) used transactions data from July to September 2007 to show that the unwinding likely began in July and centered on securities that shared certain common traits such as high or low book-to-market ratios, because such factors were used by many quantitative portfolio managers attempting to exploit the same empirical anomalies.

In retrospect, we now realize that the Quant Meltdown of August 2007 was only one of a series of crises that hit financial markets during the 2007–2008 crisis period. In fact, after the close of trading on August 9, 2007, central banks from around the world engaged in a highly unusual coordinated injection of liquidity in financial markets, not because of equity markets, but because of a so-called “run on repo” when the interbank short-term financing market broke down (Gorton and Metrick 2012). The summer of 2007 ushered in a new financial order in which the “crowded trade” phenomenon—where everyone rushes to the exit doors at the same time—now applied to entire classes of portfolio strategies, not just to a collection of overly popular securities. In much the same way that a passing speedboat can generate a wake with significant consequences for other ships in a crowded harbor, the scaling up and down of portfolios can affect many other portfolios and investors. Algorithmic trading greatly magnifies the impact of these consequences.

### **May 6, 2010: The Perfect Financial Storm**

In the course of 33 minutes starting at approximately 1:32 pm central time, US financial markets experienced one of the most turbulent periods in their history. The Dow Jones Industrial Average experienced its biggest one-day point decline on an intraday basis in its entire history and the stock prices of some of the world’s largest companies traded at incomprehensible prices: Accenture traded at a penny a share, while Apple traded at \$100,000 per share. Because these dramatic events happened so quickly, the events of May 6, 2010, have become known as the “Flash Crash.”

The subsequent investigation by the staffs of the Commodity Futures Trading Commission (CFTC) and Securities and Exchange Commission (SEC) concluded that these events occurred not because of any single organization’s failure, but rather as a result of seemingly unrelated activities across different parts of the financial system that fed on each other to generate a perfect financial storm (CFTC/SEC 2010). An automated execution algorithm on autopilot, a game of “hot potato” among high-frequency traders, cross-market arbitrage trading, and a practice by market makers to keep placeholder bid–offer “stub quotes” all conspired to create a breathtaking period of extreme volatility.

Kirilenko, Kyle, Samadi, and Tuzun (2011) analyzed the Flash Crash and found that a rapid automated sale of 75,000 E-mini S&P 500 June 2010 stock index futures contracts (worth about \$4.1 billion) over an extremely short time period created a

large order imbalance that overwhelmed the small risk-bearing capacity of financial intermediaries—that is, the high-frequency traders and market makers. After buying the E-mini for about 10 minutes, high-frequency traders reached their critical inventory levels and began to unwind their long inventory quickly and aggressively at a key moment when liquidity was sparse, adding to the downward pressure. High-frequency traders rapidly passed contracts back and forth, contributing to the “hot potato” effect that drove up trading volume, exacerbating the volatility.

Meanwhile, cross-market arbitrage trading algorithms rapidly propagated price declines in the E-mini futures market to the markets for stock index exchange-traded funds like the Standard & Poor's Depository Receipts S&P 500, individual stocks, and listed stock options. According to the interviews conducted by the SEC staff, cross-market arbitrage firms “purchased the E-Mini and contemporaneously sold Standard & Poor's Depository Receipts S&P 500, baskets of individual securities, or other equity index products” (CFTC/SEC 2010). As a result, a liquidity event in the futures market triggered by an automated selling program cascaded into a systemic event for the entire US financial market system.

As the periods during which short-term liquidity providers are willing to hold risky inventory shrink to minutes if not seconds, Flash-Crash-type events—extreme short-term volatility combined with a rapid spike in trading volume—can easily be generated by algorithmic trading strategies seeking to quickly exploit temporarily favorable market conditions.

### **March and May 2012: Pricing Initial Public Offerings in the Digital Age**

On Friday, May 18th, 2012, the social networking pioneer, Facebook, had the most highly anticipated initial public offering in recent financial history. With over \$18 billion in projected sales, Facebook could easily have listed on the NYSE along with larger blue-chip companies like Exxon and General Electric, so Facebook's choice to list on NASDAQ instead was quite a coup for the technology-savvy exchange. Facebook's debut was ultimately less impressive than most investors had hoped, but its lackluster price performance was overshadowed by an even more disquieting technological problem with its opening. An unforeseen glitch in NASDAQ's system for initial public offerings interacted unexpectedly with trading behavior to delay Facebook's opening by 30 minutes, an eternity in today's hyperactive trading environment.

