From: Susan Von Struensee [redacted]
Sent: Saturday, June 19, 2021 12:22 PM
To: Comments
Subject: [EXTERNAL MESSAGE] RIN 3064-ZA24
Attachments: SSRN-id3622468.pdf; SSRN-id3864965.pdf

COMMENTS OF SUSAN VON STRUENSEE, JD, MPH

to the

Request for Information and Comment on Financial Institutions' Use of Artificial Intelligence, Including Machine Learning

86 FR 16837-38 (March 31, 2021)

Agency/Docket Numbers:

Docket ID OCC-2020-0049

Docket No. OP-1743

Docket No. CFPB-2021-0004

Docket No. NCUA-2021-0023

Please see attached file

A Survey of Fintech Research and Policy Discussion

The intersection of finance and technology, known as fintech, has resulted in the dramatic growth of innovations and has changed the entire financial landscape. While fintech has a critical role to play in democratizing credit access to the unbanked and thin-file consumers around the globe,

those consumers who are currently well served also turn to fintech for faster services and greater transparency. Fintech, particularly the blockchain, has the potential to be disruptive to financial systems and intermediation. Our aim in this paper is to provide a comprehensive fintech literature survey with relevant research studies and policy discussion around the various aspects of fintech.

The topics include marketplace and peer-to-peer lending, credit scoring, alternative data, distributed ledger technologies, blockchain, smart contracts, cryptocurrencies and initial coin offerings, central bank digital currency, robo-advising, quantitative investment and trading strategies, cybersecurity, identity theft, cloud computing, use of big data and artificial intelligence and machine learning, identity and fraud detection, anti-money laundering, Know Your Customers, natural language processing, regtech, insuretech, sandboxes, and fintech regulations.

Fintech activities have been progressing quickly and penetrating all areas of the financial system. Under the new landscape where financial firms use cloud storage and cloud computing to achieve high speed and efficiency, it is no longer true that larger firms are more efficient, and it is no longer true that larger firms beat smaller firms (but faster one beat those that are slower). In other

words, large firms that do not fully use the technology could potentially fall behind. Cloud computing may also introduce new exposure to cyber-risks that did not exist before, while it has greatly benefited traditional financial institutions in terms of efficiency, resiliency, and flexibility.

BigTech firms are playing increasing roles in financial services and real-time payments and providing cloud computing services to large and small financial institutions.

It remains unclear, however, how loans originated by fintech firms will perform relative to traditional loans in an extreme environment, and they are going through a real test now under the COVID-19 crisis. It is also unclear how the complex algorithms with alternative data, which worked well earlier, would continue to perform in the new landscape after the COVID-19 crisis. The AI models may need to be retrained with new data to reflect the "new normal."

In the meantime, fintech lenders have a role to play in the background (white label services) to assist small banks in screening and processing a large number of loan applications under the Coronavirus Aid, Relief, and Economic Security Act (CARES Act) to support small businesses.

Blockchain technology and smart contracts have attracted a lot of interest, with thousands of fintech start-ups around the globe working on resolving blockchain scalability, with the potential of allowing consumers to have control over their own information and less reliance on third-party intermediation. Blockchain technology has also made possible for real-time payment and settlement, eliminating frictions in the current payment system, especially for cross-border payments. The success of stablecoin could potentially contribute to more widespread use of blockchain technology. Central banks around the globe are considering whether to issue central bank digital currency (CBDC) to stabilize the value of digital assets or to take a supporting role for the public sector to operate in the new cash-lite economy. The Chinese government is also experimenting with blockchain technology, introducing the Blockchain-based Services Network (BSN) in April 2020, intended to serve as the backbone infrastructure technology for massive interconnectivity throughout the mainland China.

How should fintech firms be regulated? Many tech firms, especially the payment platforms, have access to unique data about their customers, and they have comparative advantage in making small business loans. These nonbank lenders are providing similar financial products and services outside the banking regulatory framework. There have been discussions around activity-focused regulations so that all lenders (banks and nonbanks) would be subject to the same regulations. It is debatable whether the future mainstream financial technology would be blockchain and DLTs, quantum computing, or something else — and how the industry and policymakers could be best prepared to keep pace with the evolving technologies and the new adoption.

While the advanced technology has delivered vast benefits, it has also allowed for more sophisticated cyberattacks. In addition, third-party vendor risk management has become more important than ever as loans that were approved on fintech platforms could end up on banks' balance sheet. Banks also increasingly rely on cloud computing services provided by nonbank BigTech firms.

Given all these innovations and rapid digital transformation, the existing regulations need to adapt to keep up with the new financial landscape — to protect consumers and financial systems while continuing to promote responsible fintech innovations.

Going forward, it's uncertain how fintech as the mainstream technology for the financial services industry would evolve. Economists have been trying to predict what would happen to fintech when the next recession comes. As we are finalizing this draft, the entire fintech industry is facing a real test for the first time under the current COVID-19 crisis. It has been evident that collaboration among fintech firms and traditional lenders have been critical in pushing the funding relief to those targeted recipients (mainly small businesses) in a timely fashion. There may be evidence of potential complementarity leading to a wave of mergers among financial firms and tech firms, thus further reducing the distinction between fintech firms and traditional firms.

Once again, this is an opportune time for researchers to further explore the impact of fintech on consumers and the overall financial systems and stability and to design appropriate financial regulations for the new landscape.

Citation

Allen, Franklin and Gu, Xian and Jagtiani, Julapa A., A Survey of Fintech Research and Policy Discussion (May, 2020). FRB of Philadelphia Working Paper No. 20-21, Available at SSRN: https://ssrn.com/abstract=3622468 or http://dx.doi.org/10.21799/frbp.wp.2020.21

Modelling in finance is a challenging task: the data often has complex statistical properties and its inner workings are largely unknown. Deep learning algorithms are making progress in the field of data-driven modelling, but the lack of sufficient data to train these models is currently holding back several new applications. Generative Adversarial Networks (GANs) are a neural network architecture family that has achieved good results in image generation and is being successfully applied to generate time series and other types of financial data. The purpose of this study is to present an overview of how these GANs work, their capabilities and limitations in the current state of research with financial data, and present some practical applications in the industry. As a proof of concept, three known GAN architectures were tested on financial time series, and the generated data was evaluated on its statistical properties, yielding solid results. Finally, it was shown that GANs have made considerable progress in their finance applications and can be a solid additional tool for data scientists in this field.

Citation:

Eckerli, Florian, Generative Adversarial Networks in finance: an overview (June 11, 2021). Available at SSRN: https://ssrn.com/abstract=3864965 or http://dx.doi.org/10.2139/ssrn.3864965

Conclusion

I recommend that the Comptroller of the Currency, the Federal Reserve System, the Federal Deposit Insurance Corporation, the Consumer Financial Protection Bureau, and the National Credit Union Administration take steps to ensure compliance in FinTech with the OECD AI Principles, and the OSTP/OMB Guidance on Regulation of Artificial Intelligence Applications. I specifically recommend that the industry limit the scope of defenses for negligent and fraudulent parties whose actions have a legal or significant effect on an individual and discredit the commitment toward trustworthy AI.

Respectfully Submitted,

Susan von Struensee, JD, MPH

Keywords: fintech, marketplace lending, P2P, alternative data, DLT, blockchain, robo advisor,

regtech, insuretech, cryptocurrencies, ICOs, CBDC, cloud computing, AML, KYC, NLP, fintech regulations. GANS

Federal Reserve Bank of Philadelphia

Working Papers
Research Department

WP 20-21
June 2020
https://doi.org/10.21799/frbp.wp.2020.21

A Survey of Fintech Research and Policy Discussion

Franklin Allen
Imperial College London

Xian Gu
Central University of Finance and Economics and the University of Pennsylvania

Julapa Jagtiani
Federal Reserve Bank of Philadelphia Supervision, Regulation, and Credit Department

Consumer Finance Institute

ISSN: 1962-5361

A Survey of Fintech Research and Policy Discussion*

Franklin Allen
Imperial College London

Xian Gu
Central University of Finance and Economics and
the University of Pennsylvania

Julapa Jagtiani
Federal Reserve Bank of Philadelphia

First Draft: April 21, 2020
Current Draft: May 28, 2020

Abstract

The intersection of finance and technology, known as fintech, has resulted in the dramatic growth of innovations and has changed the entire financial landscape. While fintech has a critical role to play in democratizing credit access to the unbanked and thin-file consumers around the globe, those consumers who are currently well served also turn to fintech for faster services and greater transparency. Fintech, particularly the blockchain, has the potential to be disruptive to financial systems and intermediation. Our aim in this paper is to provide a comprehensive fintech literature survey with relevant research studies and policy discussion around the various aspects of fintech. The topics include marketplace and peer-to-peer lending, credit scoring, alternative data, distributed ledger technologies, blockchain, smart contracts, cryptocurrencies and initial coin offerings, central bank digital currency, robo-advising, quantitative investment and trading strategies, cybersecurity, identity theft, cloud computing, use of big data and artificial intelligence and machine learning, identity and fraud detection, anti-money laundering, Know Your Customers, natural language processing, regtech, insuretech, sandboxes, and fintech regulations.

Keywords: fintech, marketplace lending, P2P, alternative data, DLT, blockchain, robo advisor, regtech, insuretech, cryptocurrencies, ICOs, CBDC, cloud computing, AML, KYC, NLP, fintech regulations

JEL Classification: G21, G28, G18, L21

---

*Author contacts: Julapa Jagtiani, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106, julapa.jagtiani@phil.frb.org; Franklin Allen, Imperial College London, f.allen@imperial.ac.uk; and Xian Gu, Central University of Finance and Economics, and the University of Pennsylvania; xiangu@wharton.upenn.edu. Comments are welcome. The authors thank Mitchell Berlin and Bill Wisser for their comments, and thanks to Erik Dolson, Adam Lyko, Dan Milo, and Andes Lee for their research assistance.

1. Introduction

The rapid advance in financial technology (fintech) in recent years has played an important role in how financial products and services are produced, delivered, and consumed. Fintech has become one of the most popular discussion topics recently, primarily because of its potential disruption to the entire financial system. There has been a dramatic digital transformation in the financial landscape. The term *fintech* is, however, a broad term, and it tends to mean different things to different people. The goal of this paper is to describe the various aspects of fintech and its role in each segment of the financial market and the associated impact on consumers and the financial system overall.

A great deal of data have been collected in recent years. For example, as of 2016, IBM estimated that 90 percent of all the global data was collected in the past year. The amount of data collection accelerated even more between 2016 and 2020. There have been new opportunities for data to be monetized, such as through data aggregation. Big data (including data from nontraditional sources and trended data) have been collected and used widely, in conjunction with advances in artificial intelligence (AI) and machine learning (ML) for digital identity and fraud detection, sales and marketing, security trading strategies, risk pricing and credit decisions, and so forth. More than 2 billion consumers are currently excluded from financial systems around the globe (especially in less developed countries such as Bangladesh, Nigeria, and Pakistan) who could potentially benefit from the use of more data and complex algorithms to access credit. There also have been new questions related to data ownership and the ethical use of data, such as who should have control over the ability to aggregate, use, and share data to safeguard consumer privacy and to avoid systemic misuse of consumer data.

Cloud storage and cloud computing have also played increasing roles in payment systems, financial services, and the financial system overall. Financial data and payment data have been stored in the cloud, and cloud computing has made it possible for many fintech innovations, such as real-time payment and instantaneous credit evaluations/decisions. Firms no longer need to commit a large investment (usually unaffordable for smaller firms) to in-house technology, but they could outsource to the cloud computing service providers and share the cost with other firms. This leveled the playing field; size is no longer the most important determinant for success. Consumers' preferences have also adapted to prioritize faster services and greater convenience and transparency through online services and applications. There have been concerns among regulators about the impact on the safety and soundness and stability of the financial systems (e.g., the impact on the payment system when a cloud service platform is rendered nonoperational, the exposure to a greater risk of cyberattack, and other similar events).

Blockchain and smart contracts are the buzzwords in the fintech community, partly because blockchain is the technology underlying bitcoin transactions. Blockchain and other digital ledger technologies (DLT) have also been used in creating various cryptocurrencies, initial coin offerings (ICOs), other payment applications, and smart contracts — thus, leading some to believe that blockchain has the potential to become the mainstream financial technology of the future. There has been some disappointing evidence on the role and potential of blockchain in that it may not be as disruptive as initially expected, and one of the main obstacles seems to be its scalability. For example, bitcoin transactions take about 10 minutes to clear, and it is expected to take longer as the block length gets longer over the years. While thousands of tech start-ups and other tech experts have been working to resolve the issue, permissioned blockchain platforms have benefited some segments of the economy through their use for identity detection, supply chain management, digital-asset-backed lending, and securitization.

Fintech activities have been progressing quickly, penetrating all areas of the financial system. Fintech has produced great benefits to a large number of consumers around the world and has made the financial system more efficient. The rapid growth of bank-like services provided by fintech firms has raised potential concerns among bank supervisors. There have also been legal challenges and concerns associated with fintech around consumer privacy and the potential fintech disruption to overall financial stability. While fintech could greatly improve credit access and enhance efficiencies (providing faster, better, or cheaper services) in the financial system, risk cannot be completely eliminated. In this paper, we provide a comprehensive summary of what research studies have found so far, what the experts (academic, industry, and regulators) are working on, and the potential evolving nature of fintech's impact on consumer privacy and well-being, the structure of the financial and payment systems, the role of financial intermediation, and the effectiveness of existing regulatory policies. The rest of the paper is organized as follows.

In Section 2, we discuss recent enhanced systems for credit scoring using AI/ML and alternative data, the roles of marketplace lending and peer-to-peer (P2P) lending, and digital banking and investment services. Section 3 discusses how fintech has played a big role in digital payment, such as e-wallet and allowing a large number of the unbanked population around the world to be included in financial systems for the first time. The roles of alternative data in financial inclusion, improving credit access, and more accurate risk pricing will also be discussed.

Section 4 describes the roles of blockchain, other distributed ledger technologies (DLTs), and smart contracts. As mentioned earlier, these have been the underlying technologies for cryptoassets and initial coin offerings (ICOs), which will be discussed in Section 5. There are frictions in the current payment system, especially cross-border payments. Consumers have come to expect faster or real-time payments with minimal fees. Digital currencies could potentially deliver these, and the payment processes have been involving rapidly toward a cash-lite (or potentially cashless) economy. Section 5 will also discuss the developments around the potential for central banks to issue fiat digital currencies, so-called central bank digital currency (CBDC). This idea of CBDC acknowledges that trust is the most important factor in payments, and private sectors may not be able to accomplish the goal of originating and supporting the value of the digital currencies it issues. There are also fears around CBDC: Several key considerations need to be incorporated into CBDC's design to avoid adverse impact on the financial system and the ability to conduct effective monetary policy.

Section 6 deals with fintech's roles in securities trading and markets, such as the high-frequency trading or program trading that uses big data and ML algorithms to deliver superior performance. Section 7 discusses the impact of fintech on cybersecurity, which has been one of the top concerns among corporate CEOs and senior management teams. While the advanced technology has delivered vast benefits, the technology has also allowed for more sophisticated cyberattacks. Given all these innovations and rapid digital transformation, the existing regulations need to adapt to keep up with the new financial landscape.

The increasing roles of BigTech and cloud computing in financial services, their potential impact on interconnectedness between financial institutions, and how these activities are likely to evolve in the near future are discussed in Section 8. There are just a handful of providers for all financial institutions, and these providers are currently not subject to supervision by bank regulators. There have been concerns about quality control, data security, and a possible conflict of interest that need to be addressed in the new fintech regulatory framework. Some of the technologies have also been used to assist regulators in regulatory compliance examination, such as the natural language processing (NLP) and the ML techniques used in RegTech, which will be discussed in Section 9, along with the various factors to be considered in designing fintech regulations to protect consumers and the financial systems while continuing to promote responsible fintech innovations.

Finally, Section 10 provides conclusions and policy implications, such as those related to open banking policy, ethical use of consumer data, and whether a cashless economy is expected in the near future. Quantum computing has also been transitioning from theory into practice, with potential implications/disruptions in the financial services industry and the overall economy in the coming decade. It is debatable whether the future mainstream financial technology will be blockchain and DLTs, quantum computing, or something else — and how the industry and policymakers can best be prepared to keep pace with evolving technologies and the new adoption. We will also discuss potential directions for future fintech research.

2. Credit Scoring, Digital Banking, and Marketplace Lending

2.1 Credit Scoring Using AI/ML and Alternative Data

Credit scores, such as FICO scores (or Vantage Scores), have served as the primary factors in credit decisions, especially for credit card applications. Previous studies, such as Mester, Nakamura, and Renault (2007) and Norden and Weber (2010), have documented the importance of consumer credit history and other financial and accounting data in credit risk evaluation by lending institutions. However, about 26 million American consumers have thin credit files or do not have bank accounts (unbanked); thus, they do not have FICO scores because of an insufficient credit history. More recently, there has been a breakthrough in which consumers' default probability could be estimated not only from their official credit history or credit ratings but rather from more complex statistical methods using AI and ML techniques, along with (nontraditional) alternative data. These big data and complex algorithms have been rapidly adopted by fintech lenders to overcome the limitations of traditional models and data in evaluating borrowers' credit risk and their ability to pay back loans.

Fintech lending, which started in personal lending after the recent financial crisis, has expanded to cover small business lending and mortgage lending in recent years. Previous research studies that compare traditional default prediction models with more advanced techniques using AI/ML seem to suggest that there are significant lifts in predictive ability. Jagtiani and Lemieux (2019), Goldstein, Jagtiani, and Klein (2019), and Croux, Jagtiani, Korivi, and Vulanovic (2020) have documented that the information asymmetry, which used to be one of the critical issues in evaluating borrower risks, could be overcome through AI/ML and alternative data, especially for borrowers with limited credit history and the unbanked population. Soft information about borrowers can be obtained without being in close proximity to the borrowers in peer-to-peer (P2P) lending. This includes information on friendship and social networks, online footprints, and text-based analysis; see Iyer, Khwaja, Luttmer, and Shue (2016); Hilderbrandt, Puri, and Rocholl (2017);

Lin, Prabhala, and Viswanathan (2013); Gao, Lin, and Sias (2017); Dorfleiter et al. (2016); and Berg, Burg, Gombovic, and Puri (2018).

In addition, Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2018) examine the mortgage market by comparing the traditional logit models with ML techniques in predicting borrowers' default probability. They find that using ML models would result in a slightly larger number of borrowers who would have access to credit, and ML models would also marginally reduce disparity in the acceptance ratios across racial and ethnic groups.[1] Overall, they find evidence of the higher out-of-sample predictive accuracy for default when using ML technology compared with simpler logistic models.

Similarly, Gao, Lin, and Sias (2018) use a natural language processing (NPL) technique, a form of AI, to analyze textual information from borrowers' writing on online lending platforms. This allows them to efficiently quantify the informational content of lengthy personal text descriptions, which were not previously possible. The authors use a combination of well-developed computational linguistics measures and advances in ML to examine the roles of linguistic style for borrowers on the Prosper (fintech lending) platform. They find that lenders tend to bid more aggressively, are more likely to grant credit, and charge lower interest rates to borrowers whose writing is more readable, more positive, and contains a lower level of deception cues. Again, the evidence supports an argument that advanced technology used in credit decision could expand credit access to more consumers based on nontraditional data. Lenders, however, are required to demonstrate that the use of alternative data does not result in biased credit decisions against certain segment of the borrowers, such as race and gender.

Likewise, Iyer, Khwaja, Luttmer and Shue (2016) use data from the same Prosper lending platform, and they find that nontraditional information tends to play an important role in credit risk evaluation and allow credit access for borrowers who are less creditworthy by traditional measures, such as those with credit scores below prime.[2] The authors cited the following basic proxies for soft information — whether the borrower posts a picture, the number of words used in the listing text descriptions, friend endorsements, etc. Their findings suggest that, while traditional information such as credit score, requested loan amount, and current delinquencies are important in credit decisions, lenders also use alternative data (nonstandard information about the borrowers),

[1] However, they also note that the cross-group disparity of equilibrium interest rates increases when using ML models. They attribute these changes primarily to greater flexibility.

[2] However, due to difficulties in coding, they do not fully quantify the large selection of soft information available in Prosper listings.

especially for less creditworthy and borrowers with thin files who have fewer alternate funding options.

Consistent with these findings, Berg, Burg, Gombovic, and Puri (2018) analyze the information content of borrowers' digital footprints, such as activities people do online. The proprietary data set is from an e-commerce company located in Germany. This includes a basic set of variables, such as whether their email contains their real name, whether they make purchases at night time, and the number of typing mistakes, all of which are found to be important. Their empirical investigations suggest that even the simple, easily accessible variables from the digital footprint could be valuable for consumers' default prediction. In fact, they find that these variables alone are as good as the Equifax Risk Score in predicting loan outcomes. In addition, using these online behavioral variables together with the Equifax Risk Score would increase the model's predictive power overall, suggesting that the digital footprint variables complement (rather than substitute) standard traditional information from consumer credit bureaus. Therefore, lenders could enhance their credit and risk pricing decisions by looking at both the traditional risk scores and alternative data.

More recently, fintech giants Alibaba and WeChat Pay Points Credit by Tencent in China have built their new credit scoring system based on alternative data they collect from nontraditional sources, including social media, online shopping, payment applications, cell phone accounts, and others. This type of scoring provides a more comprehensive view of consumers' financial lives, and it is meant to help fill the credit gap for people who cannot get a loan because of their lack of credit history.

In addition to using alternative data and AI/ML to potentially expand credit access, another important question is the price of credit. Jagtiani and Lemieux (2019) use loan-level data from the LendingClub personal lending platform, and they compare interest rates (APRs that account for all the origination fees) charged by LendingClub with the interest rate borrowers would have to pay by carrying a credit card balance (i.e., contractual interest rate from Y-14M stress test data reported by CCAR banks). They find that the use of alternative data by LendingClub has allowed some below-prime consumers to receive credit at much lower cost. The rating grades assigned by LendingClub (using information not related to FICO scores) also perform better in predicting loan outcomes. Related to this, Jagtiani, Hughes, and Moon (2019) find that LendingClub also became more efficient than other lenders in its peer group size as of 2016 (after more alternative data has been incorporated into credit scoring and pricing) compared with 2013.

2.2. Internet-Based Banking and Investment Services

Recent retail banking services innovations primarily rely on technological advances such as faster Internet access and improved payment process. The fast-growing Internet access since the late 1990s has spurred the adoption of online banking. In 1995, the Security First Network Bank was the first online-only bank created, and Wells Fargo was the first brick-and-mortar bank to launch online banking (checking account) websites around the same time (Hernandez-Murillo, Llobet, and Fuentes, 2010). By 2001, eight banks in the U.S. had more than 1 million online customers. Banks in other countries around the globe also started to develop software applications to allow customers to access their accounts online. Digital banking was being adopted along with some new regulatory frameworks to address the associated new risks (Arner, Barberis, and Buckley, 2016a).

The determinants of the adoption of Internet-based banking has been well documented by existing literature. Furst, Lang, and Nolle (2002) use a cross-section sample of banks in 1999 and show that important factors that determine a bank's decision to adopt new Internet technologies for online banking services are those that are more profitable, have a larger asset size, have presence in urban markets, and are a subsidiary of a bank holding company (BHC). Hernandez-Murillo, Llobet, and Fuentes (2010) use a more recent sample of U.S. banks and confirm that, in addition to bank characteristics (branching intensity, capital-to-asset ratio, nonperforming loan ratio, being subsidiary of a BHC), bank customers' demographic factors (household income, education, Internet access) are also important determinants of online banking adoption.

The adoption of online banking also affects bank performance. Previous studies have documented that, over time, Internet adoption reduces the bank's operating costs and improve its profitability. Interestingly, DeYoung, Lang, and Nolle (2007) and Hernando and Nieto (2007) find that Internet banking services would not replace bank branches because online banking seems to be complementary rather than a substitute for physical and personal banking services. Goddard, McKillop, and Wilson (2009) find that U.S. credit unions that do not provide Internet banking services to their members are more likely to fail or to be acquired by another institution that provides online banking.

More recently, advances in financial technology also improved digitization in payments — both retail and wholesale payments. Rysman and Schuh (2016) summarize recent developments in mobile payments and a faster payments system. Technology firms, rather than banking firms, have been leaders in providing mobile payments in the U.S., suggesting potential benefits from partnerships between banks and fintech firms. Several fintech firms have also been providing white-label technological services for their bank partners. Partnership alliance between banking firms and technology companies would also help banks to digitalize their credit decision process and their risk management.[3] Klus, Lohwasser, Hornuf, and Schwienbacher (2018) investigate the drivers and the extent to which banks interact with fintech start-ups, using detailed information on strategic alliance made by the 100 largest banks in Canada, France, Germany, and the UK. They find that larger banks are better able to integrate start-ups into their own business model and strategy than smaller banks. They also find significantly positive market reactions in response to announcements of alliance formation between digital banks and fintech firms.

Technological innovations in recent years have allowed for the rapid adoption of mobile payment and fast payment systems around the globe, especially in China. The adoption has been slower in the U.S.[4] Crowe, Rysman, and Stavins (2010) find that standard concerns such as the cost of adoption and network effects could play a critical role in the speed of adoption. The cost of adoption may be higher in the U.S. because of the fragmented nature of the telecommunications system and the banking industry and the legacy payments system in the U.S. To implement a single national mobile payment mechanism would require an agreement between all the mobile carriers and the banks that issue cards to consumers. Greene, Rysman, Schuh, and Shy (2014) and Rysman and Schuh (2016) discuss the pros and cons of implementing a faster payments system in the U.S., suggesting that the net potential benefits could be large and could extend beyond faster speed. It would, however, be expensive to build a completely new payments system in the U.S. than in countries that did not have an established legacy payments system. There are also related questions about how such a new system should be managed and funded, the fee structure, and new regulatory regimes associated with the new payments systems.

2.3 Peer-to-Peer (P2P) Lending

Marketplace lending (MPL), commonly referred to as peer-to-peer (P2P) lending platforms, has emerged as an appealing new channel of financing for consumers and small businesses over the last decade. P2P lending platforms are designed to match lenders and borrowers and to eliminate the intermediary middleman. The platforms connect investors (funding supply) and borrowers (funding demand) directly to facilitate the transaction. The patterns showing that new traditional

[3] For example, China Construction Bank (CCB), one of the Big-4 state-owned banks, announced a strategic pact with Alibaba in 2017, which allowed Alibaba to sell CCB's financial products through its Ant Financial platforms (https://www.spglobal.com/marketintelligence/en/news-insights/trending/hsmye-hebytelrkvaqiqtg2).

