DOI: 10.1111/fima.12297

ORIGINAL ARTICLE

Financial Management WILEY

# The future of fintech

**Sanjiv R. Das**

Leavey School of Business, Santa Clara University, Santa Clara, California

**Correspondence**
Sanjiv R. Das, Leavey School of Business at Santa Clara University, Santa Clara, CA 95053.
Email: rsdas@scu.edu

Sanjiv R. Das is the William and Janice Terry Professor of Finance in the Leavey School of Business at Santa Clara University in Santa Clara, CA.

## Abstract

This article describes the growing field of financial technology (fintech) and the different financial paradigms and technologies that support it. Fintech is primarily a disintermediation force where disruptive technologies are the drivers. This framework discusses 10 primary areas in fintech comprising a taxonomy, which categorizes research in the field and also proposes a pedagogical structure. Pitfalls of fintech are also analyzed. Overall, the great strides made in computing technology, mathematics, statistics, psychology, econometrics, linguistics, cryptography, big data, and computer interfaces have combined to create an explosion of fintechs.

KEYWORDS

fintech, blockchains, trading, payments, robots, AI, regtech

Financial firms are rapidly using technology to transform their businesses. In this article, I survey 10 areas in which fintech is poised to deliver high value to firms, markets, and regulators. Much of this value is created through the use of machine learning technology, big data, cloud computing, and cryptographic methods.

Fintech refers to various financial technologies used to automate processes in the financial sector, from routine, manual tasks to nonroutine, cognitive decision making. Various areas of finance are subject to disruption, such as payment systems, contract checking, trading, risk management, quantitative asset management, lending, mobile banking, customer retention, and investment banking. Annual fintech financing in 2018 was $112 billion, comprised of 2,196 deals, doubling over that of the previous year (2017: $51 billion).[1] Figure 1 presents one representation of the fintech landscape.

Fintech may be characterized by technological change in three broad areas of finance: (a) raising capital, (b) allocating capital, and (c) transferring capital. Fintech is disrupting all three. For example, payment systems efficiently transfer capital. Firms such as CommonBond (https://commonbond.co/) are using technology to revolutionize how capital is supplied. Likewise, robo-advising platforms are changing the way capital is allocated.

**Definition.** Fintech is any technology that eliminates or reduces the costs of financial intermediation.

Definitions of fintech vary depending on source, and the definition here is a general approach to capturing the main flavor of what is currently understood as fintech. For example, the Bank for International Settlements (BIS) defines

© 2019 Financial Management Association International

[1] https://assets.kpmg/content/dam/kpmg/xx/pdf/2019/02/the-pulse-of-fintech-2018.pdf.

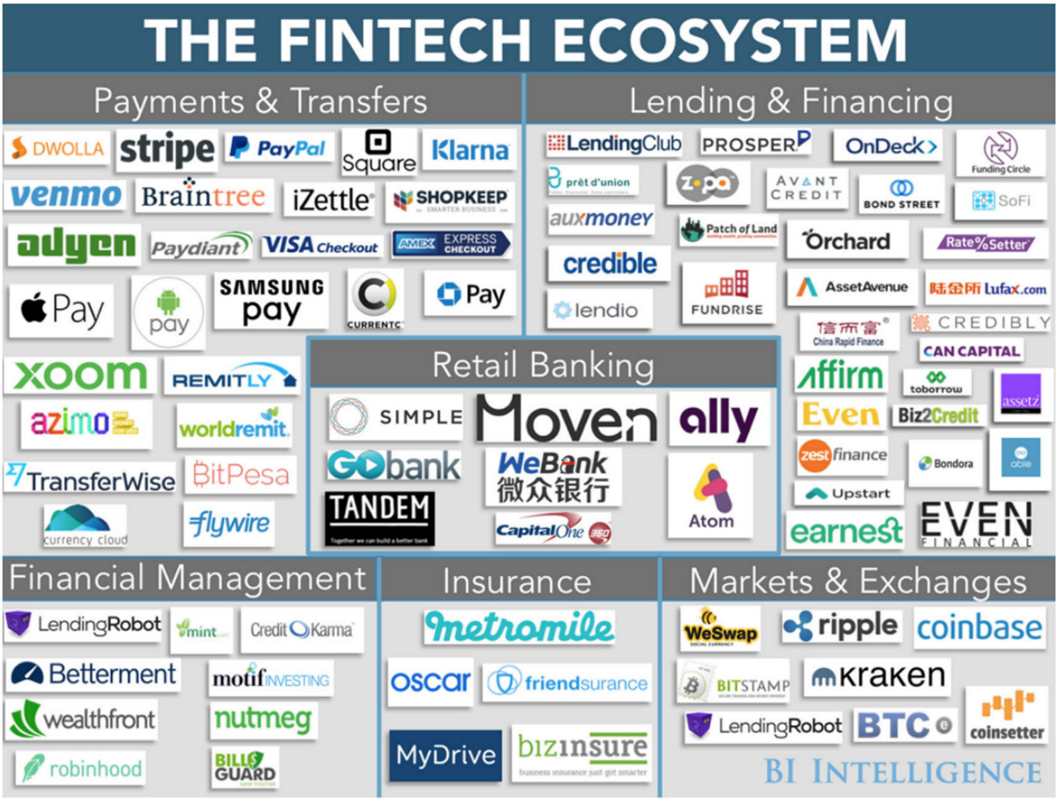


**FIGURE 1** The fintech landscape. These "lumascape" diagrams show a sampling of a business domain and firms within it. See https://pbs.twimg.com/media/CYOHdfgW8AERjdO.png [Color figure can be viewed at wileyonlinelibrary.com]

fintech in credit to "broadly include all credit activity facilitated by electronic (online) platforms that are not operated by commercial banks" (see *BIS Quarterly Review*, September 2018, p. 31).

What is the long-run driver for fintech as a disruptor? The main factor is the high cost of financial intermediation. This has historically always been so. In an interesting study, Philippon (2016) shows that the average cost of intermediation has held steady at around 2% of transaction amounts. Figure 2 shows that this cost has hovered around this level since 1880 until current times, for an astonishingly extended period of more than a century. While one may speculate about the reasons for this high cost of financial intermediation, such as lack of competition on the supply side, or ignorance of consumers on the demand side, the fact remains that these rents accruing to large financial institutions are now ripe for the picking by smaller, agile, fintech players. Technology is often a cheaper intermediary and a driver of competition. Over the last two decades, employment in the financial sector has expanded from 5% to 6.5% of the workforce, as shown in Figure 3. Fintech may very likely reduce this number dramatically.

Fintech applications range from simple automation to complex decision making. Many rely on big data, and necessitate investments in cloud infrastructure and analytics. Successful fintech applications display some common characteristics. First, it is valuable to develop models from a theoretical foundation before bringing in data. This helps both in preparing the data based on extant theory, and makes interpretation of the results facile as they are seen from the backdrop of a theoretical foundation. For example, automated lending models are based on a theoretical foundation of financial concepts, such as leverage, and customer behavior, which suggests the econometric specification for the data. The variables used in the model, denoted as "features" in machine learning (ML) models are better chosen using theoretical ideas, especially when causal models are intended. Second, sharp definition of the problem statement

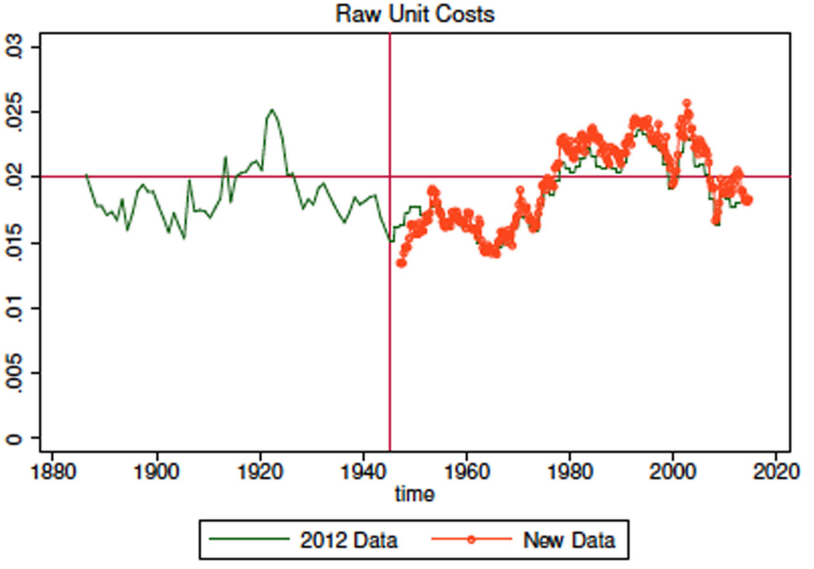


**FIGURE 2** The cost of financial intermediation has held steady at 2% over time [Color figure can be viewed at wileyonlinelibrary.com]
*Source*: Philippon (2016).

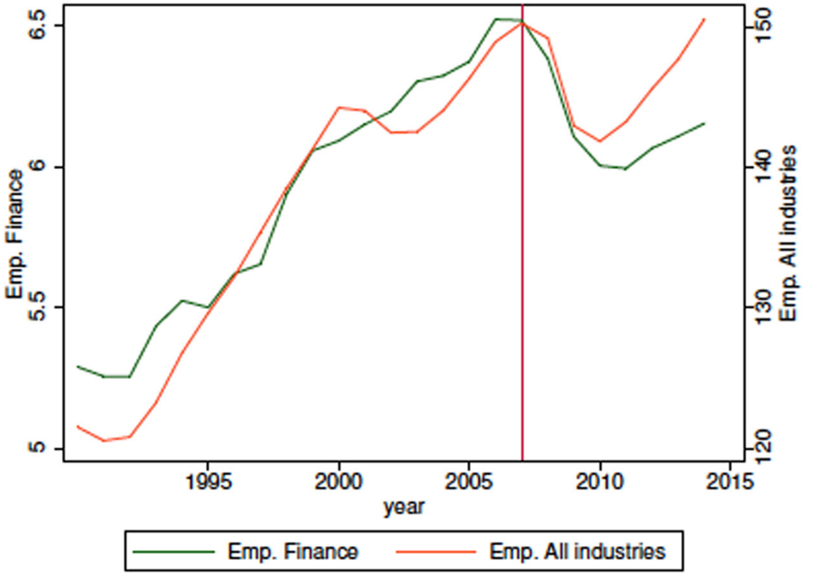


**FIGURE 3** Employment in the financial sector has increased its share from 5.0% to 6.5% in the last two decades [Color figure can be viewed at wileyonlinelibrary.com]
*Source*: Philippon (2016).

is critical—the question is primary and data is secondary. To consider the problem of lending again, there are many different firms with varied offerings, as may be seen in Figure 1. What differentiates them from one another is the specific pain point in the lending process that they seek to ameliorate. Any entrant in this congested area of fintech must sharply define the niche problem it is solving so as to create a unique value proposition. Third, many fintech

**TABLE 1** A sampling of benefits for banks from implementing fintech using analytics

| |
|---|
| – Monitoring corporate buzz |
| – Analyzing data to detect, analyze, and understand the more profitable customers or products |
| – Targeting new clients |
| – Customer retention |
| – Automated lending activity |
| – Market prediction and trading |
| – Risk management |
| – Automated financial analysts |
| – Financial forensics to detect rogue employees |
| – Credit cards: optimizing use, marketing offers |
| – Fraud detection |
| – Detecting market manipulation |
| – Social network analysis of clients |
| – Measuring institutional risk from systemic risk |

offerings confront big data and computational bottlenecks. They face problems with data extraction, integration, curation, quality, and analytics. How well these are handled will determine the success of a fintech venture. For a detailed description of these issues, see Alexander, Das, Ives, Jagadish, and Monteleoni (2017). Fourth, fintech tends to be multidisciplinary, calling for teams that cut across many fields. A case in point is robo-advising, where an amalgamation of experience from portfolio management, behavioral finance, user interface design, risk management, legal aspects of consumer finance, and so forth, are all required. Another example is in cryptocurrencies, where specialized knowledge of cryptography, monetary systems, payment methods, distributed computing, and so on, come into play. Applications that are based on natural language processing also entail many different sorts of expertise in linguistics, computer science, artificial intelligence (AI), and economics. ML and analytics figure prominently in many fintech applications. Table 1 is a sampling of areas in which large banks may benefit from fintech-implemented analytics.

