Are Fintech Mortgage Loans Riskier?

A study on the risk-impact of fintech lenders in the U.S market

Zhao Zhang

Dissertation presented for MA (Hons) Accounting and Finance 2019-20

DISSERTATION ADVISOR: Dr. Srivastav Abhishek

WORD COUNT: 13747

Table of Contents

Abstract

Financial technology has a growing impact on the mortgage lending market. Fintech lenders have absorbed market share from their traditional counterparts and changed the landscape of the mortgage origination process. However, the risk impact of fintech lending is rarely studied. Providing an insight into the value and risk of fintech features, this dissertation examines the performance of loan-specific risk indicators and default rate among fintech-issued loans from 2010 to 2017, in the U.S fixed-rate mortgage market. In addition, a logit regression analysis is conducted to understand the default risk impact of the fintech status. The result suggests that fintech lenders can perform well in certain aspects of the screening process and the short-term default window, but fail to outperform traditional lenders overall. In addition, fintech status has a significantly negative risk impact in long-term default windows. Although fintech lenders have the merits in terms of efficiency and market friction, the risk associated is more persistent and detrimental. As such, the findings of this dissertation contribute to the relative research gap regarding fintech-issued mortgages. They also provide implications for the overall market regulatory environment and present opportunities for future research.

1. Introduction

Financial technology-based companies have a rapidly increasing impact on the residential mortgage market, and particularly in the origination process. Automation and algorithm-based online origination systems have a competitive advantage in aspects such as client experience, regulatory compliance, efficiency, and processing cycle time (Oliver Wyman, 2016). These factors can transform consumer experience while elevating firms' operational efficiency, which translates to reduced friction costs and agency costs (Fuster et al., 2019). As a result, Fintech mortgage lenders are gaining market share rapidly (Buchak et al., 2018). According to Fuster et al. (2019), fintech lenders have an increased market share from 2% in 2010 to 8% in 2016 in originations. However, it also poses a disruptive force over traditional lending institutions, such as depository-based corporate and local banks. For instance, the stringent regulatory statues such as capital requirements for traditional banks can be circumvented by fintech lenders. While the key advantages can accelerate the growth of fintech lending and change consumer experience for the better, financial disruptions such as the subprime mortgage crisis has taught us the importance to understand the potential defects and risks when transformative business models emerge. Therefore, it is crucial to understand the value and the risk-impact of fintech features.

1.1 Research Gap

A substantial amount of recent literature captures the essence of the business model and technological traits of fintech lending activities (Philippon, 2015; Lee and Shin, 2018; Gomber et al., 2018; Buchak et al., 2018). Moreover, research evidence in credit rating (Braggion et al., 2018), borrower screening (Maggie and Yao, 2018), and lender discrimination (Bartlett et al., 2019) helps us better understand and identify the distinctive features and differences of fintech lenders compared to traditional lenders. On the other hand, residential mortgage default has been extensively studied for many years. Ample empirical evidence exists for identifying mortgage risk by participants such as lender and borrower, and its determinants such as interest rate and home equity (Quercia and Stegman, 1992). However, to the best of my knowledge, few articles have specifically addressed the impact of fintech features on mortgage default risk. In addition, evidence on default rate performance associated with fintech lenders is also scarce. This dissertation aims to fulfill these research gaps.

1.2 Research Objective

This dissertation first examines the differences in the origination process to distinguish fintech lenders from traditional lenders. However, the core focus of the dissertation is the mortgage default performance and the risk impact made by the fintech lenders. To achieve this objective, an overview of the fintech features and business models is first conducted to form a theoretical basis for lender classification. After identifying and separating loans by lender type, it examines the differences and default performance associated with widely used and academically accepted risk measures, such as loan to value ratio and FICO credit score, in each lender group. Most importantly, a regression analysis is employed to understand how fintech status impacts mortgage default. Through both descriptive statistics and logistic regression analysis, this study seeks to have a comprehensive understanding of the loan quality and default risk associated with fintech lenders. Building on existing research of the fintech characteristics and mortgage defaults, it also seeks to explore potential explanations and to provide informed inference on competitive advantages and underlying systemic risks of the fintech lending model. Ultimately, based on the definition of the fintech status, this dissertation explores the potential advantages and defects of the fintech system in the context of mortgage lending, thereby unveiling practical implications at both the firm and market level.

1.3 Structure

Chapter two conducts a literature review on related topics. It first examines the fintech growth and provides background knowledge at the industry and market level. Next, it delves into the evidence on the distinctive features of the fintech lending model, which establishes the basis for the lender classification problem and the interpretation of the fintech status used in empirical tests. This is followed by a review of evidence on mortgage default, which provides an academic foundation for variable selection and model development used in the empirical studies.

Chapter three describes the methodology and research design for the empirical tests. Specifically, issues encountered in the process of data collection and processing are identified and explained. In addition, the approach to the fintech lender classification problem, which is a unique and crucial aspect of this study, is addressed and explained. Next, the limitations section regarding the model and data is discussed. Based on the research approach and evidence explored in the literature review, a hypothesis is established at the end of the chapter.

Chapter four presents and discusses the results and findings. The descriptive statistics section analyses and compares the performance and distribution of basic risk indicators between fintech lenders and traditional lenders. It is followed by the interpretation and inference of the regression results, on coefficients and statistical significance of fintech status, as well as the pattern of changes under different models. A robustness section is also presented to examine the model performance under different conditions and with different parameters.

Lastly, the dissertation concludes on the key contributions and discusses the practical and theoretical implications. A suggestions section is also included for potential future research and testing.

2. Literature Review

2.1. The Evolution of Fintech

The world is no stranger to fintech. In fact, the term was first used in 1972, as an abbreviation for "financial technology, combining bank expertise with modern management science techniques and the computer" (Bettinger, 1972). Patrick Schueffel (2017), in an updated and more comprehensive interpretation, defines fintech as "a new financial industry that applies technology to improve financial services." From online stock trading in the early 1990s to a worldwide fervour for crowdfunding projects in the 2010s, to the evolving dominance of mobile payments such as PayPal and AliPay, innovative applications of technology have been influential and transformative to different aspects of financial services. The financial technology industry now covers an entire array of digital services and solutions that extend to virtually every branch of traditional financial services: payment transactions, insurance, wealth management, banking, lending, and security.

More importantly, the growth of fintech market is now at a revolutionary pace: the fintech market value across 48 fintech unicorns is worth over $187 billion as of the first half of 2019, or slightly over 1% of the global financial industry (CB Insights Research, 2019). Such growth of the fintech industry is only accelerating: Global investment in fintech ventures more than doubled in 2018, amounting to 55.3 billion USD (Accenture, 2019). Dominant markets include the Americas, the Asia Pacific, and Europe are producing hundreds of fintech-oriented start-ups and pouring venture capital into the industry. Statistics in early 2019 suggest that Americas have 5779 fintech start-ups, followed by 3583 in Europe, Africa, and the Middle East, and 2849 in the Asia Pacific (Statista, 2019). The global presence of fintech entrepreneurship indicates a significant trend of innovation, and thus an inevitable process of transformation for the financial services industry in the future.

While developed nations in America and Europe are leading the fintech progress, developing regions of the world have become early adopters of financial technology and seen transformative impacts: China is the global leader in the fintech payment sector. It has demonstrated a rare case where the country transitions from a cash-based economy to a cashless economy without the popularization of credit card payments: its cash withdrawals from ATMs are declining, and third-party payment transactions (e.g., via QR code scan) are forecasted to be at 354 trillion yuan in 2020 (Bloomberg, 2019). Consequently, global fintech giants such as Alibaba are seizing the opportunity and reaching a status of dominance in both the domestic and global stage.

It is evident that the fintech potential is tremendous: Goldman Sachs estimates that fintech may cause a market disruption worth 4.7 trillion USD to the traditional financial services in the future (Goldman Sachs, 2015). Such threat imposed onto the traditional financial service sector can be attributed to the advantages of integrating cutting-edge technology, such as reduced cost of transactions, increased accessibility, and convenience. However, it is crucial to understand that such growth is subject to the restraints of regulations, the government, and the general economic environment. Next, I focus on the business model and characteristics of the fintech lending market, and specifically the mortgage market, to obtain a comprehensive overview of the fintech model in the focused field.

2.2 The Fintech Lending Model

The mortgage market seems to have a slower transition towards digitization in comparison with markets such as consumer lending. Human interaction-based applications and processing are still dominant in the industry for the purpose of risk management. However, we have seen a decline in the function of traditional banks and the rise of shadow banks. In the post-financial crisis period, mortgage originations by shadow banks increased from 30% in 2007 to 50% of the market share in 2015 (Buchak et al., 2018). A growing proportion of non-traditional financial institutes (i.e., shadow banks) are fintech firms with technology-oriented business models. There is compelling evidence for the blossom of the fintech mortgage market: fintech lenders expanded their market share in mortgage lending from 2% in 2010 to 8% in 2016 (Fuster et al., 2019), and that four out of twenty largest mortgage originators are fintech firms (Fannie Mae, 2019). Such growth is not unwarranted: fintech firms have demonstrated crucial competitive advantages over traditional business models, enabling lenders and consumers a more convenient, efficient, and accessible channel to the financing process.

For mortgage lenders, the defining features of the fintech model are an end-to-end online platform and centralized processing, augmented by automation (Fuster et al., 2019). Until traditional lenders adopt fintech-based business models, they will continue to bear the disadvantage of reduced variable costs from its competitors' agile and efficient processes (Fannie Mae, 2019), because such features can generate tremendous value for both mortgage lenders and consumers. For lenders, digitization and implementation of cutting-edge technology can improve crucial functions or provide capabilities in the mortgage origination process, in areas such as data aggregation and synchronization, automated documentation and archiving, and automated compliance checks (Oliver Wyman, 2016), which are heavily focused expenses for mortgage lenders (Fannie Mae, 2019). For instance, data mining technology has considerable applications in the credit checking and approval process of mortgage applications, and we have seen fintech lenders such as SoFi and QuickenLoans adopt a data-oriented process. In contrast, traditional banks are still dominated by a standard and manual approach (Gomber et al., 2018). An integrated fintech model, therefore, would intuitively reduce variable costs, and enhance the efficiency in aspects such as administration and processing speed in the loan origination process. Notably, many fintech lenders securitize mortgages without a depository function or branch network to support their financing (Fuster et al., 2019). In other cases, innovative business models are provided: Loftium provides down payment to borrowers in exchange for letting income (e.g., AirBnB) in the new home with certain expected occupancy rates (Gomber et al., 2018). Although innovative solutions can present risks of liquidity and the absence of safety net, such diversion from traditional funding and business models also enables fintech lenders to circumvent the regulatory pressure such as capital requirements (e.g., requirements to retain loan on the balance sheet), thereby granting firms freedom and manoeuvrability for accelerated growth and expansion. At the same time, without a balanced regulatory environment or adequate policies in place, the market is exposed to risks in factors such as liquidity. As a result, the emergence and accelerated growth of fintech lenders may can pose a systematic risk, and potentially cause detriment to the consumers as well as the market.

Traditional mortgage lenders have identified “consumer-facing technology” and “online business-to-consumer lenders” as their most important business priority and biggest threat: A Fannie Mae survey (2019) suggests that the borrower-centric model is gaining momentum and will become predominant in the next decade. Fintech lenders may have an advantage in such aspects: for instance, their end-to-end online platform means paperless applications and increased freedom with regard to timing and location, which translate to customer experience, convenience, and easy accessibility. For instance, the application programming interfaces (API) technology implemented in the mobile apps or online portals of fintech lenders enables users to easily connect their borrowing accounts to their bank accounts, for quick extraction of data regarding income deposits, and streamline the verification process (Gomber et al., 2018). It not only saves customers from the paperwork and the back-and-forth between different organization and branches, but also provide essential speed and security for lenders. Such features can substantially reduce business risks through cost reduction in overheads and minimize frictions from human errors and procedural inefficiencies.