[4] While consumers perceive debit and credit card payments fairly quickly as immediate, the settlement step often takes several business days. Payments through the automated clearing house (ACH) can also take several business days to complete.

bank loans are trending downward, while new P2P lending are trending upward, raise a natural question whether these two types of lenders complement or substitute each other in the credit market. Balyuk (2017) explores the effect of P2P lending on consumers' access to credit and finds that fintech lenders have improved credit access to consumers who cannot access credit from traditional banks — see also Chava and Paradkar (2018). Similarly, Jagtiani and Lemieux (2018) find that, using data from the LendingClub consumer platform, fintech lenders have penetrated areas that may be underserved by traditional banks, such as in highly concentrated markets and areas that have fewer bank branches per capita. Other studies, such as Tang (2018) and de Roure, Pellizon, and Thakor (2018), show that P2P lending could be a substitute for or complement bank lending, depending on the situation. If assuming banks face an exogenous regulatory shock, P2P lending could act as a substitute for bank lending for serving inframarginal borrowers; however, on small-scale loans, P2P lending would be more likely to complement bank lending.

There have also been questions whether the P2P lending platform could uncover the "correct" pricing for the transactions. There are two market mechanisms to pricing: through an auction or a more common approach through posted prices. An auction process typically relies on the relative strength of lenders and borrowers to determine the price; whereas, posted prices are predetermined by complex algorithms used by the lending platforms.[5] Wei and Lin (2017) show that under posted prices, borrowers are more likely to obtain credit, but the default probability is also higher. Franks, Serrano-Velarde, and Sussman (2018) use a peer-to-business lending platform and find that a 1 percent increase in the interest rate corresponds to a less than a half percent increase in the default probability, implying that information efficiency was not reached in the pricing process. Participation of sophisticated investors can improve screening outcomes but also could create adverse selection among investors. Vallee and Zeng (2018) and Liskovich and Shaton (2017) find that sophisticated investors tend to outperform less experienced ones, especially when platforms provide more sufficient information to investors through websites or application programming interfaces (API).

Loan pricing seems to be affected by soft information about borrowers. Butler, Cornaggia, and Gurun (2017) and Crowe and Ramcharan (2012) show that in addition to standard financial variables (such as credit scores, employment history, homeownership), P2P platforms also provide nonstandard soft information about the borrowers that is useful for lenders to make credit decisions. Previous research shows that the borrowers' soft information, including personal

[5] One of the lending platforms, Prosper, switched from the auctions approach to the posted prices approach in December 2010 (see Prosper.com).

characteristics such as race, age and beauty (Ravina, 2018), social capital (Lin and Pusianinen, 2018; Hasan, He, and Lu, 2019), characteristics of listing text (Iyer, Khwaja, Luttmer, and Shue, 2016), hometown (Lin and Viswanathan, 2016), social network (e.g., friend endorsements) (Lin, Prabhala, and Viswanathan, 2013; Feedman and Jim, 2014) affects lenders' decision in terms of ex-ante loan pricing and loan amount. These factors also improve the ex-post lending outcomes. Many of these factors, however, are prohibited in credit decisions, according to the fair lending and consumer privacy regulations. Jagtiani, Vermilyea, and Wall (2018) describe how big data and complex algorithms from AI/ML are being used appropriately by lenders and other market participants and how regulators are adapting their supervisory approach accordingly.

3. Fintech and Financial Inclusion

3.1 Digital Wallet and Credit Access for Unbanked Individuals

Financial inclusion, typically defined as the use of formal financial services, especially by the disadvantaged, has become a subject of growing interest. Indeed, the positive relationship between financial development and economic development, documented by literature on financial-growth nexus, is suggestive of a positive association between finance and poverty alleviation (Beck, Demirguc-Kunt, and Levine, 2007). A central question is how to promote financial inclusion for regions endowed with asymmetric information, weak institutions, and lack of basic infrastructure necessary for banking. According to the Global Finance database by the World Bank, 31 percent of adults do not have a bank account as of 2018, a decline from 38 percent in 2014 and 49 percent in 2011. The numbers also show that there are large disparities between males and females and between rich and poor.[6] Kendall, Mylenko, and Ponce (2010) make a rough estimation of banked and unbanked individuals around the world using a new set of financial access indicators for 139 countries as of 2003. They estimate that people have on average 3.2 financial accounts per adult, and 81 percent of the adults have at least one bank account in developed countries. In contrast, people have on average less than one account (specifically only 0.9 account) in developing countries, and only 28 percent of the adults in developing countries have a bank account.

What are important factors that determine the level of financial access? Beck et al. (2007) show that the overall level of economic development, the quality of the institutional environment, the degree of credit information sharing, the level of initial endowments, and the development of

[6] For more details, see the article "Financial Inclusion on the Rise, But Gaps Remain, Global Findex Database Shows" by the World Bank, https://www.worldbank.org/en/news/press-release/2018/04/19/financial-inclusion-on-the-rise-but-gaps-remain-global-findex-database-shows.

physical infrastructure are positively associated with financial outreach and depth; whereas, the cost of enforcing contracts and the degree of government ownership of banks are negatively associated with the financial outreach and depth. Kendall, Mylenko, and Ponce (2010) also suggest a similar conclusion. In addition, legal origin and religion have a less consistent impact on financial inclusion. Allen et al. (2016) explore the factors underpinning financial inclusion; they show that policies to promote financial inclusion are especially effective among the most commonly excluded groups of individuals from access to finance: the poor, those living in rural areas, females, and young individuals.

Fintech can promote financial inclusion in different ways. Allen et al. (2019) examine bank branch penetration and financial access in Kenya, a country that has made significant strides in financial inclusion. The emergence of Equity Bank, a pioneering and private institution that devised a banking service targeting low-income and less-educated customers and underserved regions, has had a positive and significant impact on households' use of bank accounts and credit access in Kenya during 2006 to 2010. The number of deposit and loan accounts of Equity Bank represents around 50 percent and 30 percent of the total number of deposit and loan accounts in Kenya, respectively. Their finding suggests that the successful business model of Equity Bank in Kenya, provides a solution to the financial access problem that has hindered real growth in many African countries. Similarly, Hau et al. (2018), using a comprehensive loan data set from Alibaba, a lending online e-commerce platform, show that fintech credit promotes financial inclusion in China. Fintech helps to mitigate local credit supply frictions in the credit market and extend the "frontier" of credit availability to small businesses with low credit scores. In addition, these online e-commerce platforms have promoted a self-selection process in which more funding tends to be channeled to those online merchants that receive better rating by their customers; see Huang, Li, and Shan (2019).

3.2 The Roles of Alternative Data in Improving Credit Access and Risk Pricing

The recent rapid development in fintech and use of alternative data in evaluating credit risk have also raised questions about whether the use of alternative data and advanced algorithms has improved credit assessment and enhanced access for the underbanked. A related issue is whether these innovations carry a risk of disparate treatment and violate fair lending and consumer privacy regulations. Schweitzer and Barkley (2017) examine the characteristics of online small business borrowers using the Federal Reserve's 2015 Small Business Credit Survey data and find supportive results that businesses denied access to credit by banks turned to fintech lenders to arrange credit for their business. Similarly, Jagtiani, Lambie-Hanson, and Lambie-Hanson (2019) explore the mortgage market and find evidence suggesting that more borrowers turned to mortgage fintech lenders in areas where there were higher denial rate by traditional lenders.

Consistent findings have been documented by other studies. Ahmed et al. (2016) find that borrowers would have been unable to secure external financing without the fintech online lending platform, despite being creditworthy. Jagtiani and Lemieux (2019), using account-level data from LendingClub and Y-14M bank stress test data, find that for the same risk of default, consumers pay smaller spreads on personal installment loans from a fintech platform than from (traditional) credit card borrowing. Berg, Burg, Gombovic, and Puri (2018) suggest that digital footprints can help boost financial inclusion, allowing unbanked consumers to have better access to finance. Similarly, Frost et al. (2019) show that fintech firms often start as payment platforms and later use consumer data to expand into some provisions of credit, insurance, and savings and investment products. These firms lend more in countries with less competitive banking systems and less stringent regulations. In the case of Argentina, fintech lenders have information advantages in credit assessment using alternative data and advanced algorithms.

The use of alternative data has also raised issues related to fair lending. As mentioned earlier, previous studies show that soft information such as applicants' look, relationship, and network has been widely used in evaluating credit availability and pricing. Fu, Huang, and Singh (2018) find that, despite the great benefits that big data and complex ML algorithms can bring, these predictions of creditworthiness could be subject to bias, and lenders may unknowingly make credit decisions based on factors that are related to protected class (such as race and gender), even though these sensitive attributes are not used as inputs to the model. Padhi (2017) show that the discrimination and competition in online alternative lending could result in a riskier pool of potential borrowers for banks and, thus, could impact safety and soundness in the banking system and threaten the financial stability overall. Jagtiani and John (2018) provide an overview of how alternative data and other fintech innovations have changed the entire financial landscape.

4. Distributed Ledger Technology (DLT), Blockchain, and Smart Contracts

Distributed ledger technology (DLT) is a term widely used to describe various record-keeping technologies, such as decentralized data architecture and cryptography, which allow the keeping and sharing records in a synchronized way while ensuring their integrity through the use of consensus-based validation protocols. Blockchain is a specific type of DLT, containing blocks of records that are linked using cryptography. Each block contains a cryptographic hash of the previous block, a timestamp, and transaction data. By design, a blockchain is intended to be resistant to modification of the data; however, there have been concerns around cybersecurity related to blockchain.[7] In general, blockchain is designed to be an open, distributed ledger that can record transactions between two parties efficiently and in a verifiable and permanent way. Blockchain has become a buzzword mainly because it is the main technology underlying bitcoin transactions. Since late 2015, blockchain in general (not bitcoin blockchain) attracted explosive interest from the industry as a new way to create, exchange, and track ownership of financial assets on P2P platforms. Blockchain technology has also facilitated the creation of smart contracts; these are computerized protocols allowing terms contingent on decentralized consensus that are tamper proof and self-enforcing via automated execution; see Szabo (1994) and Cong and He (2018).

The idea of blockchain was initially introduced by Haber and Stornetta (1991) to authenticate authorship of intellectual property. Much later in 2008, Nakamoto (2008) reintroduced it as a method of validating ownership of the cryptocurrency bitcoin. It is viewed as a potential mainstream financial technology of the future to eliminate a trusted third party in financial intermediation. Szabo (2005) proposes a similar idea to overcome the problem of the dependence on a trusted third party. The proposed "bit gold" was a protocol whereby costly bits could be created online with minimal dependence on trusted third parties and then securely stored, transferred, and assayed with similar minimal trust. The creation of such bit gold would be based on various functions including "client puzzle function," "proof of work function," or "secure benchmark function."

There are three main types of blockchain. The first type, a private blockchain, comes originally from the chain proposed by Haber and Sornetta (1991). In this chain, there needs to be an entity with authority, identified as sponsor or gatekeeper, taking complete control over what is written on the ledger. Such sponsors or gatekeepers can restrict entry into a market, access monopolistic user fees, edit incoming data, or limit users' access to market data. The second type, a permissioned blockchain, is one in which the write privilege is granted to a consortium of entities. These entities govern the policies of the blockchain and take control of propagating and verifying transactions. The third type is a public blockchain, in which the write privilege is completely unrestricted. Because the writers are allowed to be anonymous in the public blockchain, there needs to be an efficient, fair, and real-time mechanism to ensure that all participants agree on a consensus on the status of the ledger. Well-known types of consensus mechanism algorithms are:

[7] Gervais et al. (2016) studies the security issues existing in blockchains, focusing on how to defend against "double-spend" attacks or other types of attacks that could be undertaken by an entity control over a large portion of the computing power of the network.

(1) proof of work (POW), which is used by major cryptocurrency networks such as Bitcoin and Litecoin; and (2) proof of stake (POS), which is used by Peercoin.[8] Yermack (2017) and Abadi and Brunnermeier (2018) offer an introduction of the development and different types of blockchain.

Major central banks and stock exchanges have been exploring the usage of DLT in payments, clearing and settlements; see Mills et al. (2016) and Benos, Garratt, and Gurrola-Perez (2017). The driving force behind the efforts in deploying DLT in financial transactions and settlement is an expectation that this technology will reduce (or even eliminate) inefficiencies and related frictions that currently exist for storing, recording, transferring, and exchanging digital assets through financial markets. Decentralization, which is the core concept of the DLT, has both pros and cons. A growing literature discusses the advantages that a blockchain can offer compared with the traditional ledgers. The blockchain-based smart contract is encoded to assure one party that its counterparty will fulfill the promise with certainty; therefore, it can eliminate some contracting frictions like the need for costly verification or enforcement, in an automated and conflict-free way; see Cong and He (2018) and Harvey (2016). Meanwhile, through decentralized consensus, blockchain improves the ability to bootstrap and operate a marketplace without the need for a traditional intermediary, which further lowers the cost of networking and improves the competition. Overall, participants can make investments to support and operate a shared infrastructure without assigning market power to a platform authority; see, e.g., Catalini and Gans (2018) and Abadi and Brunnermeier (2018). More specifically, Yermack (2017) overviews the impact of blockchain on corporate governance and argues that, in addition to resulting in lower cost and more accurate record keeping, a blockchain can also bring greater liquidity and improve transparency of ownership. Similarly, Ma, Gans, and Tourky (2018) argue that regulation of bitcoin mining can reduce the overall costs of the system and improve welfare.

Despite these great benefits, there are also potential negative features around blockchain technology. For example, Cong and He (2018) argue that, through distributing all transaction information, business privacy can be an issue when pushing for real-world blockchain application. Similar findings are also documented in Malinova and Park (2017), as traders want to hide their

[8] The proof of work (POW) requires a participant node to prove that it has accomplished a computationally difficult task before getting the rights to write on the ledger. Therefore, one disadvantage of POW is that it needs high-energy consumption and longer processing time. The proof of stake (POS) emerges and evolves as a low-cost and low-energy consuming alternative to POW algorithm, attributing mining power to the proportion of coins held by each miner. It is reported that one of the major networks, Ethereum, is transiting to adopt the POS blockchain instead of POW soon (see, e.g. https://www.coindesk.com/testnet-gorli-ethereum-serenity). Other types of consensus algorithms include proof of capacity (POC), allowing the sharing of memory space of the contributing nodes on the blockchain network.

identities to prevent front running. Lack of full transparency, the net aggregate welfare is believed to be weakly higher if investors are allowed to split their holdings among many identifiers. Abadi and Brunnermeier (2018) propose a Blockchain Trilemma, suggesting that no ledger can satisfy the three properties simultaneously: correctness, decentralization, and cost efficiency. Distributed ledgers promote competition but eliminate rents through "fork competition," endangering instability, and miscoordination. Tinn (2018) documents that blockchain technology facilitates faster learning and more frequent effort decisions, which in turn changes the type of financing contracts that are the most efficient or could even make traditional debt and equity contracts more costly. Easley, O'Hara, and Basu (2019) theoretically analyze how equilibrium transaction fees evolve in the bitcoin blockchain, from a mining-based structure to a market-based ecology, and demonstrate that transaction fees are not welfare improving.

5. Cryptocurrencies, Initial Coin Offerings, and Central Bank Digital Currencies

In recent years, the initial coin offerings (ICOs) have exponentially grown as a new form of financing for start-ups. There are 4,136 cryptocurrencies in existence as of May 2020.[9] This number does not include many that have failed. One of the most successful ICOs was from Ethereum in July 2014 that collected in total $18.4 million USD in 42 days.[10] The market capitalization of Ethereum reached more than $30 billion in 2017.[11] These successes led to great excitement about cryptocurrencies as a new form of financing for innovations in the upcoming digital age but, in the meantime, created concerns about rampant speculation and financial instability.

Given the extreme volatility of cryptocurrencies,[12] stablecoins have been developed in the crypto world. The central feature of stablecoins is that they are pegged to another asset, like the U.S. dollar, national currencies, or commodities, and their issuers back up the value of coins through holding sufficient reserves to support the value of the stablecoins. In June 2019, Facebook announced its Libra 1.0, which would be built upon a "permissioned" blockchain platform: the Libra Association. Unlike some other types of stablecoins (e.g., Tether, True USD), Libra's value would be

[9] See website https://www.coinlore.com/.

[10] Antonio Madeira, "How Does an ICO Work," Cryptocompare , January 2018, https://www.cryptocompare.com/coins/guides/how-does-an-ico-work/.

[11] For developments in pricing information, market caps, and coin issuance, see Top 100 Cryptocurrencies by Market Capitalization, Coinmarketcap, http://coinmarketcap.com/ (last accessed December 19, 2018). See also ICO Tracker, Coindesk, https://www.coindesk.com/ico-tracker.

[12] For example, bitcoin once dropped from nearly $20,000 to around $6,000 from December 2017 to March 2018.

backed 1:1 by a basket of fiat currency bank deposits, including the U.S. dollar, Euro, Pound Sterling, Japanese Yen, and government securities, rather than any single physical currency. There have been serious criticisms and concerns around Libra, and several of the initial 28 members of the Libra Association have withdrawn themselves, including PayPal, Visa, Mastercard, Stripe, and eBay. Major financial institutions such as JPMorgan Chase and Goldman Sachs have also announced they do not plan to join the Libra Association. On April 16, 2020, the Libra Association published a revised version of its 2019 white paper. The revisions attempt to address major concerns raised by the international regulatory community. The updated Libra 2.0 white paper makes several key changes, but financial regulators still have concerns.[13] These are evidence of the importance of trust in the payment system; thus, there may be a role for the central banks, "trusted institutions," to issue stable digital currencies.

5.1 Cryptocurrencies: Pricing, Impact on Central Banking and Regulations

As a new form of electronic money, cryptocurrency is part of the latest innovations in the financial system. However, the idea of virtual money is not new; electronic payment systems have been growing steadily for decades. Online fantasy games provided a platform for issuing virtual currencies in the 1980s, which have been regarded as predecessors for the current cryptocurrencies such as bitcoin. Later, the M-pesa, a currency denominated in mobile phone minutes, was introduced in Kenya by Safaricom in 2007 (Raskin and Yermack, 2016). M-pesa greatly facilitates money transfers between individuals and therefore increases the probability of being banked for the poor (Allen et al., 2019). Bitcoin, originally proposed by Nakamoto (2008), and introduced into circulation in 2009, is the most successful cryptocurrency so far in terms of volume but with high price volatility. The price peaked at about $19,000 per bitcoin in late 2017 and then declined by 82 percent in 2018 (a 38 percent decline in November alone), to a price of around

[13] The changes in Libra 2.0 include: 1) offering single-currency stablecoins in addition to the multicurrency coin (Libra Coins). Specifically, the design of the Libra payment system is envisioned to start with single-currency (USD, EUR, GBP, and SGD) coins in addition to the basket currency version; 2) enhancing the safety of the Libra payment system with a robust compliance framework. The association plans to develop strong Anti-Money Laundering and Combating the Financing of Terrorism (AML/CFT) standards and establish a Financial Intelligence Function to help support and uphold operating standards for network participants; 3) forgoing the future transition to a permissionless system, while maintaining its key economic properties. Regulators have been concerned about the risk in such an open system of unknown participants taking control of the system and removing key compliance provisions. Instead, network participants (service providers) will be selected through a competitive process based on a set of published criteria; and 4) building strong protections into the design of the Libra Reserve. Libra will still be fully backed by reserves, of which 80 percent will be invested in liquid short-term government securities, and the remainder will be held in cash. In addition, with input from regulators, the association will develop a regulatory capital framework to ensure it maintains an appropriately sized, loss-absorbing capital buffer.

$3,200 as of December 2018.[14] The major difference between bitcoin and M-pesa is that bitcoin has been used all around the world and is circulated over an open Internet network, rather than tied to any traditional banking system. As of March 11, 2020, there are 5,183 cryptocurrencies in existence, with the total market capitalization of $218.4 billion (USD).[15]

It has remained unclear what type of asset cryptocurrency is — whether it should be treated as a currency, a commodity, or a security. Another related concept is that of a token. The Securities and Exchange Commission (SEC) classifies the tokens into three categories: cryptocurrency tokens, security tokens, and utility tokens.[16] Cryptocurrency tokens (or coins) are a means of exchange and a store of value; security tokens, determined by the *Howey test*, represent a conventional security that is recorded and exchanged on a blockchain; and utility tokens give the holder the right to access a product or services on platforms.[17] Proponents of cryptocurrencies highlight various potential benefits, which can improve overall welfare; critics note that those features may facilitate illicit financial activities.

An important question raised by the recent literature on cryptocurrencies is what determines the fundamental value. In addition to being an alternative to traditional currency as a means of exchange, bitcoin and other digital currencies have also been treated as investment vehicles, and the price has been extremely volatile, further leading to potential risks in financial markets. Several recent studies provide theoretical pricing models of cryptocurrencies.

Schilling and Uhlig (2019) develop a pricing model of bitcoin in a simple endowment economy, in which they assume there are two types of monies for transactions: bitcoin and dollars. They show that in the "fundamental" cases, in which bitcoin is used in transactions, the price of bitcoin follows a martingale pattern, meaning that today's price is the best forecast for tomorrow's price; while in the "speculative" cases, in which the buyers hold bitcoin for speculative purposes, the dollar price of bitcoin is expected to rise, and the agents start hoarding bitcoin with the expectation of a price increase. In equilibrium, if the bitcoin prices and marginal utility of consumption are

---

[14] See Coinbase website: https://www.coinbase.com/

[15] For the up-to-date market capitalization of all types of cryptocurrencies, see https://coinmarketcap.com/.

[16] See the SEC statement on cryptocurrencies and initial coin offerings by Chairman Jay Clayton at https://www.sec.gov/news/public-statement/statement-clayton-2017-12-11; see also a reported investigative report on the DAO (decentralized autonomous organization) hack at https://www.sec.gov/litigation/investreport/34-81207.pdf.

[17] The SEC has jurisdiction over securities pursuant to the Securities Act and Exchange Act. A token is a "security" if it meets the definition of an investment contract as outlined in *SEC v. W.J. Howey Co*. Under the Howey test, an investment contract exists when it is: (1) an investment of money; (2) in a common enterprise; (3) with a reasonable expectation of profits; (4) from the managerial or entrepreneurial efforts of others. *Securities and Exchange Commission v. W. J. Howey Co.,* 328 U.S. 293, 301 (1946).

negatively correlated, then agents are willing to hold bitcoin with an expectation of bitcoin price appreciation.

Sockin and Xiong (2018) propose a pricing model for cryptocurrencies. In their model, a cryptocurrency serves two functions — one as membership in a platform with access to goods and services provided by the platform — and the other as the initial pricing for the platform. In their model, when the demand for the platform is publicly observable, there exists either two or no cutoff equilibria, suggesting that one may observe two entirely different dynamics of cryptocurrencies in practice. And the cryptocurrency price would be volatile in this case unless there is government intervention/regulations. On the other hand, when the transaction demand for the platform is unobservable, the trading price of the cryptocurrency would serve as an important channel for aggregating private information and serve to facilitate coordination between the two equilibria (the high- and low-price equilibria). As the high- and low-price equilibria are disparate, it would be difficult for outsiders to diagnose the health of the cryptocurrency simply based on its observed price.

Cong, Li, and Wang (2018) provide a dynamic asset pricing model for cryptocurrencies/tokens on blockchain platforms. Their model allows for both user-base externality and endogenous user adoption. The expected price appreciation makes tokens attractive to early users, allowing them to capitalize future prospect of the platform and hence accelerating adoption. They find a positive but nonlinear relationship among token price and platform productivity, users' heterogeneous transaction needs, and network size.

Pagnotta and Buraschi (2018) provide an equilibrium model of the bitcoin market in a decentralized network. They show that the overall production (capacity) and the price of bitcoin are jointly determined. Their calibration shows that the bitcoin price is very susceptible to the fundamental properties of its demand and supply. Tripling the current network size raises the equilibrium price of bitcoin from $14,200 to $77,627 (a 546 percent increase).

In addition to these theories, empirical studies also document the volatility of cryptocurrency prices. For example, Kroeger and Sarkar (2017) show that the price of bitcoin varies significantly in the various exchanges around the globe. For example, the maximum absolute price difference ranges between 17 percent for the Coinbase-BTC-e exchange pair and as high as 41 percent for the Bitstamp-BTC-e pair during the same time period of 2013–2016. They attribute such violation of the law of one price to two main reasons: One is the microstructure frictions such as bid-ask spread, order book depth, and volatility; the other is the use of bitcoin to avoid foreign exchange restrictions, money laundering, and so on. Makarov and Schoar (2018) construct an arbitrage index and calculate the arbitrage profits from bitcoin trading across different exchanges and regions. They show that the arbitrage profits from December 2017 to February 2019 was more than $1 billion. In addition, the recurrent arbitrage opportunities in cryptocurrency prices relative to fiat currencies are much larger across regions (the U.S., Japan, and Korea) than within the same region.

The sudden rise of cryptocurrencies may pose challenges to central banks at many levels. Fernández-Villaverde and Sanches (2016) analyze the impact of cryptocurrencies on monetary policy effectiveness. Their game theory model shows that there exists several equilibria, in which only one of the several equilibria would be stationary, such that the value of all privately issued cryptocurrencies would be stable over time, while other equilibria may be undesirable. Even the best equilibrium (stationary) cannot deliver the socially optimum amount of money supply. Therefore, if the cryptocurrencies become widespread, central banks would have difficulty finding appropriate intermediate targets for the monetary policy. Unless central banks could find ways to stabilize the supply of cryptocurrencies as individuals, corporations, and financial institutions increase their holdings of cryptocurrencies, the financial system might be less stable. As an option, central banks could consider creating their own central bank digital currencies (CBDCs) to control money supply. There has been wide discussion on this. We leave it for Section 5.3.

5.2 Initial Coin Offering: Structure, Valuation, and Regulation

Initial coin offerings (ICOs) are mechanisms to raise funds by selling coins or tokens, using blockchain technology, to support a product launch or a new virtual currency. ICOs are a conjunction of crowdfunding and blockchain. Tokens purchased in an ICO give the participant certain rights, most frequently the right to use the platform services that are being developed or ownership rights. The coins can also be exchanged for other cryptocurrencies (and even potentially fiat currencies) on secondary markets. They operate similarly to initial public offerings (IPOs) but typically skirt the usual regulations and restrictions on IPOs. While start-ups have traditionally relied on venture capital to raise funds and grow, ICOs present a more decentralized and democratic alternative. The total capital raised through ICOs was only $16 million across two deals in 2014 and $6.1 million (USD) across three deals in 2015, but spiked exponentially in the following two years.[18] The biggest ICO until mid-2018 was priced by EOS and raised $4.1 billion (USD).[19]

---

[18] For more data, see the ICODATA website, https://www.icodata.io/stats/2014.

[19] See the report from PwC at https://cryptovalley.swiss/wp-content/uploads/20180628_PwC-S-CVA-ICO-Report_EN.pdf.

Reaching a consensus on a formal and complete definition of ICOs is challenging, given a large number of different forms that current ICOs take. Through an ICO, an individual or group of founders can offer a stock of specialized crypto tokens for sale, usually promising that those tokens will operate as the only medium of exchange when assessing the venture's future products to avoid running afoul of securities laws. The structure of ICOs is based on the offer of digital tokens or coins that use blockchain technology. A typical ICO process starts with a white paper, similar to a prospectus, which describes the financing project and the rights given to the investors. Very often the white paper determines a minimum and a maximum amount of coins that need to be subscribed in order for the financing project to go live.