The incredible effectiveness of ML in fintech relies on advances in computational algorithms, such as matrix factorization, deep learning, classification methods, and so forth, and a proliferation of special purpose hardware (cloud platforms, graphics processing units [GPUs], etc.) that exploit vast quantities of data. The confluence of these scientific developments has led to huge advantages for firms that have invested in ML technologies. Figure 4 shows that the performance of hedge funds that use ML outstrips that of funds that do not.

The scope of fintech is huge. In a recent paper, Srinivasan (2016) states that as of 2015, Ernst & Young reports that there were 1,400 fintech firms, with more than $33 billion in funding. In comparison, credit card fraud accounted for $31 billion in losses in a year. Therefore, even small improvements in just this area will generate substantial value compared to the investments being made in fintech. Furthermore, modern hardware, cloud infrastructure, and software tools have made possible the rapid development of sophisticated fintech platforms by very small teams, enabling entrants with minimal funding to compete in this space. Rapid development also has a dark side. Lo (2017), p. 17, points out that “the unintended consequences of technology-leveraged finance include firesales, flash crashes, botched initial public offerings, cybersecurity breaches, catastrophic algorithmic trading errors, and a technological arms race that has created new winners, losers, and systemic risk in the financial ecosystem.” He suggests a strong focus on robust technology to manage this problem.

The distribution of fintech is naturally uneven. Interestingly, one of the countries with high adoption rates for fintech is Hong Kong, with the United States coming second, followed by Singapore. The Chinese government is now one of the biggest spenders on deep learning technologies.

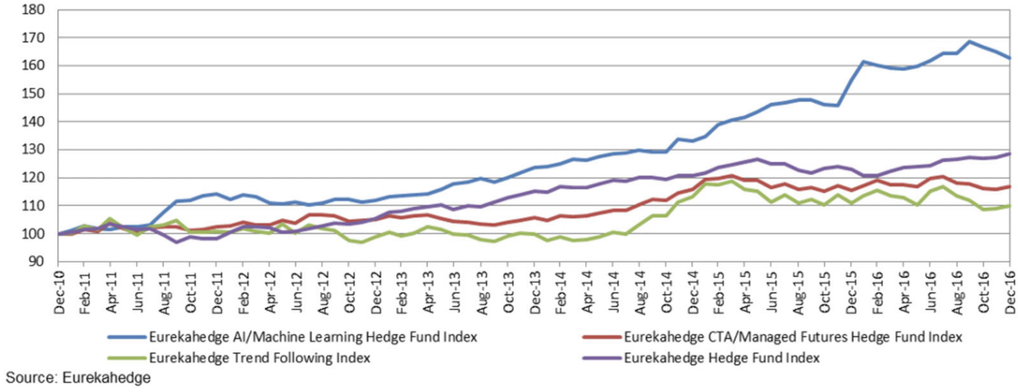


**FIGURE 4** Hedge fund performance is driven by ML. AI/machine learning hedge fund index versus quants and traditional hedge funds. See http://www.eurekahedge.com/Research/News/1614/Artificial-Intelligence-AI-Hedge-Fund-Index-Strategy-Profile [Color figure can be viewed at wileyonlinelibrary.com]

In the following sections, I present a taxonomy of 10 areas where I expect fintech to grow rapidly and act as both an innovation and disruption. Such a taxonomy is also useful in designing academic curricula for fintech. After these discussions, we will explore the use of ML, AI, and deep learning through some use case examples. These examples will also highlight the tremendous advance in ML technology that has made these disruptions possible. In the following sections, we will examine some of the ideas and underpinnings of various uses of mathematics and computer science, married with economic models in the space of Fintech solutions.

## 1 | MACHINE LEARNING, ARTIFICIAL INTELLIGENCE, AND DEEP LEARNING

As described in Culkin and Das (2017), AI has come of age through implementing "deep learning" neural networks[2] because of the confluence of three important factors: (a) the efficacy of mathematical analysis for calibrating neural nets; (b) improvements in hardware and software that allow very large (deep) neural nets to be computed efficiently; and (c) the availability of big data with which to train these models. Deep learning models have been proven to uncover subtle nonlinearities in data that are not discoverable using standard, more or less, linear econometric models. These models implement pattern recognition to high levels of accuracy. By casting finance problems as pattern recognition problems, we are able to avail of higher predictability than before. We can also use deep learning to train pricing models from data on inputs and prices in the markets, bypassing the need for theoretical models, such as those used for option pricing. In the words of Peter Norvig, "All models are wrong, and increasingly you can succeed without them." He makes the case that data are the new theory.[3]

In order to explore this idea, we revisit work from the 1990s. Hutchinson, Lo, and Poggio (1994) explored using neural nets to learn the Black and Scholes (1973) option pricing formula, also developed by Merton (1973). In Culkin and Das (2017), they created data for the assessment of how a deep neural net would learn this equation, by randomly simulating a range of call option prices. The data were divided into two random sets, one for training, and the other for validation. Before passing the prices to the deep learning net, the data were normalized by exploiting a facet

[2] A neural net is a highly nonlinear fitting model made up of several simple nested nonlinear functions that take inputs and train them to generate a discrete or continuous output, through training on large data sets. Neural nets are used, for example, in natural language translation, and in autonomous vehicles. This is a very simplistic definition of neural nets, and nets with several 10s of layers are known as "deep-learning" nets.

[3] https://www.wired.com/2008/06/pb-theory/.

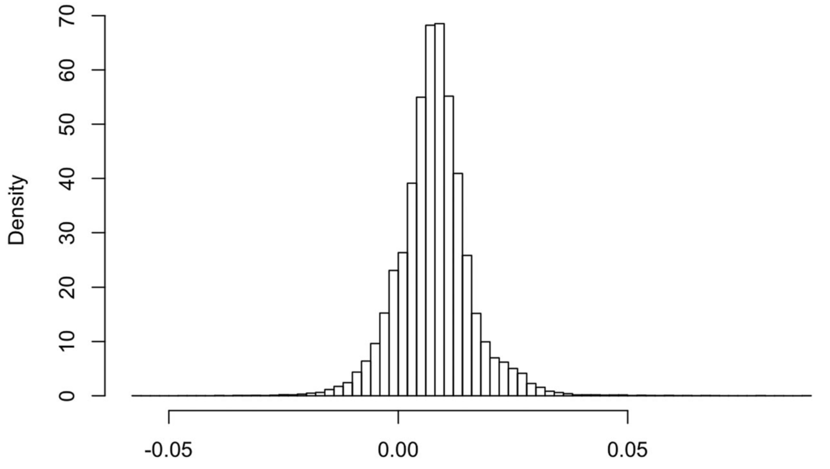


**FIGURE 5** Call option pricing errors from the fitted deep learning neural net
*Source*: Culkin and Das (2017).

of the Black–Scholes call option function—that is, that the pricing function is linear homogeneous in $(S, K)$, that is, $C(S, K) = K \cdot C(S/K, 1)$. Therefore,

$$C(S, K)/K = C(S/K, 1).$$

The data were normalized by dividing both stock price *S* and call price *C* by strike price *K*. These normalized data were applied to the deep learning net to fit the input variables $S, K, T, q, r$, and $\sigma$ (the feature set) to the output prices *C*. The out-of-sample root mean-squared error (RMSE) is 0.0112, with an average error of $\pm 1\%$ of the strike. The percentage pricing error (error divided by option price) is 0.0421—that is, 4% on average. The distribution of pricing errors is shown in Figure 5. A regression of the model values on the true values has very high $R^2 = 0.9982$.

Culkin and Das (2017) also found that with the exception of very small option prices, where the percentage error tends to be magnified, moneyness is not correlated with pricing error.

This simple example supports Norvig's assertion that ML has the potential to replace theoretically derived models with data-driven models. AI and ML models are now being implemented widely across the financial services industry. Here are some examples. Bridgewater Associates, the world's largest hedge fund, has a project to automate decision making to save time and eliminate human emotion volatility.[4] Goldman Sachs now has 2 out of 600 equity traders left in one part of its business. It found that a few traders can be replaced by one computer engineer.[5] It is estimated that by 2020, at least 5% of all economic transactions will be handled by autonomous software.[6]

AI will process payment functions and learn from customer behaviors, through intelligent payments management (IPM). The potential savings are huge: AI will help consumers make daily financial decisions and monitor spending. New personal financial management apps use contextual awareness, which measures spending habits and online footprints to create personalized advice. Combining pooled financial data with end user control to offer tailor-made services is a classic AI solution that we should expect to see plenty of. Algorithms will automatically mine customer data and undertake cross-selling of financial products. Automation of cognitive tasks is now occurring rapidly. For example, J.P. Morgan has a software called COIN that interprets commercial loan agreements and has resulted in a saving of 360,000 hours of lawyer time.[7]

[4] https://www.theguardian.com/technology/2016/dec/22/bridgewater-associates-ai-artificial-intelligence-management.

[5] http://www.zerohedge.com/news/2017-02-13/goldman-had-600-cash-equity-traders-2000-it-now-has-2.

[6] http://www.gartner.com/smarterwithgartner/gartner-predicts-our-digital-future/.

[7] https://futurism.com/an-ai-completed-360000-hours-of-finance-work-in-just-seconds/.

There are several interesting applications under way. In the retail banking arena, Mizuho Financial Group sent Pepper, its humanoid robot into its Tokyo branch to handle customer inquiries.[8] They partnered with IBM to enable Pepper to understand human emotions, and build interaction into apps. The Royal Bank of Scotland is trial testing Luvo AI, a customer service assistant, to interact with staff and customers.[9] AXA (the insurer) has an app-based bot called Xtra, which engages in bespoke conversations with customers about healthy living.[10] AI is used in peer-to-peer (p2p) lending.[11] The role of chatbots in changing the interface with banking customers is growing rapidly.[12]

AI is also permeating the operations of hedge funds. BlackRock is replacing human stock pickers with machine algorithms, using deep learning neural nets, as described earlier.[13] Sentient Technologies is a hedge fund run entirely using AI.[14] It is supposed to have a secret algorithm with adaptive learning that uses 1000s of machines. Numerai is a hedge fund that makes trades by aggregating trading algorithms submitted by anonymous contributors; prizes are awarded in cryptocurrency called Numeraire, which reside on the Ethereum blockchain.[15] Numerai open-sources transformed big data, which is not revealed in pure form because it scrambles the data using homomorphic encryption, a form of encryption where the data can be used for pattern recognition and analysis without being able to extract any original information from them.

A skeptical viewpoint is that there are very little data about the track record of these AI-driven hedge funds, as the business remains secretive. It is also argued that such funds will fail because investors will remain reluctant to turn over their money completely to a machine. Yet, this is being disproved by the little data we have. Figure 4 shows that funds using ML outperform others quite handily. And money is marching into AI-driven funds, as Numerai Fund 1 LP raised more than $1 million in short order.[16]

Does the fact that hedge funds are successfully beating the market imply an unanticipated failure of market efficiency? In order to explore this proposition, Das, Mokashi, and Culkin (2018) conducted a simple experiment using pattern recognition. Employing deep learning neural nets, they trained an algorithm on data from all stocks in the S&P 500 index, and attempted to predict, using a look-back number of days (e.g., 30, 60, 90 days), the direction of movement (up or down) in the index the following day. Undertaking different experiments, the results, using daily data from 1963, show that it is difficult to train an algorithm to guess the sign of the index return with marginally better-than-the-average up days in the markets (52.7%). Accuracy levels are low and suggest that, even with a greatly expanded information set, markets are still efficient. The availability of data and open-source software packages such as TensorFlow make such tests of market efficiency easy to run. Indeed, this is probably one of the first tests of market efficiency, where the conditioning information set is as large as all stocks in the S&P 500 index.

## 2 | NETWORK MODELS: FINTECH FOR SYSTEMIC RISK MODELING

In this section, I explore how graph theory is being applied to understanding the risks of the financial system, also known as "systemic" risk. Systemic risk has some universally accepted characteristics. It is a risk that has (a) a large impact, (b) is widespread, and (c) creates a ripple effect that endangers the viability of the economic system.

---

[8] https://www.mizuhobank.com/mizuho_fintech/news/pepper/index.html.

[9] http://www.businessinsider.com/royal-bank-of-scotland-launches-ai-chatbot-luvo-using-ibm-watson-2016-9?r=UK&IR=T&IR=T.

[10] https://www.the-digital-insurer.com/dia/xtra-by-axa-ai-driven-personal-wellness-coaching-app/.

[11] http://www.nanalyze.com/2017/04/ai-fintech-startups-loans-new-credit/.

[12] https://www.americanbanker.com/news/this-is-how-financial-services-chatbots-are-going-to-evolve.

[13] http://fortune.com/2017/03/30/blackrock-robots-layoffs-artificial-intelligence-ai-hedge-fund/.