2.3. An overview of Fintech Lending vs. Traditional Lending

2.3.1 The Fintech Way

The algorithmic processing of mortgage applications allows significant advantages in speed and efficiency for fintech lenders. Consumers do not have to physically attend to bank branches nor engage in the step-by-step process directed by underwriters and consultants. As Fuster et al. (2009) suggest, fintech lenders, on average, reduce the processing time by ten days, or 20% over traditional lenders. Through reduced frictions and increased efficiency, value is generated for both consumers and lenders, which creates a win-win situation: Buchak et al. (2018) estimate a 4% increase in premium for fintech firms and illustrates that the highest credit score demographics of fintech borrowers pay 0.6 basis points more than ordinary credit score population. Borrowers are willing to pay for convenience, and this trade-off increases the profitability for lenders as well.

On the other hand, traditional lenders are often depository banks facing heavy internal and external regulations that cause friction in areas such as processing speed and transaction volume. Once a dominant force in the mortgage market, large depository lenders now only occupy seven seats of the twenty largest mortgage originators in the U.S market (Fannie Mae 2019). Fuster et al. (2019) point out the complex legacy processes, information systems, and organizational structures of bank lenders are a significant disadvantage in terms of nimbleness. Through examinations on the regulatory impact of capital requirements, mortgage servicing rights, mortgage-related lawsuits, empirical evidence illustrates that regulatory burden constitutes roughly 60% of the shadow-banking market share growth. Moreover, these regulatory burdens are often lopsided – 98% percent of observed mortgage lawsuit settlements target traditional banks, and as a result, shadow banks gained most market share where regulation exceedingly pressure traditional banks (Buchak et al., 2018).

Evidence also suggests that fintech firms do not provide significant advantages over cost for consumers. Philippon (2015) states that despite advances in information technology in financial services, the cost of intermediation has not decreased in years. While some evidence suggests that fintech lenders do not provide economically different interest rates (Fuster et al., 2019), others suggest that fintech firms offer higher interest rates (Buchak et al., 2018), but applies data and information technology to set their prices. In fact, traditional banks still hold crucial advantages: first-time borrowers are generally inclined to choose a face-to-face mortgage origination process with a traditional bank for security and a sense of education (Fuster et al., 2019; Buchak et al., 2018), and there are arguments favouring an advantage of the marginal costs of funding and higher product quality for traditional banks (Buchak et al. 2018). Fintech firms' growth in market share may be attributed to the competitive pricing it offers, but it is influenced by factors such as increased mortgage demand or the value of the convenience and fluid experience from technological advantages. Nevertheless, the overall risk-impact of cost is a complicated issue. For instance, fintech lenders may sacrifice price for growth and market share, and offer competitive prices for suboptimal borrowers, who would not obtain such price from traditional lenders. As a result, fintech-issued loans may perform worse compared to traditional lenders, but its profitability may be unaffected due to the growth in transaction volume. More importantly, such strategy-induced potential trade-offs may attract a significant number of high-risk borrowers into mortgage origination decisions and increase the overall default risk at the market level.

2.3.2 Seizing the Market Gap

Fintech lenders also seem to fulfil the market gaps where traditional banks fail to optimize. Evidence suggests that borrowers fail to optimize their opportunity to refinance mortgage (Andersen et al., 2015; Keys et al., 2016) or that they make inefficient decisions and significant mistakes in household financing activities (Campbell et al., 2006). The refinance process can utilize the information of the initial purchase, therefore becomes a more standardized process, which is captured by the technological advantages of fintech firms: they demonstrate a reduced processing time of 14.6 days for refinancing processes (in comparison with an average of 10-day reduction overall and a 9.2-day reduction for purchase mortgages) (Fuster et al., 2019); on the other hand, borrowers are refinance more with fintech lenders, and that a mortgage refinance is 20% more likely to be from fintech firms (Buchak et al., 2018). Although there is a lack of survey or data evidence for better refinancing decisions resulted from employing fintech options, it is speculated that fintech firms grant the convenience for consumers to act swiftly and efficiently, therefore at least creating better refinancing opportunities in terms of timing.

In their study targeted at fintech firms' role of filling the credit gap of mortgage market, Allen et al. (2019) points out that in the wake natural disasters, traditional banks manage risks by requiring higher income and lower risk measures, such as credit score and debt to income ratios. On the contrary, fintech lenders do not exhibit the same management measures, nor any reduced performance from doing so. It is consistent with the findings of Fuster et al., (2019), who demonstrate that encountering demand shocks, fintech Lenders demonstrate more elasticity and a similar advantage of processing time over traditional lenders. However, these measures may also create a credit boost effect in periods of high risk and volatility. Although empirical evidence lacks in this regard, the risk-impact of fintech-issued loans generated by demand shocks and market gaps may well be underperceived. As a result, such loans may be a double-edged sword to fintech lenders and the market.

Finally, fintech lenders seem to prosper where traditional banks do not deliver: fintech lenders absorb market share where few local banks exist, or when banks do not provide adequate pricing or demonstrate insufficient processing speed (Fuster et al., 2019). It is consistent with the study of Jagtiani and Lemieux (2018), which suggests that fintech lending activities fill the market gap of areas underserved by traditional banks. A plausible explanation is that fintech lenders are able to materialize its information and technological advantage by reducing cost-ineffective risk-management measures such as tightened lending standards, thereby capturing consumers that traditional banks fail to attend to under different circumstances.

2.3.3 Demographics and Consumer Characteristics

Another critical aspect of fintech lending growth is user demographics. Key user demographics (i.e., the millennials) becoming the primary consumer base can provide favourable growth environment for the fintech industry, as characteristics such as mobile-using habits align with the fintech features (International Trade Administration, 2016). The acceptance, adoption, and adeptness for technology are intuitively the most fundamental consumer traits for fintech mortgage borrowers, as they are expected to make critical financial decisions through navigating a complete online system with an adequate sense of trust and security. Empirical evidence identifies urban, educated, and financially experienced consumers as driving forces of growth (Fuster et al., 2019). Notably, this pattern of user traits illustrates a win-win situation in the fintech lending growth environment. For instance, educated earners are likely to adopt fintech lending instruments for a higher premium or price rates.

2.3.4 Customer Screening

The traditional mortgage screening process is a face-to-face, human interview-based process that involves the evaluation of hard information (e.g., credit score history, income level) as well as judgment on soft information. Community banks, for example, thrived through establishing personalized lender-borrower relationships and evaluating soft information (DeYoung et al., 2004). Morse (2015) reviews technologies employed in fintech lending and suggest that capturing and evaluating soft information can result in a more effective screening process and reduce information frictions, thereby advocating for the importance of soft information in lending activities. While fintech firms can speed up the evaluation process of hard information, their automated, centralized, and algorithmic approach also strips the firms of the capacity to evaluate applicants' soft information, potentially resulting in ineffective screening or cherry-picking the applicants that may disrupt the market in a detrimental and underperceived manner.

Mixed evidence exists regarding the effectiveness of the fintech screening process. Fintech lenders demonstrate significantly less discrimination regarding pricing and approvals than their counterparts and exhibit continuously declining discriminatory behaviour over time, which is attributable to the removal of face-to-face interactions (Bartlett et al. 2019). It demonstrates the potential for fintech firms to reduce market frictions imposed by socio-political issues, and in a way, echoes its capability of filling market gaps. Fuster et al., (2019) find robust evidence on consistent over-performance in terms of default rates for fintech lenders, suggesting that not only do fintech lenders exhibit effective screening outcomes, but they also present a more effective screening process, perhaps due to elimination of error and fraud. On the other hand, fintech lenders exhibit the least selective screening process among traditional banks, non-fintech shadow banks, and fintech lenders (Allen et al.,2019). Although algorithmic-based screening process is well capable to reduce human errors in areas such as hard information evaluation, it can also be detrimental in areas due to the lack of human involvement. For instance, Braggion et al. (2018) study the effectiveness of fintech origination on the peer-to-peer lending market in China, and find evidence suggesting that fintech lenders fails to adjust their screening process despite higher loan defaults and delinquency, and that P2P lending undermines traditional risk-management mechanisms such as loan-to-value caps. Maggio and Yao (2018) find limited evidence of ex-ante based adverse selection but identify that fintech borrowers can use fintech lenders as a channel to boost credit scores. These borrowers subsequently exhibit a deteriorating credit performance after a certain period. Such evidence suggests that the hard information-based, algorithmic screening mechanism of fintech lenders can potentially direct consumer behaviour for the worse or even provide opportunities for exploitations and arbitrage.

To a certain degree, the issues revolving the fintech screening process represent both the potential benefits and threats for the market and economy. The automated approach to screening may reduce market frictions caused by human biases and errors, promote market competition, and stimulate economy in a beneficial manner. In contrast, Fintech lenders and fintech services alike are not necessarily subject to the same level of stringent regulatory requirements as banks, which provide malign opportunities of arbitrage and encourage immediate and potentially excessive spending by consumers.

2.4 Mortgage Default

This study aims to analyse the risks associated with the fintech lending model, and particularly the isolated effect of fintech status on default. Therefore, we examine evidence associated with mortgage default. Existing literature has extensively studied the causes and trigger events that lead to default decisions and contribute to default risk. In this regard, the focus on both the endogenous risks of loan quality and exogenous factors help better understand the factors that lead to higher mortgage default rates. Specifically, endogenous risks are embodied in front-end consumer characteristics, through the screening process and evaluation of soft and hard information. Exogenous factors include general macro-economic conditions and shocks such as unemployment, which can trigger unanticipated default decisions, and require predictive capabilities to manage the associated risks. In addition, potential implications on the effectiveness of information processing, risk mitigation, and predictive capability for both fintech lenders and traditional lenders are also discussed.

Home equity is widely studied and recognised as an essential factor in default decisions. Early empirical research suggests that among studied factors, home equity consistently and significantly explains default decisions the best (Vandell, 1978; Campbell and Dietrich, 1983; Case and Shiller, 1996). Campbell and Dietrich (1983) find evidence that default risk is magnified by higher negative amortization risk, which is represented by current and original loan to value (LTV) ratios. Quigley and Order (1995) also highlight the importance of the equity factor in default explanations, but argue that personal characteristics also contribute to the final default decision. Because the studied loans in this dissertation are from Fannie Mae, who resells and packages loans as mortgage-backed securities (MBS), they generally follow stringent requirements and reflect the general underwriting guidelines for safe investments. As a result, the loan-specific data employed by this dissertation, including LTV, is not representative of the high-risk loan population, and unlikely to have outlier effects.

Another branch of studies focuses on the impact of macroeconomic conditions and shock events that lead to default decisions. Case and Shiller (1996) observe that higher default rates follow real estate price declines and downswings in periods of growth, suggesting price movement can be a valuable tool for risk hedging in the derivatives market. Campbell and Dietrich (1983) provide evidence of statistical significance in the relationship between default and most relevant economic variables in the period of the 1960s and 1970s. More importantly, they were among the first to discuss the impact of regional unemployment rates on default decisions, stating that "there is a continuing need for geographic diversification in mortgage default risk. At the firm level, fintech lenders may have a general advantage in mitigating such risk: its online presence and easy accessibility allow for geographic diversification, capturing undervalued and unattended areas in the process. This resonates with the evidence that fintech lenders capture market gaps in underserved areas (Jagtiani and Lemieux, 2018) and seize value in shock events such as natural disasters (Allen et al., 2019). As a result, fintech lenders may be able to diversify the risks associated with location, thereby establishing an advantage over traditional banks.