The startling growth of ICOs has raised concerns about irrational exuberance. Two central questions are related to how ICOs could be accurately evaluated and priced, and what would be appropriate regulatory responses. A growing literature studies the structure of ICOs. One consistent finding is that the ICO structure could help to solve the coordination problems among investors; see Li and Mann (2018); Cong, Li, and Wang (2018); Bakos and Halaburda (2018); and Catalini and Gans (2018). For example, Li and Mann (2018) present an economic mechanism through which token and ICO structures create values for both entrepreneurs and platform users. In their model, by transparently distributing tokens before the launch of the platform, an ICO can overcome coordination failures during platform operations, in which each user of the platform cares about the activities of his transaction counterparty on the other side. Put differently, in the typical structure of an ICO, when a user purchases a token, the volume of the subscription is publicly observable, given the transparency of the smart contract implementing the ICO. This allows platform users to communicate with other users and in turn motivate them to participate as well.

Sockin and Xiong (2018) document the dual roles of cryptocurrency in the platform: as a member for the transactions and as an initial financing for the platform, which makes ICOs different from traditional project financing mechanisms in the way that investors and business customers are not separated. The outcome is that the trading price and volume of the cryptocurrency not only provide financing of the cryptocurrency, but also directly impact the business operations of the platform. In addition, Bakos and Halaburda (2018) find that token issuance could serve as a low-cost alternative of financing when the platform faces credit constraints, as its cost of capital increases over time. In doing so, the platform trades off its future revenue for the present. On the other hand, an agency conflict might exist between entrepreneurs and investors, as documented in Chod and Lyandres (2018).

In terms of ICOs evaluation and pricing, there have been limited studies so far. The overall empirical evidence on the valuation of ICOs suggest that the success of ICOs is determined by an assortment of characteristics of the tokens, entrepreneurs, as well as the underlying projects. Initially, there is ex-ante information asymmetry between an entrepreneur and outside investors, resulting an adverse selection process that could potentially hamper a successful fundraising through ICOs. Howell, Niessner, and Yermack (2018) use a sample of 453 completed ICOs and find that liquidity and trading volume are higher for tokens that offer voluntary disclosure and provide a signal of quality and potential value of the project. The professional background of an entrepreneur is also strongly associated with ICO success. In addition to these determinants, Amsden and Schweizer (2018) document that the tradability of tokens or coins as well as the uncertainty and quality of venture (e.g., whether being on Github or Telegram, the length of white papers) also matter. Lee, Li, and Shin (2018) use a sample of 1,500 ICOs and find that the diverse opinions from a number of online analysts form an aggregate signal that predicts the quality of the underlying projects as well as the subsequent token sales. Other related studies on the returns and success of ICOs include Adhami, Giudici, and Martinazzi (2017); Hu, Parlour, and Rajan (2018); Momtaz (2018); and Benedetti and Kostovetsky (2018).

Given the nature of entrepreneurs financing through ICOs, particular risks of an ICO include potential fraud such as issuance of scam coins. Scam coins usually establish a Ponzi scheme and rob investors of their currencies.[20] Some smart contracts for ICOs are artifactual, which does not match with the description in the white papers. Cohney et al. (2019) compare smart contracts with the associated white papers for 50 ICOs and find that some popular ICOs have retained the power to modify their tokens' rights in the smart contract but failed to disclose this possibility in the white papers. More frequent risky events surrounding ICOs are cyberattacks. For instance, the DAO (decentralized autonomous organization) was hacked only one month after it was launched.[21] The DAO, built on Ethereum, was an investment entity that aims to raise money through the DAO ICO to fund projects. The DAO was so popular that it raised over $150 million and was the largest

[20] According to *The Economist*, in Vietnam, police are investigating two ICOs (Pincoin and Ifan), run by a firm named Modern tech, that are accused of duping investors out of $660 million. See "Initial Coin Offerings Have Become Big Business," *The Economist,* https://www.economist.com/technology-quarterly/2018/09/01/initial-coin-offerings-have-become-big-business. See also Kai Sedgwick, "46% of Last Year's ICOs Have Failed Already," Bitcoin (February 23, 2018), https://news.bitcoin.com/46-last-years-icos-failed-already/; Nathaniel Popper, "SEC Issues Warning on Initial Coin Offerings," *New York Times* (July 25, 2017), https://www.nytimes.com/2017/07/25/business/sec-issues-warning-on-initial-coin-offerings.html.

[21] Shendra Kumar Sharma, "Details of the Dao Hacking in Ethereum in 2016," Blockchain Council (August 2017), https://www.blockchain-council.org/blockchain/details-of-the-dao-hacking-in-ethereum-in-2016/.

crowdfunding project in history.[22] The holders of DAO tokens had the power to select which project to invest in by voting, which was administered by the DAO codes. However, one of the flaws is how the DAO acted as a factory for creating child "smart contracts" that split off from the main DAO to a "child-DAO." Shortly after the DAO was launched, a hacker drained a total of 3.6 million Ether (worth around $70 million) into a "child DAO" that was in the same structure as the "parent DAO" in a few hours by exploiting the codes repeatedly. Other potential risks of ICOs are associated with the institutional and legal environment in different countries where the ICOs are launched as well as the potential systemic risks (e.g., Zetzsche et al., 2019).

In the U.S., the SEC has classified nearly all ICOs as securities, within their competency and authority.[23] Although the SEC has carved out the "utility token" from the security category, certain definitions and requirements for utility tokens remain vague and need clarification. Zetzsche et al. (2019) give a review of financial law and regulation on ICOs and crowdfunding in different jurisdictions and explore an appropriate approach to regulate the ICOs. In other jurisdictions, China and Korea have both banned ICOs outright, while Switzerland has provided an ICO-friendly model. Singapore, Hong Kong, the UK, and Australia have proposed regulatory warning for ICOs.

The overall results from the literature (theoretical studies) are consistent with an argument that, while it may be necessary for ICO activities to be regulated, it would not be desirable to completely ban all ICO activities (an approach adopted in China and Korea). Regulators and academics have been working to design the appropriate regulatory framework for ICOs — to curtail ICO scams while continue to encourage fintech innovations. One possible approach to reducing ICO scams is to follow a protocol similar to the Rule 144A restrictions on private placements — ensuring that only qualified investors can participate in ICOs.

5.3 Central Bank Digital Currency (CBDC)

The recent advances in cryptographic and DLTs, along with extreme volatility in private digital currencies, have led to a discussion around the possibility of central banks issuing their own central bank digital currencies (CBDCs); see Mancini-Griffoli et al. (2018); Jagtiani et al. (2020); Engert and Fung (2017); Bordo and Levin (2017); Ali et al. (2014); and Agur, Ari, and Dell'Ariccia (2019). The threats to financial stability by private cryptocurrencies have led to intensive debates among

[22] See DAO White Paper, https://daostack.io/wp/DAOstack-White-Paper-en.pdf.

[23] The key documents that provide guidance on how the SEC will likely classify any ICO are "The DAO Investigative Report," released in July 2017 (DAO Report) and "Cease and the Desist Order Against Munchee," issued on December 11, 2017, (Munchee Order).

policymakers and monetary economists about whether and how central banks play a role in CBDCs — whether to issue the CBDC or play a supporting role in issuing CBDCs. Based on a BIS survey of central banks, with 63 central bank respondents representing jurisdictions covering close to 80 percent of the world population, Barontini and Holden (2019) show that many central banks are progressing from conceptual work on CBDCs into experimentation and proofs-of-concept, including cooperating with other central banks. Jagtiani, Papaioannou, and Tsetsekos (2019) and Jagtiani et al. (2020) provide a review of CBDCs and highlight several important country experiences with CBDCs.

One of the concerns among central banks in issuing CBDCs is the impact on the banking sector (the risk of a bank run as people would prefer an account with a central bank) and the stability of the financial systems overall. Andolfatto (2018) develops models in which the introduction of CBDCs would lead to an expansion of bank deposits if CBDCs compel banks to raise their deposit rates. They find that interest-bearing CBDCs generally promote financial inclusion as they diminish the demand for cash. Brunnermeier and Niepelt (2019) model the relationship between bank panics and CBDCs and suggest that whether CBDCs would undermine financial stability depends on the monetary policy accompanying the issuance of CBDCs and on the strength of a central bank's commitment to serve as a lender of last resort. They show that with a strong commitment, the issuance of CBDCs, which are equivalent to a transfer of funds from deposit to CBDC accounts, would give rise to an automatic substitution of deposits by central bank funding (the "pass-through" mechanism). Furthermore, in this case, the issuance of CBDCs would not undermine but rather strengthen financial stability, because a depositor run into CBDCs would trigger a pass-through funding automatically and turn the central bank into a large depositor. Similarly, Keister and Sanches (2019) find that issuing CBDCs would enhance efficiency in exchange at the expense of crowding out deposits.

Raskin and Yermack (2016) explore the impact of digital currency on the future of central banking and argue that a sovereign digital currency could narrow the relationship between citizens and central banks and remove the need for the public to keep deposits in fractional reserve at commercial banks. According to Fernández-Villaverde, Sanches, Schilling, and Uhlig (2020), assuming that competition among commercial banks is allowed and that depositors do not panic, account-based CBDCs would allow central banks to engage in large-scale intermediation with private financial institutions for deposits, and this would give consumers the possibility of holding a bank account with the central bank directly. Brunnermeier, James, and Landau (2019) review the ongoing digital revolution as well as its impact on the traditional model of monetary exchange. They suggest that in a digital economy in which most activities are conducted through networks with their own monetary instruments, CBDCs would open up a direct channel by which monetary policy could be transmitted to the public and permit the central bank's unit of account to remain relevant in a fast-changing digital economy. Overall, CBDC proponents highlight that CBDCs would improve the safety and efficiency of the financial and payment systems as well as increase central banks' controlling power over monetary policy. Critics, however, raise concerns about the significant risks that digital currencies might bring to the rest of the financial system.

6. Fintech, Trading, and Algorithmic Investment Strategies

6.1 High-Frequency Trading

Over the last decade, the forces of technology have increasingly shaped the market structure and trading behavior. One type of computerized trading that attracted the most attention is high-frequency trading, which was largely absent in 2001 but participated in about 50 percent of the trading activities by 2010 (SEC, 2010). In the meanwhile, liquidity has improved, and the transaction cost of end users has declined substantially in this period. The bid-ask spread that investors paid on the market order, which is a measure of the trading cost for retail investors, was reduced. And the implementation shortfall, which is a measure of the transaction cost for large institutional investors was also largely reduced from 2001 to 2011; see Menkveld (2016), Jones (2013), and Linton and Mahmoodzadeh (2018).

While there is no formal definition for high-frequency trading (HFT), the SEC (2010) identifies some features that often are attributed to high-frequency trading (i.e., the use of extraordinarily high-speed and sophisticated algorithms, the use of co-location services and individual data feeds, the very small and short-lived positions in margin accounts, and the submission of numerous orders and cancelling shortly). HFT represents a large subset but not all of algorithmic trading, which is generally defined as the use of computer programming to make decisions about order submissions and cancellations.[24] Menkveld (2016), Jones (2013), O'Hara (2015), and Kirilenko and Lo (2013) give a review of the types of high-frequency traders and trading strategies.

[24] In SEC (2010) and SEC (2014), algorithmic trading (AT) encompasses a broad range of activity that includes particularly the large-order execution algorithms often used by or on behalf of institutional investors. These algorithms slide the large orders into many small pieces that are fed into the marketplace over time. This type of algorithmic trading is not classified as HFT as the time horizon of trading is far beyond the short time frame of HFT.

O'Hara (2015) makes a point about the complexity of information in a high-frequency age. In a high-speed and co-located world, information is not just asset related but also order related. Being informed is multidimensional, meaning that it can be seeing and acting on market prices faster than competitors or knowing more about the assets or the markets. A key question is whether HFT can benefit market quality or hurt it. Theories seem to suggest that if HFT is a faster acting agent as it can trade on new information instantly before it becomes available to others, adverse selection is more severe and market quality decreases (Biais, Fousault, and Moinas, 2015; Foucant, Hombert, and Rosu, 2016; Du and Zhu, 2014); however, if HFT is a better informed agent, then liquidity improves (e.g., Ait-Sahalia and Saglam, 2013; Goettler, Parlour, and Rajan, 2009). A tradeoff is documented in Jovanovic and Menkveld (2016) that, if HFT is both an informed and fast agent, it can update limit orders quickly based on new information as it has both information and speed; therefore, HFT can avoid some adverse selection and further provide some benefit to uninformed traders who need to trade. An implication is that we need to be careful when interpreting lower bid-ask spreads as better market quality because this may be driven by other traders who are not being able to earn the spread through posting limit orders.

On the empirical side, a number of existing studies provides evidence on the effects of HFT. Some papers use specific market structure changes to identify the causal and positive effects of HFT on market quality. For example, Hendershott, Jones, and Menkveld (2011) employ the implementation of an automated quote at the New York Stock Exchange in 2003 and show that at least for large-cap stocks in that period, an increase in algorithmic trading causes an improvement in stock market liquidity. Menkveld (2013) examines the entry of high-frequency market-makers into the trading of Dutch stocks in July 2007 and find that because of the higher degree of competition, compared with untreated Belgian stocks, the bid-ask spreads of Dutch stocks are about 15 percent narrower, and adverse selection is also less. Gai, Yao, and Ye (2014) use data from two Nasdaq technology upgrades in 2010 and find that there is little change in bid-ask spreads and market depths, suggesting that there could be diminishing liquidity benefits.

One of the most salient trends in securities markets has been the fragmentation in the trading system (Menkveld, 2016). Without a central market, traders need to search for liquidity across different venues, which the high-frequency traders are specialized in. Existing studies find that an increase in trading fragmentation is associated with lower costs and faster execution speeds in a given asset class (Foucault and Menkveld, 2008; O'Hara and Ye, 2011; Degryse, De Jong, and van Kervel, 2015). Pagnotta and Philippoin (2018) find that competition among venues increase investor participation, trading volumes, and allocative efficiency; but in the meantime, there could be excessive fragmentation and inefficient execution speed, suggesting that the welfare consequences depend critically on the ability to impact entry decisions and that sound regulations would be important.

Overall, evidence strongly indicates that HFT is good for average market quality. However, its impact on market quality in stressed conditions is of particular concern to regulators and investors. On May 6, 2010, in the course of 33 minutes, starting from 2:32 p.m. EST, the U.S. stock market experienced one of the most turbulent periods in the history, known as the Flash Crash. Kirilenko, Kyle, Samadi, and Tuzun (2011) analyze the Flash Crash using the audit trail data for all 15,000 accounts that traded the E-mini that day. They find that HFT did not trigger the Flash Crash, but their responses to the extreme selling pressure exacerbated the price decline. Basically what HFTs did was buy contracts from fundamental sellers in the E-mini during the price decline and then proceeded to sell contracts and compete for liquidity with fundamental sellers. However, the SEC (2014) points out that Kirilenko, Kyle, Samadi, and Tuzun (2017) did not capture more than one-third of total HFT activity in the E-mini during the Flash Crash. Easley, Lopez de Prado, and O'Hara (2011, 2012) use intraday data and find that order imbalance was severe in the minutes just before the Flash Crash, confirming that HFT and other trading intermediaries were overwhelmed by selling pressure. Another recent market glitch associated with HFT is trading errors by Knight Capital, one of the largest market-makers for U.S. equities, on August 1, 2012. A new trading algorithm introduced by this HFT firm rushed into service without sufficient testing and caused a large accumulation of stock positions over about 45 minutes at the start of the trading day. Knight Capital lost $456.7 million from liquidating the positions in the end, wiping out its capital.

These types of trading errors and extreme events call for more HFT regulations. The benefits of automation and other new technology in financial markets are indisputable, but they have to be evaluated and tested sufficiently. Several initiatives have been proposed by regulators and academics, including a real-time consolidated order-level audit trail (Jones, 2013), a required minimum time-in-force for orders (SEC, 2010), a small securities transaction taxes; see Mehta (2013) and Kirilenko and Lo (2013). Kirilenko and Lo (2013) further propose some principles for an advanced financial regulatory framework in the Digital Age that they refer to as Financial Regulation 2.0. We will discuss the regulatory issues further in Section 8.

6.2 Quantitative Investment Strategies

In recent years, there has been considerable hype concerning the use of sophisticated computer-based and AI/ML techniques for quantitative investment strategies. Many have argued that this would be the future of investment management. We have been here before. For example, Allen and Karjalainen (1999) use a genetic ML algorithm to learn technical trading rules for the S&P 500 index using daily prices from 1928 to 1995, and they confirm that the stock markets are efficient in the sense that it is not possible to make money after transaction costs when using technical AI/ML trading rules. The rules are able to identify periods to be in the market when returns are high and volatility is low and out of the market when the reverse is true, consistent with Brock, Lakonishok, and LeBaron (1992).

Empirical evidence suggests that quantitative strategies can work well in stable environments but often work very poorly when there is a crisis or some other unexpected events. An example is the Quant Meltdown in the week of August 6, 2007 (e.g., Khandani and Lo, 2007; Kirilenko and Lo, 2013), when some of the most successful hedge funds suffered unprecedented losses. It appeared that the loss was concentrated among “quantitative equity market neutral” or “statistical arbitrage” hedge funds (Khandani and Lo, 2011). More recently, the investment field has been leveled in computer resources, data, and statistical tools such as the AI/ML algorithms. However, as with most quantitative applications in finance, there is a dark side of these powerful techniques. When misapplied, these techniques can lead to false discovery and disappointment. For example, Arnott, Harvey, and Markowitz (2019) point out that, unlike in physical and biological sciences, data are much more limited in scope in investment finance, which might be far too small for most AI/ML applications and impossibly small for advanced deep learning approaches. Lopez de Prado (2018) raises 10 pitfalls and proposes solutions to the application of AI/ML in investment and finance.

An interesting question regarding the application of computer and advanced technologies in investment is why quantitative investment strategies produce alpha, a measure of excess returns. One of the most important aspects of such investment strategies for both buyers and sellers of the products is to understand whether and how they produce excess returns. Are they simply reallocating returns across states so they look good most of the time but melt down occasionally? Are there missing risk factors that are not modelled properly so it seems there is an excess return but in fact there is not? Financial crises, which have low probability but high impact, are another possible form of risk of this type. Risks can also be endogenous if a new strategy appears promising and is simultaneously adopted by a large number of investment firms so that the environment changes and new types of risk arise. When they all try to exit together, prices change in unexpected ways. Are there market imperfections that the strategy is helping to get around so the excess returns come from the increased efficiency? Quantitative investment strategies using AI/ML could help to overcome market imperfections through reduced transaction costs and improved diversification or carry trades.

6.3 Investment Advisors, Wealth Management, and InsurTech

Online wealth management platforms providing investment advice driven by algorithms (more commonly called *robo-advisors*) have emerged to replace or complement traditional (human) investment advisors. Robo-advisors leverage data provided by investors to construct and manage a tailored appropriate investment portfolio for them. Since the first major service launched by Betterment in 2010, this market has grown rapidly, accumulating nearly $45 billion in assets under management (AUM).[25] Later, witnessing the success of independent robo-advisors, traditional money managers have also started launching such services that work in tandem with the products that they conventionally offer.[26] A typical robo-advisor platform consists of three phases: (1) the initial investor screening, (2) investment strategy implementation, and (3) monitoring and rebalancing the strategy.

For the initial client intake and screening, robo-advisors provide automated in-app or online questionnaires, including general questions about clients' financial preferences and goals, upon which the robo-advisors could build a profile of the client's financial goals, investment strategies, and risk tolerance. Then, robo-advisors will select targeted assets and curate a personalized portfolio based on the client profile. Currently, most robo-advisors implement passive investment strategies, including strategies using ETFs as baseline investments.[27] Finally, robo-advisors would monitor and rebalance a client's investment portfolio, to ensure the portfolio risk does not exceed clients' preferences while simultaneously taking advantage of price changes in the

[25] Since launched, Betterman has raised over $200 million in equity funding and has over $9 billion in AUM. The second largest independent robo-advisor to date, Wealthfront, has raised about $130 million in equity and has over $6 billion in AUM. It is widely expected that the robo-advice market will continue to skyrocket. A particularly aggressive projection predicts that robo-advisors will have $5.0 trillion to $7.0 trillion in AUM by the year 2025 (https://hbr.org/2018/01/robo-advisers-are-coming-to-consulting-and-corporate-strategy).

[26] For example, Charles Schwab and Vanguard were the two pioneers in providing robo-advising services in line with their other advising products (see, e.g., https://smartasset.com/retirement/the-top-10-robo-advisors).

[27] It is reported that active management robo-advising strategies (typically hybrid solutions) have a growing presence in the marketplace (see, e.g., https://money.usnews.com/investing/investing-101/articles/2018-10-17/should-i-use-an-actively-managed-robo-advisor).

market. Compared with the traditional human financial advising, the key promise of robo-advising is evident: delivering convenient and unbiased financial advice at significantly lower cost.[28]

Based on an automated portfolio optimizer introduced by a brokerage firm to its clients in India, D'Acunto, Prabhala, and Rossi (2019) examine the determinants of using robo-advising and the effects. While users and nonusers are indistinguishable along several demographic characteristics, including gender, age, and trading experience, users are generally more sophisticated and usually have a larger amount of investment funds than nonusers. In addition, undiversified investors could benefit more from robo-advising, and their market-adjusted investment performance could improve significantly. Robo-advisors can reduce pervasive behavioral bias and cognitive litigations that traditional human financial advisors might have (Linnainmaa, Melzer, and Previtero, 2017), such as the disposition effect, trend chasing, and the rank effect. Hodge, Mendoza, and Sinha (2018) examine the investment judgment of humanizing robo-advisors and find that investors are more likely to follow the advice of a robo-advisor when the advisor exhibits fewer human features. Moreover, investors are more likely to follow the advice of a named human advisor compared with an unnamed human advisor; while investors are less likely to follow the advice of a named robo-advisor compared to an unnamed robo-advisor. Overall, advisors' credibility will mediate how likely investors are to follow their recommendations.

Robo-advising makes access to financial advice available at lower cost, mitigates underdiversification, and improves investment performance; however, it might be subject to conflict of interests and other potential risks. Baker and Dellaert (2018) argue that because humans design and implement robo-advisors, we cannot always expect an ideal robo-advisor in terms of honesty, competence, and suitability. Despite the difference between robo-advisors and human advisors, they are currently regulated under the same statutory framework — under the Investment Advisers Act of 1940 (in the U.S.) and the Market in Financial Instruments Directive (MiFID) framework (for the EU). A narrative, driven by academia and lawyers, argues that robo-advisors are inherently incapable of meeting the Investment Advisers Act standards; i.e., the fiduciary duty (duty of care and duty of loyalty obligations).[29] For example, some critics argue that

[28] According to Ringe and Ruof (2019), because of making use of the economies of scale, robo-advisors charge average fees of between 0.4 percent (mainly in the U.S.) and 0.8 percent (mainly in Europe), whereas the fee for human financial advice usually amounts to 1 percent to 2 percent.

[29] The fiduciary duty required by the Investment Advisers Act of 1940, means that advisors must (1) take the reasonably necessary steps to avoid misleading clients, (2) provide full and fair disclosure to their clients and prospective clients of all material information, and (3) eliminate, or at least disclose all conflicts. See https://www.sec.gov/investment/im-guidance-2017-02.pdf.

using an electronic questionnaire to gather information from clients without further confirming the accuracy are far from enough to satisfy the investment advisor duty of care (e.g., Fein, 2015). Others criticize that robo-advisors lack human perceptions and are not equipped to address market failures (Fein, 2015).[30]

Ji (2018) argues that the conflicts of interest concerns associated with robo-advising could be quite severe because of bias in the algorithms that reflects a firm's existing conflicts of interest. However, traditional human advisors may also be influenced by individual outside incentives. Recently, both the SEC and the Financial Industry Regulatory Authority (FINRA) have directly addressed some of the ambiguity, with respect to how robo-advisors fit in the current regulatory structure, through a series of guidelines. For example, the SEC recommends that robo-advisors provide a concise explanation of their business model because of the unique nature of their business as well as fully disclose all fees and other charges.[31] It remains unclear whether the current rules are effective in dealing with the risks and challenges of robo-advising. Ringe and Ruof (2019) and Financial Conduct Authority (2017) propose a regulatory sandbox, an experimentation space, which would allow market participants to test robo-advising services in the real market, with real consumers, but under close scrutiny of the supervisor. Such a sandbox can allow for mutual learning for both the firms and regulators and reduce regulatory uncertainty for all the market participants.[32]

Insurance technology, manifested as InsurTech, is a subset of fintech dedicated to the insurance industry. The insurance industry has access to a large amount of data, which would be critical for the application of complex algorithms like AI/ML methods. With big data and advanced techniques, more precise measurements of underlying insurance risk can be estimated. Insurers could have better protections against operational risks, such as preventing insurance fraud or money laundering. Customers can also have a wider range of better tailored products and services. Massive investment funding in the form of mergers and acquisitions, venture capital, and private equities, have flowed into the InsurTech industry. Over $35 billion was raised by InsurTech

---

[30] See also, Samantha Sharf, "Can Robo-Advisors Survive a Bear Market?" available at https://www.forbes.com/sites/samanthasharf/2015/01/28/can-robo-advisors-survive-a-bear-market/#243f981ce7ec; Michael Wursthorn and Anne Tergesen, "Robo Advisor Betterment Suspended Trading During 'Brexit' Market Turmoil," available at https://www.wsj.com/articles/robo-adviser-betterment-suspended-trading-during-brexit-market-turmoil-1466811073.

[31] SEC, *Investment Management Guidance*: *Guidance Update*, No. 2017-02 (Feb. 2017) https://www.sec.gov/investment/im-guidance-2017-02.pdf

[32] See Section 8 for more about the regulatory sandbox.

companies from 2016 to 2019, through 852 deals.[33] Geographically, the InsurTech industry has been dominated by the U.S. since 2012. China, Germany, the UK, and France are the other top InsurTech markets.[34]

Lin and Chen (2019) examine the potential risks associated with the application of InsurTech, as well as the possible regulatory solutions. The risks come from the general use of technology, such as cybersecurity, data protection, consumer privacy issues, and other special features related to the insurance industry. Addressing fraud risk is key to online business models for InsurTech providers. They propose a "meta-regulation" approach: regulated firms write a set of rules tailored to the firms themselves, and these rules are subject to further scrutiny and approval by a regulatory agency. Thakor (2019) briefly discusses the development in InsurTech. Through applying the new methods, overall, InsurTech firms are believed to be able to provide better calibrated risk assessment and pricing, with less pooling across customers with heterogeneous but ex-ante indistinguishable risk profiles.