[14] https://en.wikipedia.org/wiki/Sentient_Technologies As of 2016, it is the most well-funded AI company.

[15] https://numer.ai/.

[16] https://www.sec.gov/Archives/edgar/data/1667103/000166710316000002/xslFormDX01/primary_doc.xml.

Systemic risk is an attribute of the economic system and not that of a single entity. Its measurement should have two important features: (a) quantifiability—it must be measurable on an ongoing basis and (b) decomposability—aggregate system-wide risk may be broken down into risk contributions from all financial entities in the system. Financial institutions (FIs) that have large risk contributions to aggregate systemic risk may be deemed "systemically important."

The Dodd–Frank Act of 2010 and Basel III regulations characterize a systemically risky FI as one that is (a) large, (b) complex, (c) interconnected, and (4) critical—that is, provides hard-to-substitute services to the economy. Failure of such an institution is potentially disruptive to the financial system. The Dodd–Frank Act does not offer quantification specifics.

Recently, Das (2016) proposed a metric for systemic risk that has both the attributes of system risk, captures the features of systemically important FIs, and is consistent with the three universal characteristics of systemic risk. In order to accurately characterize systemic risk, graph theory is used, and a network of banks is constructed. To see an example of this network construction, see Burdick et al. (2011), where text analytics was used to extract interbank loan transactions from Securities and Exchange Commission (SEC) filings, and construct a co-lending network of money flows between banks. Coupling this interconnectedness information with credit quality information from banks leads to a new measure of systemic risk, that has attractive properties, as formalized in Das, Kim, and Ostrov (2019). There is now a vast literature speaking to this problem, as in papers by Espinosa-Vega and Sole (2010), Billio, Getmansky, Lo, and Pelizzon (2012), Chan-Lau, Chuang, Duan, and Sun (2018), to cite just a few. It takes a careful combination of graph theory and economics, using large quantities of data, to generate a single metric for measuring systemic risk. The requirement to do this in real time makes this a particularly interesting fintech problem. For example, downloading and text mining all SEC filings relating to interbank loans in order to visualize the interbank lending network is an especially interesting problem, requiring terabytes of data combined with graph theory. See Figure 6 for the networks plotted on an annual basis for 2005, and for 2006–2009.

The mathematics for this model in its stochastic form are created by generalizing the Merton (1974) single-firm credit risk model. This extended model derived in Das et al. (2019) allows us to construct stochastic risk networks in a structural credit risk modeling framework. The model's data are standard Merton model inputs for each firm:

- Equity price $= \mathbf{s} = \{s_1, s_2, \ldots, s_n\}$
- Equity volatility $= \boldsymbol{\sigma} = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$
- Number of shares $= \mathbf{m} = \{m_1, m_2, \ldots, m_n\}$
- Risk-free rate $= r$

The model variables are all derived from the Merton model):

- $n$ = number of banks in the system
- $\mathbf{a}$ = $n$-vector with components $a_i$ that represent the assets in bank $i$ (derived from $s, \sigma, m, r$).
- $\lambda$ = $n$-vector with components $\lambda_i$ that represent the average yearly chance of bank $i$ defaulting (from $s, \sigma, r$).
- $\mathbf{E} = n \times n$ matrix with components $E_{ij}$ that represent the probability that if bank $j$ defaults, it will cause bank $i$ to default (from $s, \sigma, r$). This matrix represents the risk connectedness of the banks.

The following calculations lead to a single metric for systemic risk that captures both, interconnectedness and credit quality of all banks. Define $\mathbf{c}$ to be an $n$-vector with components $c_i$ that represent bank $i$'s credit risk. More specifically, I define

$$\mathbf{c} = \mathbf{a} \odot \lambda,$$

Financial Management WILEY


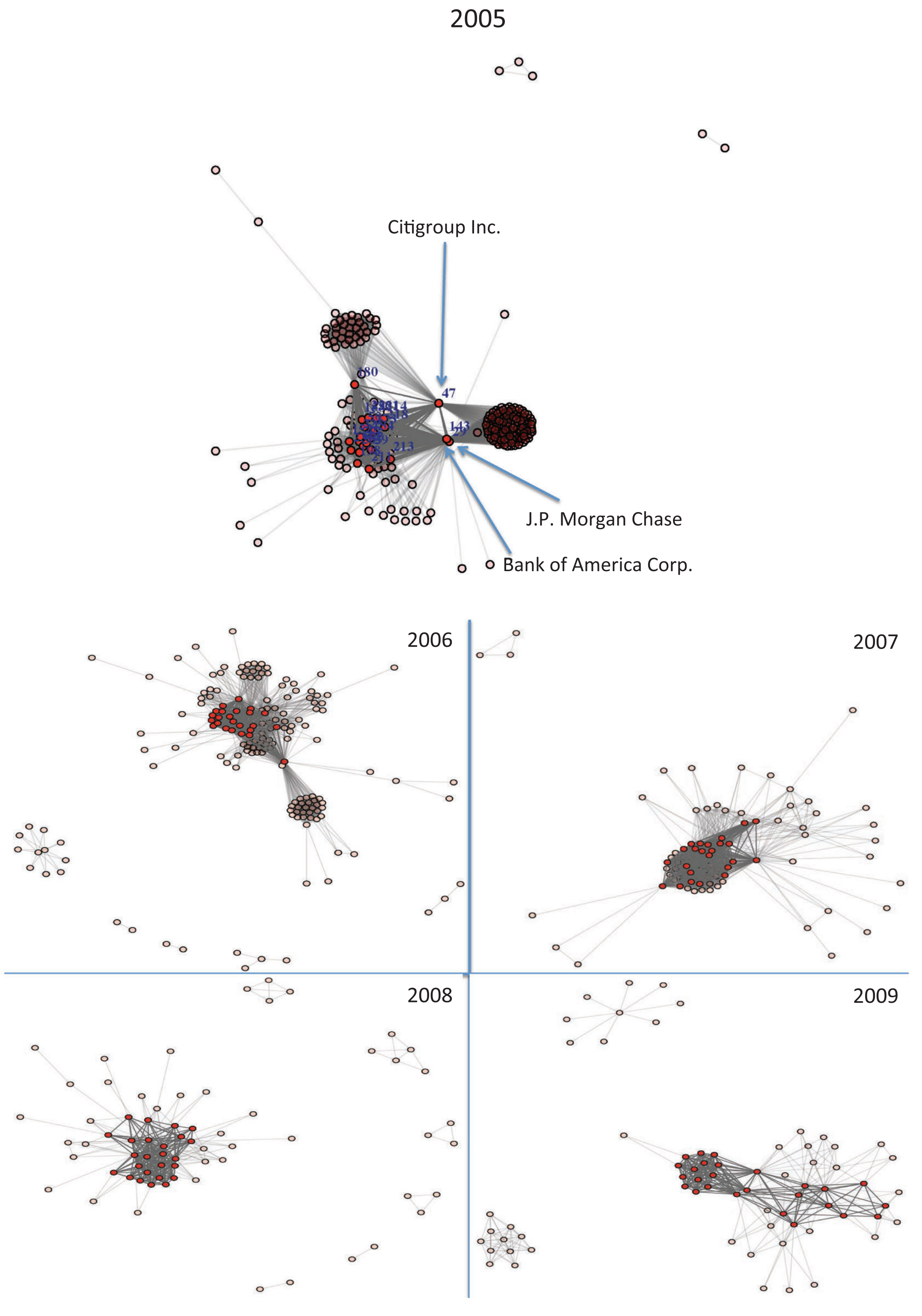


**FIGURE 6** Lending networks between banks for five years 2005–2009 [Color figure can be viewed at wileyonlinelibrary.com]
*Source*: Burdick et al. (2011).

where $\odot$ represents component multiplication (the Hadamard product)—that is, $c_i = a_i \lambda_i$. The aggregate systemic risk created by the $n$ banks in the system is

$$R = \frac{\sqrt{\mathbf{c}^\top \cdot \mathbf{E} \cdot \mathbf{c}}}{\mathbf{1}^\top \mathbf{a}}, \tag{1}$$

where $\mathbf{1}$ is an $n$-vector of ones, so the denominator $\mathbf{1}^\top \mathbf{a} = \sum_{i=1}^{n} a_i$ represents the total assets in the $n$ banks. We note that $R$ is linear homogenous in $\lambda$.

The metric developed has four valuable properties, derived in Das et al. (2019):

1. All other things being equal, $R$ should be minimized by dividing risk equally among the $n$ financial institutions, and maximized by putting all the risk into one institution.
2. $R$ should increase as the financial institutions become more interconnected.
3. If all the assets, $a_i$, are multiplied by a common factor, $\alpha > 0$, it should have no effect on $R$.
4. Substanceless partitioning of a bank into two banks should have no effect on $R$.

Implementation of this simple system on a large-scale across all banks is an interesting challenge in the fintech space.

The metric also leads to other useful metrics, which are easy to compute. The metric $R$ is linear homogeneous in $\lambda$. Let $\alpha$ be any scalar constant. If we replace $\lambda$ with $\alpha\lambda$, it immediately follows that $\mathbf{c}$ is replaced by $\alpha\mathbf{c}$, and, by the equation for $R$, we see that $R$ is replaced by $\alpha R$.

We can also calculate the sensitivity of $R$ to changes in $\lambda$: Differentiating our equation for $R$ with respect to $\lambda$

$$\frac{\partial R}{\partial \lambda} = \frac{1}{2} \frac{\mathbf{a} \odot [(\mathbf{E} + \mathbf{E}^T)\mathbf{c}]}{\mathbf{1}^T \mathbf{a} \sqrt{\mathbf{c}^T \mathbf{E} \mathbf{c}}}$$

whose components represent the sensitivity of $R$ to changes in each bank's value of $\lambda$. This is the basis of risk decomposition, equal to $D = (\frac{\partial R}{\partial \lambda} \odot \lambda)$, a vector containing each bank's contribution to $R$.

A version of this model was implemented in India, with support from the Reserve Bank of India. Metrics may be produced daily. More details are presented in Das (2016), and for illustration Figure 7 shows the Indian network, systemic score $S$, and risk decomposition $D$. The metric is therefore easy to implement and offers a real-time systemic risk management tool for regulators.

Using publicly derived data for the U.S. banking system, Das et al. (2019) also calculated the top risky banks and top risky links (Figure 8). On the technical side, we have 3-D visualizations of financial networks that support better decision making using graph databases such as `neo4j`.[17]

## 3 | PERSONAL AND CONSUMER FINANCE

Households are consumers and allocators of capital. On the consumption side, they borrow money to finance consumption and investments in capital assets. They also earn incomes, generate savings, and allocate their wealth into various asset classes.

In recent times, households have faced several hurdles in the consumption-investment cycle. First, asset returns have become exceedingly low, and this has made retirement targets more elusive. For older investors, who rely on safe income streams after completing their work lives, decent risk-free returns are no longer available, as risk-free interest rates have dropped to near zero. Further, risk premiums on speculative assets have seemingly shrunk, though there are differences of opinion on the expected equity risk premium. Second, medical advances have made longevity risk severe.

[17] https://neo4j.com/users/kineviz/.

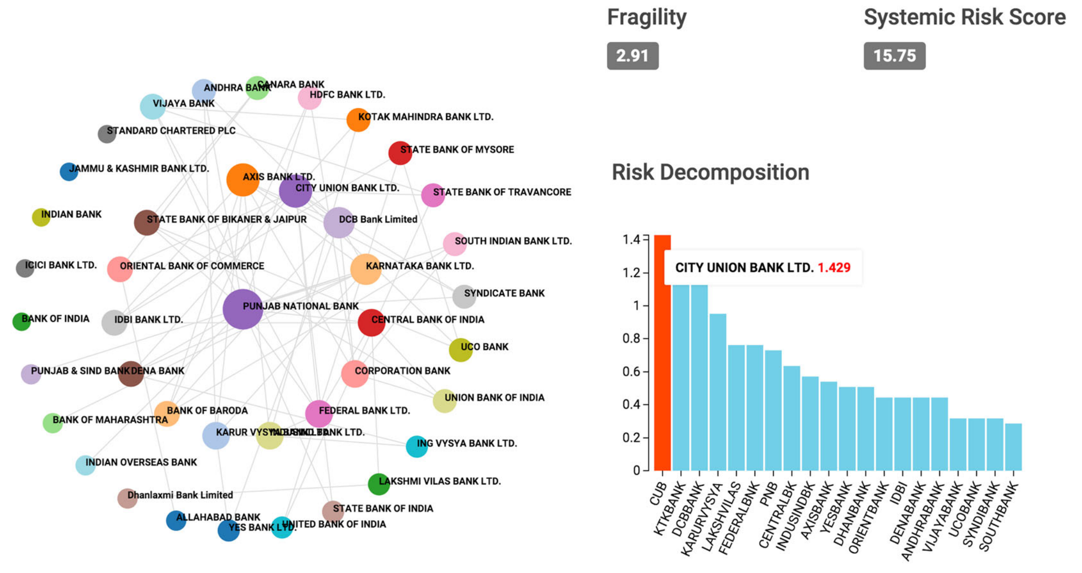


**FIGURE 7** Network of banks in India and risk decomposition, December 3, 2015 [Color figure can be viewed at wileyonlinelibrary.com]
*Source*: Das (2016).