As the hottest initial public offering of the last ten years, Facebook's opening attracted extraordinary interest from investors and was expected to generate huge order flows, but NASDAQ prided itself on its ability to handle high volumes of trades so capacity was not a concern. NASDAQ's IPO Cross software was reportedly able to compute an opening price from a stock's initial bids and offers in less than 40 microseconds (a human eyeblink lasts 8,000 times as long). However, on the morning of May 18, 2012, interest in Facebook was so heavy that it took NASDAQ's computers up to five milliseconds to calculate its opening trade, about 100 times longer than usual. While this extended calculation was running, NASDAQ's order system allowed investors to change their orders up to the print of the opening trade on the tape. But these few extra milliseconds before the print were more

than enough for new orders and cancellations to enter NASDAQ's auction book. These new changes caused NASDAQ's initial public offering software to recalculate the opening trade, during which time even more orders and cancellations entered its book, compounding the problem in an endless circle (Schapiro 2012). As the delay continued, more traders cancelled their previous orders, "in between the raindrops," as NASDAQ's CEO Robert Greifeld rather poetically explained. This glitch created something software engineers call a "race condition," in this case a race between new orders and the print of the opening trade, an infinite loop that required manual intervention to exit, something that hundreds of hours of testing had missed.

Though the initial public offering was scheduled to begin at 11:00 am that morning, delays caused trade opening to occur a half an hour late. As of 10:50 am, traders had not yet received acknowledgements of pre-opening order cancellations or modifications. Even after NASDAQ formally opened the market, many traders still had not received these critical acknowledgements, which created more uncertainty and anxiety (Strasburg, Ackerman, and Lucchetti 2012). By the time the system was reset, NASDAQ's programs were running 19 minutes behind real time. Seventy-five million shares changed hands during Facebook's opening auction, a staggering number, but orders totaling an additional 30 million shares took place during this 19-minute limbo. Problems persisted for hours after opening; many customer orders from both institutional and retail buyers went unfilled for hours or were never filled at all, while other customers ended up buying more shares than they had intended (Strasburg and Bunge 2012; McLaughlin 2012). This incredible gaffe, which some estimates say cost traders \$100 million, eclipsed NASDAQ's achievement in getting Facebook's initial public offering, the third largest IPO in US history.

Less than two months before, another initial public offering suffered an even more shocking fate. BATS Global Markets, founded in 2005 as a "Better Alternative Trading System" to NASDAQ and the NYSE, held its initial public offering on March 23, 2012. BATS operates the third-largest stock exchange in the United States; its two electronic markets account for 11–12 percent of all US equity trading volume each day. BATS was among the most technologically advanced firms in its peer group and the envy of the industry. Quite naturally, BATS decided to list its initial public offering on its own exchange. If an organization ever had sufficient "skin in the game" to get it right, it was BATS, and if there were ever a time when getting it right really mattered, it was on March 23, 2012. So when BATS launched its own initial public offering at an opening price of \$15.25, no one expected its price to plunge to less than a tenth of a penny in a second and a half due to a software bug affecting stocks with ticker symbols from A to BFZZZ, creating an infinite loop that made these symbols inaccessible on the BATS system (Oran, Spicer, Mikolajczak, and Mollenkamp 2012; Schapiro 2012). The ensuing confusion was so great that BATS suspended trading in its own stock, and ultimately cancelled its initial public offering altogether.

As isolated incidents, both the Facebook glitch and the BATS fiasco can be explained as regrettable software errors that extensive testing failed to catch, despite the best efforts of engineers. But two similar incidents in the space of two months

suggest that the problem is more general than a few isolated computer errors. More worrisome is the fact that these glitches are affecting parts of the industry that previously had little to do with technology. After all, initial public offerings have been a staple of modern capitalism since the launch of the Dutch East India Company in 1602. But apparently, launching an initial public offering in a world with microsecond algorithmic trading has become an extremely challenging technical enterprise.