7. Cybersecurity

Cyber-risk has become one of the most important sources of risks for corporations. Recent surveys by PricewaterhouseCoopers (PwC) from 2017 to 2019 show that more than half of the CEOs expect cybersecurity and data breach incidents to threaten stakeholder trust in their industries in the near future.[35] On top of that, the financial industry has witnessed the most incidents with data losses.[36] As documented in previous sections, while the advanced financial technology brings efficiency gains for financial institutions and improves financial inclusion, it has also increased the sophistication of the attack, making it harder and more complex to prevent. The increased cyber-risks because of the advanced financial technology can take multiple forms, including cyber-risks related to the data-sharing process such as data aggregation through an open API. This allows third-party vendors to access consumer data directly from their bank account. There has been increased vulnerability, especially when the interfaces between the two systems are not compatible (e.g., not

[33] Willis Towers Watson, 2020, Quarterly InsurTech Briefing Q4 2019, https://www.willistowerswatson.com/en-GB/Insights/2020/01/quarterly-insurtech-briefing-q4-2019.

[34] In the third quarter of 2019, China contributed 13 percent of the total InsurTech deal activity, driven by an increase of interest in start-ups contributing to the growth of China's health insurance industry. See Willis Towers Watson, 2019, Quarterly InsurTech Briefing Q3 2019, October, https://www.willistowerswatson.com/en-US/Insights/2019/10/quarterly-insurtech-briefing-q3-2019

[35] For more details, please see "Risk in Review Study" by PwC from 2017 to 2019.

[36] Verizon, 2019, *Data Breach Investigations Report 2019*, Verizon, https://enterprise.verizon.com/resources/reports/dbir/.

designed around the same time period by different developers and often because of limitations of banks' legacy technology. It is increasingly difficult to thoroughly identify all potential sources of vulnerability in the systems as the processes are usually more time-consuming and expensive.

The various financial applications that allow consumers to transfer funds and make payments through voice or facial recognition also increase cyber-risk; for example, Amazon Alexa and CapOne, Bank of America's ERICA app. There have been several incidences of successful cyberattacks, including an incident in February 2016, when hackers stole $81 million from the Central Bank of Bangladesh, through transfer fraud via compromised SWIFT servers.[37] The SEC's EDGAR system was also hacked by a Ukrainian hacker using deceptive hacking techniques in 2016. The extracted EDGAR files that contain nonpublic earnings information was passed to individuals who used it to trade in the narrow window before the companies released to the public, resulting in at least $4.1 million of illegal profits.[38]

Another well-known cyberincident is the Equifax hack in September 2017, in which the hackers exploited a vulnerability in the software that Equifax uses to build its websites to steal customer names, Social Security numbers, birthdates, and addresses, affecting 147.7 million Americans. Later in July 2019, the Amazon Alexa and CapOne data breach exposed credit card application data for those who applied between 2005 and 2019, affecting roughly 100 million individuals in the U.S., and 6 million customers in Canada. Many cryptocurrency exchanges have also been shut down because hackers were able to steal assets from the exchange systems. Finally, cloud computing, which is an enabler of the fintech ecosystem (payment gateways, digital wallets, secured online payments rely on cloud-computing services) also resulted in new types of risk, especially cyber-risks.

While the direct costs related to cyberincidents (such as the cost of forensic investigation, legal assistance, customer notification and post-breach customer protection, and other measures) can be relatively well understood, the indirect costs (such as the reputation risks of brand names, negative shock to existing customer relationships, or depreciation of intellectual property value), on the other hand, are far less visible, more long term and more difficult to be quantified. Therefore, there can be significant uncertainty surrounding the potential impact of cyberincidents, as documented in Kopp, Kaffenberger, and Wilson (2017).

[37] For more details, see https://en.wikipedia.org/wiki/Bangladesh_Bank_robbery and https://www.reuters.com/investigates/special-report/cyber-heist-federal/.

[38] In January 2019, the SEC charged the nine defendants in the EDGAR hacking case, a Ukrainian hacker, six individual traders in California, Ukraine, and Russia and two entities. For more details, see https://www.sec.gov/news/press-release/2019-1.

Who are more likely to be the targets of cyberattacks? Using the data breach events caused by cyberattacks reported to the Privacy Rights Clearinghouse (PRC) over the period 2005 to 2017, Kamiya et al. (2018) find that firms are more likely to be attacked when they are larger, less financially constrained, more highly valued, have more intangible assets, and in less competitive industries. Firm-level corporate governance characteristics, such as features of the board, are not found to predict the likelihood of cyberattacks.

Conventional wisdom suggests that hacking events would have some negative and sticky influence on firms' reputations and henceforth on growth prospects. Recent studies confirm this prediction and find that a firm's stock price drops upon the announcement of a serious cybersecurity breach. For example, Lin, Sapp, Rees-Ulmer, and Parsa (2018) show that there is an abnormal return of -1.44 percent in the five-day window surrounding the public announcement of a data breach, and such price decline does not reverse over the following month. Amir, Levi, and Livne (2018) document that firms may underreport the cyberattacks and investors cannot discover most attacks independently. For the withheld attacks, the negative abnormal return in the stock market (3.6 percent) in the month when the attack is discovered is higher than that of the disclosed attacks (0.7 percent). Bianchi and Tosun (2019) find that upon corporate announcements on cybersecurity breach, daily excess returns drop and trading volume increases in the short run. Kamiya et al. (2018) document a significant mean cumulative abnormal return of -0.84 percent during the three-day window around cyberattack announcements, which can be translated into an average value loss of $495 million per attack. Lin, Sapp, Rees-Ulmer, and Parsa (2018) find significant evidence of opportunistic insider trading in the three months prior to cybersecurity breach announcements.

In a relative long-run scenario, existing studies find that firms adjust their investment and financial policies in response to cyberattacks. Kamiya et al. (2018) document that attacks in which personal financial information is appropriated are associated with a decrease in sales growth for large firms and retail firms, with an increase in leverage and debt maturity, but with a weak deterioration in financial deficit and capital expenditure. Firms also further respond by reducing CEO risk-taking incentives and strengthening their risk management. Consistently, Bianchi and Tosun (2019) also find weak evidence of negative impact on cash flows and operating performance upon cybersecurity incidents.

The true costs of cybersecurity breach could be huge, yet the market could fail to provide a social optimal level of security because of information asymmetries, externalities, coordination failures, and barriers to entry (Kopp, Kaffenberger, and Wilson, 2017). It remains unclear what the best policy response to cyber-risk should be, including how to design ex-ante regulation and assign ex-post liability. Moreover, cyberattacks raise issues about jurisdiction as warfare in cyberspace pays little regard to national boundaries. So far, there remains neither clear nor meaningful international consensus on the sovereignty and jurisdiction issues relating to cyberattacks and nor on the governance of cyberattacks (Lin, 2016).[39]

8. Cloud Computing and the Roles of BigTech

A key driver behind the recent explosion of data is the concomitant growth of computing power as well as cloud servicing. Cloud storage and cloud computing have allowed large volumes of information to be stored cheaply on external, third-party, Internet-based servers. By using cloud services, incumbent systems are increasingly open and operate on a shared rather than closed basis. The Financial Conduct Authority (FCA) in the UK issued guidance on financial institutions' use of cloud storage facilities in 2016.[40] The UK challenger bank, OakNorth, became the first bank in the UK to transfer its core systems into the cloud by using Amazon's Web Services (AWS) in 2016.[41]

Fintech was previously defined by the Financial Stability Board as "technologically enabled financial innovation that could result in new business models, applications, processes, or products with an associated material effect on financial markets and institutions and the provision of financial services."[42] After 2008, a new era of fintech has been marked by the arrival of wave of the new fintech start-ups, including LendingClub, Prosper, Paypal, Venmo, Square, Plaid, Credit Karma. A new term, *BigTech*, created later, refers to large existing firms with the primary business in providing digital services rather than mainly in financial services (Frost et al., 2019). In other words, fintech companies operate primarily in the financial sector, while BigTech firms offer financial products only as a subsector of their main businesses. Amazon, Google, Microsoft, Uber, Tencent, and Alibaba are examples of BigTech firms that offer various forms of payment, lending, or other financial services. The activities of BigTech firms in credit provision are so far the most

[39] For example, the U.S. generally prefers a multiple stakeholder model of cyber governance in which states, international organizations, and private actors all play a role in governance, while China and Russia generally prefer a sovereign-oriented model that gives individual states most of the power.

[40] Financial Conduct Authority (FCA), *Finalized Guidance for Firms Outsourcing to the "Cloud" and Other Third-Party in IT Services*, July 2016, file:///C:/Users/c1jaj01/AppData/Local/Microsoft/Windows/Temporary%20Internet%20Files/IE/9U2E1YMD/fg16-5.pdf

[41] E. Dunkley, "OakNorth Takes UK Banking into the Cloud," *Financial Times*, May 26, 2016, https://www.ft.com/content/36c4eba2-2280-11e6-9d4d-c11776a5124d

[42] See Financial Stability Board, 2017. *Financial Stability Implications from FinTech*, https://www.fsb.org/wp-content/uploads/R270617.pdf.

pronounced in China and have been growing in other jurisdictions as well, though on a smaller scale. Frost et al. (2019) show that BigTech firms usually start with payments, and then expand into credit provision, insurance, and savings and investment products.

Stulz (2019) reviews the recent development in finTech and BigTech and compare them with traditional banks. He defines fintech as the use of digital technologies and big data to enhance existing financial services. Given this definition, fintech firms compete with banks for specific financial services; fintech firms are relatively lightly regulated, not part of inflexible organizations, and not saddled with legacy IT systems. BigTech firms are technology companies whose business model is focused on exploiting digital technologies. They have potential big advantages compared with banks and fintech firms; they not only have all the technical knowhow and up-to-date systems that fintech companies aspire to, but also the scale that large banks have. Other than these benefits, they also have access to a wide range of data that banks and fintech firms do not have access to, but they do not the legacy or the organizational issues. All these advantages might allow BigTech firms to replace traditional banks (see also Frost et al., 2019).

Other than competing with incumbent financial institutions, there are also other forms of interactions between BigTechs and traditional banks. In many cases, BigTech firms are important third-party service providers to financial institutions. Amazon AWS, Microsoft, and Google are all large providers of cloud services to many financial institutions. Ali Cloud, an affiliated firm of Ant Financial, is a dominant player in the Asian market. There has been evidence that BigTech firms have helped to promote financial inclusion. For example, using the proprietary (alternative) data on small businesses' credit and payment activities through BigTech firms (BigTech Credit) in Argentina and China, Frost et al. (2019) find that firms that have access to credit through BigTech firms were able to expand their product offering resulting in significant growth. There have been, however, concerns around the impact on financial stability as a large number of financial institutions start to rely more and more on a few large cloud service providers. Efforts have been devoted into building infrastructure in a more coordinated fashion (for compatible systems) allowing firms to switch instantaneously from one cloud service provider to another provider without major interruption to the financial and payment systems. However, the future remains uncertain regarding the overall impact — much depending on the regulatory design to protect financial stability and consumers while also promoting efficiency through fintech innovations.

9. Regulation of Fintech

Prior to the global financial crisis in 2008, financial innovation was viewed very positively, resulting in a laissez-faire and deregulatory approach to financial regulations. After the crisis, fintech and data-driven financial services providers profoundly challenge the current regulatory paradigm. Financial regulators are seeking to balance the competing objectives of promoting innovations, financial stability, and consumer protection. In this section, we review recent development in regulating fintech, especially how technology is playing an ever-increasing role in financial regulation itself.

9.1 RegTech

RegTech, a contracted term of *regulator*y and *technology*, describes the use of technology in the context of regulatory monitoring, reporting, and compliance (e.g., Zetzsche et al., 2017; Arner, Barberis, and Buckley, 2016b). A defining feature of RegTech is that it shifts from supervision by humans to supervision by machines and analysis of data. After the 2008 global financial crisis, the cost of regulatory obligations has dramatically increased, providing a strong economic incentive for more efficient reporting and compliance systems to better control risks and reduce compliance costs. For example, in 2009, the SEC created the division of Economic and Risk Analysis, to use data insights for better regulation. The Bank of England highlighted some specific technologies currently used in financial regulations, such as pattern analysis, which can be used to identify unusual patterns of activity such as "spoofing," front running, and wash trades; big data analysis, which typically uses a far larger number of inputs than standard surveillance techniques; predictive coding, which looks to identify patterns of activity; or digitalization of voice communications, which could be more effective than written communications.[43]

Looking forward, the next stage of the development of RegTech is likely to be driven by regulators' effort to increase their supervisory capacity, using techniques such as AI/ML and deep learning (Arner, Barberis, and Buckley, 2016b). For instance, a close to real-time surveillance system is proposed by the Bank of England when discussing the future of regulation.[44] Auer (2019) proposes a new concept, Regulation Automata (RA), and shows that distributed ledger technologies (DLT) could be used to ensure high-quality and low-cost compliance. The RA would allow

[43] C. Roxburgh, M. Shafik, and M. Wheatley, 2015. *Fair and Effective Market Review: Final Report, Bank of England*, June 10, 2015, https://www.bankofengland.co.uk/report/2015/fair-and-effective-markets-review---final-report.
[44] A. Haldane, 2014. *Managing Global Finance as a System*, Bank of England, October 29, 2014, https://www.bankofengland.co.uk/-/media/boe/files/speech/2014/managing-global-finance-as-a-system.pdf?la=en&hash=93BF6D650AAE5D055618D2D2DBC5870DC0580FA7.

supervisors to verify compliance with regulatory goals by reading the market's distributed ledger without the need for businesses to actively collect, verify, and deliver the data, therefore, reducing the costs of regulations. In the DLT-based market, data credibility is assured by economic incentives, whereas in today's compliance process, trustworthiness is guaranteed by relevant authorities and the threat of legal penalties.

RegTech has also been developed within major financial institutions and fintech firms. Recent leading examples include compliance requirements in the financial industry such as anti-money laundering (AML) and know your customers (KYC). AML regulations are a legislative attempt to prevent the practice of generating income through illegal actions or to conceal criminal origin and true ownership. Back to the historical legislation, while not mentioning money laundering specifically, the Bank Secrecy Act of 1970 essentially laid the groundwork for what would become the "anti-money laundering complex" that governs a substantial portion of financial law enforcement today. For example, for insured banks and financial institutions, the act imposed a duty to maintain records of "the identity of each person having an account in the U.S. with the bank and of each individual authorized to sign checks, make withdrawals, or otherwise act with respect to any such account." The decades afterward saw the Money Laundering Control Act of 1986, Annunzio-Wylie Anti-Money Laundering Act of 1992, the Money Laundering and Financial Crimes Strategy Act of 1998, and the USA PATRIOT Act of 2001. Internationally, an intergovernmental organization called the Financial Action Task Force (FATF) on Money Laundering was launched in 1989, establishing international soft law standards that address money laundering and terrorist financing.[45] The FATF rules were implemented by financial institutions in most jurisdictions as well as infrastructure providers such as SWIFT and CLS.

From a standpoint of compliance, AML requires that every client or potential client of a financial institution must be reviewed under the central elements, usually known as KYC due diligence. Hence, KYC is an intensive process that requires documentation of clients' identity, income, and sources of funds at a deeper than surface level. Identity today can be analogue (paper documents), digitized (e.g., scan of an ID document), or digital (e.g., online footprint). Identity verification, both on the acceptance of a new customer (onboarding) and on an ongoing basis, is essential to protect against fraud and crime, and henceforth is fundamental to market integrity as

[45] The 40 Recommendations of the FAFT, originally drawn up in 1990, revised in 1996, and again reviewed and updated in 2003, are universally recognized as the international standards for AML.

well. At the same time, identification and KYC rules can be major barriers to accessing financial services for millions of individuals and small businesses.

Arner, Zetzsche, Buckley, and Barneris (2018) argue that technology can present an opportunity to solve this challenge and ensure the objectives of financial inclusion and market integrity through the development of digital identity infrastructure and electronic KYC (e-KYC) infrastructure, but not at the cost of financial stability. A good example is India Stack, a collection of open application programming interfaces (APIs) that provides a paperless e-KYC service to instantly establish the identify of prospective banking customers in India.[46] An estimation by the Ministry of Finance of India shows that moving from paper-based KYC to e-KYC in India reduced the average cost of verifying customers from roughly $15 to $0.50, and Indian banks that make the shift can shorten the time spent on verifying customers from more than five days to seconds.[47]

In the age of cryptocurrency and blockchain, while financial institutions gradually developed tools to combat money laundering over the last decades, AML and KYC have become more intricate and complex because of the difficulty to verify identity and IP addresses, as well as the growing theft of cryptocurrencies.[48] Lawmakers around the world in 2018 added a broader scope to AML, such as the adoption of the European Union's Fifth Anti-Money Laundering Directive (AMLD5) that includes cryptocurrency exchanges. In February 2019, the FATF published a draft of an Interpretive Note to Recommendation 15, which was formally adopted as part of the FATF Standards in June 2019, to guide regulatory authorities in member countries when identifying risk, sharing information, and monitoring virtual asset service providers. Virtual asset service providers will need to be registered or licensed and monitored by authorities.[49]

---

[46] India Stack is an idea originated by a group of Indian IT entrepreneurs and supported by the government and the Reserve Bank of India. The e-KYC service allows customers to electronically provide their demographic and personal information, including proof of identity, address, date of birth and gender, to financial providers, who can verify it in real time. More introduction of India Stack, https://indiastack.org/about/.

[47] For more details, see https://www.cgdev.org/blog/overcoming-know-your-customer-hurdle-e-kyc.

[48] In virtual money laundering, the first step of the cleaning process is called *layering*; the criminals move money into the cryptocurrency system and move it around by using mixers, tumblers, and chain hopping. The pseudo-anonymous nature of virtual currencies makes it exponentially difficult to trace funds as compared with cash. The second step is burst-out integration; criminals can obtain fiat currency through withdrawing a basic currency to a connected bank account or transferring to real estate (see, KPMG report, *Anti-Money Laundering in Times of Cryptocurrencies*, June 2018, https://assets.kpmg/content/dam/kpmg/ch/pdf/anti-money-laundering-in-times-of-cryptocurrency.pdf).

[49] FATF, *Public Statement: Mitigating Risks from Virtual Assets*, February 22, 2019, http://www.fatf-gafi.org/publications/fatfrecommendations/documents/regulation-virtual-assets-interpretive-note.html.

In the U.S., to identify and address the risk of identity theft, the SEC and the Commodity Futures Trading Commission (CFTC) jointly issued rules (SEC's Regulation S-ID and CFTC's Subpart C) requiring certain regulated entities that qualify as either financial institutions or creditors to adopt, implement, and update periodically a written program on identity safety known as the Red Flag Rules in 2013. In September 2018, the SEC brought its first enforcement action under the Identity Theft Red Flag Rules against a registrant, Voya Financial Advisor, because it failed to adequately protect customer information following a six-day cyberattack in 2016.[50]

9.2 Regulation of Fintech

The goals in fintech regulation has been to design a policy framework that would encourage and support disruptive innovation to enhance financial inclusion and economic growth while at the same time provide protection to the safety and soundness of the banking system and financial stability overall. Brummer and Yadav (2019) summarize market integrity, rules simplicity, and financial innovations as the three-legged foundation to accomplish these goals. They point out the tradeoffs and the potential interplay that could interfere with one another. In designing fintech regulations, regulators seek to provide clear rules, maintain market integrity, and encourage fintech innovations, but they are likely to achieve only two of the three objectives. For example, if regulators prioritize market safety, financial stability, and simple and transparent rulemaking, the rules would likely impose broad prohibitions, which can largely inhibit fintech innovations. Alternatively, if regulators wish to prioritize encouraging fintech innovations and provide clarity of rules, they might have to use a simple and low-intensity regulatory framework that may not ensure safety, soundness, and stability in the financial system. Finally, if regulators look to enable innovations and promote market integrity, they would likely impose a set of more complex rules (and with specific cases of exemptions), which are less transparent and not easy to understand.

With the use of big data and AI/ML, the current wave of fintech has created informational and regulatory gaps and loopholes that need to be closed. Rapid advance in the technologies have also increased the uncertainties and difficulty in evaluating the impact of new technologies on consumers, the market, and financial systems overall. Jagtiani and John (2018) provide an overview of fintech innovations and regulatory considerations, especially discussion around consumer protection in response to fintech growth. Brummer and Yadav (2019) propose several regulatory strategies in dealing with the potential risks posed by fintech, including informal guidance, no-

---

[50] SEC, Advisers Act Rel. 5048 (September 26, 2018), https://www.sec.gov/litigation/admin/2018/34-84288.pdf.

action letters, regulatory sandboxes, and other pilot programs, licensing versus chartering forms of organization.[51]

A fintech regulatory sandbox is a regulatory innovative method in fintech, based on informal guidance. Specifically, it is a mechanism for firms to conduct tests of new fintech products and services in a live environment, with regulatory oversight and subject to certain conditions and safeguards. The innovators would be provided with an environment within which they could experiment and try out their new ideas — under a more relaxed regulatory environment but would cause no harm to the general population or the financial system. For example, Ringe and Ruof (2019) propose a regulatory sandbox for robo-advice, with which market participants test robo-advice services in the real market, with real consumers, but under close scrutiny of the supervisor.

The first fintech regulatory sandbox was proposed and implemented by the UK's Financial Conduct Authority (FCA) in November 2015. The concept was then followed by other countries such as Canada, Malaysia, Singapore, Australia.[52] The FCA accepted applications for its first three cohorts in June and December 2016 and June 2017. For the three cohorts, the FCA received a total of 146 applications and admitted 68 of them. Nine companies, which had been accepted into the first two cohorts, were unable to test their solutions for a variety of reasons.[53]

The benefits of regulatory sandboxes include not only signaling a friendly general regulatory approach to fintech innovations but also reducing fintech regulatory uncertainty and increasing regulatory and supervisory capacity.[54] Jagtiani and John (2018) point out how the lack of clarity around which alternative data variables could be used in credit risk modeling has created regulatory uncertainty. The fintech industry seems to be more concerned about fintech regulatory uncertainties and the lack of clarity than regulation itself, as they set their strategies and business model. Fintech firms are willing to be regulated in a level playing field with traditional lenders, for example, as evident in several recent applications by a fintech lender to become a bank: LendingClub acquiring Radius Bank in February 2020, and Square applied and received approved

---

[51] Regulators can also design or oversee tests involving new innovations or techniques, observe outcomes, and then tailor rulemaking to its most efficient and effective form. In other words, pilots provide regulators with a way to generate information on the likely effects of particular products or services. As an example, China has been using frequent pilots when making new rules tied to liberalizing financial markets.

[52] The UK's sandbox is one of five initiatives carried out under the umbrella of the FCA's Project Innovate, which include regulatory sandbox, innovative hub, advice unit, RegTech initiative, and engagement initiative.

[53] FCA, *Regulatory Sandbox Lessons Learned Report*, October 2017, https://www.fca.org.uk/publication/research-and-data/regulatory-sandbox-lessons-learned-report.pdf.

[54] J. Crane, L. Meyer, and E. Fife, 2018. *Thinking Inside The Sandbox: An Analysis of Regulatory Efforts to Facilitate Financial Innovation*, RegTech Lab, https://www.regtechlab.io/report-thinking-inside-the-sandbox.

to become an Industrial Loan Corporation (a depository institution that is insured by the FDIC) in March 2020.

The U.S. regulatory authorities, including the SEC, CFPB, OCC, and the FDIC, have also implemented some pilot programs to further understand the various aspects of fintech to consumers and the financial systems.[55] Zetzsche, Buckley, Arner, and Barberis (2017b) summarize the current regulatory approaches: doing nothing (being permissive to highly restrictive), case-by-case approach (such as no-action letters in the U.S.), and structured experimentalism (such as sandboxes, testing, and piloting). They propose a new smart regulation approach, which comprises regulatory design and implementation in stages as follows: testing and piloting environment, conducting a regulatory sandbox, issuing a restricted licensing or a special charter scheme, and finally when size and income permits, moving to operating under a full license. From each stage to the next, regulatory complexity and costs increase, as does the scope of fintech innovations.

While traditional firms are increasing investments in technology to keep up with the new consumer preferences and to sustain in the new tech landscape, tech start-ups and BigTech firms are rapidly getting involved in payments and providing other financial services. In the long run, traditional financial institutions, fintechs and BigTechs in financial services may converge, as large international banks may buy big data sets from various sources and compile these with their own proprietary data. In addition, some BigTechs may ultimately apply for full financial services licenses and become global financial conglomerates, as evident in the Chinese financial system with Alibaba BigTech and its Ant Financial that owns the largest mutual fund in the world.

The opportunities that BigTechs can provide are obvious in reducing transaction costs, as well as improving risk management and financial inclusion. From the regulatory point of view, both BigTechs and fintechs seek to minimize regulatory constraints and costs. However, they are still different in terms of clients' trust and potential systemic risks. First, BigTechs create trust in a world unrelated to financial services (such as trust in the payment platforms operated by Amazon and PayPal) and leverage clients' trust in the financial sphere (such as Amazon and PayPal using data from their own payment platforms to make loans to small businesses). Second, BigTechs are often significant firms prior to stepping into the financial sector, while fintech firms usually start small as problem-driven firms. Hence, a central regulatory issue is the tradeoff between improving innovation/financial inclusions and reducing systemic risks. One possible way to respond,

[55] The OCC launched an innovation pilot program in April 2019, https://www.occ.gov/news-issuances/news-releases/2019/nr-occ-2019-42.html

proposed in Zetzsche, Buckley, Arner, and Barberis (2017a), is requiring authorization for data gathering and analytics when used for financial services, if the size exceeds certain thresholds.

10. Concluding Remarks

Fintech has the potential for disruption to the entire financial system, and it has played an important role in rapid digital transformation and how financial products and services are produced, delivered, and consumed in recent years. The current COVID-19 pandemic may also expedite us toward fully digital lifestyle and cashless economy, as digital banks are seeing increased traffic while people embrace digital and contactless technologies. In this paper, our aim is to provide a comprehensive review of research studies and policy discussion around fintech, covering the various aspects of fintech and its role in each segment of the financial market and the associated impact on consumers, other market participants, and the financial system overall.

So much data has been collected and monetized in recent years, leading to an interesting fintech mantra: "Data is the new currency." The vast amount of data and the fast and complex computing algorithms have become key factors that drive innovations in recent years, and billions of consumers around the globe have benefited from these changes. The new generation of consumers are less tolerant to a mediocre quality of services and opaque fine prints in their financial relationship, while BigTech firms are exploring opportunities to serve. Many believe that firms relying on the fine print, such as the various hidden fees (like credit card fees), will not survive the new landscape. It is important to recognize that these advanced technologies that improve our lives also come with new set of risks, such as consumer privacy risk and greater cybersecurity risk.