There are heightened concerns about senior citizens outliving their savings. Third, there is substantial volatility risk when reaching for yield. Seeking higher returns in alternative asset classes comes with substantial risk, not just in the second moment of returns but also from the presence of negative skewness and excess kurtosis. Fourth, whereas risk-adjusted returns keep falling, high-cost providers of financial services have retained their pricing schedules. As shown in Figure 2, the average unit cost of financial intermediation is 2%. Fifth, inequality is also rising, and the financial system operates adversely against the poor, in a growing cycle of bias (Piketty, 2014). The share of income of the top decile of the population has grown from 35% in the 1940s to around 50% today. These factors scream out for a fintech solution to raise risk-adjusted returns so that investors may seek a less anxious retirement, with lower costs and inequality.

Various fintech initiatives are addressing these issues. Robo-advising now enables investors to use technology to place their money in well-diversified asset pools at much lower cost, while offering solutions to the retirement problem. Firms such as Wealthfront (www.wealthfront.com) and Betterment (www.betterment.com) are cutting out high-cost retirement providers using technology, while also educating naive, small investors. They have been followed by the major players, such as Vanguard and Schwab. On the lending side, we have better credit models that enable firms to segment customers whose FICO scores otherwise excluded them from accessing financing. Not all subpar FICO scores come from low-quality borrowers, and the ability to tease out the subset that might be of better quality offers fintech lenders an opportunity to open up a new lending sector. Interesting approaches are explored in this space: for example, social media interactions are used to identify good customers, as in Wei, Yildirim, Bulte, and Dellarocas (2016). Digital footprints may be used to assess individual default risk, as shown in Berg, Burg, Gombović, and Puri (2019). Friendship networks may be exploited in peer-lending schemes, as demonstrated in Lin, Prabhala, and Viswanathan (2013). Firms such as PayActiv[18] are disintermediating the payday lending market, dropping the costs of borrowing by around 90%. Big data helps remove biases that often arise with small data, and these may be eliminated as discussed in Chowdhry, Das, and Hartman-Glaser (2016). However, big data needs to be handled carefully, as biases within may end up permeating fintech algorithms (O'Neill, 2016). Overall, consumer finance is ripe for new, refreshing improvements, driven by fintech innovation.

[18] www.payactiv.com.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 20051230 | ABNYY | FNMA | FMCC | BCS | SAN | CS | LEHMQ | GS | C | JPM | |
| 20060630 | FNMA | LEHMQ | DB | GS | SAN | C | JPM | | | | |
| 20061229 | LEHMQ | BSC.1 | ABNYY | SAN | BBVA | MS | C | FNMA | DB | BAC | BCS |
| 20070629 | MTU | LEHMQ | MS | GS | ABNYY | FNMA | | | | | |
| 20071231 | ABNYY | SAN | FNMA | FMCC | MTU | | | | | | |
| 20080630 | MTU | FNMA | FMCC | DB | C | MS | | | | | |
| 20081231 | MTU | FMCC | FNMA | C | BAC | DB | | | | | |
| 20090630 | MFG | FNMA | MTU | C | BAC | DB | | | | | |
| 20091231 | MFG | MTU | NMR | FNMA | FMCC | C | | | | | |
| 20100630 | MTU | MFG | NMR | CS | BBVA | DB | SAN | | | | |
| 20101231 | MFG | MTU | BBVA | SAN | CS | DB | | | | | |
| 20110630 | MFG | MTU | MS | BBVA | SAN | | | | | | |
| 20111230 | BAC | C | MS | DB | BBVA | WFC | | | | | |
| 20120629 | BAC | MS | C | JPM | BBVA | | | | | | |
| 20121231 | BBVA | DB | BAC | C | MS | SAN | | | | | |
| 20130628 | LYG | MFG | SMFG | RBS | MS | BBVA | | | | | |
| 20131231 | LYG | SMFG | MFG | SAN | BCS | HSBC | JPM | DB | BAC | C | |
| 20140630 | LYG | BCS | SMFG | BBVA | BAC | | | | | | |
| 20141231 | RBS | LYG | MFG | MTU | SMFG | | | | | | |
| 20150630 | RBS | LYG | SAN | MTU | SMFG | | | | | | |
| 20151231 | RBS | SAN | BCS | DB | MS | BAC | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| 20051230 | FMCC:ABNYY | ABNYY:FMCC | ABNYY:BCS | ABNYY:JPM | ABNYY:CS |
| 20060630 | FNMA:DB | DB:FNMA | LEHMQ:GS | GS:LEHMQ | FNMA:C |
| 20061229 | BSC.1:LEHMQ | LEHMQ:BSC.1 | LEHMQ:MS | DB:C | DB:ABNYY |
| 20070629 | LEHMQ:MTU | MS:MTU | MTU:MS | LEHMQ:GS | MTU:LEHMQ |
| 20071231 | SAN:ABNYY | ABNYY:SAN | FMCC:ABNYY | FNMA:ABNYY | FNMA:SAN |
| 20080630 | FNMA:MTU | FMCC:MTU | MTU:FNMA | MTU:FMCC | FNMA:FMCC |
| 20081231 | FMCC:MTU | FNMA:MTU | C:MTU | MTU:DB | FMCC:DB |
| 20090630 | MFG:MTU | FNMA:MTU | FNMA:MFG | MTU:MFG | C:MTU |
| 20091231 | MFG:MTU | NMR:MTU | FNMA:MTU | FMCC:MTU | MTU:MFG |
| 20100630 | MFG:MTU | MTU:MFG | NMR:MTU | NMR:MFG | BBVA:MTU |
| 20101231 | MFG:MTU | MTU:MFG | BBVA:MTU | BBVA:MFG | SAN:MTU |
| 20110630 | MFG:MTU | MTU:MFG | MS:MTU | MS:MFG | BBVA:MTU |
| 20111230 | C:BAC | BAC:C | MS:BAC | BAC:DB | MS:DB |
| 20120629 | MS:BAC | C:BAC | BAC:C | BAC:JPM | JPM:BAC |
| 20121231 | BBVA:BAC | C:DB | MS:BAC | DB:C | BAC:BBVA |
| 20130628 | LYG:MFG | SMFG:MFG | MFG:SMFG | LYG:SMFG | MFG:LYG |
| 20131231 | LYG:SMFG | SMFG:LYG | SMFG:MFG | LYG:MFG | MFG:SMFG |
| 20140630 | LYG:BCS | SMFG:BCS | LYG:SMFG | BCS:SMFG | BCS:LYG |
| 20141231 | RBS:MTU | LYG:MTU | MFG:MTU | SMFG:MTU | RBS:MFG |
| 20150630 | LYG:RBS | SAN:RBS | RBS:SAN | LYG:SAN | RBS:LYG |
| 20151231 | RBS:SAN | SAN:RBS | SAN:BCS | RBS:BCS | BCS:SAN |

**FIGURE 8** Top systemically risky banks in United States and top risky links (banking relationships). The figure shows the ticker symbols of the financial institutions. The period covered for each row is the last six months leading up to the date shown in the leftmost column. Dates are shown in YYYYMMDD format [Color figure can be viewed at wileyonlinelibrary.com]
*Source*: Das et al. (2019).

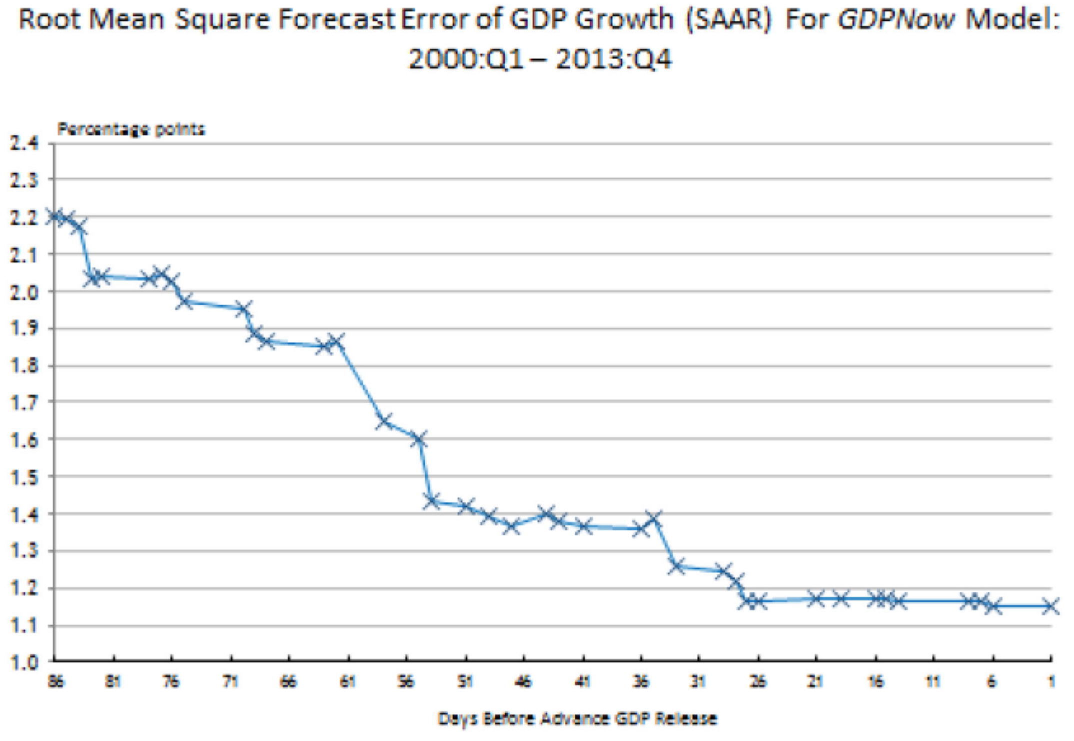


**FIGURE 9** Accuracy of forecasts made by the `GDPNow` model, as a function of the nearness of the date of GDP statistics release [Color figure can be viewed at wileyonlinelibrary.com]
*Source*: Higgins (2014).

## 4 | NOWCASTING

Forecasting has traditionally been a punctuated process with annual or, at best, quarterly forecasts. Macroeconomic forecasts are rarely able to use current data, as statistics for gross domestic product (GDP), inflation, unemployment, and so forth, are usually available only with a lag. Moreover, these reported statistics are often updated and corrected over time, so they are being revised just as they are being used to make forecasts. This is a shifting target, comprised of very low-frequency data.

The availability of new sources of almost streaming time series that are correlated with macroeconomic data has opened up the possibility that forecasting may be undertaken at high frequency in real time. This approach is known as nowcasting and has been gaining popularity as a new area in fintech.

The Federal Reserve Bank of Atlanta has implemented a nowcasting system known as `GDPNow` (Higgins, 2014). This model forecasts GDP growth by aggregating 13 subcomponents of the GDP with a chain-weighting methodology used by the U.S. Bureau of Economic Analysis. The model beats other forecasting approaches and delivers more timely forecasts. Figure 9 shows that the forecasts become more and more accurate as the date of the GDP number release nears. For a survey of nowcasting methods, see Bańbura, Giannone, Modugno, and Reichlin (2013).

Nowcasting generally involves combining data of different frequencies to create as high-frequency a forecast as possible, see Giannone, Reichlin, and Small (2008). In this sense, it is akin to ensemble modeling, where different data are combined to improve prediction accuracy. In Aastveita, Gerdrupa, Jore, and Thorsrudb (2014), U.S. real-time data are used to produce combined density nowcasts of quarterly GDP growth, based on a system of three commonly used model classes. Nowcasts are updated whenever new information is released. Combined density nowcasts always perform well relative to the model classes in terms of both logarithmic scores and calibration tests. The density combination approach performs more effectively than standard approaches.