### **August 2012: Trading Errors at the Speed of Light**

On August 1, 2012, a broker-dealer in securities, Knight Capital Group, Inc. experienced what it later called “a technology issue at the open of trading at the NYSE related to a software installation that resulted in Knight sending erroneous orders into the market.” These orders and the unintended trades resulted in a rapid accumulation of positions “unrestricted by volume caps” and, between 9:30 am and 10:00 am eastern time, created significant swings in the share prices of almost 150 stocks (McCrank 2012; see also Telegraph 2012; Schapiro 2012). Unable to void most of these trades by classifying them as “erroneous,” Knight Capital had no choice but to liquidate them in the open market. This liquidation resulted in a \$457.6 million loss for the company, effectively wiping out its capital, causing its stock to lose 70 percent of its value, and forcing it to seek rescuers. After a few nerve-racking days, Knight Capital announced that it had “secured \$400 million in financing,” allowing it to survive. However, the stock of Knight Capital never really recovered, and in December 2012, the company was acquired by GETCO.

Just 42 days prior to the incident, Knight’s chairman and chief executive officer, Mr. Thomas M. Joyce, while testifying before the US House of Representatives Committee on Financial Services, strongly argued in favor of a practice known as *internalization*, in which broker-dealers like Knight are permitted to post prices that are fractions of a penny better than prevailing quotes which are denominated in increments of a penny. For example, if the best bid and offer prices on an organized exchange are \$100.01 and \$100.02, respectively, internalization would allow Knight to post a bid at \$100.011 or an offer at \$100.019. Retail brokers can then legally send a retail customer’s order (like “buy 500 shares”) to Knight rather than to an organized exchange because most markets offer participants “price priority,” which means that a buyer can step to the front of the order queue if that buyer is willing to pay a higher price than all other market participants, including the designated market maker. Sometime during the course of the day, often within seconds, the internalizer would find the inventory it owes to the customer by buying 500 shares of the stock at a lower price, say \$100.001, from another retail customer or at another trading venue such as a dark pools, another internalizer or an organized exchange. It would then pocket the 1 penny difference between the two prices. Internalizers must use their own capital to fill customers’ orders and, due to the Securities and Exchange Commission rule that came out in December 2011 in the wake of the Flash Crash, must have prudent risk management safeguards in place.

The losers from internalization are the organized exchanges that lose order flow and its associated fees to the internalizers. In October 2011, exchanges operated by the NYSE Euronext filed with the Securities and Exchange Commission

proposed a rule to establish a “Retail Liquidity Program,” a way to attract retail order flow to the New York Stock Exchange by allowing them to execute retail orders at sub-penny prices. Several broker-dealers, including Knight Capital, sent comment letters to the SEC arguing against the Retail Liquidity Program. However, after a prolonged comment period, the SEC concluded that “[t]he vast majority of marketable retail orders are internalized by [over-the-counter] market makers, who typically pay retail brokers for their order flow,” while “[e]xchanges and exchange member firms that submit orders and quotations to exchanges cannot compete for marketable retail order flow on the same basis” (SEC 2013). Consequently, on July 3, 2012, the SEC approved the introduction of the Retail Liquidity Program to “promote competition between exchanges and [over-the-counter] market makers.” On July 5, 2012, the NYSE Euronext issued a press release stating that the Retail Liquidity Program would be offered on some of its exchanges for one year on a pilot basis starting on August 1, 2012.

On August 2, 2012, in an interview on Bloomberg TV, Knight’s CEO Joyce stated: “We put in a new bit of software the night before because we were getting ready to trade the NYSE’s Retail Liquidity Program. This has nothing to do with the stock exchange. It had to do with our readiness to trade it. Unfortunately, the software had a fairly major bug in it. It sent into the market a ton of orders, all erroneous, so we ended up with a large error position which we had to sort through the balance of the day. It was a software bug, except it happened to be a very large software bug, as soon as we realized what we had we got it out of the code and it is gone now. The code has been restored. We feel very confident in the current operating environment we’ve reestablished.”