Fintech activities have been progressing quickly and penetrating all areas of the financial system. Under the new landscape where financial firms use cloud storage and cloud computing to achieve high speed and efficiency, it is no longer true that larger firms are more efficient, and it is no longer true that larger firms beat smaller firms (but faster one beat those that are slower). In other words, large firms that do not fully use the technology could potentially fall behind. Cloud computing may also introduce new exposure to cyber-risks that did not exist before, while it has greatly benefited traditional financial institutions in terms of efficiency, resiliency, and flexibility. BigTech firms are playing increasing roles in financial services and real-time payments and providing cloud computing services to large and small financial institutions.

It remains unclear, however, how loans originated by fintech firms will perform relative to traditional loans in an extreme environment, and they are going through a real test now under the COVID-19 crisis. It is also unclear how the complex algorithms with alternative data, which worked well earlier, would continue to perform in the new landscape after the COVID-19 crisis. The AI models may need to be retrained with new data to reflect the "new normal." In the meantime, fintech lenders have a role to play in the background (white label services) to assist small banks in screening and processing a large number of loan applications under the Coronavirus Aid, Relief, and Economic Security Act (CARES Act) to support small businesses.

Blockchain technology and smart contracts have attracted a lot of interest, with thousands of fintech start-ups around the globe working on resolving blockchain scalability, with the potential of allowing consumers to have control over their own information and less reliance on third-party intermediation. Blockchain technology has also made possible for real-time payment and settlement, eliminating frictions in the current payment system, especially for cross-border payments. The success of stablecoin could potentially contribute to more widespread use of blockchain technology. Central banks around the globe are considering whether to issue central bank digital currency (CBDC) to stabilize the value of digital assets or to take a supporting role for the public sector to operate in the new *cash-lite economy*. The Chinese government is also experimenting with blockchain technology, introducing the Blockchain-based Services Network (BSN) in April 2020, intended to serve as the backbone infrastructure technology for massive interconnectivity throughout the mainland China.

How should fintech firms be regulated? Many tech firms, especially the payment platforms, have access to unique data about their customers, and they have comparative advantage in making small business loans. These nonbank lenders are providing similar financial products and services outside the banking regulatory framework. There have been discussions around activity-focused regulations so that all lenders (banks and nonbanks) would be subject to the same regulations. It is debatable whether the future mainstream financial technology would be blockchain and DLTs, quantum computing, or something else — and how the industry and policymakers could be best prepared to keep pace with the evolving technologies and the new adoption. While the advanced technology has delivered vast benefits, it has also allowed for more sophisticated cyberattacks. In addition, third-party vendor risk management has become more important than ever as loans that were approved on fintech platforms could end up on banks' balance sheet. Banks also increasingly rely on cloud computing services provided by nonbank BigTech firms. Given all these innovations and rapid digital transformation, the existing regulations need to adapt to keep up with the new financial landscape — to protect consumers and financial systems while continuing to promote responsible fintech innovations.

Going forward, it's uncertain how fintech as the mainstream technology for the financial services industry would evolve. Economists have been trying to predict what would happen to fintech when the next recession comes. As we are finalizing this draft, the entire fintech industry is facing a real test for the first time under the current COVID-19 crisis. It has been evident that collaboration among fintech firms and traditional lenders have been critical in pushing the funding relief to those targeted recipients (mainly small businesses) in a timely fashion. There may be evidence of potential complementarity leading to a wave of mergers among financial firms and tech firms, thus further reducing the distinction between fintech firms and traditional firms. Once again, this is an opportune time for researchers to further explore the impact of fintech on consumers and the overall financial systems and stability and to design appropriate financial regulations for the new landscape.

References

Abadi, J., and M. Brunnermeier. 2018. "Blockchain Economics," NBER Working Paper 25407.

Adhami, S., G. Giudici, and S. Martinazzi. 2017 "Why Do Business Go Crypto? An Empirical Analysis of Initial Coin Offerings." *Journal of Economics and Business*, 100: 64-75.

Agur, I., A. Ari, and G. Dell'Ariccia. 2019. "Designing Central Bank Digital Currencies." IMF Working Paper 19/252, November.

Ahmed, U., T. Beck, C. McDaniel, and S. Schropp. 2016. "Filling the Gap: How Technology Enables Access to Finance for Small- and Medium-Sized Enterprises." MIT Press Journals: *Innovations*, 10 (3-4): 35–48.

Ait-Sahalia, Y., and M. Saglam. 2013. "High Frequency Traders: Taking Advantage of Speed." NBER Working Paper 19531.

Ali, R., J. Barrdear, R. Clews, and J. Southgate. 2014. "The Economics of Digital Currencies," *Bank of England Quarterly Bulletin.*

Allen, F., E. Carletti, R. Cull, J. Qian, L. Senbet, and P. Valenzuela. 2019. "Improving Access to Banking: Evidence from Kenya." IDB Working Paper.

Allen, F., A. Demirguc-Kunt, L. Klapper, M. Soledad, and M. Peria. 2016. "The Foundations of Financial Inclusion: Understanding Ownership and Use of Formal Accounts." *Journal of Financial Intermediation* 27: 1–30.

Allen, F., and R. Karjalainen. 1999. "Using Genetic Algorithms to Find Technical Trading Rules." *Journal of Financial Economics*, 51: 245–271.

Amir, E., S. Levi, and T. Livne. 2018. "Do Firms Underreport Information on Cyber-Attacks? Evidence from Capital Markets." *Review of Accounting Studies*. Forthcoming.

Amsden, R., and D. Schweizer. 2018. "Are Blockchain Crowd Sales the New 'Gold Rush'? Success Determinants of Initial Coin Offerings." Concordia University Working Paper.

Andolfatto, D. 2018. "Assessing the Impact of Central Bank Digital Currency on Private Banks." Federal Reserve Bank of St. Louis Working Paper 2018-25.

Arner, D., J. Barneris, and R. Buckley. 2016a. "The Evolution of Fintech: A New Post-Crisis Paradigm?" *Georgetown Journal of International Law*, 1272–1319.

Arner, D., J. Barneris, and R. Buckley. 2016b. "FinTech, RegTech and the Reconceptualization of Financial Regulation," *Northwestern Journal of International Law and Business* 37(3): 371–413.

Arner, D., D. Zetzsche, R. Buckley, and J. Barneris. 2018. "The Identity Challenge in Finance: From Analogue Identity to Digitized Identification to Digital KYC Utilities," EBI Working Paper.

Arnott, R., C. Harvey, and H. Markowitz. 2018. “A Backtesting Protocol in the Era of Machine Learning.” *Journal of Financial Data Science* 1(1): 64–74.

Auer, R. 2019. “Regulation Automata: Using the Blockchain to Efficiently Supervise Financial Markets.” BIS Working Paper.

Baker, T., and B. Dellaert. 2018. “Regulating Robo Advice Across the Financial Services Industry,” *Iowa Law Review* 103: 713–750.

Bakos, Y., and H. Halaburda. 2018. “The Role of Cryptographic Tokens and ICOs in Fostering Platform Adoption.” New York University Working Paper.

Balyuk, T. 2018. “Financial Innovation and Borrowers: Evidence from Peer-to-Peer Lending.” University of Toronto Working Paper.

Barontini, C., and H. Holden. 2019. “Proceeding with Caution— A Survey on Central Bank Digital Currency.” BIS Working Paper 1010. January.

Beck, T., A. Demirguc-Kunt, and R. Levine. 2007. “Finance, Inequality and the Poor.” *Journal of Economic Growth* 12: 27–49.

Beck, T., A. Demirguc-Kunt, and M. Peria. 2007. “Reaching Out: Access to and Use of Banking Services Across Countries.” *Journal of Financial Economics* 85: 234–266.

Benedetti, H., and L. Kostovetsky. 2018. “Digital Tulips? Returns to Investors in Initial Coin Offerings.” Boston College Working Paper.

Benos, E., R. Garratt, and P. Gurrola-Perez. 2017. “The Economics of Distributed Ledger Technology for Securities Settlement.” Bank of England Staff Working Paper 670.

Berg, T., V. Burg, A. Gombovic, and M. Puri. Year? “On the Rise of Fintech — Credit Scoring Using Digital Footprints.” Working Paper.

Biais, B., T. Foucault, and S. Moinas. 2015. “Equilibrium Fast Trading.” *Journal of Financial Economics* 73: 3–36.

Bianchi, D., and O. Tosun. 2019. “Cyber Attacks and Stock Market Activity.” Working Paper.

Bordo, M., and A. Levin. 2017. “Central Bank Digital Currency and the Future of Monetary Policy.” NBER Working Paper 23711.

Brock, W., J. Lakonishok, and B. Lebaron. 1992. “Simple Technical Trading Rules and the Stochastic Properties of Stock Returns.” *Journal of Finance* 47: 1731–1764.

Brummer, C., and Y. Yadav. 2019. “Fintech and the Innovation Trilemma. *Georgetown Law Journal* 107: 235–307.

Brunnermeier, M., and D. Niepelt. 2019. “On the Equivalence of Private and Public Money.” NBER Working Paper 25877, May.

Brunnermeier, M., H. James, and J. Landau. 2019. "The Digitalization of Money." NBER Working Paper 26300, September.

Butler, A., J. Cornaggia, and U. Gurun. 2017. "Do Local Capital Market Conditions Affect Consumers' Borrowing Decisions?" *Management Science* 63(12): 4175–4187.

Catalini, C., and J. Gans. 2018. "Some Simple Economics of the Blockchain." NBER Working Paper 22952.

Chava, S., and N. Paradkar. 2018. "Winners and Losers of Marketplace Lending: Evidence from Borrower Credit Dynamics." Georgia Tech Working Paper.

Chod, J., and E. Lyandres. 2018. "A Theory of ICOs: Diversification, Agency and Information Asymmetry." Boston College Working Paper.

Cohney, S., D. Hoffman, J. Sklaroff, and D. Wishnick. 2019. "Coin-Operated Capitalism." *Columbia Law Review*. Forthcoming.

Cong, L., Y. Li, and N. Wang. 2018. "Tokenomics: Dynamic Adoption and Valuation." Columbia Business School Research Paper 18-46.

Cong, L., and Z. He. 2018. "Blockchain Disruption and Smart Contracts." NBER Working Paper 24399.

Croux, Christophe, Julapa Jagtiani, Tarunsai Korivi, and Milos Vulanovic. 2020. "Important Factors Determining Fintech Loan Default: Evidence from LendingClub Consumer Platform." *Journal of Economic Behavior & Organization* 173, May, 270–296.

Crowe, C., and R. Ramcharan. 2012. "The Impact of House Prices on Consumer Credit: Evidence from an Internet Bank," Federal Reserve Board, FEDS Working Paper.

Crowe, D., M. Rysman, and J. Stavins. 2010. "Mobile Payments in the United States at Retail Point of Sale: Current Market and Future Prospects." Federal Reserve Bank of Boston Public Policy Discussion Paper 10-2.

D'Acunto, F., N. Prabhala, and A. Rossi. 2019. "The Promises and Pitfalls of Robo-Advising." *Review of Financial Studies* 32(5): 1983–2020.

Degryse, H., F. De Jong, and V. Kervel, 2015. "The Impact of Dark Trading and Visible Fragmentation on Market Quality." *Review of Finance* 19(4): 1587–1622.

de Roure, C., L. Pelizzon, and A. Thakor. 2018. "P2P Lenders versus Banks: Cream Skimming or Bottom Fishing?" SAFE Working Paper 206.

DeYoung, R., W. Lang, and D. Nolle. 2007. "How the Internet Affects Output and Performance at Community Banks." *Journal of Banking and Finance* 31: 1033–1060.

Dorfleitner, G., C. Priberny, S. Schuster, J. Stoiber, N. Weber, I. de Castro, and J. Kammler, 2016. "Description-Text Related Soft Information in Peer-to-Peer Lending: Evidence from Two Leading European Platforms." *Journal of Banking and Finance* 64: 169–187.

Du, S., and H. Zhu. 2014. “Welfare and Optimal Trading Frequency in Dynamic Double Auctions.” NBER Working Paper 20588.

Easley, D., M. Lopez de Prado, and M. O’Hara. 2011. “The Microstructure of the ‘Flash Crash’: Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading. *Journal of Portfolio Management* 37: 118–128.

Easley, D., M. Lopez de Prado, and M. O’Hara. 2012. “Flow Toxicity and Liquidity in a High Frequency Trading World.” *Review of Financial Studies* 25: 1457–1493.

Easley, D., M. O’Hara, and S. Basu. 2019. “From Mining to Markets: The Evolution of Bitcoin Transaction Fees.” *Journal of Financial Economics*134(1): 91–109.

Engert, W., and B. Fung. 2017. “Central Bank Digital Currency: Motivation and Implications.” Bank of Canada Staff Discussion Paper.

Fein, M. 2015. “Robo-Advisors: A Closer Look.” Working Paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2658701.

Fernández-Villaverde, F., and D. Sanches. 2016. “Can Currency Competition Work?” NBER Working Paper 22157.

Fernández-Villaverde, F., D. Sanches, L. Schilling, and H. Uhlig. 2020. “Central Bank Digital Currency: Central Banking for All?” NBER Working Paper 26753, February.

Financial Conduct Authority, 2017. *Regulatory Sandbox Lessons Learner Report*. April. https://www.fca.org.uk/publication/research-and-data/regulatory-sandbox-lessons-learned-report.pdf.

Foucant, T., J. Hombert, and I. Rosu. 2016. “New Trading and Speed.” *Journal of Finance* 71: 335–382.

Foucant, T., and A. Menkveld. 2008. “Competition for Order Flow and Smart Order Routing Systems.” *Journal of Finance* 63: 119–158.

Franks, J., N. Serrano-Velarde, and O. Sussman. 2018. “Marketplace Lending, Information Aggregation and Liquidity.” London Business School Working Paper.

Freedman, S., and G. Jin. 2017. “The Information Value of Online Social Networks: Lessons from Peer-to-Peer Lending.” NBER Working Paper 19820.

Frost, J., L. Gambacorta, Y. Huang, H. Shin, and P. Zbinden. 2019. “BigTech and the Changing Structure of Financial Intermediation.” BIS Working Paper 779.

Fu, R., Y. Huang, and P. Singh. 2018. “Crowd, Lending, Machine and Bias.” Carnegie Mellon University Working Paper.

Furst, K., W. Lang, and D. Nolle. 2002. “Internet Banking.” *Journal of Financial Services Research* 22: 95–117.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther. 2018. "Predictably Unequal? The Effects of Machine Learning on Credit Markets." Imperial College London Working Paper. July.

Gai, J., C. Yao, and M. Ye. 2014. "The Externalities of High Frequency Trading." UIUC Working Paper.

Gao, Q., M. Lin, and R. Sias. 2017. "Word Matters: The Role of Texts in Online Credit Markets." Working Paper.

Gervais, A., G. Kharame, K. Wust, V. Glykantzis, H. Ritzdorf, and S. Capkun. 2016. "On the Security and Performance of Proof of Work Blockchains." In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security.

Goddard, J., D. McKillop, and J. Wilson. 2009. "Which Credit Unions Are Acquired?" *Journal of Financial Services Research* 36: 231–252.

Goettler, R., C. Parlour, and U. Rajan. 2009. "Informed Traders in Limit Order Markets." *Journal of Financial Economics* 93: 67–87.

Goldstein, I., J. Jagtiani, and A. Klein. 2019. "Fintech and the New Financial Landscape." *Bank Policy Institute (BPI): Banking Perspectives*, Q1:7, March 2019.

Greene, C., M. Rysman, S. Schuh, and O. Shy. 2014. "Costs and Benefits of Building Faster Payment Systems: The U.K. Experience and Implication for the United States." Federal Reserve Bank of Boston Research Paper Series Current Policy Perspectives14-5.

Habor, S., and S. Stornetta. 1991. "How to Time Stamp a Digital Document," Lecture Notes in Computer Science 537: 437–455.

Harvey, C. 2016. "Cryptofinance." Duke University Working Paper.

Hasan, I., Q. He, and H. Lu. 2019. "Stereotypes in Person-to-Person Lending: Evidence from Debt Crowdfunding." Fordham University Working Paper.

Hau, H., Y. Huang, H. Shan, and Z. Shen. 2018. "FinTech Credit, Financial Inclusion and Entrepreneurial Growth." University of Geneva Working Paper.

Hendershott, T., C. Jones, and A. Menkveld. 2011. "Does Algorithmic Trading Improve Liquidity?" *Journal of Finance* 66: 1–33.

Hernandez-Murillo, R., G. Llobet, and R. Fuentes. 2010. "Strategic Online Banking Adoption" *Journal of Banking and Financ*e 34: 1650–1663.

Hernando, I., and M. Nieto. 2007. "Is the Internet Delivery Channel Changing Banks' Performance? The Case of Spanish Banks." *Journal of Banking and Finance* 31: 1083–1099.

Hildebrandt, T., M. Puri, and J. Rocholl. 2017. "Adverse Incentives in Crowdfunding." *Management Science* 63(3): 587–608.

Hodge, F., K. Mendoza, and R. Sinha. 2018. "The Effect of Humanizing Robo-Advisors on Investor Judgments." University of Washington Working Paper.

Howell, S., M. Niessner, and D. Yermack. 2018. "Initial Coin Offerings: Financing Growth with Cryptocurrency Token Sales." NBER Working Paper 24774.

Hu, A., C. Parlour, and U. Rajan. 2018. "Cryptocurrencies: Stylized Facts on a New Investible Instrument." University of California, Berkeley, Hass Working Paper.

Huang, Y., Y. Li, and H. Shan. 2019. "The Distributional Effect of Fintech Credit: Evidence from E-commerce Platform Lending." University of Geneva Working Paper.

Jagtiani, J., and C. Lemieux. 2018. "Do Fintech Lenders Penetrate Areas That Are Underserved by Traditional Banks?" *Journal of Economics and Business* 100, November–December 2018, 43–54.

Jagtiani, J., and C. Lemieux. 2019. "The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence from LendingClub Consumer Platform." *Financial Management* Winter 2019, 48(4), 1009–1029.

Jagtiani, J., L. Lambie-Hanson, and T. Lambie-Hanson. 2019. "Fintech Lending and Mortgage Credit Access." Federal Reserve Bank of Philadelphia Working Paper 19-47, November.

Jagtiani, J., M.G. Papaioannou, and G. Tsetsekos. 2019. "Cryptocurrencies in the Global Economy." *The Journal of Investing*. Cryptocurrency Special Issue 2019, 28 (3), April, 7–12.

Jagtiani, J., M. Papaioannou, G. Tsetseko, E. Dolson, and D. Milo. 2020. "Cryptocurrencies: Regulatory Perspectives and Implications for Investors." in *Handbook of Alternative Finance*, edited by Raghavendra Rau, Robert H. Wardrop, and Luigi Zingales, Palgrave MacMillan, forthcoming.

Jagtiani, J., T. Vermilyea, and L. M. Wall. 2018. "The Roles of Big Data and Machine Learning in Bank Supervision." *The Clearing House: Banking Perspectives*, Quarter 1, 2018, 6(1), 44–51.

Ji, Megan. 2018. "Are Robots Good Fiduciaries? Regulating Robo-Advisors Under the Investment Advisers Act of 1940." *Columbia Law Review* 117: 1542–1584.

Jones, C. 2013. "What Do We Know About High-Frequency Trading?" Columbia Business School Research Paper 13-11.

Jovanovic, B., and J. Menkveld. 2016. "Middlemen in Limit-Order Markets." NYU Working Paper.

Iyer, R., A. Khwaja, E. Luttmer, and K. Shue. 2016. "Screening Peers Softly: Inferring the Quality of Small Borrowers." *Management Science* 62(2): 1554–1577.

Kaymiya, S., J. Kang, J. Kim, A. Milidonis, and R. Stulz. 2018. "What is the Impact of Successful Cyberattacks on Target Firms." NBER Working Paper 24409.

Keister, T., and D. Sanches. 2019. "Should Central Banks ISSUE Digital Currency?" Federal Reserve Bank of Philadelphia Working Paper 19-26.

Kendall, J., N. Mylenko, and A. Ponce. 2010. "Measuring Financial Access Around the World." World Bank Policy Research Working Paper 5253.

Khandani, A., and A. Lo. 2007. "What Happened to the Quants in August 2007?" *Journal of Investment Management* 5(4): 5–54.

Khandani, A., and A. Lo, 2011. "What Happened to the Quants in August 2007? Evidence from Factors and Transactions Data." *Journal of Financial Markets* 14(1): 1–46.

Kirilenko, A., A. Kyle, M. Samadi, and T. Tuzun. 2017. "The Flash Crash: The Impact of High Frequency Trading on an Electronic Market." *Journal of Finance* 72(3): 967–998.

Kirilenko, A., and A. Lo. 2013. "Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents." *Journal of Economic Perspectives* 27(2): 51–72.

Klus, M., T. Lohwasser, and L. Hornuf. 2018. "How Do Banks Interact with Fintechs? Forms of Strategic Alliance and Their Economic Impact." CESifo Working Paper Series 7170.

Kopp, E., L. Kaffenberger, and C. Wilson. 2017. "Cyber Risk, Market Failures, and Financial Stability." IMF Working Paper 17/185.

Kroeger, A., and A. Sarkar. 2017. "The Law of One Bitcoin Price?" Federal Reserve Bank of New York Working Paper.

Lee, J., T. Li, and D. Shin. 2018. "The Wisdom of Crowds in FinTech: Evidence from Initial Coin Offerings." Working Paper.

Li, J., and W. Mann. 2018. "Initial Coin Offerings and Platform Building." University of Californa, Los Angeles Anderson Working Paper.

Lin, L., and C. Chen. 2019. "The Promise and Perils of InsurTech." NUS Law Working Paper 21.

Lin, M., N. Prabhala, and S. Viswanathan. 2013. "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending." *Management Science* 59(1): 17–35.

Lin, M., and S. Viswanathan, 2016. "Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market." Management Science, 62(5): 1393-1414.

Lin, T. 2016. "Financial Weapons of War." *Minnesota Law Review* 100: 1377–1440.

Lin, T., and V. Pursiainen. 2016. "Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market." *Management Science* 62(5): 1393–1414.

Lin, T,. and V. Pursiainen. 2018. "Fund What You Trust? Social Capital and Moral Hazard in Crowdfunding." University of Hong Kong Working Paper .

Lin, Z., T. Sapp, J. Rees-Ulmer, and R. Parsa. 2018. "Insider Trading Ahead of Cyber Breach Announcements." Iowa State University Working Paper.

Linnainmaa, J., B. Melzer, and A. Previtero. 2017. "Retail Financial Advice: Does One Size Fit All?" *Journal of Finance* 72: 1441–1482.

Linton, O., and S. Mahmoodzadeh. 2018. "Implications of High-Frequency Trading for Security Market." *Annual Review of Economics* 10: 237–259.

Liskovich, I., and M. Shaton. 2017. "Borrowers in Search of Feedback: Evidence from Consumer Credit Marketts." Federal Reserve Board, FEDS Working Paper.

Lopez de Prado, M. 2018. "The 10 Reasons Most Machine Learning Fund Fail." Working Paper.

Ma, J., J. Gans, and R. Tourky. 2018. "Market Structure in Bitcoin Mining." NBER Working Paper 24242.

Makarov, I., and A. Schoar. 2018. "Trading and Arbitrage in Cryptocurrency Markets." MIT Sloan Working Paper.

Malinova, K., and A. Park. 2017. "Market Design with Blockchain Technology." University of Toronto Working Paper.

Mancini-Griffoli, T., M. Peria, I. Agur, A. Air, J. Kiff, A. Popescu, and C. Rochon. 2018. "Casting Light on Central Bank Digital Currency." IMF Staff Discussion Note.

Mehta, N. 2013. "Pros and Cons of 'Tobin Tax' Divides EU." *Financial Times*. March 3: https://www.ft.com/content/ce94512e-81a3-11e2-ae78-00144feabdc0.

Menkveld, A. 2013. "High Frequency Trading and the New Market Makers." *Journal of Financial Markets*, 16: 712–740.

Menkveld, A. 2016. The Economics of High-Frequency Trading: Taking Stock." *Annual Review of Financial Economics* 8: 1–24.

Mester, L., L. Nakamura, and M. Renault. 2007. "Transaction Accounts and Loan Monitoring." *Review of Financial Studies,* 20: 529–556.

Mills, D., K. Wang, B. Malone, A. Ravi, J. Marquardt, C. Chen, A. Badev, T. Brezinski, L. Fahy, K. Liao, V. Kargenian, M. Ellithorpe, W. Ng, and N. Baird. 2016. "Distributed Ledger Technology in Payments, Clearing and Settlement," Federal Reserve Board FEDS Working Paper 095.

Momtaz, P. 2018. "Initial Coin Offering." University of California, Los Angeles, Anderson Working Paper.

Nakamoto, S. 2008. "Bitcoin: A Peer-to-Peer Electronic Cash System." Unpublished manuscript. https://bitcoin.org/bitcoin.pdf.

Norden, L., and M. Weber. 2010. "Credit Line Usage, Checking Account Activity, and Default Risk of Bank Borrowers. *Review of Financial Studies* 23: 3665–3699.

O'Hara, M. 2015. High Frequency Market Microstructure. *Journal of Financial Economics* 116: 257–270.

O'Hara, M., and M. Ye. 2011. "Is Market Fragmentation Harming Market Quality?" *Journal of Financial Economics* 100: 459–474.

Padhi, M. 2017. "The Effects of Peer-to-peer (P2P) Lending on Competition, Discrimination, and Financial Stability," University of Maryland Working Paper.

Pagnotta, E., and A. Buraschi. 2018. "An Equilibrium Valuation of Bitcoin and Decentralized Network Assets." Imperial College London Working Paper.

Pagnotta, E., and T. Philippon. 2018. "Competing on Speed." *Econometrica* 86(3): 1067–1115.

Raskin, M., and D. Yermack. 2016. "Digital Currencies, Decentralized Ledgers, and the Future of Central Banking." NBER Working Paper 22238.

Ravina, E. 2018. "Love & Loans: The Effect of Beauty and Personal Characteristics in Credit Markets." Working Paper.

Ringe, W., and C. Ruof. 2019. "A Regulatory Sandbox for Robo Advice." University of Hamburg Working Paper.

Rysman, M., and S. Schuh. 2016. "New Innovations in Payments." NBER Working Paper 22358.

Schilling, L., and H. Uhlig. 2019. "Some Simple Bitcoin Economics." *Journal of Monetary Economics* 106: 16–26.

Schweitzer, M., and B. Barkley. 2017. "Is Fintech Good For Small Business Borrowers? Impacts on Firm Growth and Customer Satisfaction." Federal Reserve Bank of Cleveland Working Paper.

Securities and Exchange Commission. 2010. "Concept Release on Equity Market Structure." Release 34-61358, File No. S7-02-10, https://www.sec.gov/rules/concept/2010/34-61358.pdf.

Securities and Exchange Commission. 2014. *Equity Market Structure Literature Review*, Part II: High Frequency Trading.
https://www.sec.gov/marketstructure/research/hft_lit_review_march_2014.pdf.