Another facet of nowcasting is the creation of real-time indices that are nontraded. Of course, for traded indices, the process of trading itself generates the indicators of value in real time. But, for nontraded indices, nowcasting using other traded variables and streaming data is feasible. For example, Chacko, Das, and Fan (2016) provide a theoretical model in which it is possible to generate an illiquidity index from any market sector for which an exchange-traded fund (ETF) exists. In this model, liquidity in any market sector is modeled as a set of options related to immediacy of trading. When immediacy is hard to come by in illiquid markets, the option to trade perforce has higher value. Converting this option value into basis points delivers the price of liquidity as a trading spread. The model is easy to implement and requires only two inputs at any point in time—the price of the ETF and the net asset value (NAV) of its underlying holdings, both of which are reported daily by ETFs. The formula for the illiquidity index is as follows:

$$BILLIQ = -10,000 \times \ln\left[\frac{NAV}{NAV + |ETF - NAV|}\right], \tag{2}$$

where $NAV$ is the net asset value of the exchange traded fund and $ETF$ is its price. In liquid markets, both $ETF$ and $NAV$ trade very close to each other. However, in illiquid markets, it is easier to trade the ETF rather than the underlying components (i.e., the NAV). Therefore, the two diverge in price, and the formula above picks up this difference and converts it into a basis points spread.

Inflation forecasting has leapfrogged traditional consumer price index (CPI) indicators that have been criticized for two reasons. One, they are delayed and infrequent, and two, they are often based on outdated baskets of goods and therefore, do not represent current inflation levels. Using easily scraped prices from internet shopping sites, various efforts are under way to create real-time inflation indices. The Billion Prices project at MIT is a fascinating example of this new phenomenon.[19]

These examples suggest that the future of nowcasting is a direct consequence of the growing availability of streaming data, and the technologies that enable its use. These platforms will eventually provide feature sets that will form the kernel of deep learning trading systems considered in Section 1.

## 5 | CYBERSECURITY AND CLANDESTINE FINTECH

People trust financial institutions to keep their wealth and information secure. They also trust banks to operate responsibly, but that impression of trust erodes with every financial crisis. The increasingly digital manifestation of banks has opened them up to hacking, and the nature of theft has changed dramatically. No longer is information security a matter of simply maintaining a good firewall. Rogue agents, in both digital and human form, reside deep inside bank systems, and chief information security officers (CISOs) at banks no longer assume that their firewalls are failsafe. PwC (2014) reports that 45% of financial sector companies experience crime versus an average of 34% in other industries.

Cyberthreats to financial institutions may be categorized into three major forms. First, threats come from state-organized actors. The news is replete with fact and conjecture about the role of state-sponsored hacking. Second, organized crime has discovered that hacking is a source of easy profits, at much less risk. When you can break and enter and loot assets or information without leaving the comfort of your home, it is easy to see why criminals see this as a far better (often transnational) enterprise than old-style smash and grab.[20] Cybercrime has resulted in stealing information as in the well-known scandal involving Target.[21] The "Bangladesh Bank cyber heist" in February 2016 entailed five fake

---

[19] http://www.thebillionpricesproject.com/.

[20] Despite this, physical bank robberies remain numerous; the FBI reported 3,033 in 2018, down from 4,251 in 2016. See https://www.fbi.gov/investigate/violent-crime/bank-robbery.

[21] The cyber breach of Target resulted in the theft of information of 40 million customers. See: http://www.reuters.com/article/us-target-breach-idUSBRE9BH1GX20131219.

money transfers through the SWIFT network, totaling $101 million, of which only $38 million was recovered.[22] The bank's account at the Federal Reserve Bank of New York was hacked, and the thefts were traced to Sri Lanka and the Philippines, evidence that cybercrime in the financial sector operates easily on a global scale. The third actor in financial cybercrime is rogue bank employees. These are sleeper agents who gain access to servers as employees and wait for an opportune moment to operationalize their malicious intent. Internal bank security systems are being geared up to detect suspicious activity inside firewalls.

The recent dissemination of cybersecurity protocols by the SANS Institute has laid the groundwork for a formalized approach. The essential idea is that there are 20 critical security controls (CSCs) that may be measured, and then aggregated into a cybersecurity score.[23] This score may now be integrated into a bank's risk management practices. Figure 10 shows the different CSCs and the overall framework (upper panel). It also shows the various tools that are widely used to manage compliance with the CSCs (lower panel).

Financial cybercrime has been detailed in popular books. Poulsen (2011) recounts how a hacker in the United States, Max Butler, stole 1.8 million credit card accounts from two Russian crime syndicates, and was eventually caught by the Federal Bureau of Investigation (FBI). (This was lucky for him, as he may easily have fallen into the hands of the Russians, who might have meted out a much worse sentence than the 13 years Butler received.) Menn (2010) tells the tale of a dyslexic young Californian who worked with the Feds to counteract hacking and financial ransom by the Russian mob and the mafia. It also describes how United Kingdom's cybercrime units proved to be effective against Russian hackers. In the area of cybercrime in finance, fact reads like thriller fiction.

Finally, a discussion of financial cybercrime would be incomplete without mentioning the role of the "dark web." A large amount of clandestine fintech occurs on the dark web, which resides on the dark net—that is, networks on the web that require special handshaking software to connect servers (closed versus open protocols). Most of our normal daily experience involves the complement of the dark net, that is, the clear net. The dark net is a subnet of the internet where people connect servers to each other via private networks. The deep web is a subset of the internet that is not indexable by search engines. While different from the dark web, there is some overlap of dark and deep webs. The role of the dark web is to make agents who interact safe from detection—that is, it ensures anonymity and nontraceability of the source of activities on the web. As one might imagine, money laundering relies heavily on the dark net. For a detailed treatment of the dark net, see Bartlett (2014).

An example of the dark side of fintech evolving on the dark net is the creation of "tumblers," which are services to hide the source of transactions in Bitcoin or other cryptocurrencies. By sending transactions through a tumbler, one can mask origination. Tumblers mix up transactions by acting as intermediaries and take a fee for doing so. These are algorithmic services, and may also be bundled automatically with some cryptocurrencies such as Cloakcoin (appropriately named!). Because the blockchain is a public, distributed, and immutable transaction ledger, it makes tracing a transaction possible, rather than otherwise. It is the fact that it is decentralized and does not need a single recording agency who controls transactions and can change rules at a whim that make it attractive. Contrary to popular belief, it is because the blockchain does *not* guarantee anonymity, that the role of tumblers has become widespread.

## 6 | FRAUD DETECTION AND PREVENTION

An interesting aspect of financial crime is that it mostly allows the perpetrators to be removed from the scene of the crime. Online fraud is an excellent example of this phenomenon. Much of credit card fraud now occurs on shopping sites, and the old business of making copies of charge cards with stolen credit card numbers is quickly turning

[22]https://en.wikipedia.org/wiki/Bangladesh_Bank_robbery.

[23]https://www.sans.org/media/critical-security-controls/critical-controls-poster-2016.pdf.

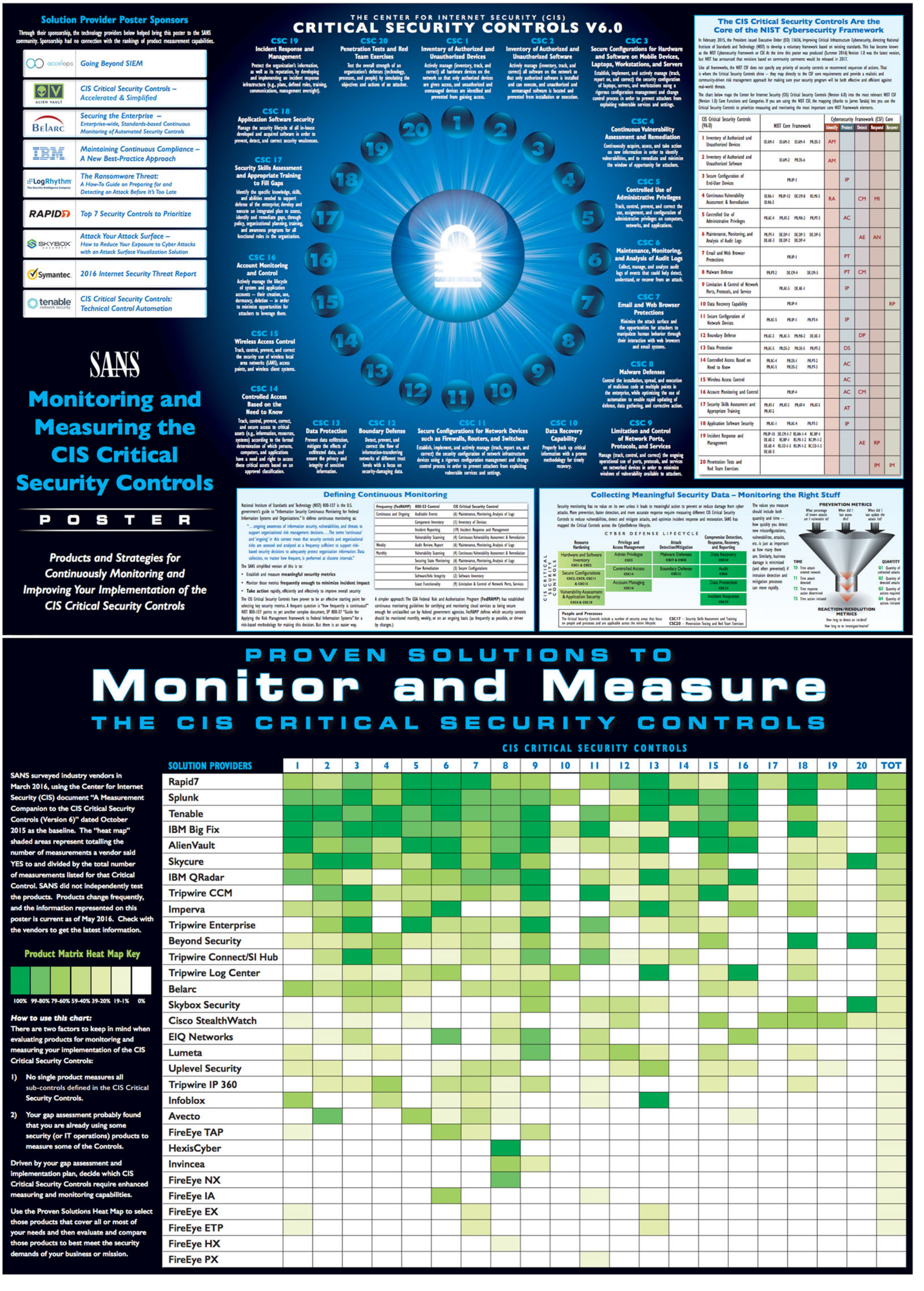


**FIGURE 10** SANS Institute Critical Security Controls (CSCs), shown in the top diagram. The bottom diagram shows the various tools used for compliance with the CSCs. See https://www.sans.org/media/critical-security-controls/critical-controls-poster-2016.pdf. See also https://www.sans.org/security-resources/posters/20-critical-security-controls/55/download [Color figure can be viewed at wileyonlinelibrary.com]

irrelevant.[24] Much of the fraud occurs in three ways: account takeovers, synthetic ID use, and business e-mail compromise, and the number of successful attempts has risen 34% from 2013 to 2016 (Hasham, Hayden, & Wavra, 2018).

Online financial fraud management begins with record keeping, and it is essential that all online activity be logged so that traceback is possible. Strict authentication is important, too. Many financial platforms are designed around multifactor authentication, as in the popular two-step verification implemented by many banks and asset management firms. End-to-end encryption is necessary to prevent a man-in-the-middle attack. Authentication is implemented in many ways, using tokens, passwords, PIN codes, digital keys, biometrics, and so forth. Many digital wallets today embed multiple digital passwords, often three, and can be unlocked with at least two of these, which offers a high level of security.[25]

However, authentication is a futile process if someone has stolen information through a data breach, because the authentication system will not detect a malicious use of validly accessed data. Activity analysis is needed, and additional data might be brought to bear. For example, stolen credit card numbers offer an online thief full access to purchasing power, so they need to be detected through recognizing that an attempted purchase does not follow the user's standard behavior patterns. This is the realm in which ML has proven to be extremely effective, especially when alternative data are used. Solutions in this space come from using social media to detect anomalous behavior, noticing different devices, unexpected patterns in e-mail use, unusual locations of use of the credit card from standard ones, and so on. This is also known as the field of adaptive behavioral analytics, epitomized by firms such as Bionym, EyeVerify, and BioCatch.[26]

As we have seen, the first layer in fraud prevention is authentication, which is related to cybersecurity. As is well known, this is proving to be a hard problem, as firewalls and accounts are breached in large numbers. A second layer of fraud detection lies in using varied data and conducting behavioral analysis, which is not easy to do, and requires extensive data curation. A third aspect of the problem lies in the difficulty of training ML algorithms to detect fraud. Algorithms that use big data in fraud detection are trained on highly imbalanced data sets. For example, detecting credit card fraud is fraught with imbalanced categories. The rate of fraudulent transactions in this realm is ~0.10%—that is, 1 in 1,000 transactions is fraudulent, which means the data have 999 zeros for every 1 binary outcome. Precision and recall are very difficult to control in such data settings, and false positives are rampant. These "anomaly detection" problems have been addressed through new techniques such as topological data analysis (TDA), popularized by the firm Ayasdi.[27] Other techniques have been developed by firms such as Simility.[28] Various ML techniques are used from oversampling to boosting, to deep learning, to random forests. Overall, fraud detection is addressed with an interesting cocktail of authentication, behavioral modeling, device monitoring, and anomaly detection using ML on big data.