However, the decision of default is ultimately exercised at the household and individual level. Campbell and Cocco (2015) suggest that the duality of long-term default costs and immediate budget relief have a significant impact on default decisions based on personal utility maximisation. Therefore, borrowers with immediate financial concerns may exercise default to relieve budget constraints, whereas less restricted borrowers may delay and cancel their default as housing pricing increases. Notably, the utility-maximisation line of reasoning also applies to the studied FRM category. Cunningham and Cappone (1990) find empirical evidence that expectations on interest rate affect FRM defaults more than adjusted-rate mortgage defaults. Campbell and Cocco (2015) further explain that for adjusted-rate mortgages (ARM), default risk increases as interest rates and inflation drives up payment requirements, whereas the reverse is true for FRM. As a result, FRM default risk is positively related to the level of the interest rate set at origination. Interestingly, previous research on the difference of interest rate return mixed messages for fintech lenders: While Buchak et al. (2018) find evidence that fintech lenders set a higher interest rate, Fuster et al. (2019) criticize the observation and state fintech-issued mortgages are at a lower rate by 2.3 base-point in their data sample. Subsequently, this dissertation also examine data on the difference of interest rate, providing implications in aspects such as pricing and risk expectation, and risk management[1].

There is also the technical aspect that impacts the effectiveness of default risk assessment: Brunnermeier (2009) highlights the importance of price movement, suggesting that the failure of statistical models in the build-up to the subprime mortgage crisis is partially attributable to the inability to process unforeseen house pricing fluctuations. In terms of predictive capabilities, it is

[1] See section 4.1

suspected that fintech lenders should have an overall advantage over traditional lenders. Although both groups should and do have statisticians and computational power to capture, analyze, and predict macroeconomic movements, the technology-driven core functions of fintech firms should also enable them to have an edge at innovative and frontier solutions on information processing. Moreover, Rajan et al. (2012) attribute such failure of statistical models to the incapability of capturing changes in the relationships between variables when causal factors change. It suggests that there is an inefficiency in information processing, which results in delayed or omitted responses and solutions. Again, in this regard, fintech firms should perform better, since it is able to process loans faster and more effectively (Fuster et al. 2019), which reduces the processing cycle. At the same time, algorithmic-based systems employ tools such as machine learning (Jagtiani et al., 2019) that provide constant and rapid adaptiveness to address changes in relationships and shifts in patterns.

Elul et al. (2020) use credit card utilization as a proxy to measure illiquidity as the consequence of shock events and conclude that both negative equity and illiquidity contribute to default. However, shocks such as unemployment are associated with higher default risks in terms of triggering default decisions and serve a better role of default trigger than negative home equity. Notably, they emphasize the potential benefit of using broader balance sheet information in borrowers' default risk assessment. Because fintech lenders use automated and algorithmic-driven systems, its input information is conceivably limited. Subsequently, it may result in inadequate judgment and evaluation compared to a human officer, who is able to require a more comprehensive set of documents, evaluate soft information, and observe information that's undervalued by the standard requirements.

Soft information is another key issue that contribute to the differences between fintech and traditional lenders. Stein (2002) emphasizes that loan quality relates to both hard and soft information. Rajan et al. (2012) echo the sentiment. They also observe that statistical default models fail after transitioning from a low securitization to a high securitization period, since they underpredict defaults for borrowers whose soft information is more valuable. Consequently, omission of soft information and human judgment may result in inadequate evaluation of loan quality. Because fintech firms generally minimize the human interactions in their origination process, the fintech system may be exposed to more screening risks. Furthermore, in their study on the subprime mortgage, Mian and Sufi (2009) find nationwide evidence in the relationship between the increase in defaults and the relative growth in mortgage credit in the United States. The potential credit-expansion effect of fintech lenders, therefore, can increase the risk of default and become a detriment to businesses.

At the same time, such risk of misinformation (i.e., incompleteness, omission, misjudgment) is an endogenous one within the mortgage industry, and therefore largely universal across both fintech and traditional lenders. Rajan et al. (2012) state that lenders set interest rates based on a selected set of reported variables. As a result, lenders can omit or misjudge relevant information associated with default risk. Consequently, among borrowers with similar reported information, approved borrowers demonstrate worse performance in unreported criterion compared to unaccepted borrowers. Notably, such an effect may be inflated among fintech lenders, since their automated approach may magnify the flaws and imperfection that persist in the industry-level current credit evaluation and approval processes. As a result, it also exposes the opportunities of arbitrage and direct consumer behaviour as mentioned earlier, thereby increasing risk exposure.

2.5 Literature Review Summary

Lee and Shin (2018) capture five fundamental blocks of a fintech ecosystem: fintech start-ups, technology developers, government, financial customers, and traditional financial institutions. Our literature review finds interconnected effects among these elements, especially among fintech lenders, financial customers, technological differences, and traditional lenders. The intertwined relationship of these features can serve a complementary effect that has continuously driven the growth of the sector in recent years. For instance, highly educated and financially experienced consumer demographics contribute to the increasing market presence of fintech lenders and are willing to pay higher premium in exchange for convenience and technological advancement; government regulations exert capital pressure and regulatory restraints on traditional lenders, which create market gaps and enable fintech lenders to flourish where traditional banks perform inadequately. Similarly, the human-based, soft information-oriented screening approach of traditional banks create market frictions such as discrimination or human errors, which is compensated by the algorithmic and automated process of fintech lenders. On the other hand, the lack of soft information screening in the fintech origination process may fail to account for valuable information related to credit risks. There is also evidence suggesting that the automation and online process can induce arbitrage behaviour or direct consumer behaviour to the detriment. In addition, significant regulatory advantage over traditional banks may cause disruptive market and economic frictions, which implies potential market-level risk until a more mature, established regulatory system on the fintech industry is in place. Next, based on the unique features of fintech lenders and the empirical evidence in mortgage default from exiting research, an empirical analysis is conducted to differentiate fintech lenders from traditional lenders and evaluate the mortgage default performance.

3. Data and Methodology

Chapter 3 states the research philosophy, design, and data collection process. It also focuses on the core issue relevant to the empirical study, such as the fintech lender classification problem. In addition, based on research evidence examined in the literature review, I develop a hypothesis that is addressed in the results and interpretation section.

3.1 Research Philosophy

Since the dependent variable for our regression analysis, default status, is a binary variable, a logistic regression model is employed. Logistic models are used for projection of probability and inference of impact on the pass/fail occurrence of an event. In the case of this dissertation, it corresponds to whether the loan will default or not. A variety of factor can contribute to the default decision, including loan-specific factors, shock events, and general macroeconomic conditions. In addition, there's no consensus on the determinants or magnitude of impact for mortgage default. As a result, this study employs a selected set of control variables and specifically focus on the controlled average impact of fintech status, which provides implications to the risks associated with the fintech features entailed by the used method of fintech lender classification.

3.2 Research Approach

This empirical study intends to analyse the impact of fintech lenders, and thereby fintech features, on the mortgage quality and default risk. Firstly, I compare and analyze traditional loan-specific risk indicators between fintech lenders and non-fintech lenders. Through descriptive statistics analysis, I try to understand whether fintech-issued mortgages are subject to more risks in terms of borrower characteristics and loan quality. Next, I employ the following logistic regression model to estimate the impact of fintech lenders on a loan's default status and to consequently infer the default risk associated with fintech status:

$$\textbf{default status}_{ijt} = \alpha + \beta_1 \textbf{ fintech status}_{j} + \beta_2 \textbf{ loan-specific controls}_{ijt} + \beta_3 \textbf{ macroeconomic controls}_{ijt} + \varepsilon_{ijt}$$

Null Hypothesis: the originator's fintech status does not impact a mortgage loan's log likelihood of default

Default status of loan *i* sold by lender *j* at the time *t* is 1 if the loan has been delinquent over a specified time span based on different models. The fintech indicator is 1 for fintech lenders, and 0 for non-fintech lenders. Loan Controls contain loan-specific variables that are core risk measures of a mortgage loan, including the borrower's FICO credit score, debt-to-income ratio (DTI), loan-to-value ratio (LTV) and the original interest rate. Lastly, a selected set of macroeconomic variables at the time of origination, including the annual real median house income growth rate, annual real GDP growth rate, unemployment rate, and the all-transaction housing price index. I employ loan-specific risk controls to account for the risk carried by the borrower, and macroeconomic controls to account for the risk of macroeconomic conditions at the time of origination. For the logistic regression analysis, I regress default status of different time span (i.e., 1-month delinquent, 3-months delinquent) against the fintech indicator and the control variables for the isolated impact of fintech status. Employing different time horizons also allow observations under loosely defined and strictly defined default status. It offers an opportunity to understand if the patterns of influence shift under different conditions, thereby providing implications to the subject population as well as the lender group. For instance, short-term delinquency may result from temporary inconvenience or unforeseen circumstances. However, a large percentage of borrowers that enter a short-term delinquency status can make subsequent payments. The implications during this period, therefore, can offer insights on the lender's business risk, such as operating efficiency or user convenience. On the other hand, loans that enter long-term delinquency status can be forced into foreclosure procedures and are considered to have material consequences. Long-term default models, therefore, can help provide a more accurate understanding on the serious default risk and overall loan quality.

3.3 Fintech Lender Classification

The lender classification problem is central to this empirical analysis. I follow the empirical studies of Buchak et al. (2018) and Fuster et al. (2019) and identify FinTech lenders by business model. Buchak et al. (2018) subjectively identify fintech lenders based on online presence and the lack of human involvement. Similarly, Fuster et al. (2019) echo the approach and classify fintech lenders if mortgage borrowers can reach the preapproval process online. This paper approaches classification based on their combined results, since the core of their approach and the majority of the classified firms overlap.

The implementation of online and automation technology is universal across major originators, and fintech lenders can involve human interactions as well. However, this study focuses on the consumption-end of fintech features. An originator is classified as a fintech Lender if the origination process is automated and mostly online. Specifically, fintech firms should possess the streamlined and automated capability for the processing and verification of information, rates offering, and approval in the initial stages (Burns, 2015). It suggests that the screening and selection process is algorithm-based and independent of human judgment. Non-fintech lenders, on the other hand, generally requires human evaluation of applicant criterion (hard the distinction is fundamental to the empirical study, as the riskiness of fintech-issued mortgages can derive from such differences in business models. Subsequently, final interpretation of results can be misleading if lenders are misidentified.

I employ the Fannie Mae acquisition dataset for lender classification. However, there are missing values for this information, and sellers with small market shares are uniformly categorized as "Other" in the dataset. I exclude these loans from the analysis. As such, the fintech lenders in this dissertation is rather loosely defined. However, it still captures the core features of fintech lenders in the mortgage industry, and therefore should be interpreted as a close but not fully accurate representation of fintech firms. See Appendix 1. for complete and classified list of lender firms.

3.4 Data

3.4.1 Data description

The quantitative approach of this dissertation research determines the importance of data collection and processing. The employed database, Fannie Mae, is an open-access government-backed organization that records single family, loan-level acquisition and performance data in the U.S conventional fixed-rate mortgage market. The selected data sample covers the period from 2010 to 2017. Separate raw datasets for acquisition and performance data are downloaded, processed, and merged to form a final dataset for empirical testing.

Fannie Mae acquires mortgage loans from financial institutions and other lenders and repackages them into mortgage-backed securities (MBS) in the secondary market. As a government sponsored enterprise, Fannie Mae securitizes a substantial amount of its acquired loans and return interest to investors. The securitization process also allows Fannie Mae to provide stability and liquidity for payment and pricing in the mortgage market (Fannie Mae, 2019). Moreover, the loans acquired and securitized by Fannie Mae closely follow the underwriting guidelines and naturally eliminate the outliers in terms of creditworthiness, such as high LTV loans. Consequently, the quality of loans in the Fannie Mae database is reasonably presumed to be above-average quality. Therefore, I also forego the data normalization and outlier elimination measures such as winsorization. Fannie Mae also presents unedited and raw sets of quarterly data with guidelines of processing, basic data format, and glossary terms for each variable. In addition, it also provides standard coding for raw data processing in R. However, due to the fintech classification problem, convenience, and software adeptness, this thesis uses python instead and conducts a manual process of data wrangling.

Furthermore, the Federal Reserve Economic Database (FRED) is used for national and state-level macroeconomic control variables. FRED is a U.S government-based research facility with time-series economic data from 91 sources. It is a widely used and accepted resource for academic research. For national data, a dictionary dataset is created, and the date-assorted macroeconomic control variables are mapped accordingly based on each loan's origination date; for state-level data, datasets are manually aggregated for each state and an identical process is then performed. State-level macroeconomic data also excludes the U.S Virgin Islands and Puerto Rico region.