Sockin, M., and W. Xiong. 2018. "A Model of Cryptocurrencies." Princeton University Working Paper.

Stulz, R., 2019. "FinTech, BigTech, and the Future of Banks." Fischer College of Business Working Paper 2019-03-020.

Szabo, N. 1994. "Smart Contract." Unpublished manuscript.
http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart.contracts.html.

Szabo, N. 2005. "Bit Gold." Unpublished manuscript. https://nakamotoinstitute.org/bit-gold/.

Tang, H. 2018. "Peer-to-Peer Lenders versus Banks: Substitutes or Complements?" *Review of Financial Studies*. Forthcoming.

Tinn, K. 2018. “Blockchain and the Future of Optimal Financing Contracts.” Imperial College London Working Paper.

Thakor, A. 2019. “FinTech and Banking: What Do We Know?” *Journal of Financial Intermediation*. Forthcoming.

Vallee, B., and Y. Zeng. 2018. “Marketplace Lending: A New Banking Paradigm?” HBS Working Paper.

Wei, Z., and M. Lin. 2017. “Market Mechanisms in Online Crowdfunding.” *Management Science* 62(12): 3999-4446.

Yermack, D. 2017. Corporate Governance and Blockchains, *Review of Finance*. 21(1): 7-31.

Zetzsche, D., R. Buckley, D. Arner, and J. Barberis. 2017a. “From Fintech to Techfin: The Regulatory Challenges of Data-Driven Finance.” University of Hong Kong Working Paper.

Zetzsche, D., R. Buckley, D. Arner, and J. Barberis. 2017b. “Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation.” University of Hong Kong Working Paper.

Zetzsche, D., R. Buckley, D. Arner, and L. Fohr. 2019. “The ICO Gold Rush: It’s a Scam, It’s a Bubble, It’s a Super Challenge for Regulators.” *Harvard International Law Journal* 63(2).

GENERATIVE ADVERSARIAL NETWORKS IN FINANCE: AN OVERVIEW

A PREPRINT

Florian Eckerli *
School of Engineering
Zurich University of Applied Sciences
Winterthur, Switzerland
eckerflo@students.zhaw.ch

Joerg Osterrieder*
School of Engineering
Zurich University of Applied Sciences
Winterthur, Switzerland
joerg.osterrieder@zhaw.ch

The Hightech Business and Entrepreneurship Group
Faculty of Behavioural, Management and Social Sciences
University of Twente
Enschede, Netherlands
joerg.osterrieder@utwente.nl

June 11, 2021

ABSTRACT

Modelling in finance is a challenging task: the data often has complex statistical properties and its inner workings are largely unknown. Deep learning algorithms are making progress in the field of data-driven modelling, but the lack of sufficient data to train these models is currently holding back several new applications. Generative Adversarial Networks (GANs) are a neural network architecture family that has achieved good results in image generation and is being successfully applied to generate time series and other types of financial data. The purpose of this study is to present an overview of how these GANs work, their capabilities and limitations in the current state of research with financial data, and present some practical applications in the industry. As a proof of concept, three known GAN architectures were tested on financial time series, and the generated data was evaluated on its statistical properties, yielding solid results. Finally, it was shown that GANs have made considerable progress in their finance applications and can be a solid additional tool for data scientists in this field.

***K*eywords** Generative Adversarial Networks, GANs, Time Series, Synthetic Data

1 Introduction to the modelling problem in finance

The finance industry is one of the most influential fields impacted by new developments in AI. More specifically machine learning has been deployed to forecasting, customer service, risk management and portfolio management, among others. The finance industry generates huge amounts of data which offers new opportunities for models to be applied and used to obtain better knowledge and spot opportunities. It also has the resources to apply into research of the most novel techniques.

Financial markets, like the economy, are highly complex systems where it is often impossible to explain macro phenomena by a simple summation of micro processes or events. It's hard to predict the results of actions in the system

*This article is based upon work from COST Action 19130 Fintech and Artificial Intelligence in Finance, supported by COST (European Cooperation in Science and Technology), www.cost.eu. (Action Chair: Joerg Osterrieder)

Financial support by the Swiss National Science Foundation within the project "Mathematics and Fintech - the next revolution in the digital transformation of the Finance industry" is gratefully acknowledged by the corresponding author. This research has also received funding from the European Union's Horizon 2020 research and innovation program FIN-TECH: A Financial supervision and Technology compliance training programme under the grant agreement No 825215 (Topic: ICT-35-2018, Type of action: CSA)

and sometimes impossible to find the causes of large abnormalities, or even to single out factors that influenced an event. In the 2010 flash crash[1] for instance, plausible explanations only surfaced years after the episode, and even in specific segments, studies[2] have found that tools for market analysis show a incomplete picture. Modelling in this setting involves dealing with a great deal of uncertainty.

While there are large amounts of data generated by the banking and financial sectors, it's possible to notice the shortage of publicly available datasets originating from these environments. Daily stock price data, for instance, is openly available, but many other types of data are protected by regulations and privacy laws. This scarcity of historical and other types of data has been a limiting barrier to achieve better models. The data-intensive nature of ML models generates the situation where a lack of data limits the potential of data-driven methods.

This paper aims to show an overview of the novel framework for generating synthetic data called Generative Adversarial Networks, also known as GANs, a family of generative models that is being used to produce diverse types of synthetic financial data. GANs are being deployed to solve some of the scarcity of real data, optimize portfolios and trading strategies, among other uses. The goal is to provide a basic understanding of how a GAN works, present the current state of research and current practical applications. Additionally, a proof of concept model was built to generate some synthetic financial time series with a data-driven approach.

1.1 Stylized facts of financial time series

The dynamics of financial markets can be defined as examples of complex phenomena[3], since the underlying mechanics of the processes are largely unknown. However, years of research in financial time series have shown that although they may look completely random, variations of asset prices share several significant statistical properties, which are common across a multitude of markets and timeframes. These properties are known as stylized empirical facts, or just stylized facts.

Stylized facts were gathered by taking statistical features of asset returns found in studies about many markets and instruments. A causal analysis of distinct economic scenarios over the globe may expect that being influenced by different events and environment would result in different statistical properties of the said returns. Still, a result of decades of empirical studies of financial markets has shown that some key characteristics are constantly present across markets with different assumed characteristics. Due to these generalizations, the obtained features end up losing in precision, giving the stylized facts a more qualitative aspect. Is important to note that these qualitative properties are not easy to reproduce via modelling of stochastic processes and end up being essential in modelling asset price dynamics, as it is expected that these models can capture/replicate these statistical features.

From the proposed stylized facts[4] a selection of five is portrayed here based on their relevance and use in the modelling literature:

• **Linear unpredictability** or **absence of autocorrelations**: Except for short intra-day time scales, it is expected that asset returns show minimal linear autocorrelations.

• **Fat-tailed distribution** or **heavy tails**: the distribution of returns seems to display a power-law or Pareto-like tail, which in practical terms means that the probability of extreme values occurring is much larger than in a normally distributed dataset.

• **Volatility clustering**: different measures of volatility display a positive autocorrelation over several days. Meaning that there is a tendency of high volatility events to cluster together. Simplifying, large price changes tend to be followed by large changes and small price changes tend to be followed by small changes.

• **Gain/loss asymmetry**: Observations have shown that large plunges in stock prices and index values do no share equally large upward movements.

• **Aggregational Gaussitanity**: Over larger time scales, the distribution of returns looks like a normal distribution.

1.2 Models used in finance

Presently there are a few approaches to financial time series modelling, notably there are the **ARCH/GARCH** family of models, which rely on classical statistics and model the change in variance over time in a time series by describing the variance of the current error term as a function of past errors. Often the variance is related to the squares of the previous innovations. As AR (auto regressive) models, they heavily rely on past information to construct a prediction. They are commonly used in time series modelling where time-varying volatility and volatility clustering are present. They don't have a stochastic component since volatility is completely dependent on previous values. ARCH/GARCH models are used to predict the variance at future time steps.

Another used family of models are the Agent-based models (**ABM**s). In agent-based models, the agents are entities, typically represented computationally as objects. These agents hold a state, which can be any data that describes the agent. ABMs simulate interactions of multiple agents to re-create and predict the behaviour of complex phenomena, where the goal is to generate higher level properties from the interaction of lower level agents.

Modelling financial time series is a big challenge, since the dynamics of financial markets are very complex in nature, and the mechanism generating the data, and thus its original distribution, is unknown. Adopting a purely data-driven modelling approach to this problem could provide new solutions or alternative paths by removing a source of bias from modelling.

Artificial intelligence in finance
Making use of newly available modelling techniques and processing power, financial analysis has been evolving constantly[5] into more complex modelling methods, such as the deep learning approaches and most recently, Generative Adversarial Networks. GANs may provide good results, since they can generate data by sampling only from real data, often with no additional assumptions or inputs. Avoiding assumptions may be a very important aspect to this type of modelling due to the largely empirical aspect of financial data, the avoidance of possible human bias being infiltrated into the modelling process could possibly be a step forward in this field of study.

2 Generative Adversarial Networks

Generative adversarial networks, GANs, are a series of generative machine learning frameworks first introduced by Ian Goodfellow and his collaborators in 2014[6]. They gained a lot of attention due to its simplicity and effectiveness. In a short time span, considerable progress was made to the initial application of GANs - image generation, but also hundreds of different types of GANs were created to optimize for various tasks, from computer vision to fraud detection in banks, the framework presented in 2014 has come a long way. From getting good images on the MNIST dataset and recognizable human faces in the original paper, as of 2021, GANs can generate near perfect human faces with StyleGAN[7], as seen in figure 1, and have a constantly expanding portfolio of applications.

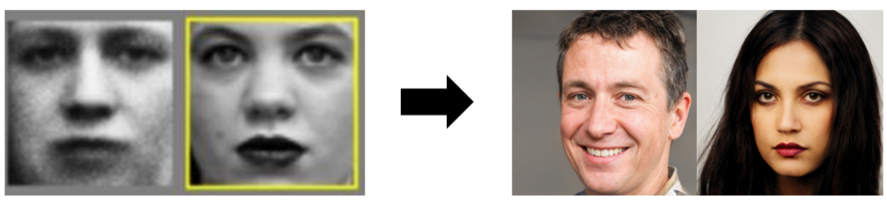

Figure 1: 2014: first generated faces[6] vs. 2020: current state of the art[7]

There are also several applications to finance and financial time series analysis, although compared to other fields, these frameworks are still considered novelties, which makes for an exciting area of research. Since the research is still in its infancy, its reasonable to assume that further applications are yet to be developed and improved upon. Presently, much of the research is still in development and concrete applications are limited, but this paper will show that the current work already shows great potential.

GANs belong to the family of generative models in machine learning ML, generative models are processes that can generate new data instances, more formally, given a observable variable X and a target variable Y, a generative model is a statistical model of the joint probability distribution on P(X|Y). Generally, the process involves discovering patterns in the input data and learning them in a way that the model can then generate new samples that retain characteristics of the original dataset.

2.1 The GAN framework

The original GAN framework estimates generative models through an adversarial process, where two models (usually neural networks) are trained in parallel as represented in figure 2 and described below:

There is a generative model, the Generator (G) that captures the data distribution and generates new data. The second model is a classifier, the Discriminator (D) which estimates the probability that a sample came from the training data and not from G. The training process for the Generator is to maximize the probability of its output being misclassified by the Discriminator.

The Generator is responsible for the generation of data, and the Discriminator has the task of assessing the quality of the generated data and providing feedback to the Generator. These neural networks are optimised under game-theoretic conditions: The Generator is optimised to generate data to fool the Discriminator, and the Discriminator is optimised to detect the source of the input, namely the Generator or the real data set.

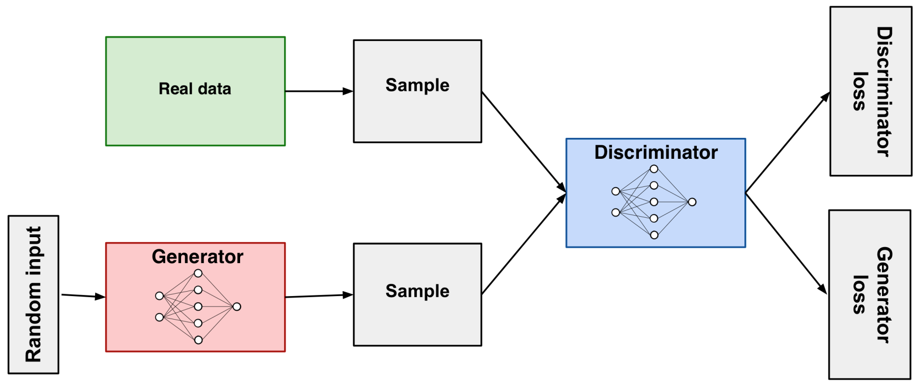

Figure 2: The originally proposed architecture for Generative Adversarial Networks

2.1.1 The Discriminator

The Discriminator is a classifier. It receives Real and Synthetic (from G) data and attempts to distinguish them. It can use several different network architectures, depending on the type of data being classified. The Discriminator network connects to two loss functions that are used in different parts of training. After classifying real/synthetic data, the Discriminator loss penalizes it for misclassifications, and its weights are updated via **backpropagation** from its loss through the network.

2.1.2 The Generator

The Generator network G uses feedback from the Discriminator D to learn to generate synthetic data that ideally resembles the original data in key aspects. Its goal is for the created data to be classified as real by the Discriminator.

The network receives a random input, some type of noise, from which it generates some output, this output is then evaluated by the Discriminator and results in a Generator loss, which then penalizes the Generator for not deceiving the Discriminator. By introducing noise, a GAN can potentially produce a wide variety of outputs by sampling from different places in the target distribution. Usually noise is introduced by sampling from the uniform distribution.

2.1.3 Loss function

The GAN training process uses loss functions that measure the distance between the distributions of the generated and real data to assess their similarity. There are many proposed methods to solve this challenge. In the original "vanilla" GAN, a so-called minimax loss was introduced.

The formulation of the minimax loss is derived from the cross entropy between the real and generated distributions by the JS divergence when the Discriminator is optimal. For the Generator, minimizing the loss is equivalent to minimizing $\log(1 - D(G(\mathbf{z})))$ since it can't directly affect the $\log D(\mathbf{x})$ term in the function. The Jensen-Shannon - JS - divergence measures the similarity between two probability distributions, it is based on the entropy of a discrete random variable being a measurement of the amount of information required on average to describe that variable.

In the original framework, the Generator and Discriminator losses come from a single measure of distance between probability distributions. The two terms are updated in an alternating fashion, depending on which network is being trained.

$$\min_G \max_D V(D,G) = \mathbb{E}_{\mathbf{x}}[\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z}}[\log(1 - D(G(\mathbf{z})))]$$

• $D(\mathbf{x})$ is the Discriminator's estimate of the probability that real data instance $\mathbf{x}$ is real.
• $\mathbb{E}_{\mathbf{x}}$ is the expected value over all real data instances.
• $G(\mathbf{z})$ is the Generator's output when given noise $\mathbf{z}$.
• $D(G(\mathbf{z}))$ is the Discriminator's estimate of the probability that a fake instance is real.
• $\mathbb{E}_{\mathbf{z}}$ is the expected value over all random inputs to the Generator (in effect, the expected value over all generated fake instances $G(\mathbf{z})$).

2.1.4 Optimizers

Optimizers update the model in response to the output of the loss function. In essence, they have control over the learning process of a neural network by finding the values of parameters such that a loss function is at its lowest. The learning rate is a key hyperparameter that scales the gradient and sets the speed at which the model is updated.

Most models use a gradient descent-based optimizer. Gradient descent is the direction of steepest descent of a function, these algorithms are used to find the local minimum of differentiable function by small iterations.

Adaptive moment estimation (Adam) is an optimization algorithm used to update network weights iteratively, it is an extension to stochastic gradient descent and has seen broad adoption in deep learning applications from computer vision to natural language processing. Adam works by storing both the exponentially decaying average of past squared gradients and exponentially decaying average of past gradients. There are a few types of optimizers being used in GAN literature, but Adam is currently among the most popular choices.

2.1.5 Training

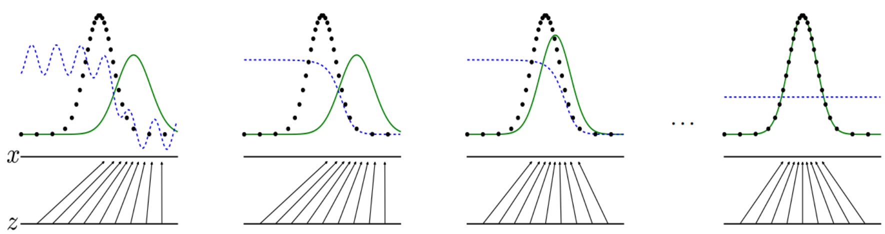

Figure 3: Theoretical evolution of GAN training[6], as the Generative distribution (green line) approaches the real data distribution, the Discriminative distribution (dotted blue line) will be unable to distinguish them and stabilizes at $D(\mathbf{x})$ = ½

The Discriminator (D) and Generator (G) networks are trained separately, by different alternating processes.

Generator - G
Sample from random noise z -> new sample is produced in G -> D classifies the sample as "Real" or "Fake" -> Loss calculated from D classification -> Backpropagate through Discriminator and Generator to obtain gradients -> Gradients used to change G weights

Discriminator - D
D classifies real data and G fake data -> Discriminator loss penalizes G for misclassifying real and fake instances -> D updates its weights through backpropagation from the Discriminator loss through the network D.

Adversarial training
Following the description above, both models are competing against each other in a way called adversarial in game theory, and they are playing a zero-sum game. This means that when the Discriminator successfully identifies a sample, it is rewarded or no update is done to the model parameters, whereas the Generator is penalized with large changes to model parameters. From the other perspective, when the Generator tricks the Discriminator, it is rewarded, or no update is done to its parameters, but the Discriminator is penalized, and its model parameters are changed.

Using the Discriminator to train the Generator
When training a neural net, weights are altered to reduce the error or loss of the output. Generative Adversarial Networks are more complex since the Generator is not directly connected to its loss function. It is the Discriminator that produces

the output that will affect the Generator (Generator loss). Backpropagation then adjusts each weight by calculating the weight's impact on the output.

The impact of a Generator weight depends on the impact of the Discriminator weights it feeds into. Backpropagation starts at the output and flows back through the Discriminator into the Generator. In an optimal case, where both the Discriminator and Generator evolve at the same pace, the Generator would end up generating samples indistinguishable from those drawn from real data, in a process shown in 3.

2.1.6 Challenges

Although capable of generating very accurate synthetic samples[7], GANs are also known[8] to be hard to train. Training two networks simultaneously means that when the parameters of one model are updated, the optimization problem changes. This creates a dynamic system that is harder to control. Non convergence is a common issue in GAN training. Deep models are usually trained using an optimization algorithm that looks for the lowest point of a loss function, but in a two player scenario, instead of reaching an equilibrium, the gradients may conflict and never converge, thus missing the optimum minima. In other words, if the Generator gets too good too fast, it may fool the Discriminator and stop getting meaningful feedback, which in turn will make the Generator train on bad feedback, leading to a collapse in output quality.

Mode collapse is a failure mode of GANs that happens due to a deficiency in training. It can happen when the Generator maps several noise input-values to the same output region, or when the Generator ignores a region of the target data distribution. This means that if the Generator gets stuck in a local minimum generating limited samples, the Discriminator will eventually learn to differentiate the Generator's fakes, ending the learning process and leading to an undiversified output. This is a problem because in generative modelling, the goal is not only to create realistic looking samples, but also to be able to produce a wider variety of samples. Several adaptations of the original model try different adjustments to mitigate these problems.

The ongoing GAN research has shown that when the Discriminator gets too good, the training of the Generator can fail. The reason is that an optimal Discriminator doesn't provide sufficient feedback for the Generator to properly learn. This is called the **vanishing gradient** problem, when the gradient gets so small that in backpropagation it does not change the weight values of the Generator's initial layers, so their learning can get very slow and eventually come to a halt.

2.1.7 Lack of a proper evaluation metric

Another issue concerning the development of GANs, is how to evaluate their training accuracy. Since GANs have been first developed around image generation, the most used evaluation metric is currently the Fréchet Inception Distance (FID). It is based on the Fréchet Distance, a measurement that compares the statistics of two multivariate normal distributions – mean and covariance matrix - to quantify how far apart the distributions are from each other. It uses the features extracted from the imageNet dataset by the Inception v3 network. Thus, limiting this metric to image generating GANs.

2.1.8 What is being done

To solve the main training issues, vanishing gradient and mode collapse, research has developed into two main approaches: The proposal of new network architectures and the use of different loss functions. Recent research[8] shows that the performance of GANs is related to the network and batch sizes, indicating that well designed architectures have critical effect on output quality. The redesign of loss functions, including regularization and normalization has yielded improvements on training stability. It is important to note that improvements are targeted at specific applications, so presently there is not one unique fits-all solution.

3 The evolution of GANs in finance - overview and state of research

The possibility of recreating complex distributions via GANs has captured the attention of the quantitative finance researchers, although GANs most notorious application lies in computer vision and image generation, there is some promising research on the field of data-driven modelling in finance, especially its application to time series analysis and generation.

There is currently a steady development undergoing in the broad field of finance, with several applications being researched for generative adversarial networks. Some of the principal developments, also portrayed in table 1, are in market prediction[9], tuning of trading models[10], portfolio management[11] and optimization [12], synthetic data generation[3] and diverse types of fraud detection[13].

When modelling financial time series, there are several models used and developed over the years, such as the before mentioned ARCH/GARCH models, Black Scholes (1973) and Heston (1993). Until the advent of machine learning, made possible by the ever-increasing available computational power, development of financial models has been slow. Purely data-driven modelling via neural networks and machine learning is a growing sub-field of research[14]. The application of GANs has the potential to improve the modelling of complex and unknown statistical dynamics present in financial data.

3.1 About the need for synthetic data

Among the different applications of GANs being researched in finance, the topic of generation of synthetic datasets deserves awareness. Synthetic data is important because it can be tailor-made for specific uses or conditions where real data may be lacking or unavailable. This can be useful in numerous cases such as:

- Privacy and compliance rules may severely limit data availability and application[15].
- Data is often required in product testing environments and is often limited or unavailable to testers[16].
- Machine learning requires large amounts of training data, such data can be expensive and scarce[17].

The growth of data-driven processes and the new possibilities provided by this kind of analytics and modelling created a higher demand for data and data scientists in finance. Financial datasets are often very regulated, which limits their use in developing new products and processes. As described by JP Morgans' AI Research[18], anonymization is unreliable, and encryption can cripple the use cases of data. Statistically consistent synthetic financial data can solve most of these limitations, producing high-volume artificial test data also greatly improves scalability and cooperative work with usually sensitive or limited datasets.

Table 1: GANs in finance research

Field	Application	Method
Time Series Forecasting	Market Prediction	GAN-FD [9], ST-GAN [19], MTSGAN [20]
	Fine-Tuning of trading models	C-GAN [10], MAS-GAN[21]
Portfolio Management	Porfolio Optimization	PAGAN[11], GAN-MP[22], DAT-CGAN[23], CorrGAN[12]
Time Series Generation	Synthetic time series generation and Finance Data Augmentation	TimeGAN[24], WGAN-GP[25], FIN-GAN[3], Quant GAN[14], RA-GAN[26], CDRAGAN[27], SigCWGAN[28], ST-GAN[19]
Fraud Detection	Detection of market manipulation	LSTM-GAN[13]
	Detection of Credit Card Fraud	RWGAN[29], LSTM-GAN-2[30]

3.2 Main GANs in financial research

There have been several GAN variants proposed in the literature to improve performance, these can mainly be divided into two types: Architecture and Loss variants. In the Architecture variants, structural changes were made to adapt the GAN to a certain purpose, or to improve overall performance. In Loss variants, different approaches to Loss functions try to improve stability and performance while training, often trying to solve the issue of non-convergence. Modifications have been made to tailor each network to its specific goal and used data type. Overall, the main topic of GAN research is and remains centred around image generation and computer vision. Even so, based on the continuing output of time series and finance applied models it is clear that GANs are helping to expand the field of research. There have been some milestone papers which will be discussed in the next segment.

3.2.1 FIN-GAN - 2019

FIN-GAN[3] is a proposed application of the original GAN with the goal of generating synthetic time series that replicate the main stylized facts of financial time series. There are no relevant structural changes to the first proposed framework. For the architecture of D and G, three structures were proposed, 1- MLP (multi layer perceptron) architecture with four layers of fully connected neural networks with the hyperbolic tangent activation function. 2- CNNs (convolutional neural networks) architecture with six convolutional layers and the leaky ReLU activation function. 3- MLP-CNN which combines CNN and MLP by the element-wise multiplication of their outputs.

The algorithm remains generally unchanged from the original proposition, with minimax loss and the Adam optimizer for updating D and G parameters. The main findings were that changes on the Generator's neural network architecture have a greater effect on the quality of output then changes on the Discriminator. The use of batch normalization (rescaling hidden vectors to keep mean and variance consistent during training) showed a large fluctuation in the generated time series and a strong autocorrelation. An indication that this common tool for image generation in deep learning is not ideal for modelling of financial time series. The output is set to a large value because the length of the time-series should be long enough to reliably calculate the statistical properties such as the probability distribution and the correlation functions.

FIN-GAN applied to finance
FIN-GAN can generate financial time series that are deemed realistic, due to the presence of the major stylized facts, showing the characteristics that are intrinsic to time series found in the real world. The approach marks an important point for GANs in finance, since it is a model that learns the properties of data without requiring explicit assumptions or mathematical formulations, something that stochastic process and agent-based modelling cannot do without non-trivial assumptions.

3.2.2 Conditional GAN (cGAN) - 2019

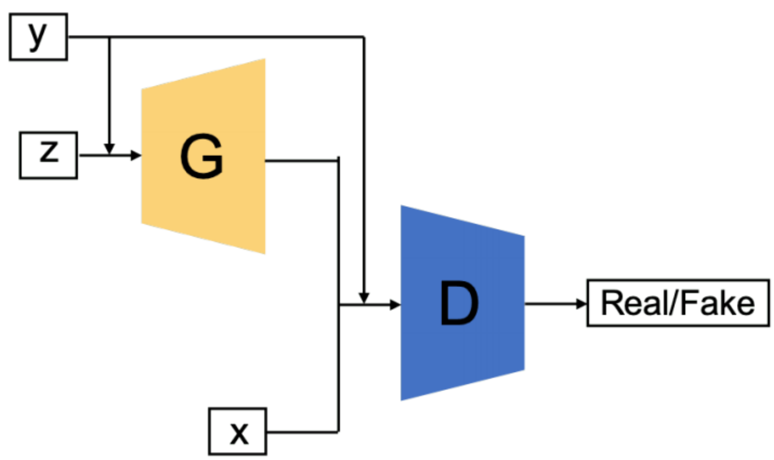

Figure 4: Conditional-GAN architecture[8]

The Conditional GAN (cGAN) was first introduced in 2015[31], and later adapted to finance by Koshiyama [10] 2019. This GAN's architecture has been structurally extended to a conditional model (figure 4) where the Generator and Discriminator are conditioned on some extra information y. This auxiliary information y could take the form of class labels or other data. Which is then fed into both D and G as additional input layers, in a process called conditioning.