## 7 | PAYMENT AND FUNDING SYSTEMS

Digital payments have changed the way in which we move money around, but they have also changed our concept of money. Starting with PayPal, we now exchange money on several platforms, and bypass the banking system. Since then, we have many such services—for example, Venmo, Apple Pay, Samsung Pay, Google Pay, and Dwolla.

This disintermediation takes many forms, but is leaving an electronic payment trail that is making black markets nervous. See, for example, the recent demonetization of the 500 and 1,000 rupee notes in India. Transacting in cash was hampered, and millions were lost in concealed currency hoards, mostly used for property transactions. However,

[24] Despite the increase in cybercrime, bank robberies continue to occur at high rates. In 2016, there were more than 4,000 bank robberies; see https://www.fbi.gov/file-repository/bank-crime-statistics-2016.pdf/view.

[25] Sometimes these levels of security may be breached when hackers find holes in the code for the wallet application itself, as occurred in July 2017 in the case of the Parity wallet, see https://www.coindesk.com/30-million-ether-reported-stolen-parity-wallet-breach/. About $30 million of Ether (ETH) was stolen.

[26] See https://techcrunch.com/2015/08/23/next-gen-cybersecurity-is-all-about-behavior-recognition/.

[27] https://www.ayasdi.com/.

[28] https://simility.com/.

it also gave rise to new transacting systems, with digital payments, such as the Paytm platform. In 2015, the Indian central bank gave Paytm authorization to open a payments bank, known as "Paytm Payments Bank Limited." As of writing, Paytm now has more than 250 million wallets in its transaction stream.

Another form of payments-driven disintermediation comes from disruption in the payday lending space from firms such as PayActiv.[29] This firm works with employers to issue advances against earned wages to low-income employees, part of a package of offerings that reduces financial stress for people who live from paycheck to paycheck. It is estimated that of a U.S. workforce of approximately 130 million people who receive a paycheck, about two-thirds live from one pay cycle to the next and are unable to withstand temporary cash flow shocks. These people are victims of payday lenders who charge them annualized funding rates of 400–500%. Disintermediation by firms such as PayActiv results in driving those borrowing costs down to one-tenth or less.

The lending space is also being disintermediated by p2p markets. These electronic markets connect individual lenders and borrowers directly, and because they bypass large institutions that have high overheads, they have the opportunity to offer cheaper loans. Interest rates are set by the p2p platform through analysis of borrower data or through competition among lenders using reverse auctions. Prosper[30] is one of the oldest such platforms, operating for more than a decade. It supports small loans, mostly less than $5,000, and make extensive use of borrower data. Most p2p lending platforms include borrower/lender anonymity, provide loan pricing support, offer lenders autonomy in picking their borrowers, and allow for both, secured and unsecured lending. Administrative services such as recording, credit checking, and online automation are standard. Targeted lending using big data and ML brings greater benefits to lenders (Crespo, Naveira, and Kwon, 2018).

Fintech also enables new venture financing, through crowdfunding. Top sites for this are GoFundMe,[31] KickStarter,[32] indiegogo,[33] Kiva.[34] Crowdfunding results in more than $35 billion of fund-raising per year. The platforms usually take 5–10% of the money raised as their fee, so in terms of profit, this is at least a $1–$2 billion business. There may also be a processing fee added, so this parallels underwriting fees charged by investment banks. Payment system technologies represent one of the fastest growing areas of financial disintermediation, but their success will eventually depend on whether they are truly able to offer payment services at cost levels much below that of traditional players. These costs will also be adapted through AI-driven pricing (Rizzi, Wang, & Zielinski, 2018).

## 8 | AUTOMATED AND HIGH-FREQUENCY TRADING

Whereas fintech is new terminology, high-frequency trading (HFT) has been around for a long time, marking some of the first high-tech advances in finance. TradeWorx[35] and Automated Trading Desk (ATD, bought by Citibank for $680 million in 2007), were pioneers in the field. Algorithmic trading now accounts for 50% of executed trades in the equity markets, down from around two-thirds of stock trades in the late 2000s. Estimates vary, of course, and Aldridge and Krawciw (2017) estimate the share of market trading at closer to 40%. The profits from algorithmic trading are under competitive pressure and regulatory oversight.

There is a vast academic literature on HFT, and several debates surround this somewhat under-the-radar field. Some of the highlights are as follows. First, since 2013, two-thirds of the top 30 cited papers on HFTs show positive market effects from high-frequency traders. Second, automated firms reduced trading costs and, contrary to popular opinion,

[29] https://www.payactiv.com/.

[30] https://www.prosper.com/.

[31] https://www.gofundme.com/.

[32] https://www.kickstarter.com/.

[33] https://grow.indiegogo.com/.

[34] http://www.kiva.org/.

[35] http://www.tradeworx.com/.

improved market depth and stability. Third, new research is possible because various forms of high-frequency, streaming fintech data have become available. Tick-by-tick data sets are now much more prevalent, and firms are working with academics much more closely than in the past. Fourth, there is now evidence that HFTs stabilize markets (Hendershott & Riordan, 2013), HFTs improve market quality and reduce bid-ask spreads (Hasbrouck & Saar, 2013), and HFTs reduce trading costs (Menkveld, 2013). Fifth, trading in dark pools has been prevalent for quite a while but has changed form many times as risks and technologies evolve. For a theoretical analysis of this activity, see Buti, Rindi, and Werner (2017).

This latest manifestation of the older field of market microstructure is here to stay. It is being reinvented through technology. It is also blamed for market disruptions such as the flash crash of May 6, 2010, which lasted all of a half hour in afternoon trading, when the indexes crashed and rebounded, hitting several market triggers, and transferring vast amounts of profit/loss between trading accounts. Markets are also being transformed as a better understanding of these trading models on new trading platforms is creating (hopefully) marketplaces that are treating small players more fairly than in the past. A good example is the creation of the Investor's Exchange (IEX), immortalized in the book *Flash Boys* by Michael Lewis (2014).

Looking ahead, the big changes anticipated in HFT are (a) an increase in regulation, (b) a reduction in profits from competition, (c) lower relevance on sheer speed of execution, (d) a greater role for the use of myriad sources of information, (e) the entry of deep learning and AI, (f) a reduced human role in favor of greater automation, and (g) an emphasis on hardware over software.

## 9 | BLOCKCHAINS AND CRYPTOCURRENCIES

On August 13, 2017, the price of cryptocurrency Bitcoin (BTC) surged above $4,000. This was a 20% increase over the previous week, after a plan to speed up trade execution was agreed upon. The new solution, denoted `SegWit2x`,[36] has been a bone of contention in the BTC community. However, trade execution under older protocols was slow, and this innovation is going to be a game changer. Nevertheless, trading BTC is still fraught with risk. The daily volatility of BTC is around 5%, much higher than gold (1.2%), major currencies (0.5–1%), or tech stocks, 1–2%. And a few months later, on November 29, 2017, the price of BTC broke the $10,000 price barrier, an astounding price rise for any currency. It ramped up to almost $20,000 in mid-December 2017, but then dropped drastically to below $7,000 by February 2018. It has remained around that level ever since, after dropping even lower towards the end of 2018, followed by a resurgence to $10,000 at the end of the third quarter in 2019. Of course, there is no clear sense as to where its eventual trading range lies. We do not have a good model for pricing BTC using fundamentals.

Blockchains and cryptocurrencies are distinct things. A blockchain is a distributed ledger that has four properties. It has decentralized validation. It is immutable—that is, the record may not be changed. It is secure—that is, tamper proof. It is trusted,[37] such that only valid transactions will enter the ledger and it will prevent double spending. It is also a ledger on a p2p network.

A cryptocurrency is a medium of exchange and a store of value, just like fiat currencies, though legally it is a security, and may be thought of as an asset class. Transactions in this currency are recorded on the blockchain. BTC's inventor is as yet unknown, though the original paper on which it is based is attributed to Nakamoto (2009). BTCs are traceable, but generally offer anonymity through masking methods. Whether BTC is a currency or a speculative asset has been debated in Yermack (2015).

Transactions in cryptocurrencies on the blockchain are secured by encryption methods, and transactions are recorded in a decentralized set of nodes after approval. Approval of a transaction occurs when a "miner" processes a transaction block by adding a number (the nonce) to a variable-length transaction, and then solves a hashing problem

[36] https://segwit2x.github.io/.

[37] Sometimes, this is called the opposite, a "trustless" system, because there is no need for a trusted player in the middle of all other transacting entities.

to generate a fixed-length (256 bits) hash with a set number of leading zeros (required to be 17 at the current time). This random guessing takes computational power (and electricity) and is rewarded with a set number of BTCs, which is BTC 12.5 as of writing. This mining process produces a "proof of work" that validates the transaction block. With today's technology, a block is solved every 10 min.

Transaction volumes in BTC are currently quite low, about 1.5–2.0 transactions per second (tps), or roughly 300,000 transactions per day, versus PayPal at 125 tps, and Visa at 4,000 tps. However, the technology is now widely accepted and way beyond mere proof of concept. It forms a core part of modern-day fintech.

A broader theme for the future is that the blockchain offers a networked platform on which decentralized and automatic contracting is made possible. This has been implemented most famously on the Ethereum platform, which is a programmable blockchain. Anyone can create a smart contract on the Ethereum chain, using the Ethereum virtual machine (EVM), and proof of work is rewarded with the cryptocurrency ETH, used for payments and fees on the platform. For example, one may establish a trading exchange on Ethereum, and contracts can be automatically settled on the platform once the program is so designed. It allows anonymous trading while still permitting a regulator to get an overall picture of risk concentration in the market, thereby allowing for systemic risk management. Another example of the use of the Ethereum platform is in real estate contracts. Several benefits exist such as title verification, settlement, shared equity in properties, and liquid trading of real estate assets.

Ethereum has also supported the growth of decentralized autonomous organizations (DAOs), where group arrangements are explicitly contracted on the Ethereum blockchain as smart contracts. One in particular, known as the DAO (same name), was instituted to be a $150 million venture fund, invested in by 1000s of people in a crowdsourced success. It was, however, hacked in 2016, and several millions of ETH were stolen but later restored by canceling the stolen ETH on the chain, and resetting the blockchain, which is in direct violation of the principle of immutability. This resulted in a hard fork in the Ethereum blockchain. However, the hack was a result of weak security on the DAO side, not on that of the Ethereum blockchain. The Parity wallet was hacked on Ethereum for $31 million (in July 2017). Another $150 million was vulnerable, but white hat hackers stepped in and hacked those accounts out to save them, in a strange situation where the disease itself provided the remedy.[38] These issues have not been addressed by regulators who are struggling to keep pace. Recently, the SEC ruled that tokens issued by DAO were to be treated as securities and are now regulated by federal law. A cynical viewpoint is that after regulation, cryptocurrencies will end up following the old rules of money in a new digital setting.

Another finance application that has grown rapidly are initial coin offerings (ICOs). These are analogous to a pre-product sale, where coins are issued on a blockchain as a store of future value. These will also attract regulation by the SEC, and there have been 46 ICOs in July 2017 alone. The SEC recently declared that ICOs are securities. Volumes in ICOs are doubling annually. In 2018, by June 30, 537 ICOs ($13.7 billion) were registered. In 2017, there were a total of 552 ICOs ($7.0 billion). The average size of an ICO has doubled from $12.8 million in 2017 to more than $25.5 million in 2018.[39]

Companies such as Numerai[40] have created their own cryptocurrency (called a Numeraire), which investors in their crowdsourced hedged fund may use for investments, withdrawals, transaction fees, and payments to developers of trading algorithms they may invest in. This cryptocurrency was made possible by hosting it on the Ethereum blockchain.