The fall of Knight that began on August 1, 2012, and ended with its firesale acquisition less than six months later was more than just a technological glitch—it was a consequence of the technological arms race that pitted electronic trading platforms against automated broker-dealers in the competition for valuable customer order flow.

### **September 2012: High-Frequency Manipulation**

On September 25, 2012, the Securities and Exchange Commission (2012) issued a cease-and-desist order against Hold Brothers On-Line Investment Services, an electronic broker-dealer who had been involved in manipulative trading activities through offshore high-frequency trading accounts. According to the SEC, from January 2009 to September 2010, these offshore entities engaged in “spoofing” and “layering,” high-tech versions of well-known techniques for manipulating prices and cheating investors. “Spoofing” involves intentionally manipulating prices by placing an order to buy or sell a security and then canceling it shortly thereafter, at which point the spoofer consummates a trade in the opposite direction of the canceled order. “Layering” involves placing a sequence of limit orders at successively increasing or decreasing prices to give the appearance of a change in demand and artificially increase or decrease the price that unsuspecting investors are willing to pay; after a trade is consummated at the manipulated price, the layered limit orders are canceled.



The difference between these scams and the more traditional “pump-and-dump” schemes is the speed and electronic means with which they are conducted. For example, the cease-and-desist order from the Securities and Exchange Commission contains the following illustration of the kind of manipulation that went on for nearly two years (SEC 2012, paragraph 25):

That day, at 11:08:55.152 a.m., the trader placed an order to sell 1,000 GWW shares at \$101.34 per share. Prior to the trader placing the order, the inside bid was \$101.27 and the inside ask was \$101.37. The trader's sell order moved the inside ask to \$101.34. From 11:08:55.164 a.m. to 11:08:55.323 a.m., the trader placed eleven orders offering to buy a total of 2,600 GWW shares at successively increasing prices from \$101.29 to \$101.33. During this time, the inside bid rose from \$101.27 to \$101.33, and the trader sold all 1,000 shares she offered to sell for \$101.34 per share, completing the execution at 11:08:55.333. At 11:08:55.932, less than a second after the trader placed the initial buy order, the trader cancelled all open buy orders. At 11:08:55.991, once the trader had cancelled all of her open buy orders, the inside bid reverted to \$101.27 and the inside ask reverted to \$101.37.

The most notable fact about this narrative is that all of the manipulative activity took place within 839 milliseconds between 11:08:55 and 11:08:56. It is a physical impossibility for any human trader to have accomplished this manually.

In this case, the guilty parties were caught and fined more than \$5.9 million by the Securities and Exchange Commission, the stock exchanges, and the Financial Industry Regulatory Authority, and permanently barred from the securities industry. However, their behavior is unlikely to be an isolated incident, which highlights the challenges facing regulators who need to revamp their surveillance and enforcement practices to be effective in catching the cyber-fraudsters of today.

## **Financial Regulation 2.0**

Although the benefits of automation in financial markets are indisputable, they must be evaluated with two considerations in mind: complexity and human behavior. The software and hardware that control financial markets have become so complex that no individual or group of individuals is capable of conceptualizing all possible interactions that could occur among various components of the financial system. This complexity has created a new class of finance professionals known as “power users,” who are highly trained experts with domain-specific technical knowledge of algorithmic trading. But because technological advances have come so quickly, there are not enough power users to go around. Moreover, the advantages that such expertise confers have raised concerns among those who do not have access to such technology that they are being unfairly and systematically exploited. And the growing interconnectedness of financial markets and institutions has created a

new form of accident: a systemic event, where the “system” now extends beyond any single organization or market and affects a great number of innocent bystanders. The cautionary tales from the previous section are potent illustrations of this new financial order and provide considerable motivation for the global policy debate on the proper market structure in an automated world.

At the heart of this debate is the question of how “continuous” automated financial markets should be and the costs and benefits to the various stakeholders of transacting at faster and faster speeds. Grossman and Miller (1988) offer a stylized equilibrium framework in which the differences in possible market structures boil down to a tradeoff between 1) the costs to different types of intermediaries for maintaining a continuous presence in a market and 2) the benefits to different types of market participants for being able to execute trades as “immediately” as possible.