3.4.2 Control Variable Selection

This dissertation employs four standard default risk variables provided by the Fannie Mae database, namely FICO credit score, interest rate, loan to value ratio, and debt to income ratio are selected. FICO credit score ranks consumers by the likelihood of meeting payment obligations and is a quality metric used by 90 out of the 100 largest U.S lending institutions (FICO, 2020). In addition, Fannie Mae also adopts FICO as the reported borrower credit ratings. Therefore, this thesis employs the FICO score as for the borrower credit risk at the time of origination.

The original interest rate measures the implied risk associated with the mortgage loan, especially in the fixed-rate mortgage market (Cunningham and Cappone 1990; Campbell and Cocco, 2015). Furthermore, it is also perceived as the competitive price offered to the mortgage applicant. The variable is used as an inclusive risk factor, as well as to account for the level risk perceived by the lender based on borrower characteristics.

Debt to income ratio measures the ratio of total monthly debt obligations to gross income. It measures borrower's ability and pressure of payment, and a lower DTI ratio indicates less credit risk. Moreover, DTI is a widely used factor in mortgage default models; see, for instance, Rauterkus et al. (2010), and Bhattacharya et al. (2019).

The importance of the loan to value ratio is emphasized in the literature review. It is a home equity indicator and widely considered a determinant of mortgage default decisions (Campbell and Dietrich, 1983; Campbell and Cocco 2015). A high LTV ratio indicates potential illiquidity for borrowers and higher loss for lenders in the event of default. High LTV loans can also pose external pressure to borrowers, such as the requirement of mortgage insurance.

This dissertation also employs several macroeconomic control variables to control for the economic risks that affect mortgage decisions. Median house income (MHI) befits the single-family loan-level data employed in this dissertation, it affects the value of default decision (Campbell and Cocco, 2015) and is a quality measure used to capture the shifts in the ability to pay. On the other hand, the all-transaction house price index is used as a proxy for housing prices. Case and Shiller (1996) points out the strong correlation between housing price and periods of high default rates. Price declines also result in negative home equity, and therefore increasing default risks. Therefore, a high housing price at the time of origination may result in increased origination decisions but also increase the subsequent risk in home equity as well as default decisions.

In addition, unemployment rate and real GDP growth rate are also selected. unemployment rate accounts for the shocks that affect borrowers' ability to pay. Existing evidence demonstrates that unemployment is a significant determinant for household default decision (Campbell and Dietrich 1983; Quigley and Order, 1995; Gerardi et al., 2018). Real GDP growth rate measures inflation-adjusted economic growth. It is also of importance to the time horizon of the study, as the data covers the 2010-2017 period. Arguments can be made that the economy is still under the residual impact of the subprime mortgage crisis, followed by a period of an economic recovery period. This perceived impact of recession/recovery will also result in the potential changes of signs and magnitudes regarding economic growth. Therefore, including theses variable will also, to a degree, adjust for the impact derived from changes in economic-cycles.

3.4.3 Data Structure

The exclusion of non-available and unspecified lender information is a prerequisite for the regression analysis. The acquisition data contains unique loan entries with lender names at the time of origination. Therefore, the acquisition dataset is first preprocessed, followed by an extraction of the data entries with specified lender names. A fintech indicator variable is then created based on the classification table. As a result, the study omits a group of mortgage lenders such as local banks and community banks.

The performance dataset, on the other hand, contains tens of millions of monthly data with repeated entries for each loan. Three variables from the performance dataset are extracted: a unique ID for loan identification, delinquency status, and reporting period. Delinquency status indicates the time period for which the loan has been delinquent. Because the analysis contains models with different timespan of default status, a different dataset with specified default status is created manually for each model. For instance, in the 3M model, a loan is considered to have defaulted once it has been delinquent for over 90 days. The loan entry is selected when it becomes defaulted for the first time based on reporting period, and merged with the acquisition dataset based on the unique ID. After the merge, the final combined dataset contains 8.6 million unique loans, providing a substantial sample size for empirical analysis. Out of the 8.6 million entries, 0.99 million are fintech-issued loans. It is a significant and sufficient sample size for the regression analysis and representative of the fintech market share shown in Fuster et al. (2019). See Appendix 4. for data processing procedures.

3.4.4. Sampling

The sampling procedure of this study can be categorized as consecutive sampling, meaning the complete accessible population is studied. First, all loan accounts that meet the predetermined criteria for the analysis are chosen in the Fannie Mae database. In this case, the main implication is that all loans without specific seller information are excluded from the study, in addition to a small population that lacks other information such as debt to income ratio. This is an intuitive approach to the study, since certain data availability is a prerequisite for empirical. Secondly, the time period covered by the study, from 2010 to 2017, follows the work of Buchak et al. (2018) and Fuster et al. (2019) in order to match the method of fintech lender classification. Although Buchak et al. (2018)

specifically defend the robustness of their classification to time, I decide to be safe and employ the proven approach.

In addition, the Fannie Mae data layout and the data structuring process largely eliminates the influence of outliers and anomalies. First, unknown and unavailable categorical values, such as default status and fintech indicator are excluded. Secondly, the loan-specific information, such as FICO credit score, falls within a specified range as a result of the organization guidelines. Therefore, techniques such as winsorization to eliminate outlier influence is not necessary in this case. Moreover, the small percentage of the population without loan-specific information is also excluded from the data sample, as normalization of loan-specific variables such as FICO-score or interest rate is not necessary or logical under such a large sample size.

Because of the substantially large population of the Fannie Mae database, the resulting sample size of my analysis contains 8.6 million data entries and therefore significantly reduces the statistical margin of error. It also allows the thesis to fairly capture the dynamics of the U.S. mortgage industry since the selected population fairly represents the entire population and fixed-rate-mortgage has a predominant presence in the U.S. mortgage market (Moench, 2010). However, the sampling technique is also largely restricted by the data availability and format. Consequently, the omission of certain populations such as adjustable-rate loans, local-bank issued loans, small fintech issued loans are unavoidable. Thus, the final analysis is only partially representative of the entire mortgage lending market.

3.5 Ethics

This dissertation employs secondary data from publicly accessible databases (i.e. Fannie Mae and the Federal Reserve Economic Databases). The access and download rights of datasets are public and presumably granted upon academic use. No other ethics issues apply.

3.6 Limitations

First, the logit regression model employed resembles Fuster et al. (2019). However, a selected set of macroeconomic variables is added to control for the year-specific macroeconomic conditions that may affect default risks at the time of origination. Moreover, this study tests the 1 to 6-month default horizon and employs the Fannie Mae database instead of the Ginne Mae FHA NW data. A robustness section for regional effect, loan purpose, and origination channel.

Secondly, our sample data spans from 2010 to 2017, an economic recovery period after the subprime mortgage crisis. As a result, the impact of macroeconomic and industry-related variables impact may be inflated and different compared to a booming or recession period. Other general issues such as volatility in LIBOR can magnify or reduce the impact and should also be considered.

Lastly, the combined approach to the fintech classification problem renders the definition of fintech a loosely defined and unoriginal one. However, neither the subjective definition by Buchak et al. (2018) or the end-to-end online origination definition by Fuster et al. (2019) precisely captures the fintech features. Until there is a scientific approach or government-established standard in place, inaccuracies and ambiguity related to this issue will remain.

3.7 Hypothesis

The existing literature entails two significant issues closely related to the research design of this thesis. First, evidence suggest a significant flaw of the fintech mortgage origination process is the lack of evaluation on soft information. Morse (2015) argues that incorporating soft evaluation can significantly boost the effectiveness of fintech screening process. Stein (2002) and Rajan et al. (2012) further emphasize the value of soft information in the context of mortgage market. The automated and online-based screening and approval processes of fintech lenders can result in undervaluation or omission of soft data. In contrast, human and interaction-based systems can seize such value and mitigate default risk in the origination process. Therefore, it is suspected that the lack of information has an overall negative impact on the default performance for fintech-issued mortgages.

Secondly, Maggio and Yao (2008) suggest that although fintech lenders do not exhibit ex-ante adverse selection, its screening process has an effect of credit boosting. As a result, borrowers behaviour and loan performance are worsened in the subsequent period. Similarly, Rajan et al. (2012) suggests that behavioural changes after the origination process substantially contribute to the subsequent underperformance of mortgage default. In addition, Mian and Sufi (2009) find a positive correlation between increase in default rates and credit growth. Therefore, although the easy accessibility and convenience of fintech lenders are widely recognized (Fuster et al. 2019), their potential impact on future consumer behaviour as well as credit expansion can seriously undermine the overall loan performance.

Because our classification of fintech revolves around the online origination, these factors are especially relevant. As a result, the following hypothesis is formed: fintech status has an overall negative default risk impact.

4. Results and Findings:

	Fintech Lenders			Non-Fintech Lenders			All Lenders		
	Mean	Median	Std	Mean	Med	Std	Mean	Med	Std
FICO Score	750	763	49.4	763	775	42.1	762	774	43.2
Interest Rate	3.992	3.99	0.554	4.09	4.125	0.623	4.079	4.125	0.616
DTI Ratio	33.6	35	9.02	32.4	33	9.53	32.6	33	9.47
LTV Ratio	69.8	74	16.7	70.3	75	17.5	70.3	75	17.4

Table 1. Loan-Specific Variable Descriptive Statistics

4.1 Loan-specific Variable Interpretation and Discussion

First, the results for loan-specific risk variables and discuss the implications for fintech lenders are discussed. Table 1. demonstrates that the mean FICO score for fintech lenders are 13 units lower than non-fintech lenders. In addition, the percentage distribution, as indicated by Figure 1, suggests that fintech lenders accept more borrowers from relatively lower credit range, and less from the higher credit range. As a well-established and used metrics, FICO is a good indicator of the quality of the borrower population. A higher proportion of low-credit borrower population implies a potentially higher default risk generated fintech user demographics. Nevertheless, under the 3-month default window, we observe a better performance in average default rate across all but 750-800 credit range for fintech firms (see Table 2). The result indicates that although fintech firms lend more frequently to low-credit applicants, default risk is not only uncompromised, but reduced. It provides support to the observation that fintech lenders seek to capture the value that traditional screening/information processes fail to address (Allen et al., 2019). More importantly, fintech lenders can capture and materialize value in the process. However, the largest demographics fall into the 750-800 credit range for both groups, and fintech performs significantly in this category. The consequence, as illustrated by Table 4, is that fintech lenders have an overall worse default rate compared to non-fintech lenders under the 3-month window.