The added input layer of one-hot-encoded labels guides the Generator to produce specific outputs. This level of control does not exist in the originally proposed GAN architecture. With the benefit from the additional encoded information, cGAN can also be trained on multimodal datasets that contain labelled data.

Like the original GAN, the Discriminator and Generator networks are MLP's where the Generator output layer activation is a linear function and the Discriminator output layer activation is a sigmoid function. The loss function has an added conditional element for both D and G, and they both are optimized via stochastic gradient descent.

Conditional GAN applied to finance
Conditioning information in finance: added vector y can represent a current/expected market condition, appropriate for modelling sequential data such as time series. It also enables the construction of "what-if" scenarios, commonly used for stress tests and other scenario-based models. The implemented cGAN has shown some new tools to improve trading strategies, specifically for finding optimal hyperparameters (fine-tuning) of trading algorithms, and combination of trading strategies, where trading models are trained on samples generated by the cGAN, and their outputs combined to get a final result.

Compared to the classical method of bootstrapping (resampling), the cGAN can generate more diverse training and testing sets. There is also the capability of drawing samples targeting stress events and an added layer of anonymization to the dataset, not achievable on other shuffling and resampling techniques.

3.2.3 WGAN-GP - 2019

The first Wasserstein GAN (WGAN[32]) proposed a new cost function using Wasserstein distance as an alternative to the original GAN Loss function, which is based on the KL and JS distances. The Wasserstein distance quantifies the distance between two probability distributions, it is also called Earth Mover's distance because it can be interpreted as the minimum energy cost of reshaping one probability distribution into another via stepwise (discrete) increments, analogous to transporting blocks of dirt.

This modification attempts to solve the problem of vanishing gradient and mode collapse which happened with BCE Loss GANs: Often when the Discriminator started to improve faster than the Generator, it would start outputting more extreme values, giving less meaningful feedback to the Generator, stopping the learning process. With W-Loss the Discriminator is replaced by a Critic. The Critic replaces the sigmoid with a linear activation function, this way the output is not limited between 0 and 1 and the cost function continues to grow, regardless of how far apart the distributions are. This way, the gradient won't approach zero, making the GAN less prone to vanishing gradient problem, and consequently, mode collapse.

There is a condition however, when using W-Loss, the Critic NN needs to be 1-Lipschitz continuous. Meaning that the norm of the gradient should be at most one at every point, this condition assures that the W-Loss function is continuous (differentiable) and that its growth is stable during training, this makes the underlying earth mover distance valid. To enforce this condition WGAN uses weight clipping, where the weights of the Critic are forced to take values between a fixed interval. After the weights are updated during gradient descent, any weights outside of the desired interval are clipped, so values that are too high or too low are set to some fixed maximum value. Forcing the weights to a limited range of values could limit the critics ability to learn and ultimately for the GAN to perform. This causes the need for a lot of hyperparameter tuning to adjust the model for training.

In WGAN-GP[33], another way to enforce the 1-L continuity was introduced, Gradient penalty or GP, which is a much softer way to enforce this continuity. With GP a regularization term is added to the loss function, it penalizes the critic when its gradient norm is higher than 1. The regularization term is achieved by getting a gradient on the interpolation between samples from the real and generated distributions.

FIN-GAN applied to finance
The finance application of WGAN-GP[25] for the one-dimensional case proposes a Wasserstein's GAN with Gradient Penalty with 1-dimensional convolutional networks to work on time series data. It compared the relevant statistics of generated versus original time series and proposed taking a rolling window before feeding the original series to the model, aiming to induce more variability scenarios. In the results, the generated data was largely able to replicate the stylized facts of the S&P 500 while also retaining some visual similarity with visible volatility clusters.

3.2.4 Corr-GAN - 2019

CorrGAN[12] is a novel approach to generate empirical correlation matrices estimated on asset returns. This GAN is based on the DCGAN[34] architecture.

DCGAN is a milestone GAN, being the first to apply deconvolutional neural networks in the Generator (figure 5). Some critical changes to the architecture of DCGAN compared to original GAN, which enabled higher resolution modelling and stabilized training. It also introduced the learning of hierarchies of representations in natural images. Deconvolutions are also known as transposed convolutions, they work by swapping the forward and the backward passes of a convolution. There are several key aspects of DCGAN:

• The DCGAN structure replaces pooling layers with strided convolutions for Discriminator and fractional-strided convolutions for Generators.

• Batch normalization in both the Discriminator and the Generator, improving location of generated real samples that center at zero.
• A ReLU activation is used in Generator for all layers except the output, which uses Tanh, while LeakyReLU activation is used in the Discriminator for all layers. The LeakyReLU activation prevents the network from being stuck in a "dying state" situation as the Generator receives gradients from the Discriminator.

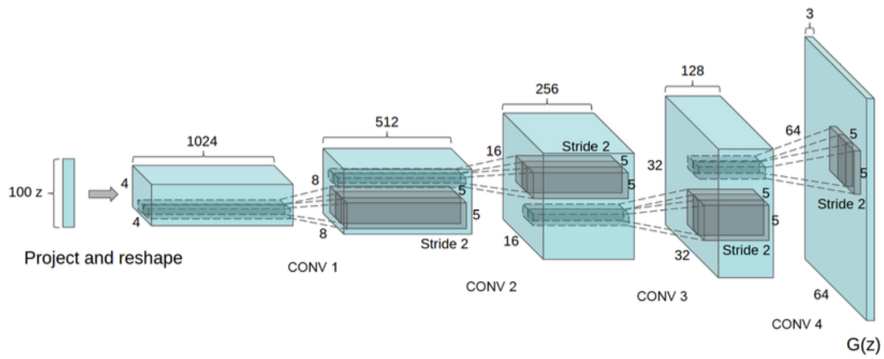

Figure 5: Generator layout of DCGAN[34]

CorrGAN applied to finance
The DCGAN architecture was originally developed to improve the quality and resolution of synthetic images. The CorrGAN paper showed an implementation of DCGAN on empirical correlation matrices estimated on S&P 500 returns. The process showed convincing results, although the generated matrices are not exactly correlation matrices (their diagonal is very close but note equal to 1), the major stylized facts were reproduced in the synthetic data. In short, the main stylized facts of financial correlation matrices are the distribution of pairwise correlations being positively shifted, the first eigenvector having positive entries and the corresponding Minimum Spanning Tree being scale free.

Practical financial applications of the CorrGAN framework could range from improving trading strategies to risk and portfolio stress testing. Some suggested applications include improving Monte Carlo backtesting and stress testing portfolios by conditioning on market regime variables of different macroeconomic scenarios. The researchers developed a website[35] to test if people can visually distinguish real from fake correlation matrices, currently they are indistinguishable to the naked eye.

3.2.5 QuantGAN - 2020

Introduced in 2019, and one of the first solid works applying GANs to financial time series generation, QuantGAN[14] consists of a GAN variation which utilizes temporal convolutional networks (TCNs) aiming at capturing long-range dependencies like volatility clusters. The objective was to approximate a realistic asset price simulator by using a neural network, data-driven concept.

The TCN are convolutional frameworks that provide a unified approach to capture low and high level information in a single model. TCNs consist of a causal 1D fully convolutional network architecture, where an output depends only on past sequence elements. Another key characteristic are more dilated convolutions. Dilation refers to the distance between elements of the input sequence that are used to compute one element of the output sequence, this means that instead of increasing the size of the kernel/filter, it introduces empty spaces for more coverage, increasing the receptive field and giving a broader view, with more context information about the input.

Empirical results suggest that TCNs are better at capturing long-range dependencies in time series than other convolutional methods. TCNs have the advantage of having more controllable gradients compared to Recurrent Neural Networks, simplifying the optimization process.

QuantGAN applied to finance
QuantGAN uses TCNs as Discriminator and Generator networks. It was trained on nine years of S&P 500 data and used the Wasserstein distance as a distributional evaluation metric. As dependence scores, it used the ACF and leverage effect scores to assess accuracy based on the stylized facts. The result was that QuantGAN generated returns closely

matched the real returns, this was corroborated by the sharp drops in the ACF and the negative correlation between squared and non-squared log returns at short time lags.

Finally, the authors evaluate that the proposed architecture generated competitive results, which can be used to approximate financial time series. They point to the lack of a unified metric to quantify the goodness of fit of two datasets, but overall the findings were a solid step in developing data-driven models in finance.

3.2.6 MAS-GAN - 2021

This new framework, developed by JP Morgan's AI Research team in a yet to be published paper[21], is a two-step method for multi-agent market-simulator calibration. It uses a GAN Discriminator to calibrate a market simulator constructed using an agent-based approach, first a calibration objective is learned with a GAN Discriminator, then the learnt calibration objective is used to simulate parameter optimization. It is the first method to use a GAN trained Discriminator as an objective function for multi-agent system optimization.

This work builds on SAGAN[36], an architectural variant developed in 2018 that introduces a self-attention layer to a Convolutional GAN. Using Convolutional Neural Networks CNNs, the Self-Attention method was designed to create a balance between efficiency and long-range dependencies (large receptive fields) by using the attention mechanism, well known from the Natural Language Processing (NLP) field of research. In a self-attention layer visualized in 6, the feature map is passed through three 1x1 convolutions, called Query, Key and Value, where the vectors generated by Query and Key multiply to create a vector of probabilities that decides how much of Value to expose to the next layer. In other words, the Q x V multiplications' output is passed through a softmax activation function that creates a so-called attention map, which is multiplied by the Value vector to create the self-attention feature map. This layer is complementary to regular convolutions, aides the network in capturing fine features from an input source.

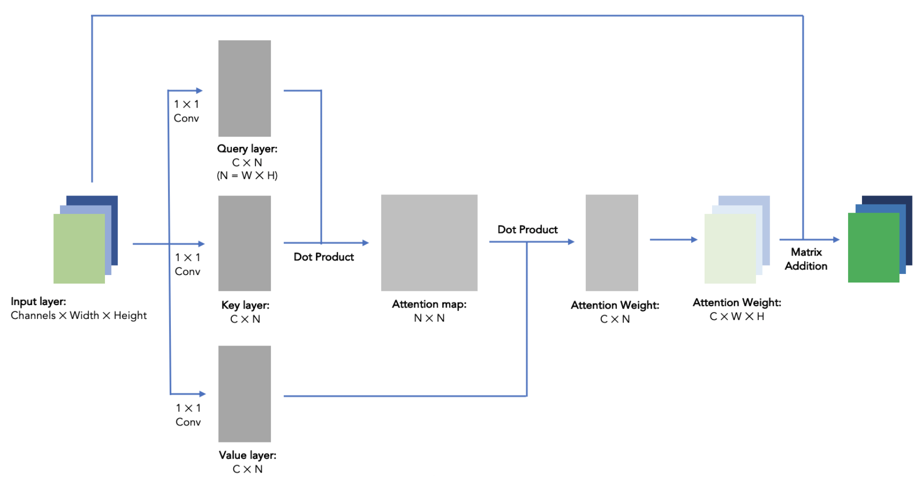

Figure 6: Self-Attention Module[37]

MAS-GAN applied to finance
The MAS-GAN model consists of a SAGAN trained on one-minute mid-price returns and cumulative one-minute traded volumes of thirty stocks traded on NASDAQ in June 2019. After the training process is done, the Discriminator is used to optimize a Discrete event model that simulates a Limit Order Book via three distinct agent behaviours. The objective is to find the most realistic agent configuration through calibration, namely, to find which configurations on a rectangular grid that can produce the simulation closest to the historical dataset. The method has shown good calibration performance for multi-market simulations, based on stylized facts and the Kolmogorov-Smirnov test. Further research into understanding behavioural explanations of the market is suggested, but not present in the paper.

3.3 Current GAN applications in the financial industry

As shown in the previous section, The topic of Generative Adversarial Nets is progression in different finance applications, but the extent of its use in practice by companies is presently limited or not fully disclosed. This section provides a short view of the publicly known implementations of GANs.

3.3.1 Generation of Synthetic data by JP Morgan AI Research

With the goal of advancing AI research and development in financial services, JP Morgan's AI Research department has a branch dedicated to generating synthetic datasets. These datasets can be requested by other research groups and comprise of 1- customer related datasets for Anti Money Laundering models, 2- customer journey events, lower level client-bank interaction dataset, 3- market execution data: limit order book data describing matches of buy and sell orders of financial instruments at a public stock exchange. 4- payment data for fraud detection: several transaction types with legitimate and abnormal activities to improve detection.

The research department proposes a framework[18] for ideal representation and transference of synthetic data. The framework suggests: 1- Privacy preserving, the specific data and context where privacy needs to be enforced. 2- Human readability, the data and its associated generative models must be readily interpretable by regulators and other agents for the sake of transparency. 3- Compactness, the representation of synthetic data should be compact and reconstructible, it should require little technical know how as to improve synthesizing it in different environments.

All the specified data is produced in-house, there is no specific information on the generating process, but there is a paper under review by JP Morgan researchers where the MAS-GAN[21] is proposed for multi-agent simulation, where the Generator is used to calibrate an agent based model as described earlier in this paper.

3.3.2 The Digital Sandbox Pilot by the FCA

The Financial Conduct Authority is a financial regulatory body from the United Kingdom. To boost innovation, they created an environment for testing of financial models, products, and services. The Digital Sandbox[38] provides an integrated, collaborative development environment for testing and scaling projects, aiming to reproduce real scenarios and perform stress tests. The initiative is in pilot stage and has already had a test run with 28 groups presenting solutions on the topics of: access to finance for SMEs, improving the financial resilience of vulnerable consumers and fraud/scam detection.

The whole Digital Sandbox heavily relies on synthetic data, as real is under strict obligations in the UK. As described in the first report[16] synthetic financial data was commissioned to leading data scientists from industry and academia. There was a two-group effort, where one group would create the synthetic data and another who could define typologies and behaviours these data was expected to have. The main approaches to this task used GANs and ABMs and the whole process was benchmarked by the Alan Turing Institute. Access to synthetic data was ranked as the most important feature of the sandbox, however, not all data was considered useful, showing that there is demand for high quality synthetic data.

The main takeaways from the first pilot were: A digital testing environment is in high demand, particularly by startups. It accelerated product development for the participating firms. The access to good synthetic data was considered extremely valuable by all participants. A second digital sandbox is planned with an expanded testing system. This promising new platform has shown potential and could be implemented in other countries if it further succeeds in the United Kingdom.

3.3.3 Fraud detection by American Express AI Labs

Fraud losses, majorly from wire transfers and credit/debit cards caused an estimated loss of 16.9 billion USD to banks, merchants, and customers[39]. Companies use their customer data to train models to predict and prevent fraud. A known issue of fraud detection on real datasets is the class imbalance: Often a dataset can represent a biased sample from reality and will inevitably lead to faulty models. Financial services multinational American Express Co. has its AI lab looking for solutions to this issue by generating synthetic data to improve their fraud detection models. Amex researchers published a paper[27] where a hybrid of Conditional and Deep Regret Analytic GANs was proposed to generate synthetic datasets.

Three tabular datasets with internal company data were used to recreate statistically similar samples. The generated data was evaluated by comparing characteristics of the real and generated data distributions, and also by the internally developed tool DataQC[40], which uses well known methods to look for anomalies in datasets and outputs a unified score of attribute anomaly levels. The generated data showed satisfactory results, but it was found that models trained on synthetic data still performed worse than those trained on real data. The research team states that further research into generating synthetic data is being done, the capacity on which the generated data is being used internally was not disclosed.

4 Application of GANs for time series generation

4.1 Experimental setup, a black-box approach

To test the performance of GANs from a data-driven perspective, a black-box framework was used to build a proof of concept model. The idea behind the black box approach is to test whether a model would show good performance with minimal user input, avoiding the complex task of tuning the model, this would greatly improve its usability by people that may not be specifically trained in deep learning or GANs, broadening the range of its application. Another reason for this approach is to evaluate the originally proposed GANs on their out-of-the-box performance on financial data, this excludes the task of tuning GANs from the scope of this project. In the experiment, the goal is to test the performance of selected GANs in capturing the structure of financial time series and generate a diverse set of output scenarios.

Setting and implementation
The framework[41] was imported to Google Colab, a hosted Jupyter notebook running the programming language python, that provides free access to computing engines, namely Google's cloud-based CPUs and GPUs. This simplifies the testing/reproducibility in collaborative environments and eliminates the need for local GPUs.

The framework consists of 12 different GANs, implemented mirroring their original proposition. The user has a choice of one GAN to be trained at a time. The framework was extended by adding a training section for inputting real time series data, since the original code was designed around image generation, the input had to be reshaped to accept uni-dimensional time series. The added modifications allow for time series to be imported from Yahoo finance, transformed to log-returns, and adjusted to fit the dimensions expected by the model. The data analysis was centred around a series of evaluation plots that were implemented to test the data generated by the model for its similarity to some of the known key statistical characteristics -stylized facts- of financial time series. The models are trained using log returns, expecting outputs with similar characteristics. The models are trained with around 5000 data points and for the sake of comparison generate outputs of the same size.

From the pool of available GANs three were used: DCGAN[34], SAGAN[36], WGAN-GP[33], the choice was made due to the differences in architecture and loss functions among the available options, and their implementation to finance data in the current literature. The used GANs follow the originally proposed structure and parameters, which were already described in the previous section.

• **DCGAN** was selected because it is a milestone GAN, the first work to apply a deconvolutional neural network architecture for the Generator which showed a bump in resolution to the generated images.
• **WGAN-GP** has shown to be successful in improving training stability by providing an alternative loss function with the use of Wasserstein's loss.
• **SAGAN** is an alternative GAN based on convolutional neural networks, with an added self-attention mechanism that improves learning on long-range dependencies for images, achieving great performance on multi-class image generation.

These GANs were chosen due to their successful past implementation on financial data. However, they were not specifically tuned for time-series generation, and their performance is therefore not comparable to papers with specific time-series implementations.

4.1.1 The S&P 500 and synthetic data generation

Instead of using prices, most studies use returns of assets. The main reason is that prices are usually non-stationary (their mean and variance change over time). Using returns makes the time series stationary, greatly facilitating the process of statistical modelling. Returns also provide a scale free summary of the investment, further improving its flexibility for analysis.

In the experiment, the dataset used to train the models, consists of a 20 year cut of the S&P 500 index (from 01.01.2001 to 01.01.2021, roughly 5000 data points). The S&P 500 is a weighted stock market index that combines 500 of the largest companies listed on US stock exchanges. It was used as it condenses the behaviour of the american economy over time to a single measurement, and is largely used to test financial time series models.

The raw S&P 500 adjusted close values were transformed to log returns for model training and later evaluation. There was no single measure of goodness of fit used, rather a qualitative analysis over the statistical properties of the generated time series where the selected stylized facts described in the first section of this paper were used as reference for performance, which in this case represents how close the stylized facts of generated returns are from those commonly found in real-world returns.

4.2 Statistical analysis of time series

In testing the three selected GANs, the main goal is to test whether the models can successfully capture the structure inherent to financial time series and generate scenarios that can be considered diverse. The main issue is the lack of a unified metric or consensus on how to quantitatively evaluate the synthetic samples. GANs do not provide an explicit representation of the generated probability distribution, making it difficult to estimate their likelihood, thus making the task of evaluating them more difficult.

So, the process of evaluation heavily relies on the key selected stylized facts, for which each model is qualitatively assessed based mostly on evaluation plots. By illustrating these key statistics, it's possible to assess some level of performance of the models. This following subsection contains the data analysis performed on the generated returns, which will give a summary of the statistics mainly concerning the distribution of the synthetic returns as well as some time-dependent properties in order to check for the similarity to the real returns.

4.2.1 Analysis of returns

By looking at the log-returns plot from the S&P 500 (figure 7), it is possible to recognize some key patterns. Firstly, the limit in which the returns fluctuate, centered at zero and staying between -10% and 10%. Second, the clusters of periods with high volatility and low volatility are usually bundled together, as asserted by the stylized fact of **volatility clustering**. This pattern is visually identifiable, and it is expected that a good synthetic sample resemble the original data by repeating this patterns of clusters.

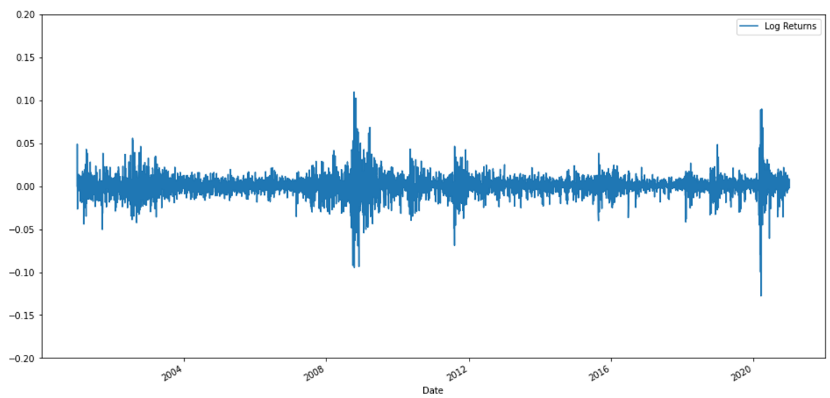

Figure 7: Log-returns of SP500 - It is possible to visually identify periods of high and low volatility

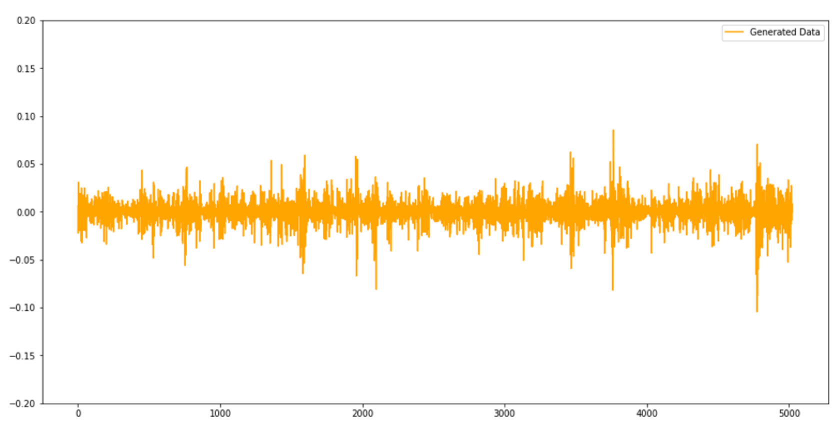

Figure 8: Returns generated by WGAN-GP

Looking at the returns generated by WGAN-GP (figure 8), it is possible to see some similarities to the real returns from S&P 500. The generated returns fluctuate in the same range as the real returns, also centered at zero, and there is some clustering going on. The issue is that the models struggled to recreate realistic clusters, as the output from WGAN-GP is clearly more homogeneous then the real data, the clusters are less pronounced.

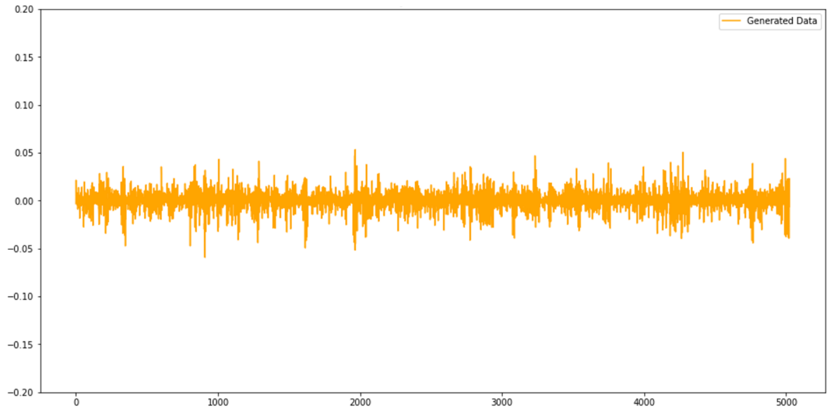

Figure 9: Returns generated by DCGAN

In the returns generated by DCGAN (figure 9), the result was less convincing than the previous models. The centering around zero is present, but there are almost no identifiable clusters, and the synthetic returns did not reproduce the larger values giving the plot a homogeneous aspect that only loosely resembles a real set of log-returns.

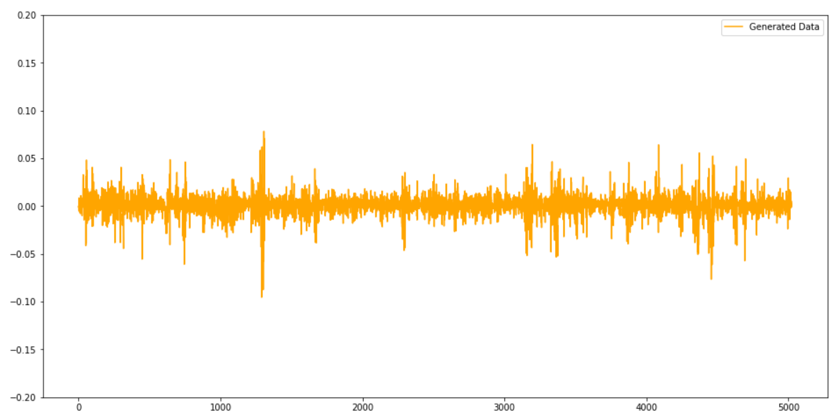

Figure 10: Returns generated by SAGAN

In the SAGAN generated returns (figure 10), the clusters are identifiable, the range and centering is correct. The issue is that the model failed to recreate realistic clusters. The synthetic clusters are too thin, diverging from their real counterparts where more activity is present around large movements.

4.2.2 Probability distribution, Kurtosis and Skewness

Here the generated and real returns were displayed to analyze their distribution and its characteristic skewness and kurtosis measurements. For a GAN to be qualified as good, it is expected that the synthetic returns distribution resemble the shape of the normal distribution (**aggregational gaussianity**) and the values for skewness and kurtosis should be close to those from real returns.

•*Skewness* is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. Here it portrays the **gain/loss asymmetry** of log returns, where it is expected that the skewness be negative, since it is more common to have large downwards movements in stock prices then the contrary.
•*Kurtosis* is a measure that defines how much the tails of a given distribution differ from the tails of a normal distribution. It identifies whether there are extreme values in the tails of a distribution. For returns, large values for kurtosis (**heavy tails**) are expected, since some low probability events have a large impact on the distribution. This behaviour does not happen in a normal distribution, so returns are expected to have the rough shape of a normal distribution, but with much heavier tails.