Automation of the trading process, including computerized algorithmic trading, has drastically reduced the costs to the intermediaries of maintaining a continuous market presence. In fact, intermediaries with the most efficient trading technology and the lowest regulatory burden realized the largest cost savings. As a result, the supply of immediacy has skyrocketed. At the same time, the frequency of technological malfunctions, price volatility spikes, and spectacular frauds and failures of intermediaries has also increased, while the net benefits of immediacy have accrued disproportionately to those who can better absorb the fixed and marginal costs of participating in automated markets. This has frustrated and disenfranchised a large population of smaller, less technologically advanced market participants who are concerned that regulators are unable to fulfill their mandate to protect investor interests, maintain fair and orderly markets, and promote capital formation.

These concerns have been met with a wide range of proposed policy and regulatory responses: do nothing; impose an outright ban on algorithmic—or at least high-frequency—trading; change the rules regarding who can be a designated intermediary and what responsibilities this designation entails; force all trading on exchanges to occur at fixed discrete intervals of time; or, instead of tinkering with “market plumbing,” just introduce a “Tobin tax” on all financial transactions. Each of these proposals contains some merit from the standpoint of at least one set of stakeholders. However, all of the proposals pose difficult tradeoffs.

Doing nothing would allow intermediaries to find more ways to reduce the costs of being continuously present in the market, leading to an even greater supply of immediacy and more efficient trading, but is unlikely to address investors’ concerns about fair and orderly markets.

Banning high-frequency trading might yield more fair and orderly markets in the short run—though the usage of “fair” in this context is somewhat strained given that a segment of market participants is being eliminated by fiat—but may also reduce market liquidity, efficiency, and capital formation as automated trading platforms have become increasingly dependent on high-frequency traders.

Changing the definition and requirements of a designated market maker to include high-frequency traders may also lead to more fair and orderly markets since such designations will prevent them from withdrawing from the market when their

services are needed most. However, such redesignation would also increase the cost to intermediaries of being present in the market due to higher capital requirements, additional compliance costs for each designated market, and greater legal costs by virtue of being a regulated entity. In the short term, this would reduce the supply of immediacy because some traders may find these costs too high to continue making markets.

Forcing all trades to occur at discrete time intervals would concentrate the supply of immediacy, not unlike the periodic batch auctions of many European stock exchanges in the 1990s. How much immediacy would be demanded by different types of market participants, how much they would be willing to pay for it, and how the costs and benefits of concentrated immediacy would be shared among them are questions that must be answered before the welfare effects of this proposal can be evaluated. However, one indication of consumer preferences is the fact that most batch-auction markets have converted to continuous market-making platforms.

Finally, the Tobin tax—a small transaction tax on all financial transactions—has become a mainstay in the public debate on financial markets. In its most recent reincarnation, a variant of the Tobin tax is set to be implemented on January 1, 2014, by 11 members of the European Union including France, Germany, Italy, and Spain (Mehta 2013). However, 15 other members, including the United Kingdom, are strongly opposed to this measure. While this tax will certainly reduce trading activity across the board, and eliminate high-frequency trading altogether in those tax jurisdictions, it will also reduce market liquidity and impair hedging activity. For example, institutional investors often rely on derivative securities such as options and swaps to hedge risk exposures to fluctuations in stock prices, interest rates, and foreign exchange rates. Intermediaries are willing to take the other side of these transactions only if they can mitigate their own risk exposures by dynamically hedging their positions in the underlying stock, bond, and foreign currency markets. Even a small transactions tax would make such dynamic hedging activity impractical (Heaton and Lo 1995). Moreover, a successful implementation of such a tax requires international coordination, otherwise trading activity and human capital will simply migrate to venues without the tax, as it did in the case of Sweden from 1984 to 1990 (Umlauf 1993; Wrobel 1996).