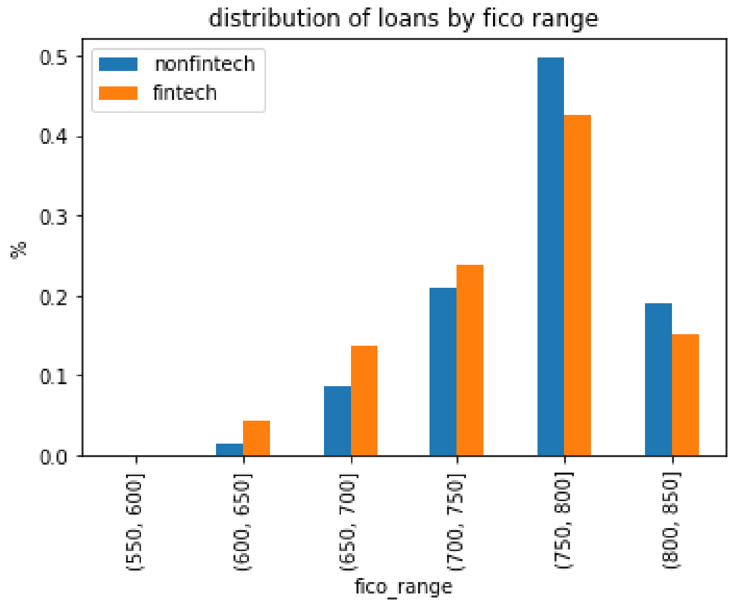


Figure 1. Loan Distribution by FICO Range and Lender Type

	Default Rate		Count	
FICO Range	Non-Fintech	Fintech	Non-Fintech	Fintech
(600, 650]	0.072488	0.052544	114598	43982
(650, 700]	0.037721	0.029863	667348	136791
(700, 750]	0.015184	0.012718	1595987	236912
(750, 800]	0.004491	0.00581	3806571	421645
(800, 850]	0.002343	0.002831	1445414	151034

Table 2. Default Rate by FICO Range and Lender Type – 3 Month Default Window

On average, a unit increase in the borrower FICO score decreases the log odds of default by a consistent level around -0.018 for the 3-month default window. This positive impact also applies to all default windows examined, as illustrated by Table 5. The coefficient and statistical significance confirm the intuitive understanding of the variable, such that higher credit ratings will yield lower default risk. Combined with the observation in descriptive statistics, it further confirms the argument that fintech lenders can capture value in low-credit borrower populations. Hence, fintech firms are able to select loans with less default rates in the low fico-range population. To a degree, the result indicates that fintech lenders can seize the market gap and use it to their advantage to a certain degree. Furthermore, it provides partial evidence against the ex-ante adverse selection conjecture, consistent with Maggio and Yao (2018). Nevertheless, since this advantage fails to compensate for the overall risk and performance of fintech firms, more pertinent endogenous risks associated with fintech status may exist.

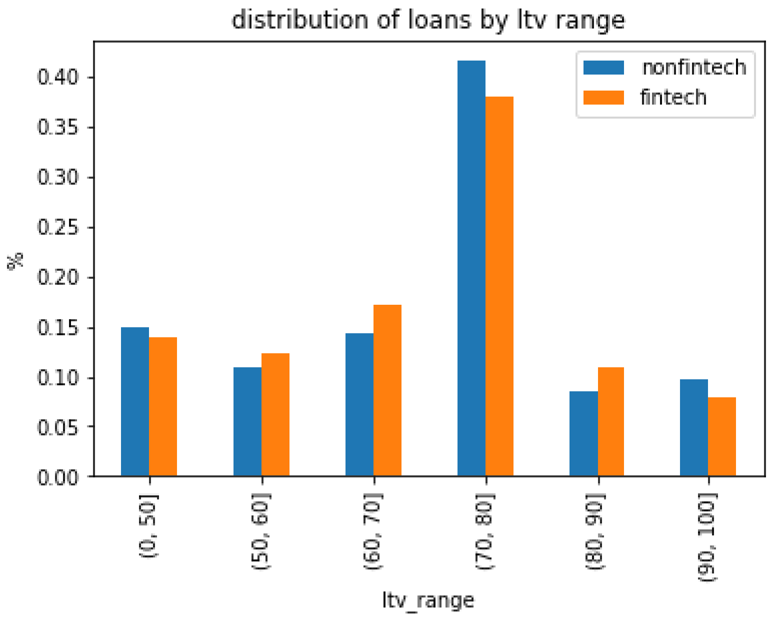


Figure 2. Loan Distribution by LTV range and Lender Type

	Default Rate		Count	
LTV Range	Non-Fintech	Fintech	Non- Fintech	Fintech
(0, 50]	0.004458	0.005858	1141504	137768
(50, 60]	0.006642	0.008992	835762	121435
(60, 70]	0.00905	0.011471	1099309	169478
(70, 80]	0.011152	0.013273	3172615	375800
(80, 90]	0.013006	0.012731	644788	107378
(90, 100]	0.018985	0.020149	741576	78613

Table 3. Default Rate by LTV Range and Lender Type

LTV measures the ratio of mortgage value to property worth. As mentioned, LTV is a significant determinant of mortgage default risk. From the lender perspective, a higher LTV amount indicates that the firm is absorbing a more considerable loss in the event of default. In order for a high LTV loan to be qualified and approved, therefore, requires more stringent guidelines and criteria in other aspects such as income and credit ratings. Figure 2. indicates no significant differences in the overall level or distribution regarding LTV. However, a smaller standard deviation of almost all loan-specific variables across fintech firms suggests that fintech firms are less likely to spend resources to prevent or compensate for higher costs and risk-bearing from such processes. Table 3. illustrates that fintech-issued loans perform worse in default rate across all LTV ranges except for the 80-90 range. Furthermore, unlike that in FICO score, the implied risk difference is more dispersed across all LTV range categories. It suggests that while fintech lenders can sometimes mitigate the risks associated with high LTV loans, the overall risk-impact is negative. Consequently, fintech lenders would bear higher losses in the event of default.

The coefficient and statistical significance of loan to value ratio also reflect its status as a significant determinant for mortgage default risk, as per previous research findings. On average, a unit increase (equivalent to a percentage increase) in LTV increases the log odds of default at the 0.015 to 0.025 range, and 0.0219 for the 3-month default window. Moreover, the magnitude of influence cross all loan-specific control variables, as indicated by the value of coefficients, are smaller when the default period is in the short-term. It suggests that the impact on short-term delinquency status is not as heavily impacted by traditional loan-specific risk measures, but instead significantly concentrated on the fintech status. In this regard, a plausible explanation fintech firms excel in areas such as accessibility and convenience of payment that they can effectively avoid short-term non-payment.

Results for both DTI and interest rate follows the same pattern (see Appendix 2). Fintech-issued loans are more likely to default across all quartiles. On average, a unit increase (equivalent to a percentage increase) in DTI increases the log odds of default at the 0.03 to 0.037 level. In addition, fintech lenders approve loans at a slightly higher DTI, indicating greater willingness to offer a higher amount of loan or to accept individuals from a lower income population, and hence bearing more risk in this regard. On the other hand, the results of descriptive statistics for interest rates are consistent with Fuster et al. (2019), however, at a larger magnitude. Compared to non-fintech lenders, fintech lenders offer interest rates at an advantage of 9.8 base points on average, an economically significant amount. A lower interest rate indicates that fintech firms are less likely to use excess interest rates to compensate or hedge riskier applications. It implies that fintech firms can offer competitive pricing through increased efficiency and reduced frictional costs.

Furthermore, a unit increase (equivalent to a percentage increase) in interest rate increase the log odds of default by 0.1750 in the 1b model; 0.5009 in the 3b model, and 0.5584 in the 6b model. The result is both economically and statistically significant, and an increasing impact is observed as the default-period expands. The result suggests that the original interest rate is a significant risk/quality indicator of FRM, as per previous research (Campbell and Cocco, 2015). Combined with descriptive statistics, arguments can be made that fintech lenders do not leverage performance well with interest rates. However, lenders set interest rate based on internal factors such as risk-aversion and firm-level return target, as well as external factors such as LIBOR rate. Because a higher interest rate is generally a risk-management measure when default risk is not adequately captured (Quercia and Stegman, 1992), fitting interest rate as a feature has other potential implications. For example, fintech companies may have lower target returns from mortgage customers, because their frictional costs are lower. For a loan of a similar price, therefore, it is expected that fintech may have higher default rates. Thus, fintech lenders may just be as effective in borrower screening, but they choose to approve riskier borrowers for growth reasons. This may also not have any substantial consequences for profitability, since origination fees and market-share growth may well compensate for the additional risks.

Overall, evidence on loan-specific variables offers limited support for our hypothesis. Fintech-issued loans perform worse in default rate across LTV, DTI, and interest rate, implying the online origination process is relatively less effective in capturing and managing risks. Contradictory to Fuster et al. (2019), the ex-post default performance of fintech lenders provides support for the lax screening theory[2]. However, the evidence is mixed regarding FICO score. Despite the fact that fintech lenders are able to manage and reduce default risks in among high-risk FICO populations. Notably, it resonates with the alternative arguments for fintech lenders, which indicates superior information processing capability. For instance, automation and algorithmic-based approval process can potentially reduce human errors and fraud and increased the elasticity of demand (Fuster et al. 2019). Along with features such as API technology and the application of machine learning technologies (Gomber et al. 2018; Jagtiani et al., 2019), it translates to increased adaptiveness to capture and manage underlying risks. As a result, fintech lenders can reduce mortgage-related risk through aspects such as reduced overheads and increased operational efficiency. Nevertheless, the evidence for such arguments are partial at best, since fintech lenders are unable to display such advantages among the dominant FICO demographics or in the overall performance.

Lender Type	Fintech	Non-Fintech	All Lender
Model	Default Rate	Default Rate	Default Rate
D1	5.178%	6.966%	6.760%
D2	1.596%	1.512%	1.543%
D3	1.189%	1.019%	1.039%
D4	0.983%	0.801%	0.822%
D5	0.810%	0.630%	0.650%
D6	0.633%	0.506%	0.520%

Table 4. Value Counts and Default Rate Calculation by Model and Lender Type

*D1-D6 Defines Default Period (from 1 months to 6 months correspondingly). Default rate calculations in the sample population suggest that fintech lenders perform well in the short-run and whereas non-fintech lenders perform well in the long-term. It is a surprising result in the sense that within serious default status (3-month delinquent and above) range, fintech lenders issued loans have a higher default rate (as indicated by D3-D6). However, fintech also performs better in the short-term default ranges, which means that

[2] That the automated online origination process will decrease the effectiveness of applicant screening compared to traditional, labour-based methods. As a result, fintech lenders may increase the overall risk at a market level.

(comparatively speaking), fintech issued firms are less likely to be short-term delinquent, but more likely to be medium and long-term delinquent.

Model	1a	1b	2a	2b	3a	3b
Fintech Indicator	-0.3854***	-0.3869***	-0.0693***	-0.0358***	0.0189***	0.0870***
FICO	-0.0075***	-0.0123***	-0.0133***	-0.0176***	-0.0143***	-0.0178***
Interest Rate	0.4705***	0.1750***	0.7533***	0.4519***	0.7614***	0.5009***
DTI	0.0201***	0.0153***	0.0351***	0.0306***	0.0366***	0.0336***
LTV	0.0051***	0.0050***	0.0169***	0.0179***	0.0199***	0.0219***
MHI		0.0542***		0.0567***		0.0504***
rGDP		-0.0376***		-0.0543***		-0.0534***
HPI		0.018***		0.015***		0.0112***
UE		0.3043***		0.2960***		0.2628***
R Squared	0.0546	0.0649	0.114	0.122	0.118	0.124

Model	4a	4b	5a	5b	6a	6b
Fintech Indicator	0.0644***	0.1812***	0.0943***	0.2833***	0.075***	0.3122***
FICO	-0.0147***	-0.018***	-0.015***	-0.018***	-0.0152***	-0.018***
Interest Rate	0.778***	0.5203***	0.7986***	0.5357***	0.8154***	0.5584***
DTI	0.0353***	0.0338***	0.0332***	0.0335***	0.0311***	0.0327***
LTV	0.0204***	0.0237***	0.0192***	0.0244***	0.0184***	0.0248***
MHI		0.0465***		0.0471***		0.0471***
rGDP		-0.0470***		-0.0431***		-0.0408***
HPI		0.0088***		0.0058***		0.0034***
UE		0.2628***		0.2777***		0.2798***
R Squared	0.117	0.123	0.113	0.122	0.110	0.120

Table 5. Logit Model Result

* Model Classification:

- 1-6 represents default period by month (e.g. 1: loans that are 1 month delinquent and above are considered default; 3: loans that are 3 months delinquent and above are considered default)
- a represents models without macroeconomic controls / b represents models with macroeconomic controls
- *=significant at 10%; **=significant at 5%; ***=significant at 1%
- Sample size: 8581298 observations

*Variable Classification:

- Default status: 0 if the mortgage has not been delinquent for the defined time span, and 1 if the loan has been defaulted for the defined time span.
- FICO (range 550-850), borrower credit score at the time of loan origination
- Interest rate: original acquisition interest rate for the fixed-rate single family loans
- DTI: debt to income ratio at the time of origination
- LTV: loan to value ratio at origination
- MHI: annual real growth rate of median house income, in percentage, at the year of origination
- rGDP: annual real GDP growth rate, in percentage, as proxy for economic growth, at the year of origination
- UE: annual unemployment rate in percentage, based on the origination period
- HPI: quarterly all-transaction housing price index, based on the origination quarter

4.2 Fintech Variable Result Interpretation and Discussion

First, I briefly discuss macroeconomic variables since they serve control purposes. Results for variables such as real GDP growth rate, housing price index, and unemployment rate that follow previous studies and common sense: real GDP growth rate has a significant positive impact on default, whereas housing price index and unemployment rate increase the log odds of default. It resonates with the argument that relevant macroeconomic variables should account for mortgage default risks to a certain degree (Campbell and Dietrich, 1983). An interesting observation on macroeconomic control variables is the negative impact of median house income growth on the default rate. A unit increase (equivalent to a percentage increase) in the MHI growth rate increases the log odds of default at around the 0.05 level, a significant and counter-intuitive result. One possible explanation is that median house income is that an increase in household income encourages families to proceed with mortgage origination decisions. In the subsequent periods of slower growth and decline, therefore, families are more likely constrained and proceed to default exercise. Another line of explanation follows the logic of the excessive spending argument (Rajan et al., 2012; Maggio and Yao, 2018) – that growth in consuming power and disposable income may encourage subsequent excessive and immediate spending, thereby increasing default risk.