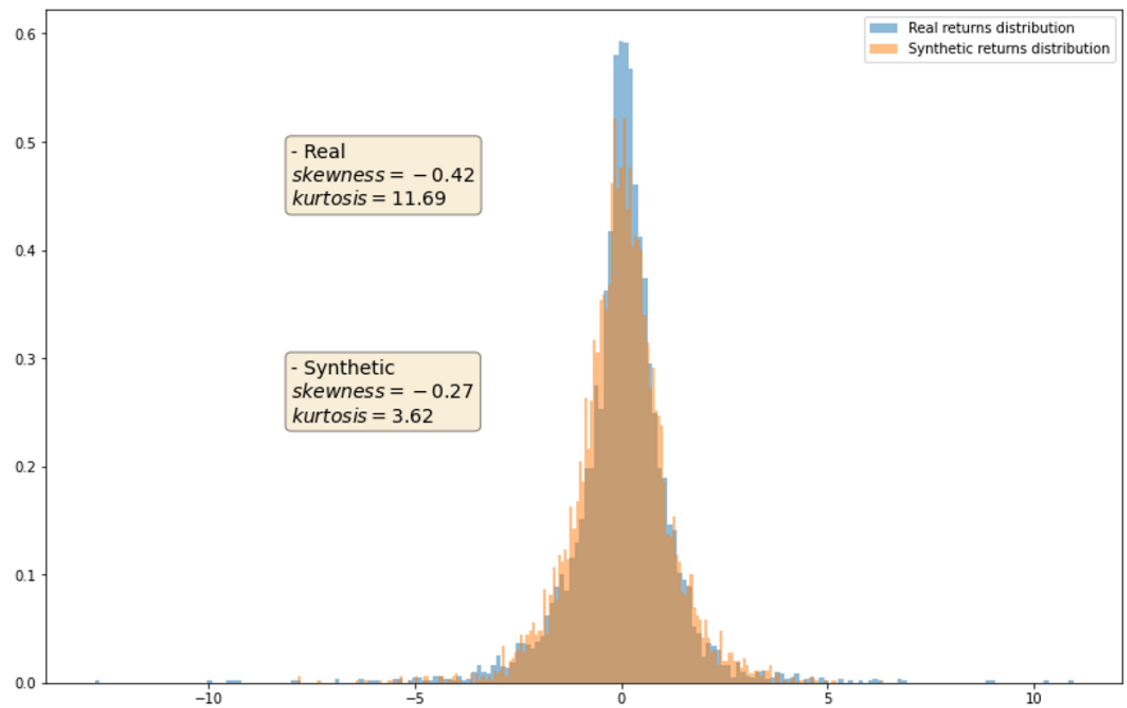

Figure 11: PDF - WGAN-GP

In WGAN-GP's distribution (figure 11), the skewness is negative and close to the value from the real samples. Kurtosis is also positive, but the overall distributions do not share the exact same format.

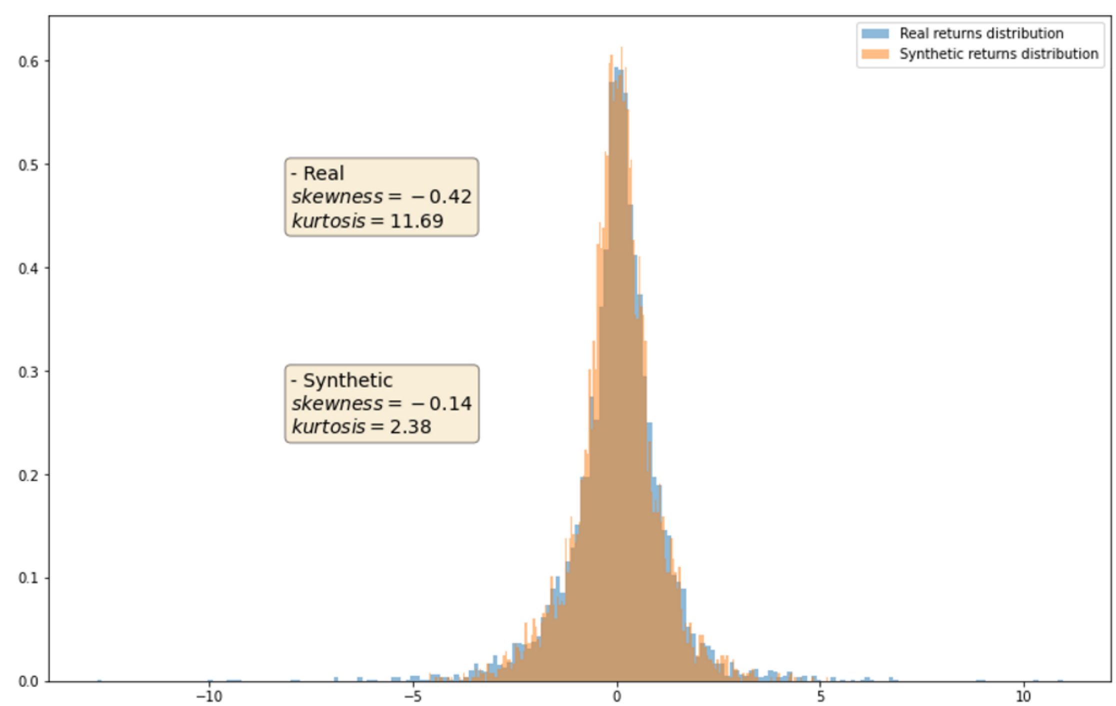

Figure 12: PDF - DCGAN

DCGAN (figure 12) and SAGAN (figure 13) show a good fit, getting close to the distribution observed in the S&P 500 dataset. The are negative and relatively close, and kurtosis is postive, although here SAGAN displays heavier tails, closer to real data. In summary, all three models produce a good skewness but fail to recreate the intensity of heavy tails present in the real distribution.

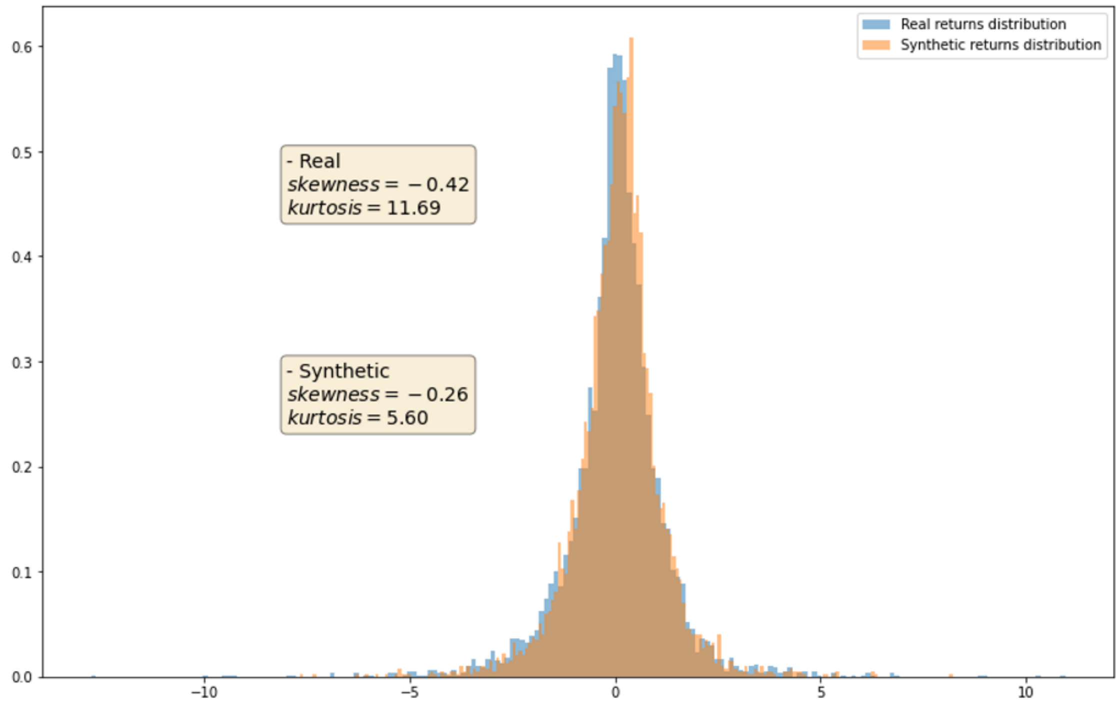

Figure 13: PDF - SAGAN

4.2.3 Autocorrelation

The auto-correlation plots (ACF) show the similarity between observations as a function of time lags between them, for financial returns the autocorrelation it is expected to be very low, given the stylized fact of **linear unpredictability**.

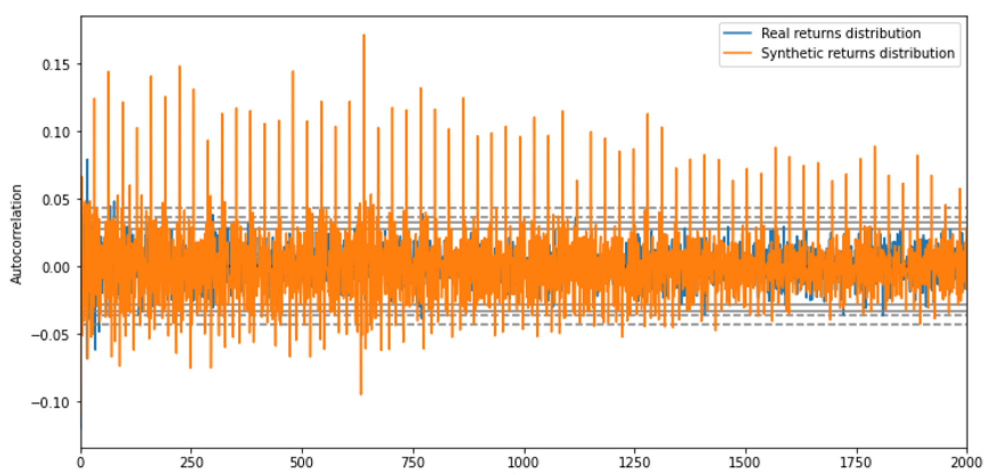

Figure 14: ACF - WGAN-GP

Here its clear that WGAN-GP (figure 14) fails to reproduce the expected linear unpredictability, as the auto-correlation is above the threshold.

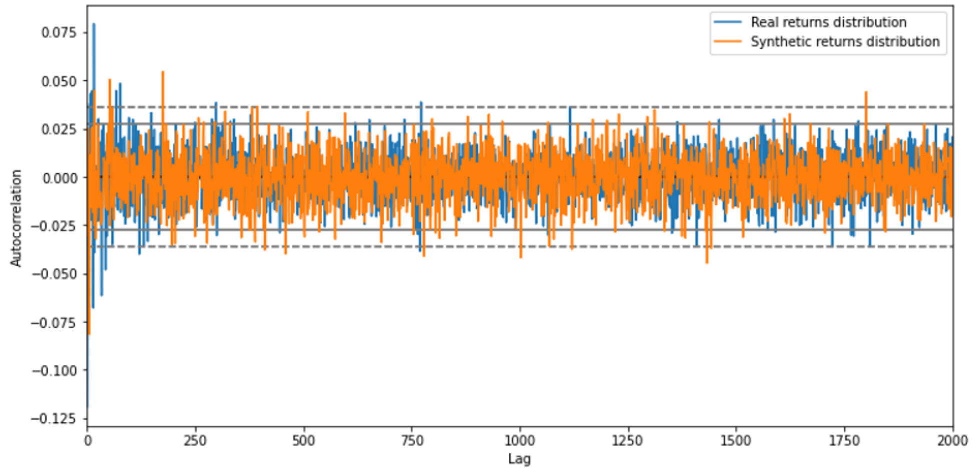

Figure 15: ACF - DCGAN

DCGAN's synthetic samples (figure 15) display a neglectful auto-correlation, precisely mimicking the behaviour of real returns.

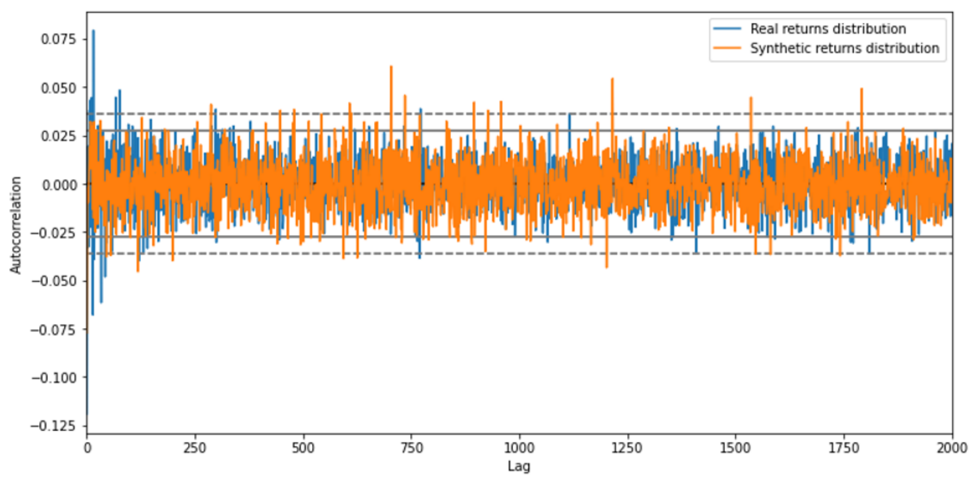

Figure 16: ACF - SAGAN

The auto-correlations in SAGAN's generated returns (figure 16) peek a little bit above the threshold, but overall its a good fit and will be considered satisfactory for this experiment.

4.2.4 Transformed prices

Transforming the generated returns back to prices, for the sake of visualization, showed a moderate amount of diversity to each GAN, with WGAN-GP (figure 17) and DCGAN (figure 18) showing realistic looking series, inside the scale of the SP500 index, and SAGAN (figure 19) reproducing a good shape but completely missing a realistic price interval.

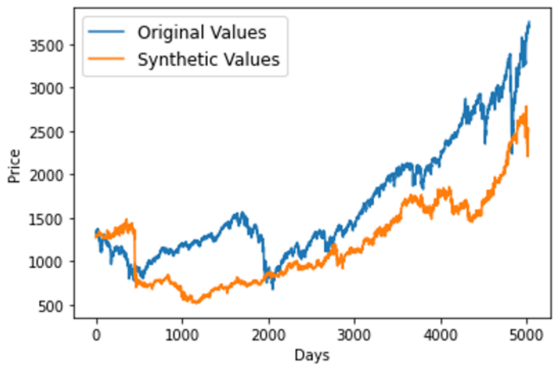

Figure 17: Prices - WGAN-GP

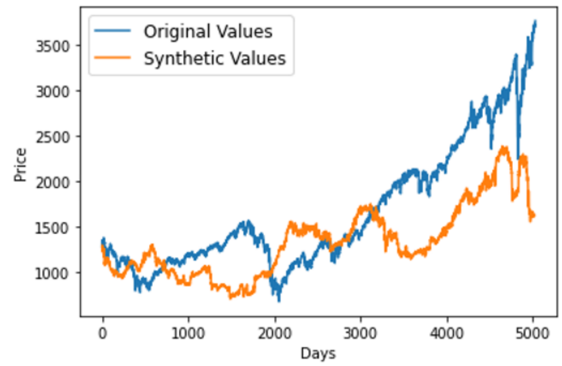

Figure 18: Prices - DCGAN

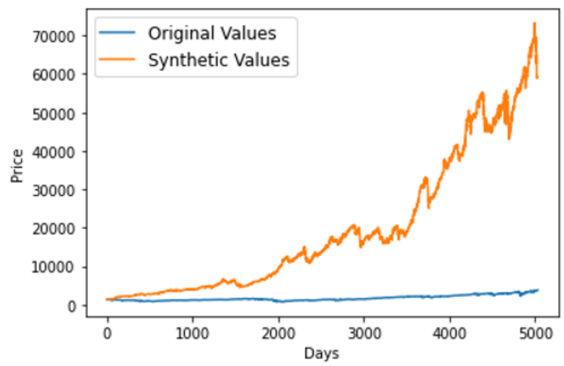

Figure 19: Prices - SAGAN

4.3 Results

Although referred to as qualitative statistics, the stylized facts and their proxy statistics can give a good perspective on the performance of the models trained in the experiment section. It was clearly possible to distinguish a best out of the box model among those tested. DCGAN performed surprisingly well considering that its original purpose was image generation it was able to capture mostly all the tested statistical properties of financial time series. Considering that the model was not tuned for this specific case, it shows that with further research into tuning and architectures, GANs have a strong potential for time series modelling.

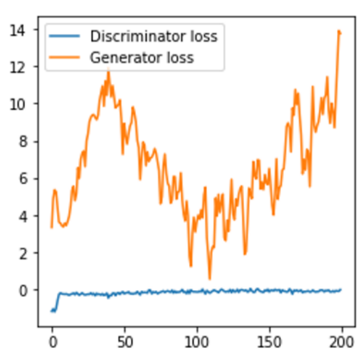

Figure 20: Loss - WGAN-GP

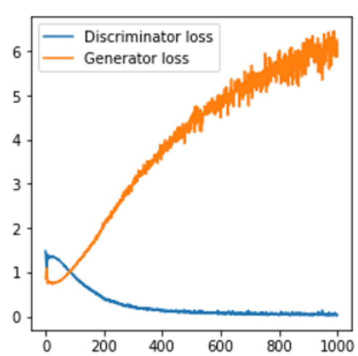

Figure 21: Loss - DCGAN

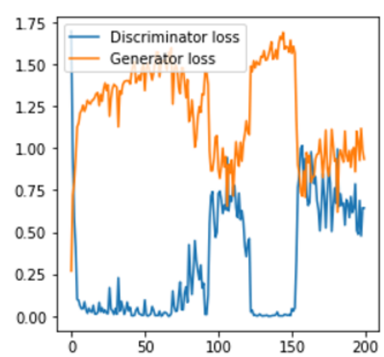

Figure 22: Loss - SAGAN

It was possible to observe some form of mode collapse in all models, it is not possible to quantify this failure mode, but visually each batch of generated data had some visual similarities in the shape of returns and prices, the produced samples did not achieve a large variety of scenarios. The issue of loss convergence was also present during training, where SAGAN notably failed to settle even after 1000 epochs, and in WGAN-GP / DCGAN the losses appeared to settle. It is important to note that this does not mean that a model that converges performs better, or that its best output is at the point of convergence. In fact, for WGAN-GP (figure 20) and SAGAN (figure 22), the model performed better before the losses appeared to settle, and only DCGAN (figure 21) performed best at the max setting of 1000 epochs.

The training was limited to 1000 epochs at fixed learning rates. With more computational power available, longer training epochs could potentially lead to better results. For a next step in these tests, more parameter tuning and better computational power would be advised to potentially get better results.

5 Conclusions - Outlook for the future of GANs in finance

5.1 Analysis of results

This paper had the goal of presenting an overview of Generative Adversarial Networks in finance, describing what they are, why they are being used in finance, followed by some key developments in research and practical applications. To test the versatility of different frameworks out of the box, the objective was to test three notorious image generation GANs on time series.

The importance and wide application of synthetic data in finance was shown as well as some of the challenges faced in the task of generating it. An overview of the finance GANs in literature was presented, along with an insight on some milestone papers in this field. The active research shows that generating synthetic financial data is achievable and viable with the use of GANs, other novel uses such as the calibration of trading models has also shown positive results. The research has refined the training of GANs with finance data, but it is still a complex task, and further stabilizing training will remain an active research topic. A remaining issue is the lack of a unified quantitative metric to assess the performance on generated financial data. Most literature relies on the reproduction of stylized facts, which is a sound metric, but a more precise score would greatly benefit the development of these models, as previously acknowledged in the literature review.

Although being a novel approach, GANs are making their way to the financial industry. Some applications were shown, specially the use of synthetic data generated by adversarial models is gaining traction. The usefulness of good synthetic data is gaining traction, since it opens the possibility of modelling information that would otherwise be regulated for privacy reasons or too scarce for deep models. It was shown that GANs can generate consistent data for various purposes, including retail banking and market data.

As a proof of concept, three known GANs were tested on time series data to assess their out-of-the-box performance. There were no clear expectations regarding the results, but the models performed well. Using selected stylized facts as qualitative metrics, the synthetic generated returns were close to their real counterparts. Considering that the models where tested out of the box, with no hyperparameter tuning except for epoch length, the results are quite positive and attest for the power of generative adversarial networks. Since the main problem with GANs is in training and financial time series generation is still an ongoing research topic, the various complex aspects of GAN training were out of the scope of the tested models.

5.2 Recommendations

Overall, GANs have a large potential in solving the specific need for synthetic financial data, along with other diverse modelling tasks. Being still a novelty in finance, they are more present in research than in practice, but their potential can increase the capabilities of data-driven deep models, reaching areas where these methods previously couldn't be applied due to limitations in data. The development of better evaluation metrics for sample comparison is a key aspect for the further advance in the field, since providing easier tools would make it easier for use by less qualified personnel. Refinements to improve stability and reliability in the training process are in active research and there are still some improvements to be made before a state-of-the-art model is presented. Taking all of this into account, Generative Adversarial Networks offer a good prospect to the modelling toolbox, adding new possibilities to the challenging world of quantitative finance.

References

[1] Katie Martin. Flash crash — the trading savant who crashed the us stock market. *Financial Times*, May 2020.

[2] Christopher B Barrett. Market analysis methods: are our enriched toolkits well suited to enlivened markets? *American journal of agricultural economics*, 78(3):825–829, 1996.

[3] Shuntaro Takahashi, Yu Chen, and Kumiko Tanaka-Ishii. Modelling financial time-series with generative adversarial networks. *Physica A: Statistical Mechanics and its Applications*, 527, Apr 2019.

[4] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 2001.

[5] Francesco Rundo, Francesca Trenta, Agatino Luigi di Stallo, and Sebastiano Battiato. Machine learning for quantitative finance applications: A survey. *Applied Sciences*, 9(24):5574, 2019.

[6] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *arXiv preprint arXiv:1406.2661*, 2014.

[7] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8110–8119, 2020.

[8] Zhengwei Wang, Qi She, and Tomas E Ward. Generative adversarial networks in computer vision: A survey and taxonomy. *arXiv preprint arXiv:1906.01529*, 2019.

[9] Xingyu Zhou, Zhisong Pan, Guyu Hu, Siqi Tang, and Cheng Zhao. Stock market prediction on high-frequency data using generative adversarial nets. *Mathematical Problems in Engineering*, 2018, 2018.

[10] Adriano Koshiyama, Nick Firoozye, and Philip Treleaven. Generative adversarial networks for financial trading strategies fine-tuning and combination. *arXiv preprint arXiv:1901.01751*, 2019.

[11] Giovanni Mariani, Yada Zhu, Jianbo Li, Florian Scheidegger, Roxana Istrate, Costas Bekas, and A Cristiano I Malossi. Pagan: Portfolio analysis with generative adversarial networks. *arXiv preprint arXiv:1909.10578*, 2019.

[12] Gautier Marti. Corrgan: Sampling realistic financial correlation matrices using generative adversarial networks. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8459–8463. IEEE, 2020.

[13] Teema Leangarun, Poj Tangamchit, and Suttipong Thajchayapong. Stock price manipulation detection using generative adversarial networks. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 2104–2111. IEEE, 2018.

[14] Magnus Wiese, Robert Knobloch, Ralf Korn, and Peter Kretschmer. Quant gans: Deep generation of financial time series. *arXiv:1907.06673v2 [q-fin.MF]*, Dec 2019.

[15] Cem Dilmegani. The ultimate guide to synthetic data in 2021, Jun 2021.

[16] Financial Conduct Authority. *Supporting innovation in financial services: the digital sandbox pilot*. FCA, Apr 2021.

[17] James Montantes. Deep learning in finance: Is this the future of the financial industry?, Jul 2020.

[18] Samuel Assefa. Generating synthetic data in finance: opportunities, challenges and pitfalls. *Challenges and Pitfalls (June 23, 2020)*, 2020.

[19] Pratyush Muthukumar and Jie Zhong. A stochastic time series model for predicting financial trends using nlp. *arXiv preprint arXiv:2102.01290*, 2021.

[20] Weijie Wu, Fang Huang, Yidi Kao, Zhou Chen, and Qi Wu. Prediction method of multiple related time series based on generative adversarial networks. *Information*, 12(2):55, 2021.

[21] Victor Storchan, Tucker Balch, and Svitlana Vyetrenko. Mas-gan: Adversarial calibration of multi-agent market simulators. Paper under review, 2021.

[22] Yerin Kim, Daemook Kang, Mingoo Jeon, and Chungmok Lee. Gan-mp hybrid heuristic algorithm for non-convex portfolio optimization problem. *The Engineering Economist*, Jun 2019.

[23] He Sun, Zhun Deng, Hui Chen, and David C Parkes. Decision-aware conditional gans for time series data. *arXiv preprint arXiv:2009.12682*, 2020.

[24] Jinsung Yoon, Daniel Jarrett, and Mihaela van der Schaar. Time-series generative adversarial networks. *NeurIPS 2019*, 2019.

[25] Fernando De Meer Pardo. Enriching financial datasets with generative adversarial networks, Aug 2019.

[26] Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard gan. *arXiv preprint arXiv:1807.00734*, 2018.

[27] Dmitry Efimov, Di Xu, Luyang Kong, Alexey Nefedov, and Archana Anandakrishnan. Using generative adversarial networks to synthesize artificial financial datasets. *arXiv preprint arXiv:2002.02271*, 2020.

[28] Hao Ni, Lukasz Szpruch, Magnus Wiese, Shujian Liao, and Baoren Xiao. Conditional sig-wasserstein gans for time series generation. *arXiv preprint arXiv:2006.05421*, 2020.

[29] Akhil Sethia, Raj Patel, and Purva Raut. Data augmentation using generative models for credit card fraud detection. In *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, pages 1–6. IEEE, 2018.

[30] Zhaohui Zhang, Lijun Yang, Ligong Chen, Qiuwen Liu, Ying Meng, Pengwei Wang, and Maozhen Li. A generative adversarial network–based method for generating negative financial samples. *International Journal of Distributed Sensor Networks*, 16(2):1550147720907053, 2020.

[31] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.

[32] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR, 2017.

[33] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. Improved training of wasserstein gans. *arXiv preprint arXiv:1704.00028*, 2017.

[34] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

[35] Gautier Marti. corrgan.io, 2019.

[36] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. In *International conference on machine learning*, pages 7354–7363. PMLR, 2019.

[37] Frank Xu. Building your own self-attention gans, Jul 2020.

[38] FCA and City of London. Digital sandbox pilot, 2021.

[39] TJ Horan. Credit card fraud: It's still a thing (and as big as ever), Jan 2021.

[40] David Love, Nalin Aggarwal, Alexander Statnikov, and Chao Yuan. An automated system for data attribute anomaly detection. In *KDD 2017 Workshop on Anomaly Detection in Finance*, pages 95–101. PMLR, 2018.

[41] Florian Eckerli. Generative adversarial nets for finance time series generation. `https://github.com/eckerli/BA_code`, 2021.