In fact, all of these proposals are addressing only the symptoms of a much deeper problem: the fact that our financial regulatory framework has become antiquated and obsolete in the face of rapid technological advances that drastically reduced costs to intermediation, but have not correspondingly increased or distributed the benefits of greater immediacy. Minimizing technical and operating errors at the level of individual trading algorithms or automated systems—which should always be encouraged—is not sufficient to minimize the incidence of disruptive market-wide events. In fact, in a competitive environment, “optimal” decisions made by subsystems (for example, at the level of individual trading algorithms or trading firms) may interact with each other in ways that make the entire financial system more prone to systemic disruptions. Therefore, Financial Regulation 2.0 necessarily involves a *systemwide* redesign and ongoing systemwide supervision and regulation.

To bring the current financial regulatory framework into the Digital Age, we propose four basic design principles that we refer to as “Financial Regulation 2.0.”

1) *Systems-Engineered*. Since most financial regulations will eventually be translated into computer code and executed by automated systems, financial regulation should approach automated markets as complex systems composed of multiple software applications, hardware devices, and human personnel, and promote best practices in systems design and complexity management. A number of these practices come from the field of systems engineering and have already been adopted in other industries such as transportation, manufacturing, and nuclear power.

2) *Safeguards-Heavy*. Financial regulation should recognize that automation and increasingly higher transaction speeds make it nearly impossible for humans to provide effective layers of risk management and nuanced judgment in a live trading environment. Thus, effective risk safeguards need to be consistent with the machine-readable communication protocols, as well as human oversight. Regulators need to encourage safeguards at multiple levels of the system.

3) *Transparency-Rich*. Financial regulation should aim to make the design and operation of financial products and services more transparent and accessible to automated audits conducted on an ongoing basis by the regulator’s own “bots.” Ideally, regulation should mandate that versions and modifications of the source code that implement each rule, as well as the data used for testing and validation of the code, are made available to the regulators and potentially the public. Regulators need to change their surveillance and enforcement practices to be more cyber-centric rather than human-centric.

4) *Platform-Neutral*. Financial regulation should be designed to encourage innovation in technology and finance, and should be neutral with respect to the specifics of how core computing technologies like operating systems, databases, user interfaces, hardware solutions, and software applications work. Doing otherwise would inevitably lock-in outdated practices, ring-fence potentially inefficient ways of doing business, and empower incumbents at the expense of potential new entrants.

Although these principles may seem unrealistic, a recent example of a regulatory initiative consistent with these principles is the set of measures surrounding the creation of “legal entity identifiers”—alphanumeric, machine-readable strings uniquely associated with each separate entity participating in a financial transaction (for example, see the legal-entity-identifier-related publications of the Financial Stability Board at [http://www.financialstabilityboard.org/list/fsb\\_publications/tid\\_156/index.htm](http://www.financialstabilityboard.org/list/fsb_publications/tid_156/index.htm)). This initiative is cyber-centric, promotes innovation, imposes system-design principles, increases transparency, enables the creation of additional risk safeguards, and encourages the implementation of risk management processes and workflows that allow human knowledge to complement the computational abilities of machines. This gives us hope that with sufficient motivation, effort, and expertise, Financial Regulation 2.0 will be achievable.

■ *Research support from the MIT Laboratory for Financial Engineering is gratefully acknowledged. We thank David Autor, Matthew Baron, Jayna Cummings, Chang-Tai Hsieh, Ulrike Malmendier, Ann Norman, Mehrdad Samadi, and Timothy Taylor for helpful comments and suggestions. The views and opinions expressed in this article are those of the authors only, and do not necessarily represent the views and opinions of any institution or agency, any of their affiliates or employees, or any of the individuals acknowledged above.*