The result regarding the fintech status is particularly interesting. When default status is defined by entering short-term delinquency (in the 30 to 89 days range), fintech lender status has a positive impact on default risk. On average, holding other variables at a fixed value, fintech lender status decreases the log-odds of loan default by -0.3869 in the complete 1-month model, and -0.0358 in the 2-month model. The result is both statistically and economically significant.

However, there is a gradual reversal in the risk-positive impact when the default period becomes prolonged. For the 3b,4b,5b, and 6b model, which corresponds to delinquency status in the 90~119 days, 120~159 days, 160~189 days, and 190~210 days range, fintech status increases the log-odds of loan default by 0.087, 0.1812, 0.2833, and 0.3122, correspondingly. It suggests that loans are increasingly likely to default in the 90 to 210 days range if it is issued by a fintech firm. Table 6. Also presents the coefficient in terms of odds ratio differences. For instance, a decrease in the log-odds of default by -0.3869 corresponds to having a 32.08% less odds to become defaulted.

Model	Coefficient	Odds Ratio Difference
1b	-0.3869	-0.3208
2b	-0.0358	-0.0352
3b	0.0870	0.0908
4b	0.1812	0.1987
5b	0.2833	0.3275
6b	0.3122	0.3664

Table 6. Log Odds to Odds Ratio Difference Conversion by Model

The default rate continuously decreases as the default-period becomes more loosely defined. This is intuitive because the majority of the loans that enter delinquency status in the short term clear the missing payment in subsequent months. In addition, the foreclosure procedure in the U.S mortgage industry typically begins after a 90-day delinquency status, during which the borrowers can be granted a grace window to clear payment before the property is withdrawn or auctioned off (Department of Housing and Urban Development, 2020). Therefore, in the analysis, the results are categorized into the short-term (1 to 2 months) impact and the long-term (3 to 6 months) impact of fintech status.

The short-term results suggest that holding other variables fixed, fintech-issued mortgages loans are less likely to become delinquent. But long-term results suggest otherwise. In addition, the value of R-squared in model 1a and 1b are significantly lower than that across other models, indicating that there is higher noise and variability in factors that impact temporary default decisions. As loans enter the serious delinquency status (above the 3-month threshold), the risk impact of fintech lender status becomes negative and magnified. Short-term delinquency can be driven from a variety of factors, such as unforeseen circumstances and simply forgetting a payment. The consequences are generally mild and non-existent, unlike serious delinquency status, which may result in lowered FICO score, charge-offs, and foreclosures. Nevertheless, a positive short-term impact indicates that fintech lenders can collect payment in a more consistent and timely fashion. It corresponds to the result from descriptive statistics of default rate, where the percentage of fintech-issued loans entering 1-month delinquency is 1.79% lower than that of non-fintech loans. Arguments can be made that fintech lenders do have a certain advantage in terms of superior ex-ante borrower selection and screening, since fintech lenders also perform better across most variables in descriptive statistics analysis and offer lower interest rate. It also provides partial evidence that advocates the fintech features, such that technological convenience and accessibility can decrease risk from a business perspective.

There are several possible reasons for the reversal and negative default impact under the serious delinquency scenarios. Fintech loans also have a higher default rate (Table 4.) in the 3 to 6-month models. Because long-term default is a continuous behaviour, the result is suggestive of a higher-risk persistence in fintech loans. One possible explanation is that traditional borrowers are more likely to miss short-term payments due to restrictions such as inconvenience and inaccessibility. Fintech borrowers, on the other hand, can mitigate such risks. Yet, the systemic risk of the fintech feature is significantly higher than the traditional models and more impactful on default status. As a result, fintech loans are less likely to be temporarily defaulted, which echoes the argument that operating efficiency that can be translated to reduced frictions and costs (Fuster et al. 2019). However, fintech loans are more likely to enter serious default status and bear detrimental consequences. Therefore, given the long-term nature of the mortgage loans themselves, fintech issued loans bear substantial risks to the firms and the market. Another explanation is that fintech borrowers may have better immediate current resources, which drives the exercise of default decisions (Campbell and Cocco, 2015). However, as Rajan et al. (2012) stated, securitization of mortgages may cause changes in behaviour and worse subsequent performance in soft and underreported information. It also resonates with the argument that the lack of soft information processing may result in ineffective screening (Morse, 2015) and potentially directed consumer behaviour (Maggie and Yao, 2018).

4.3 Robustness

For robustness testing, a set of origination channel dummy variables is added for lender classification implications and a set of loan purpose dummy variables for refinancing implications. A column also includes year dummies to control for fixed-year effect. The lender classification problem is a core feature of this study. Following and combining the established procedures by Buchak et al. (2018) and Fuster et al. (2019), the classification process is rather loosely defined and not without potentials for criticism. In addition, the Fannie Mae database only identifies its sellers of mortgages, which differs from the originator. While most lenders are retail lenders that sell the loans they originate, it is possible that wholesale activities can contaminate the result for fintech impact. Therefore, I follow Buchak et al. (2018) and add a set of original channel dummy variables (i.e. Broker, Correspondent, and Retail) to control for the wholesaling activities. In addition, evidence suggests that borrowers prefer refinancing activities through fintech lenders (Fuster et al. 2019), and refinancing loans' influence on default decisions (Capozza et al, 1998). To address such impact, I add a set of loan purpose dummies (i.e., cashout refinance, non-cashout refinance, and purchase) as well. The summarized result is as follows[3]:

Model	1-month			2-month		
Fintech Indicator	-0.3542	-0.3381	-0.4517	-0.0214**	-0.1137	-0.0961
Loan Specific Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Loan Purpose		Yes			Yes	
Origination Channel		Yes			Yes	
Year Dummy			Yes			Yes
R Squared	0.0556	0.0709	0.0713	0.1155	0.1249	0.1254

Table 7. Robustness Tests, 1-2 Months Default Window

[3] Coefficients are statistically significant at the 1% level unless indicated otherwise. ** indicates statistical significance at 5%; * indicates significance at 1%; *; ^ indicates insignificance.

Model	3-month			4-month		
Fintech Indicator	0.0739	-0.0204	0.0389	0.1386	0.0473	0.1353
Loan Specific Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Loan Purpose		Yes			Yes	
Origination Channel		Yes			Yes	
Year Dummy			Yes			Yes
R Squared	0.1193	0.1258	0.1261	0.1193	0.1250	0.1257

Table 8. Robustness Tests, 3-4 Months Default Window

Model	5-month			6-month		
Fintech Indicator	0.2029	0.1138	0.2343	0.2044	0.1171	0.2652
Loan Specific Controls	Yes	Yes	Yes	Yes	Yes	Yes
Macroeconomic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Loan Purpose		Yes			Yes	
Origination Channel		Yes			Yes	
Year Dummy			Yes			Yes
R Squared	0.1174	0.1227	0.1246	0.1149	0.1198	0.1223

Table 9. Robustness Tests, 5-6 Months Default Window

The result is consistent with the analysis in section 4.2. After adding loan purpose and origination channel, the fintech coefficient still has a sign change as default window expands, and their statistical significance remains at the 1% level in most cases. In addition, the 1-month model still has a significantly less R-squared value, indicating that the short-term default window, the impact of noise and variability are considerably higher. However, compared to the national macroeconomic control model, the fintech coefficient has a higher magnitude across the defined default horizons[4]. It suggests that fintech's explanatory power of default is partially absorbed by regional macroeconomic fluctuations, which echoes the existing research evidence that geographical factors such as unemployment contribute to mortgage default risk; see Campbell and Dietrich (1983). Nevertheless, the result is robust after controlling for fixed-year effects, and also consistent with the main model after changes in the parameters.

4.4 Final Discussion

In the descriptive statistics, fintech-originated loans demonstrate worse default performance compared to traditional lenders. In addition, fintech status demonstrates an economically and statistically significant impact on default in the logistic regression analysis. It suggests that our definition of fintech, which is the ability to originate and approve mortgage applications in an online and automated approach has a substantial impact on the loan quality and default risk. The fintech

[4] See Appendix 3. for comparison of the main model with state-level macroeconomic data

features associated with the online origination process has a significant and economic impact on default status. Therefore, the null hypothesis of regression model is rejected across all models.

Although the overall default rate is worse, fintech loans perform better in low-FICO populations, suggesting that fintech firms can capture value in traditionally conceived high credit-risk borrower populations. Furthermore, fintech lenders exhibit a positive impact in the 1-month and 2-month default window. It indicates that fintech features positively impact borrowers' short-term default decisions due to convenience and accessibility. Although loan-specific values for both lender types are at a similar level, the mean value of interest rate is 9.8 base point lower for fintech lenders, suggesting fintech lenders are less likely to use interest rate as a compensation for borrower risks. Together, there is limited evidence on the reduced business risk and frictional costs of the fintech model. However, the core of the result, as represented by models with default window above the 3-month threshold, suggests that fintech firms are not able to outperform traditional lenders. A negative default-risk impact of the fintech status is observed when default status is serious and materially detrimental.

Lastly, the analysis finds evidence in support of the hypothesis that fintech status has an overall negative risk impact. This is especially illustrated in the serious default windows. Furthermore, in all loan populations separated by defined range or quartile, fintech lenders default more across all variables but FICO score. Descriptive statistics also suggest that fintech lenders are at least able to select the same level of basic loan-risk factors from its borrowers, if not better. In addition, fintech status shifts from a positive risk impact to a negative one as default window expands. Since, fintech borrowers are also less likely to enter delinquency in the short term, it suggests that at similar level of risk represented by traditional loan-specific indicators, the default risk is more serious and pertinent for fintech firms. Thus, while factors such as credit requirement, low interest rate, accessibility, and convenience allow fintech firms to attract consumers and expand market growth, the trade-off is default performance. This also supports the argument of the negative risk impact derived of missing soft information in the screening process, as stated by Rajan et al. (2012) and Stein (2002). For instance, since fintech lenders do not sufficient address the value of soft information in their origination process, borrowers with similar hard information deteriorate in soft information performance over time, and result in an overall underperform in the long run.

5. Conclusion

The fintech status in this thesis is defined as the online presence and automated approach to mortgage origination. It is a loose definition derived from the research of Buchak et al. (2018) and Fuster et al. (2019). Based on this premise, this dissertation examines the loan quality originated by fintech lenders and explores the impact of fintech status on mortgage default. There are several key discoveries.

First, fintech-originated loans have better default performance among borrowers with relatively lower FICO score ratings. It is inferred that fintech lenders do have the capability to capture value missed by traditional lenders through their algorithmic and automated processes. In addition, no substantial differences are observed across other loan-specific risk factors. The result speaks against the ex-ante adverse selection argument, which suggests that the automated process would discriminate against high-risk borrowers.