We are no longer in the nascent stages of the blockchain revolution, but there is still large-scale innovation in this space. It is likely that there will be vast changes in financial contracting, trading, risk management, and corporate finance, all implemented on blockchain infrastructure. For a comprehensive outlook on the disruptive (and beneficial) potential of blockchains, see Harvey (2014) and Yermack (2017).

[38] https://medium.freecodecamp.org/a-hacker-stole-31m-of-ether-how-it-happened-and-what-it-means-for-ethereum-9e5dc29e33ce.

[39] https://cointelegraph.com/news/pwc-report-finds-that-2018-ico-volume-is-already-double-that-of-previous-year.

[40] https://medium.com/numerai/an-ai-hedge-fund-goes-live-on-ethereum-a80470c6b681.

## 10 | TEXT ANALYTICS

An important area of fintech lies in applications that use “alternative” data.

Textual data greatly expand the universe of available data from the merely numerical. The starting point of any text analysis is the quantification of textual data—that is, a mapping from words and documents to mathematical abstractions such as vectors, matrices, and tensors. The goal is to elucidate qualitative inferences and predictions after suitable mathematical transformations.

There are several areas in which text analytics is now being applied in finance. For a recent survey, see Das (2014). I will discuss some of these applications in this section. Sources of text fall into three categories: (a) blogs, forums, and wikis, (b) news, and (c) content generated by firms. Early work tended to focus on sentiment extraction using postings on stock message boards, such as Yahoo!, Raging Bull, Motley Fool, and Silicon Investor. Sentiment is also extracted from news sections in the *Wall Street Journal*, and the Dow Jones and Reuters news services. These sources were at least at daily frequency. More recently, text streams have become available in real time, and Twitter has become a happy hunting ground for sentiment analysts. Bollen, Mao, and Zeng (2011) state that they can predict the direction of the Dow Jones index with 87% accuracy using tweets. Since then, there have been a slew of papers using tweets, each with mixed success.

Predicting market direction is one of the obvious first uses of text mining, and considerable energy has been devoted to this goal. Early research by Antweiler and Frank (2004) and Das and Chen (2007) constructed a bullishness index,

$$B = \frac{n_B - n_S}{n_B + n_S}, \tag{3}$$

where $n_B, n_S$ is the number of messages categorized as “buy” and “sell,” respectively. This measure is also often called “polarity.” This research shows that it is hard to predict market direction using message board postings, but there is a somewhat weak prediction of volatility. Since then, much more attention has been directed towards using tweets instead. All these initial forays into sentiment analysis of stocks firmly placed textual analysis in finance onto firm footing.

Extensive use is made of corporate textual data for asset management. Data are extracted from the vast universe of regulatory corporate filings, such as 10-K and 10-Q forms. See papers by Loughran and McDonald (2011), Burdick et al. (2011), Bodnaruk, Loughran, and McDonald (2015), Jegadeesh and Wu (2013), and Loughran and McDonald (2014). These papers found that special words lists, known as lexicons, were indeed useful in determining the sentiment embedded in annual reports. For example, a count of risk-related words in annual reports is a good predictor of poor performance in subsequent quarters. Papers by Calomiris and Mamaysky (2019) and Froot, Lou, Ozik, Sadka, and Shen (2017) show that textual information from large-scale media sources coupled with market information provides strong predictability of market direction.

Another interesting development emanates from quantifying “readability” of annual reports. It turns out that metrics for readability of text are useful in predicting corporate performance, and the less readable an annual report is, the worse a firm performs. Readability was originally quantified by researchers in linguistics, and measures such as the Gunning–Fog index (Gunning, 1952) have become popular, as they are robust measures of readability. The formula for the Fog index is:

$$0.4 \times \left[ \frac{\#Words}{\#Sentences} + 100 \cdot \frac{\#ComplexWords}{\#Words} \right]. \tag{4}$$

The formula returns a number that indicates the number of years of schooling needed to be able to comprehend the text. In other words, it also renders the grade level at which the text is written. The intuition for why low readability of annual reports correlates with poor performance is that when a firm has bad news to convey, it tends to obfuscate as much as possible in its filings (e.g., 10-K). Subsequent research also found that longer annual report discussion (in number of characters) was also an indicator of poor subsequent performance; in fact, it was even better than poor readability. Once again, clearly, obfuscation comes more easily with long-winded writing than with short, pithy text. Finally, research shows that counting characters is unnecessary. Simply looking at the file size of the annual report on the SEC server is a good sorting characteristic. Bigger file sizes predict poor performance.

News analytics is also widely used to enhance asset management. This domain deals with the measurement of qualitative and quantitative attributes of unstructured news articles. Leinweber (2009) and Leinweber and Sisk (2011) offer extensive discussion on news analytics-based trading strategies. Tetlock, Saar-Tsechansky, and Macskassay (2008) show that the number of negative words forecasts negative earnings, and news stories that focus on fundamentals are more informative than other articles. For a comprehensive treatment of news analytics, see Mitra and Mitra (2011).

Prediction of corporate or banking failures has the potential to save vast sums of money. In recent work, Das, Kim, and Kothari (2019) show how early warning signals may be extracted from text by analyzing e-mails of senior management in a firm. Using Enron as a test case, they propose a “zero-revelation” technology where a software program can analyze emails for their qualitative characteristics such as sentiment, quantitative characteristics such as size and frequency, spatial aspects such as networks, or aggregate focus through extracting topics. Because the program reads and provides an aggregate analysis, without revealing the contents of the e-mails, this is a noninvasive approach that may be used by corporate management or a regulator to predict financial malaise early, thereby adding value to a firm’s risk management process.

The main features of such an approach are as follows. (a) Financials are often delayed indicators of corporate quality. (b) Internal discussion may be used as an early warning system for upcoming corporate malaise. (c) E-mails have the potential to predict such events. (d) Software can analyze vast quantities of textual data not amenable to human processing. (e) Corporate senior management may also use these analyses to better predict and manage impending crisis for their firms. (f) Most important, the approach requires zero revelation of e-mails.

The empirical analysis of Enron’s e-mails primarily covers two years, 2000 and 2001. Enron failed in the fourth quarter of 2001, so this is the period before and during the crisis the firm faced. Some clear results emanate. (a) As Enron’s condition worsens, e-mails are shorter in length. (b) Likewise, Enron’s sentiment extracted from e-mails drops a few months before the crisis, and continues into it. (c) It is interesting that the length of senior management e-mails is a better predictor of demise than is sentiment, although both are useful. (d) It is also interesting that plotting the frequency of usage of words over time also tells a story. The word *losses* sees more frequent usage than average in the crisis period, and the inverse is true for the word *profits*. For example, see the word *credit* plotted against sentiment for the two years in the data set; these track very closely, as shown in Figure 11. (e) A topic analysis also shows that negative topics become more prevalent as we approach the crisis versus positive topics. (f) Finally, the e-mail network for 2000 looks very different from that in 2001.

There are myriad applications for textual alternative data. Natural language processing using deep learning has become an important recent use of text in finance. Text is being classified to generate early warning signals about company actions, predict the effect on the stock price after an announcement. Since these new data features add value, many hedge funds add textual features as an overlay to their trading algorithms.

## 11 | CONCLUDING COMMENTS: THE GOOD AND THE BAD

In this final section, I consider two important issues that arise around fintech, labor market changes, and implementation pitfalls along the fintech path.

### 11.1 | The finance labor market

Fintech is a potential disruptor of the financial services labor market, which accounts for 6–7% of U.S. employment. AI and ML will erode many of these jobs. However, the debate on whether we will see the process evolve through intelligence augmentation (IA) versus AI (artifical intelligence) is still open. The former will lead to less job loss. It may

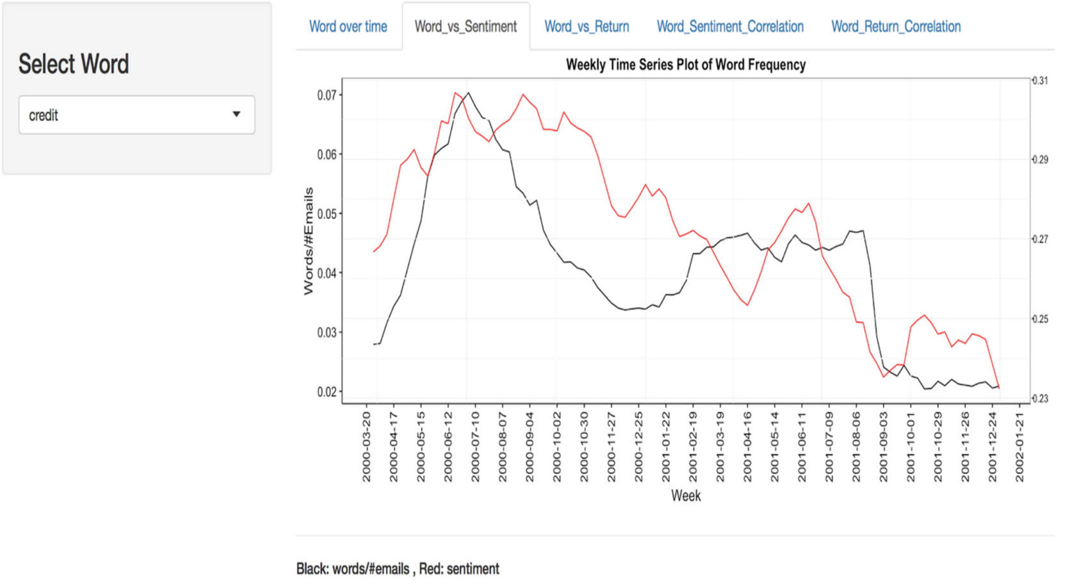


**FIGURE 11** Interactive system to plot any word over time, if the word appears in the corpus of emails from Enron for 2000 and 2001. Here, we plot the word *credit* [Color figure can be viewed at wileyonlinelibrary.com]
*Source*: Das et al. (2019).

be easier to replace high-end trading jobs (through AI) than low-end customer service roles (IA). Firms are recording traders' keystrokes and screens and using the data to train algorithms to mimic traders. These artificial agents are able to improve through techniques such as reinforcement learning. These algorithms can analyze a certain type of data far more deeply than a human can, but they are also poorer at considering a breadth of data, though they may eventually be engineered to do so.[41] In contrast, chatbots to replace customer service agents have a long way to go before they are good enough to replace humans. Middle-level jobs such as paralegal work in securities contract checking is being eliminated with algorithms that have lower error rates. This corresponds to Autor's (2015) hypothesis that while very low-end and very high-end jobs will remain, several middle-level jobs will be lost, such as financial analysts, routine trading functions, paralegals, human resources roles, loan officers, financial advisers, and so forth.

Ray Kurzweil has predicted the supremacy of machines, the "Singularity," which he suggested will occur in 2045. Norbert Wiener famously said, "We can be humble and live a good life with the aid of machines, or we can be arrogant and die."[42] Setting these ominous fears aside, we may dig deeper into the kinds of jobs that seem to be getting winnowed away. A cynical hypothesis has been put forward by Graeber (2018) that there are legions of "bullshit" jobs that are finally being eliminated such as corporate lawyers, financial advisers, and so on. These are jobs that would make no difference were they to disappear; in fact, people doing these jobs know this and bear a psychological stigma as a consequence.[43] I have a simple hypothesis for which jobs will be lost: if the job generates data, then an AI agent can be trained on that data to replace the human. This is what we see happening with trading jobs, paralegal work, and more.

Understandably, there will be jobs where fintech will keep the "human in the loop" (HITL). These are jobs where consideration of nonstandard cases is required or where a human is needed in the loop to avoid legal liability–for example, risk managers. Such jobs may escape "technological unemployment."

[41] https://www.marketwatch.com/story/ai-will-change-stock-market-trading-but-it-cant-wipe-out-the-role-of-people-2018-05-15.

[42] https://www.nytimes.com/2013/05/21/science/mit-scholars-1949-essay-on-machine-age-is-found.html.

[43] Graeber first espoused these ideas in an essay titled "On the Phenomenon of Bullshit Jobs", https://strikemag.org/bullshit-jobs/. The arrival of AI coincides now with the book on the same idea.

## 11.2 Pitfalls for fintech

While fintech heralds the new age of finance, it is by no means widely prevalent as yet, and is certainly not the panacea for all open issues. While it is both an improving force and a disruptive one, it needs to be implemented thoughtfully. Here, I briefly survey seven pitfalls to avoid when implementing fintech.