## References

- Baron, Matthew, Jonathan Brogaard, and Andrei Kirilenko.** 2012. "The Trading Profits of High Frequency Traders." [http://conference.nber.org/confer/2012/MMf12/Baron\\_Brogaard\\_Kirilenko.pdf](http://conference.nber.org/confer/2012/MMf12/Baron_Brogaard_Kirilenko.pdf).
- Bertsimas, Dimitris, and Andrew Lo.** 1998. "Optimal Control of Execution Costs." *Journal of Financial Markets* 1(1): 1–50.
- Black, Fischer, and Myron Scholes.** 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81(3): 637–54.
- Bogle, John C.** 1997. "The First Index Mutual Fund: A History of Vanguard Index Trust and the Vanguard Index Strategy." [http://www.vanguard.com/bogle\\_site/lib/sp19970401.html](http://www.vanguard.com/bogle_site/lib/sp19970401.html).
- Commodity Futures Trading Commission/ Securities and Exchange Commission (CFTC/ SEC).** 2010. *Preliminary Findings Regarding the Market Events of May 6, 2010*. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues. May 18. <http://www.sec.gov/sec-cftc-prelimreport.pdf>.
- Denning, Peter J.** 2003. "Great Principles of Computing." *Communications of the ACM* 46(11): 15–20.
- Goldman Sachs Asset Management.** 2007. "The Quant Liquidity Crunch." Goldman Sachs Global Quantitative Equity Group, August. Proprietary document for Goldman Sachs clients; not available to the general public.
- Gorton, Gary, and Andrew Metrick.** 2012. "Securitized Banking and the Run on Repo." *Journal of Financial Economics* 104(3): 425–51.
- Grossman, Sanford J., and Merton H. Miller.** 1988. "Liquidity and Market Structure." *Journal of Finance* 43(3): 617–37.
- Heaton, John, and Andrew W. Lo.** 1995. "Securities Transaction Taxes: What Would Be Their Effects on Financial Markets and Institutions?" In *Securities Transaction Taxes: False Hopes and Unintended Consequences*, edited by Suzanne Hammond, 58–109. Chicago, IL: Catalyst Institute.
- Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld.** 2011. "Does Algorithmic Trading Improve Liquidity?" *Journal of Finance* 66(1): 1–33.
- Khandani, Amir E., and Andrew W. Lo.** 2007. "What Happened to the Quants in August 2007?" *Journal of Investment Management* 5(4): 5–54.
- Khandani, Amir E., and Andrew W. Lo.** 2011. "What Happened to the Quants in August 2007? Evidence from Factors and Transactions Data." *Journal of Financial Markets* 14(1): 1–46.
- Kirilenko, Andrei, Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun.** 2011. "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market." [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1686004](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1686004).
- Lehmann, Bruce N.** 1990. "Fads, Martingales, and Market Efficiency." *Quarterly Journal of Economics* 105(1): 1–28.
- Lintner, John.** 1965. "The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics* 47(1): 13–37.
- Lo, Andrew, and Craig MacKinlay.** 1990. "When Are Contrarian Profits Due to Stock Market Overreaction?" *Review of Financial Studies* 3(2): 175–205.
- Markowitz, Harry.** 1952. "Portfolio Selection." *Journal of Finance* 7(1): 77–91.
- McCrank, John.** 2012. "Knight Glitch Likely to Lead to Regulatory Changes: CEO." Reuters, September 11. <http://www.reuters.com/article/2012/09/11/us-knight-ceo-idUSBRE88A1GW20120911>.
- McLaughlin, Tim.** 2012. "Facebook IPO

Glitch Prompts Margin Calls, Headaches.” Reuters, May 25. <http://www.reuters.com/article/2012/05/25/us-facebook-smallinvestors-idUSBRE84O18P20120525>.

**Mehta, Nitin.** 2013. “Pros and Cons of ‘Tobin Tax’ Divides EU.” *Financial Times*, March 3.

**Merton, Robert C.** 1973. “Theory of Rational Option Pricing.” *Bell Journal of Economics and Management Science* 4(1): 141–83.

**Mossin, Jan.** 1966. “Equilibrium in a Capital Asset Market.” *Econometrica* 34(4): 768–83.

**Oran, Olivia, Jonathan Spicer, Chuck Mikolajczak, and Carrick Mollenkamp.** 2012. “BATS Exchange Withdraws IPO after Stumbles.” Reuters, March 24. <http://uk.reuters.com/article/2012/03/24/us-bats-trading-idUKBRE82MOW020120324>.

**Rosenberg, Barr.** 1974. “Extra-Market Components of Covariance in Security Returns.” *Journal of Financial and Quantitative Analysis* 9(2): 263–74.