Secondly, limited evidence exists for fintech lenders' advantage in convenience and accessibility, which can translate to reduced frictional costs. The positive default risk impact of fintech status in the short-term would suggest that fintech lenders are less prone to the risk and costs of incidental default, such as temporary non-payment.

Lastly, endogenous risks associated with the fintech status are negative. Worse default performance and negative coefficients in the models of serious default suggest that the automated and online origination process cannot sufficiently evaluate or manage the default risk. Moreover, it suggests that the default risk of fintech firms is more persistent and economically detrimental.

5.1 Practical Implications

The findings of this dissertation can serve as a basis for changes in risk management practices among fintech lenders. A crucial aspect of the fintech status, for instance, is the lack of human interaction in the origination process. Subsequently, there is also a lack of evaluation of soft information related to mortgage applicants. Existing evidence has also advocated the value of soft information in the screening process. Therefore, firms should consider the inclusion of soft information in their screening process, such as modelling after the personalized screening process in community banks (DeYoung et al. 2014). It may result in the enhancement of the screening process, and thereby reducing the default risk. In addition, the result suggests that fintech lenders do try to seize value in market gaps, such as from low-FICO borrower population, but perform worse under the grand scale. It is a potential indication that the current system of mortgage evaluation may have flaws and defects. As such, the findings can also serve as the empirical evidence for potential changes in the models of mortgage default predictions and risk evaluations.

As LaCour-Little (2008) argues, understanding mortgage default is quintessential to our knowledge on the function and behaviour of the multi-trillion mortgage market. Fintech is on a path of accelerated growth in terms of market share. In addition, the future transition to a more technology-oriented origination and management process is rather inevitable for traditional lenders. Therefore, the findings of this dissertation also provide insights into potential regulatory guidelines and changes at the market level. For instance, the result suggests that although loan-specific indicators are at a similar level overall, delinquent fintech loans are more likely to default continuously and to carry more persistent risk. It offers support for the argument that the online and automated approach to mortgage origination may direct subsequent consumer behaviour for the worse. More importantly, traditional lenders are also subject to more stringent regulatory pressure such as capital requirement (Buchak et al., 2018). Aa a result, the underperformance, accelerated growth, and underregulated market environment of fintech lenders may pose significant risks at the market level. Regulatory entities and policy makers can potentially impose stringent guidelines and fintech-specific requirements such that opportunities for arbitrage and direction of consumer behaviour are minimized. After all, a more established, mature, and well-regulated market environment will limit market risk exposure as well as reducing the firm-level systemic risk from insufficient policies.

5.2 Theoretical Implications

There are a variety of traditionally accepted risk factors associated with mortgage default. The result of this study echoes the importance of such variables in the context of single-family fixed rate mortgages, including FICO, DTI, LTV, and original interest rate. The addition of macroeconomic variables also provides evidence for the impact of economic conditions, such as house price and household income. Notably, contradictory to previous findings[5], the result indicates detrimental risk carried by fintech status. A plausible explanation of such difference is the approach to fintech lender classification[6]. Therefore, a more scientific and precise approach to the classification problem may yield different results. Lastly and most importantly, this dissertation contributes to the research gap of fintech-associated mortgage risk. The findings of underperformance and implications of potential defects and endogenous risks within the fintech status can create a bridge for future studies, in areas such as analysing specific characteristics of fintech lenders or the impact of certain technological implementations.

5.3 Future Research

The data of loan-specific variables is reported at the time of origination. As a result, the study offers insight to the default risks associated with the origination and approval process. Tracking and studying the changes in the loan-specific variables should better account for the impact of fintech

[5] This study uses a similar regression approach as does Fuster et al. (2019)
[6] The lender classification used in this literature is more inclusive, or loosely defined.

status. In addition, it would also provide opportunities to study the underlying borrower population, thereby further decoding the risk impact from consumer demographics and screening process. A similar problem is encountered with macroeconomic controls. The nature of data layout dictates that mapping macroeconomic variables based on the default cut-off period would be unreasonable, since the cut-off period is inconsistent between healthy and defaulted loans. Therefore, under the context that cut-off period is consistent to all data entries, a model would account for the impact of economic conditions and shock event more directly and effectively.

Another area of research is the economic cycle impact of mortgage default and fintech growth. The lender classification problem restricts this dissertation to the period of 2010-2017, after the subprime mortgage crisis. Factors such as LIBOR rate, which has been lower in the 2010s than in the 2000s, suggest that it is a relatively low volatility period. It is also possible that due to the lack of clarity in fintech classification, there is a disproportionate representation of fintech. Consequently, fintech-issued mortgages are weighted towards a period with lower distribution of interest rate for similar risks, and thus showing a higher default likelihood compared to banks. In addition, our model result may be partially influenced by the economic and industrial consequences of recession as well as the subsequent recovery period. The shifts and changes in economic cycle can have significant revelations for the mortgage risk impact of different variables[7]. Therefore, expanding or controlling for the period of study to longer and different economic cycles may yield interesting research evidence.

Lastly, the impact of fintech status revealed in this study may well be restricted by market and mortgage type. For instance, Cunningham and Cappone (1990) find interest rate expectations have higher sensitivity to mortgage termination in fixed-rate mortgages compared to adjusted-rate mortgages. In addition, Lynn and Lyons (2019) find evidence of country-specific impact and that the double trigger impact[8] has a local context. Such phenomenon expands the possibility of exploring the fintech interactions with different market and economic settings. For instance, changing the subject of tests to adjusted-rate-mortgages in the U.K may provide a much different set of considerations regarding fintech lenders.

Acknowledgement

Thanks to Tianyu Yang, Kefan Xiong, Sean Mills and Evaline Bai for being great friends. Special thanks to Dr.Srivastav Abhishek for his guidance and suggestions throughout the dissertation project.

---

[7] Rajan et al. (2012) suggests that mortgage securitization drive borrowers to perform worse in soft information after a booming period

[8] i.e. negative home equity and shock events

6. Bibliography

Bettinger, A. (1972). FINTECH: A Series of 40 Time Shared Models Used at Manufacturers Hanover Trust Company. Interfaces, 2(4), 62-63.

Schueffel, Patrick (March 9, 2017). "Taming the Beast: A Scientific Definition of Fintech". Journal of Innovation Management. ***4*** *(4): 32–54.*

CB Insights Research. (2019). *Global Fintech Report Q2 2019 - CB Insights Research*

Accenture. (2019). *Global Fintech Investments Surged in 2018 with Investments in China Taking the Lead, Accenture Analysis Finds, UK Gains Sharply Despite Brexit Doubts*

Statista. (2019). *Number of Fintech startups globally by region*

Bloomberg. (2019). *Why Asia is leading the fintech Revolution*

Goldman Sachs (2015). *The Future of Finance Part 3: The Socialization of Finance.* Goldman Sachs Global Investment Research

Buchak, G. et al., 2018. Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, 130(3), pp.453–483.

Fuster, A. et al., 2019. The Role of Technology in Mortgage Lending. *The Review of Financial Studies*, 32(5), pp.1854–1899.

Fannie Mae. (2019). *Without Data Standards, the Mortgage Industry Doesn't Go Digital*.

Oliver Wyman. (2016). *The Future of Technology in Mortgage Originations*.

Gomber, P., Kauffman, R.J., Parker, C. and Weber, B.W., 2018. On the fintech revolution: interpreting the forces of innovation, disruption, and transformation in financial services. *Journal of Management Information Systems, 35*(1), pp.220-265.

Philippon, T., 2015. Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation †. *American Economic Review*, 105(4), pp.1408–1438.

Keys, Benjamin J, Pope, Devin G & Pope, Jaren C, 2016. Failure to refinance. *Journal of Financial Economics*, 122(3), pp.482–499.

Andersen, S., J. Y. Campbell, K. M. Nielsen, and T. Ramadorai. 2015. Inattention and inertia in household finance: Evidence from the danish mortgage market. Working Paper.

Campbell, J. Y. 2006. Household finance. *Journal of Finance* 61:1553–604.

Allen, L., Shen, Y., Shan, Y., 2018. Do FinTech Lenders Fill the Credit Gap? Evidence from Local Mortgage Demand Shocks.

Jagtiani, J. and Lemieux, C., 2018. Do fintech lenders penetrate areas that are underserved by traditional banks? *Journal of Economics and Business*, *100*, pp.43-54.

International Trade Administration. (2016). *2016 Top Markets Report - Financial Technology*.

DeYoung, R., Hunter, W. C. & Udell, G. F. (2004), `The past, present, and probable future for community banks', Journal of Financial Services Research 25(2-3), 85{133.

Morse, A. (2015) "Peer-to-Peer Crowdfunding: Information and the Potential for Disruption for Consumer Lending," NBER Working Paper 20899.

Braggion, F., Manconi, A. & Zhu, H., 2018. Can Technology Undermine Macroprudential Regulation? Evidence from Peer-to-Peer Credit in China. *IDEAS Working Paper Series from RePEc*, pp.IDEAS Working Paper Series from RePEc, 2018.

Di Maggio, M. and Yao, V.W., 2018. Fintech Borrowers: Lax-Screening or Cream-Skimming? *Available at SSRN 3224957*.

Bartlett, R., Morse, A., Stanton, R., Wallace, N., 2018. Consumer lending discrimination in the era of FinTech. Unpublished working paper. University of California, Berkeley.

Lee, I. and Shin, Y.J., 2018. Fintech: Ecosystem, business models, investment decisions, and challenges. *Business Horizons*, *61*(1), pp.35-46.

Campbell, T.S. & Dietrich, J.K., 1983. The Determinants of Default on Insured Conventional Residential Mortgage Loans. *Journal of Finance*, 38(5), pp.1569–1581.

Gerardi, K. et al., 2018. Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default. *The Review of Financial Studies*, 31(3), pp.1098–1131.

Elul, R. et al., 2010. What Triggers Mortgage Default? *The American Economic Review*, 100(2), pp.490–494.

Linn, A.C. & Lyons, R.C., 2019. Three Triggers? Negative Equity, Income Shocks and Institutions as Determinants of Mortgage Default. *Journal of Real Estate Finance and Economics*, pp.1–27.

Case, K. & Shiller, R., 1996. Mortgage Default Risk and Real Estate Prices: The Use of Index-Based Futures and Options in Real Estate. *Journal of Housing Research*, 7(2), pp.243–258.

Quigley, J. & Order, M., 1995. Explicit tests of contingent claims models of mortgage default. *The Journal of Real Estate Finance and Economics*, 11(2), pp.99–117.

Campbell, J.Y. & Cocco, J.F., 2015. A Model of Mortgage Default. *Journal of Finance*, 70(4), pp.1495–1554.

Carlos Hatchondo, J., Martinez, L. & Sánchez, J.M., 2015. Mortgage defaults. *Journal of Monetary Economics*, 76(C), pp.173–190.

Mian, Atif & Sufi, Amir, 2009. The consequences of mortgage credit expansion: evidence from the U.S. mortgage default crisis. *Quarterly Journal of Economics*, 124(4), pp.1449–1496.

Cunningham, D.F. & Capone, C.A., 1990. The Relative Termination Experience of Adjustable to Fixed-Rate Mortgages. *Journal of Finance*, 45(5), pp.1687–1703.

Rajan, U., Seru, A. & Vig, V., 2015. The failure of models that predict failure: Distance, incentives, and defaults. *Journal of Financial Economics*, 115(2), pp.237–260.

Stein, J.C., 2002. Information production and capital allocation: Decentralized versus hierarchical firms. *The journal of finance*, *57*(5), pp.1891-1921.

Moench, Emanuel, James I. Vickery, and David Aragon, 2010. Why is the market share of adjustable-rate mortgages so low? *Current Issues in Economics and Finance* 16(8).

Capozza, Dennis R., Dick Kazarian, and Thomas A. Thomson, 1998. The conditional probability of mortgage default. *Real Estate Economics* 26(3), pp.259-289.

Quercia, Roberto G., and Michael A. Stegman,1992. Residential Mortgage Default: A Review of the Literature. *Journal of Housing Research* 3(2) pp.341-79.