First, the garbage-in, garbage-out (GIGO) problem. Fintech is consuming more and more data, but it needs to be good data. Without “data curation,” financial analyses will produce poor-quality output. This issue is itself being addressed with new technologies, where ML is being used to clean data, generating better quality inputs as well as reducing the data engineering cycle drastically. This is discussed extensively in Alexander et al. (2017), and technically, companies such as Tamr[44] and Paxata[45] offer excellent solutions.

Second, we have information overload. This comprises collecting too much data and not using them properly. An antidote for this is good theoretical modeling, which suggests exactly what data should be provided to the algorithm. This approach transcribes the scope of the data needed. It may also save money as a firm may be able to move from distributed computing to using a single, large, and fast machine. Testing trading models and proper backtesting will help in improving signal-to-noise ratios. Tools for this are becoming widely available, and businesses such as Quantopian[46] and Numerai[47] are making these widely available so that trading ideas may be crowdsourced—another effective antidote to processing inhuman quantities of market information.

Third, big is not better. More data does not mean better results; quality matters. However, AI and deep learning are making it possible to let machines make sense of large-scale data. Simple applications in the consumer finance space are seeing positive results from the use of deep learning on large data, especially in the area of anomaly detection, as we have seen with firms such as Ayasdi and Simility.

Fourth, separating correlation from causality. In the search for predictive analytics, the use of big data is turning up better predictive models, but these are untraced to the cause of predictability, and are largely atheoretic. Given this, firms revisit their models on much tighter review cycles in order to ensure that they are still viable for trading. This issue becomes especially acute when considering streaming data in HFT.

Fifth, the infrastructure for fintech is expensive, and once a firm dips its toes in the water, it is required to go all in, else it becomes impossible to get good results. HFT, for example, requires an increasingly expensive investment in hardware and communications technology. Deep learning platforms require outlays for expensive chips such as arrays of GPUs. Blockchain infrastructure requires electricity-greedy special-purpose hardware. The list goes on and on. It is an arms race.

Sixth, trust is a huge issue in adoption. Fintech offerings tend to hand off previous human-driven functions to technology—for example, replacing tellers with ATMs and online banking. Cryptocurrencies are completely based on trust, where trust is transferred from a centralized and regulated repository to trust in technology and decentralization. Without trusted algorithms and data, fintech will fail (Thakor & Merton, 2018). Any firm developing a new fintech business must consider how it will implement trust through technology.

Seventh, while technology can be used to improve experience, more often than not, the opposite is the result. Excessive misdirected automation can lead to a terrible customer experience. A case in point is the widespread use of chatbots for customer service interactions, where the quality range is huge. Robo-advising is another area in which careful design of the interface becomes necessary. Firms are using design thinking to ensure that customers are well-served in fintech-driven businesses. Finance houses need to become much more consumer-centric and use design thinking as actively as tech companies like Apple do.

I end with the following prediction: Finance companies will eventually become quasi-technology firms.

[44] https://www.tamr.com/.

[45] https://www.paxata.com/.

[46] https://www.quantopian.com/.

[47] https://numer.ai/.

## ACKNOWLEDGMENTS

I am grateful to Kresimir Kalafatic, Seoyoung Kim, and Christine Parlour for discussion and comments, and for discussion on the basic taxonomy of areas in fintech. I am also grateful for comments from seminar participants at various academic and industry forums.

## REFERENCES

Aastveita, K. A., Gerdrupa, K. R., Jore, A. S., & Thorsrudb, L. A. (2014). Nowcasting GDP in real time: A density combination approach. *Journal of Business and Economic Statistics*, *32*, 48–68.

Aldridge, I., & Krawciw, S. (2017). *Real-time risk: What investors should know about fintech, high-frequency trading and flash crashes*, Hoboken, NJ: Wiley.

Alexander, L., Das, S. R., Ives, Z., Jagadish, H. V., & Monteleoni, C. (2017). Research challenges in financial data modeling and analysis. *Big Data*, *5*, 1–12.

Antweiler, W., & Frank, M. (2004). Is all that talk just noise? The information content of stock message boards. *Journal of Finance*, *59*, 1259–1295.

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, *29*, 3–30.

Bańbura, M., Giannone, D., Modugno, M., & Reichlin, L. (2013). Now-casting and the real-time data flow. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2, pp. 195–237). North Holland, The Netherlands: Elsevier.

Bartlett, J. (2014). *The dark net: Inside the digital underworld*. New York: Random House.

Berg, T., Burg, V., Gombović, A., & Puri, M. (2019). On the rise of fintechs: Credit scoring using digital footprints. *The Review of Financial Studies*, hhz099. https://doi.org/10.1093/rfs/hhz099

Billio, M., Getmansky, M., Lo, A., & Pelizzon, L. (2012). Econometric measures of connectedness econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, *104*, 536–559.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*, 637–659.

Bodnaruk, A., Loughran, T., & McDonald, B. (2015). Using 10-K text to gauge financial constraints. *Journal of Financial and Quantitative Analysis*, *50*, 623–646.

Bollen, J., Mao, H., & Zeng, X.-J. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*, 1–8.

Burdick, D., Das, S., Hernandez, M. A., Ho, H., Koutrika, G., Krishnamurthy, R., …Vaithyanathan, S. (2011). Extracting, linking and integrating data from public sources: A financial case study. *IEEE Data Engineering Bulletin*, *34*, 60–67.

Buti, S., Rindi, B., & Werner, I. M. (2017). Dark pool trading strategies, market quality and welfare. *Journal of Financial Economics*, *124*, 244–265.

Calomiris, C. W., & Mamaysky, H. (2019). How news and its context drive risk and returns around the world. *Journal of Financial Economics*, *133*, 299–336.

Chacko, G., Das, S., & Fan, R. (2016). An index-based measure of liquidity. *Journal of Banking and Finance*, *68*, 162–178.

Chan-Lau, J. A., Chuang, C., Duan, J. C., & Sun, W. (2018). Financial network and systemic risk via forward-looking partial default correlations. National University of Singapore Working paper.

Chowdhry, B., Das, S. R., & Hartman-Glaser, B. (2016). How big data can make us less racist. https://www.zocalopublicsquare.org/2016/04/28/how-big-data-can-make-us-less-racist/ideas/nexus/

Claessens, S., Frost, J., Turner, G., & Zhu, F. (2018). BIS Quarterly Review. In Borio, C., Claessens, S., Cohen, B., & Song Shin, H. (Eds.), *Bank for International Settlements.*

Crespo, I., Naveira, C. F., & Kwon, H. (2018). *Using data to unlock the potential of an sme and mid-corporate franchise* (McKinsey & Company). https://www.mckinsey.com/industries/financial-services/our-insights/using-data-to-unlock-the-potential-of-an-sme-and-midcorporate-franchise

Culkin, R., & Das, S. R. (2017). Machine learning in finance: The case of deep learning for option pricing.’*Journal of Investment Management*, *15*, 92–100.

Das, S. R. (2014). Text and context: Language analytics in finance. *Foundations and Trends in Finance*, *8*, 145–260.

Das, S. R. (2016). Matrix metrics: Network-based systemic risk scoring. *Journal of Alternative Investments*, *18*, 33–51.

Das, S. R., & Chen, M. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science*, *53*, 1375–1388.

Das, S. R., Kim, S., & Kothari, B. (2019). Zero-revelation linguistic regulation: Detecting risk through corporate emails and news. *Journal of Financial Data Science*, *1*, 8–34.

Das, S. R., Kim, S., & Ostrov, D. (2019). Dynamic systemic risk: Networks in data science. *Journal of Financial Data Science*, *1*, 141–158.

Das, S. R., Mokashi, K. M., & Culkin, R. (2018). Are markets truly efficient? experiments using deep learning algorithms for market movement prediction. *Algorithms*, *11*, 138. https://doi.org/10.3390/a11090138

Espinosa-Vega, M., & Sole, J. (2010). *Cross-border financial surveillance: A network perspective* (Working paper No. 10/105). Washington, DC: International Monetary Fund.

Froot, K., Lou, X., Ozik, G., Sadka, R., & Shen, S. (2017). *Media reinforcement in international financial markets.* (Working paper). Boston, MA: Boston College.

Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, *55*, 665–676.

Graeber, D. (2018). *Bullshit jobs: A theory*. New York: Simon & Schuster.

Gunning, R. (1952). *The technique of clear writing*. New York: McGraw-Hill.

Harvey, C. R. (2014). *Bitcoin myths and facts* (SSRN Scholarly Paper ID 2479670), Rochester, NY: Social Science Research Network.

Hasbrouck, J., & Saar, G. (2013). Low latency trading. *Journal of Financial Markets*, *16*, 646–679.

Hasham, S., Hayden, R., & Wavra, R. (2018). *Combating payments fraud and enhancing customer experience*. (McKinsey & Company). https://www.mckinsey.com/industries/financial-services/our-insights/combating-payments-fraud-and-enhancing-customer-experience

Hendershott, T., & Riordan, R. (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis*, *48*, 1001–1024.

Higgins, P. C. (2014). *GDPNow: A model for GDP "Nowcasting"* (Report 2014-7), Atlanta, GA: Federal Reserve Bank of Atlanta.

Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance*, *49*, 851–889.

Jegadeesh, N., & Wu, D. (2013). Word power: A new approach for content analysis. *Journal of Financial Economics*, *110*, 712–729.

Leinweber, D. (2009). *Nerds on wall street*. Hoboken, NJ: Wiley.

Leinweber, D., & Sisk, J. (2011). Event-driven trading and the " New News."*Journal of Portfolio Management*, *38*, 110–124.

Lewis, M. (2014). *Flash boys*, New York: Norton.

Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, *59*, 17–35.

Lo, A. W. (2017). Moore's law vs. Murphy's law in the financial system: Who's winning? *Journal of Investment Management*, *15*, 17–38.

Loughran, T., & McDonald, W. (2011). When is a liability not a liability? *Journal of Finance*, *66*, 35–65.

Loughran, T., & McDonald, W. (2014). Measuring readability in financial disclosures. *Journal of Finance*, *69*, 1643–1671.

Menkveld, A. (2013). High frequency trading and the new market-makers. *Journal of Financial Markets*, *16*, 712–740.

Menn, J. (2010). *Fatal system error: The hunt for the new crime lords who are bringing down the internet.* New York: Perseus.

Merton, R. C. (1973). Rational theory of option pricing. *Bell Journal of Economics and Management Science*, *4*, 141–183.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, *29*, 449–470.

Mitra, G., & Mitra, L. (Eds.) (2011). *Handbook of news analytics in finance*. Chichester, UK: Wiley.

Nakamoto, S. (2009). *Bitcoin: A peer-to-peer electronic cash system*. Retrieved from https://Bitcoin.org/Bitcoin.pdf.

O'Neill, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown.

Philippon, T. (2016). *The fintech opportunity*. (Working paper). Manhattan, NY: New York University.

Piketty, T. (2014). *Capital in the twenty-first century*. Cambridge, MA: Harvard University Press.

Poulsen, K. (2011). *Kingpin: How one hacker took over the billion-dollar cybercrime underground*. New York: Crown.

PricewaterhouseCoopers, (2014). *Threats to the financial services sector* (Global Crime Survey, Technical report). London, United Kingdom: Author. https://www.pwc.com/gx/en/financial-services/publications/assets/pwc-gecs-2014-threats-to-the-financial-services-sector.pdf.

Rizzi, W., Wang, Z. M., & Zielinski, K. (2018). *How machine learning can improve pricing performance.* (McKinsey & Company). https://www.mckinsey.com/industries/financial-services/our-insights/how-machine-learning-can-improve-pricing-performance.

Srinivasan, S. (2016). *Big data and financial fraud using fintech methods.* (Working paper). Houston, TX: Texas Southern University.

Tetlock, P., Saar-Tsechansky, P. M., & Macskassay, A. S. (2008). More than words: Quantifying language to measure firm's fundamentals. *Journal of Finance*, *63*, 1437–1467.

Thakor, R. T., & Merton, R. C. (2018). *Trust in lending*. SSRN Scholarly Paper ID 3210412, Rochester, NY: Social Science Research Network.

Wei, Y., Yildirim, P., Bulte, C. V. D., & Dellarocas, C. (2016). Credit scoring with social network data. *Marketing Science*, *35*, 234–258.

Yermack, D. (2015). Is bitcoin a real currency? In D. Lee, & K. Chen (Eds.), *Handbook of Digital Currency* (pp. 31–44). Cambridge, MA: Academic Press.

Yermack, D. (2017). Corporate governance and blockchains. *Review of Finance*, *21*, 7–31.

**How to cite this article:** Das SR. The future of fintech. *Financial Management*. 2019;48:981–1007. https://doi.org/10.1111/fima.12297