**Rothman, Mathew S.** 2007a. “Turbulent Times in Quant Land.” *U.S. Equity Quantitative Strategies*, August 9. Lehman Brothers Equity Research. [http://dealbreaker.com/\\_old/images/pdf/quant.pdf](http://dealbreaker.com/_old/images/pdf/quant.pdf).

**Rothman, Mathew S.** 2007b. “View from QuantLand: Where Do We Go Now?” *U.S. Equity Quantitative Strategies*, Lehman Brothers Research. Proprietary document for Lehman clients only; not available to the general public.

**Rothman, Mathew S.** 2007c. “Rebalance of Large Cap Quant Portfolio.” *U.S. Equity Quantitative Strategies*, Lehman Brothers Research. Proprietary document for Lehman clients only; not available to the general public.

**Schapiro, Mary.** 2012. “Introductory Remarks at SEC’s Market Technology Roundtable.” October, 2. <http://www.sec.gov/news/speech/2012/spch100212m1s.htm>.

**Securities and Exchange Commission (SEC).** 2012. “Order Instituting Administrative and Cease-and-Desist Proceedings Pursuant to Sections 15(B) and 21C of the Securities Exchange Act of 1934 and Section 9(B) of the Investment Company Act of 1940, Making Findings, and Imposing Remedial Sanctions and Cease-and-Desist Orders.” Administrative Proceeding, File No. 3-15046. September 25.

**Securities and Exchange Commission (SEC).**

2013. “Order Granting Approval to Proposed Rule Change, as Modified by Amendment No. 1, to Establish the Retail Price Improvement Program on a Pilot Basis until 12 Months from the Date of Implementation.” Release No. 34-68937; File No. SR-NASDAQ-2012-129, February 15.

**Sender, Henny, Kate Kelly, and Gregory Zuckerman.** 2007. “Goldman Wagers on Cash Infusion to Show Resolve.” *Wall Street Journal* (Eastern edition). August 14, p. A.1.

**Sharpe, William F.** 1964. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk.” *Journal of Finance* 19(3): 425–42.

**Strasburg, Jenny, Andrew Ackerman, and Aaron Lucchetti.** 2012. “Nasdaq CEO Lost Touch Amid Facebook Chaos.” *Wall Street Journal* (Eastern Edition). June 11, p. A.1.

**Strasburg, Jenny, and Jacob Bunge.** 2012. “Social Network’s Debut on Nasdaq Disrupted by Technical Glitches, Trader Confusion.” *Wall Street Journal* (Eastern Edition). May 19, p. A.2.

**Tabb, Larry.** 2012. “Written Testimony to the United States Senate Committee on Banking, Housing, and Urban Affairs by Larry Tabb, CEO, TABB Group.” September 20. [http://www.banking.senate.gov/public/index.cfm?FuseAction=Hearings.Testimony&Hearing\\_ID=f8a5cef9-291d-4dd3-ad3b-10b55c86d23e&Witness\\_ID=f520faa2-1cfe-48a5-b373-60bde009d3a3](http://www.banking.senate.gov/public/index.cfm?FuseAction=Hearings.Testimony&Hearing_ID=f8a5cef9-291d-4dd3-ad3b-10b55c86d23e&Witness_ID=f520faa2-1cfe-48a5-b373-60bde009d3a3).

**Telegraph, The.** 2012. “Knight Capital’s \$440m Trading Loss ‘Caused by Disused Software.’” August 14. <http://www.telegraph.co.uk/finance/newsbysector/banksandfinance/9475292/Knight-Capitals-440m-trading-loss-caused-by-disused-software.html>.

**Umlauf, Steven R.** 1993. “Transaction Taxes and the Behavior of the Swedish Stock Market.” *Journal of Financial Economics* 33(2): 227–40.

**Wrobel, Marion G.** 1996. “Financial Transactions Taxes: The International Experience and the Lessons for Canada.” Background Paper BP-419E, Research Branch, Library of Parliament, Government of Canada.

**Zuckerman, Gregory, James Hagerty, and David Gauthier-Villars.** 2007. “Impact of Mortgage Crisis Spreads, Dow Tumbles 2.8% as Fallout Intensifies; Moves by Central Banks.” *Wall Street Journal* (Eastern Edition). August 10, p. A.1.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.