The U.S. Department of Housing and Urban Development, 2020. *FORECLOSURE PROCESS*. [online] Available at: <https://www.hud.gov/topics/avoiding_foreclosure/foreclosureprocess> [Accessed 14 April 2020].

Jagtiani, Julapa, and Catharine Lemieux, 2019. "The roles of alternative data and machine learning in fintech lending: evidence from the LendingClub consumer platform." *Financial Management* ,48(4), pp.1009-1029.

Michael LaCour-Little, 2008. Mortgage Termination Risk: A Review of the Recent Literature. Journal of Real Estate Literature, 16(3), pp. 295-326.

Rauterkus, Stephanie, Grant Thrall, and Eric Hangen, 2010. Location efficiency and mortgage default. *Journal of Sustainable Real Estate,* 2(1), pp.117-141.

Bhattacharya, Arnab, Simon P. Wilson, and Refik Soyer, 2019. A Bayesian approach to modeling mortgage default and prepayment. *European Journal of Operational Research,* 274(3), pp.1112-1124.

Deng, Ran, and Shermineh Haghani, 2018. FHA Loans in foreclosure proceedings: Distinguishing sources of interdependence in competing risks. *Journal of Risk and Financial Management*, 11(1)

FICO. 2020. *FICO® Score*. [online] Available at: <https://www.fico.com/en/products/fico-score> [Accessed 12 April 2020].

LaCour-Little, Michael, 2018. Mortgage termination risk: a review of the recent literature." *Journal of Real Estate Literature,* 16(3), pp.295-326.

Lee, In, and Yong Jae Shin, 2018. Fintech: Ecosystem, business models, investment decisions, and challenges. *Business Horizons,* 61(1), pp.35-46.

Vandell, K.D., 1978. Default risk under alternative mortgage instruments. *The Journal of Finance*, *33*(5), pp.1279-1296.

Brunnermeier, M.K., 2009. Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic perspectives*, *23*(1), pp.77-100.

Federal Housing Finance Agency, 2020. *About Fannie Mae & Freddie Mac*. [online] Available at: <https://www.fhfa.gov/SupervisionRegulation/FannieMaeandFreddieMac/Pages/About-Fannie-Mae---Freddie-Mac.aspx> [Accessed 14 April 2020].

Fannie Mae, 2019. *Basics of Fannie Mae Single-Family MBS*. [online] Available at: <https://www.fanniemae.com/resources/file/mbs/pdf/basics-sf-mbs.pdf> [Accessed 14 April 2020].

Federal Reserve Bank St.Louis, 2020. *About the Research Division | St. Louis Fed*. [online] Available at: <https://research.stlouisfed.org/about.html> [Accessed 14 April 2020].

Fannie Mae, 2020. *Fannie Mae Single-Family Loan Performance Data Glossary*. [online] Available at: <https://loanperformancedata.fanniemae.com/lppub-docs/FNMA_SF_Loan_Performance_Glossary.pdf> [Accessed 14 April 2020].

Burns, M., 2015. *This Could Be The Mortgage Industry'S Iphone Moment*. [online] Verizon Media. Available at: <https://techcrunch.com/2015/11/24/this-could-be-the-mortgage-industrys-iphone-moment/> [Accessed 14 April 2020].

7. Appendix

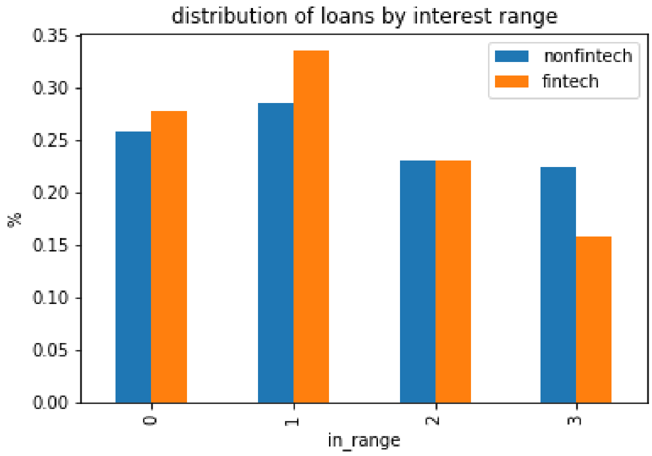
Appendix 1. Fintech Lender Classification

CALIBER HOME LOANS, INC.	Non-Fintech
FAIRWAY INDEPENDENT MORTGAGE CORPORATION	Non-Fintech
JPMORGAN CHASE BANK, NATIONAL ASSOCIATION	Non-Fintech
NATIONSTAR MORTGAGE, LLC	Non-Fintech
FREEDOM MORTGAGE CORP.	Non-Fintech
PMTT4	Non-Fintech
U.S. BANK N.A.	Non-Fintech
UNITED SHORE FINANCIAL SERVICES, LLC D/B/A UNITED WHOLESALE MORTGAGE	Non-Fintech
GUILD MORTGAGE COMPANY	Non-Fintech
FLAGSTAR BANK, FSB	Non-Fintech
AMERIHOME MORTGAGE COMPANY, LLC	Non-Fintech
DITECH FINANCIAL LLC	Non-Fintech
PENNYMAC CORP.	Non-Fintech
FRANKLIN AMERICAN MORTGAGE COMPANY	Non-Fintech
CITIMORTGAGE, INC.	Non-Fintech
FINANCE OF AMERICA MORTGAGE LLC	Non-Fintech
PROVIDENT FUNDING ASSOCIATES, L.P.	Non-Fintech
STEARNS LENDING, LLC	Non-Fintech
IMPAC MORTGAGE CORP.	Non-Fintech
FIFTH THIRD BANK	Non-Fintech
PMT CREDIT RISK TRANSFER TRUST 2016-1	Non-Fintech
PMT CREDIT RISK TRANSFER TRUST 2015-2	Non-Fintech
WELLS FARGO CREDIT RISK TRANSFER SECURITIES TRUST 2015	Non-Fintech
J.P. MORGAN MADISON AVENUE SECURITIES TRUST, SERIES 2015-1	Non-Fintech
SUNTRUST MORTGAGE INC.	Non-Fintech
HOMEBRIDGE FINANCIAL SERVICES, INC.	Non-Fintech
J.P. MORGAN MADISON AVENUE SECURITIES TRUST, SERIES 2014-1	Non-Fintech
PHH MORTGAGE CORPORATION	Non-Fintech
PNC BANK, N.A.	Non-Fintech
SIERRA PACIFIC MORTGAGE COMPANY, INC.	Non-Fintech
STONEGATE MORTGAGE CORPORATION	Non-Fintech
PROSPECT MORTGAGE, LLC	Non-Fintech
PACIFIC UNION FINANCIAL, LLC	Non-Fintech
CITIZENS BANK, NATIONAL ASSOCIATION	Non-Fintech
CHICAGO MORTGAGE SOLUTIONS DBA INTERFIRST MORTGAGE COMPANY	Non-Fintech
ROUNDPOINT MORTGAGE COMPANY	Non-Fintech
NYCB MORTGAGE COMPANY, LLC	Non-Fintech
FLAGSTAR CAPITAL MARKETS CORPORATION	Non-Fintech
USAA DIRECT DELIVERY	Non-Fintech

FEDERAL HOME LOAN BANK OF CHICAGO	Non-Fintech
COLORADO FEDERAL SAVINGS BANK	Non-Fintech
ALLY BANK	Non-Fintech
FREMONT BANK	Non-Fintech
GMAC MORTGAGE, LLC	Non-Fintech
CHICAGO MORTGAGE SOLUTIONS DBA INTERBANK MORTGAGE COMPANY	Non-Fintech
WELLS FARGO BANK, NA	Non-Fintech
BANK OF AMERICA, N.A.	Non-Fintech
REGIONS BANK	Non-Fintech
HSBC BANK USA, NATIONAL ASSOCIATION	Non-Fintech
GMAC MORTGAGE, LLC (USAA FEDERAL SAVINGS BANK)	Non-Fintech
NEW YORK COMMUNITY BANK	Non-Fintech
QUICKEN LOANS INC	Fintech
LOANDEPOT.COM, LLC	Fintech
MOVEMENT MORTGAGE, LLC	Fintech
AMERISAVE MORTGAGE CORPORATION	Fintech
CASHCALL, INC.	Fintech
HOMEWARD RESIDENTIAL, INC.	Fintech

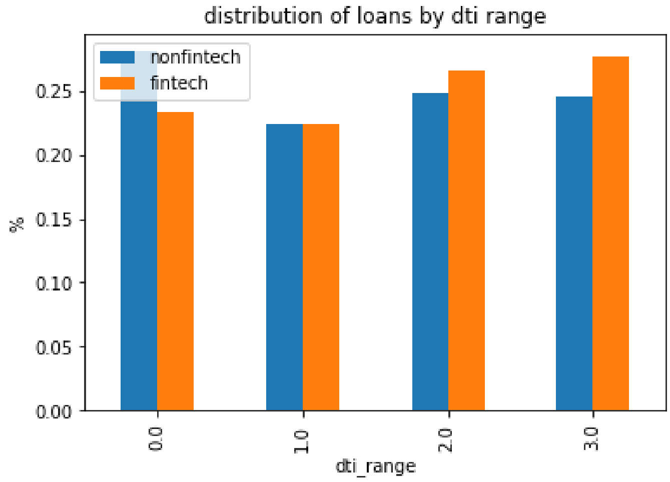
Appendix 2. Results for DTI and Interest Rate

Interest Rate (By Quartile Distribution):



	Default Rate		Count	
	Non-fintech	fintech	Non-fintech	fintech
First Quartile	0.003798	0.005186	1967552	274572
Second Quartile	0.006847	0.00926	2185611	331952
Third Quartile	0.011259	0.014979	1764497	228057
Fourth Quartile	0.021014	0.024778	1712797	155783

Debt-to-income Ratio (By Quartile Distribution)



	Default Rate		Count	
	Non-fintech	fintech	Non-fintech	fintech
First Quartile	0.004396	0.005598	2144774	231337
Second Quartile	0.007948	0.009082	1711460	221758
Third Quartile	0.012357	0.013411	1898638	262996
Fourth Quartile	0.016933	0.018004	1870264	274222

Appendix 3. Main Model Result with Regional Macroeconomic Control Variables

Model	1a	1b	2a	2b	3a	3b
Fintech Indicator	-0.3870***	-0.3524***	-0.0758***	-0.0214**	0.0103^	0.0739***
FICO	-0.0075***	-0.0078***	-0.0133***	-0.0133***	-0.0143***	-0.0142***
Interest Rate	0.4715***	0.4385***	0.7534***	0.7275***	0.7627***	0.7335***
DTI	0.0202***	0.0202***	0.0351***	0.0369***	0.0365***	0.0388***
LTV	0.0051***	0.0063***	0.0169***	0.0179***	0.0197***	0.0210***
MHI		0.6578***		0.3411***		0.0316***
rGDP		-0.2882***		-0.6073***		-1.2439***
HPI		0.001***		-0.0007***		-0.0008***
UE		0.0367***		0.0327***		0.0329***
Pseudo R Squared	0.0548	0.0555	0.114	0.116	0.118	0.119

Model	4a	4b	5a	5b	6a	6b
Fintech Indicator	0.0546***	0.1386***	0.0838***	0.2029***	0.0643***	0.2044***
FICO	-0.0147***	-0.0147***	-0.0149***	-0.0151***	-0.0152***	-0.0154***
Interest Rate	0.7808***	0.7377***	0.8207***	0.7439***	0.8197***	0.7367***
DTI	0.0353***	0.0380***	0.0332***	0.0366***	0.0312***	0.0347***
LTV	0.0201***	0.0220***	0.0189***	0.0218***	0.0181***	0.0218***
MHI		0.0465***		-0.3570***		-0.5001***
rGDP		-1.75***		-2.1439***		-3.0986***
HPI		-0.0008***		-0.0010***		-0.0010***