Do Non-Banks Bring Lower Prices for Borrowers in the Government-Sponsored Enterprise (GSE) Mortgage Market?

By

Hongyan Liang, Zilong Liu, Bingyu Zhang and Xinlei Zhao

First version: October 2021

This version: December 2023

JEL classification: G21; G23; G29; G51

Key words: Fintech; non-bank; shadow bank; pricing; GSE mortgages.

Hongyan Liang (email:Liangh@illinois.edu) Zilong Liu (email:Zilongl@illinois.edu) are from the department of Business Administration at University of Illinois Urbana-Champaign, Bingyu Zhang (email: bingyuzhang@themodeladvantage.com) is from The Model Advantage, LLC, and Xinlei Zhao (email: Xinlei.Zhao@occ.treas.gov) is from the Commercial Credit Risk Analysis Division, Office of the Comptroller of the Currency, Washington, D.C.

We thank the helpful comments from Stephen Karolyi, Lan Shi, and participants at the research seminar at the Office of the Comptroller of the Currency. The views expressed in this paper are those of the authors alone and do not necessarily reflect those of Fannie Mae, the Office of the Comptroller of the Currency or the U.S. Department of the Treasury, or any federal agency and do not establish supervisory policy, requirements, or expectations. The authors take responsibility for any errors.

Do Non-Banks Bring Lower Prices for Borrowers in the Government-Sponsored Enterprise (GSE) Mortgage Market?

Abstract

The neoclassical economic theories predict that increased competition can lead to lower prices; however, this prediction has yet to find empirical support in the consumer lending market. We find that, given the same ex post level of credit and prepayment risks, large banks charge the highest interest rates on GSE mortgages, fintech firms charge the lowest interest rates, and the interest rates charged by small banks and non-fintech non-banks fall in the middle with marginal differences between them. With the entrance of non-banks, the GSE mortgage interest rates decrease by roughly 7 basis points from 2010 to 2019.

1. Introduction

The neoclassical economic theories have predicted that increased market competition would lead to lower costs, and theoretical studies in the consumer lending market (such as Stiglitz and Weiss (1981) and Brito and Hartley (1995)) hold the same prediction. However, this prediction has yet to find collaborative evidence from empirical studies in the consumer lending market (see for example, Ausubel (1991) and Philippon (2015)).

In the past decade, many non-banks have entered the consumer lending market, [1] significantly increasing the level of market competition. Some of these new firms are called the fintech firms, since they apply new and more cost-effective technologies, and such technological innovation should make their operating costs much lower than traditional lenders. It has thus been hoped that the new entrants in the past decade, especially the fintech firms, will bring about cost

[1] As evident in Figure 1 for the GSE residential mortgage market, the market share taken up by non-banks has increased dramatically in the aftermath of the 2007-2009 financial crisis.

savings for consumers (Thakor (2021)). Unfortunately, empirical evidence supporting such a prediction has been elusive so far. Despite increased competition from new non-bank lenders in the consumer lending market and the technological innovations, studies have found that the fintech firms do not seem to generate cost savings for consumers (see for example, Tang (2019), Buchak et al. (2020a), Di Maggio and Yao (2021), and De Roure, Pelizzon, and Thakor (2021)).

To move this literature forward, we examine whether the new entrants in the trillion-dollar mortgage market and increased market competition lead to cost savings for mortgage borrowers. We argue that borrower risks may not be fully reflected in origination variables publicly available. Controlling for origination variables, non-bank loans have higher delinquency and prepayment rates, and consequently the higher interest rates might be necessary to compensate for higher levels of delinquency and prepayment risks for non-bank loans beyond the risks reflected in the mortgage origination variables. The more interesting question is thus: Accounting for comparable *ex-post* delinquency and prepayment risks, do non-banks, especially fintech firms, charge higher or lower interest rates relative to banks?

We account for the *ex-post* credit and prepayment risks when studying non-bank pricing in the GSE mortgage market, and we analyze portfolios with similar origination interest rates. To construct such portfolios, for each month, we sort all loans originated in the month into ten deciles based on origination interest rates, and bank loans thus have comparable interest rates as non-bank loans in the same decile-month. However, the delinquency and prepayment rates are higher among non-banks loans than bank loans for almost all interest rate deciles.

We then regress average portfolio interest rates on 24-months 90-day delinquency (90 DPD) rates and prepayment rates at portfolio level, the prevailing interest rates in the mortgage market, and a non-bank dummy variable to assess the magnitude of the interest rate differential among different lender types. The coefficient of the bank dummy variable turns out to be significantly positive at around 22 basis points (bps) when we consider non-banks as one group. When we separate fintech loans from non-fintech non-bank loans, we observe a more substantial pricing difference between banks and fintech companies than between banks and non-fintech firms. In particular, fintech lenders charge interest rates that are approximately 20 to 38 bps lower, while non-fintech non-banks charge interest rates roughly 8 to 18 bps lower than banks.

In addition, we find that the pricing differences between banks and non-banks in the GSE market are most prominent among large banks, the type of banks experiencing the largest regulatory burden post 2009, as well as the largest drop in residential mortgage market share (e.g., Begley and Srinivasan (2021)). Given the same level of credit and prepayment risks, large banks charge the highest interest rates on GSE mortgages, fintech firms charge the lowest interest rates, and the interest rates charged by small banks and non-fintech non-banks fall in the middle. Although small banks seem to charge marginally higher interest rates than non-fintech non-banks do in the overall sample, this evidence is not the most robust in sub-samples.

All above findings hold when we use different delinquency measures, over different horizons since loan origination, and for different loan purposes and different origination channels. These pricing differentials also hold for both high risk (such as low FICO) and low risk (such as high FICO) borrowers.

We finally evaluate the magnitude of interest rate fluctuations over time due to shifts in market shares among various types of lenders using a simulation exercise. We find that the weighted average interest rate for a hypothetical market portfolio would decline from 4.34% to 4.27% during the period 2010 - 2019, and this decline in interest rate is entirely due to the rising market share of non-banks in the GSE market.

As far as we are aware of, our paper is the first study to provide clear-cut evidence in support of the predictions of the neoclassical economic theory on lower costs brought about by higher level of competition in the consumer lending market (see for example, Stiglitz and Weiss (1981), Ausubel (1991), Brito and Hartley (1995), and Philippon (2015)).

Our findings also add to the discussion on the factors driving non-banks' successful penetration in the residential mortgage market. The explanations for the rise of non-banks in this market include non-banks' expansion into the riskier residential mortgage market, technological innovations, and speedier loan approval process by fintech mortgage lenders, etc. Nonetheless, none of the reasons can well account for the mounting market share captured by non-banks in the GSE mortgage market (see the discussions in Section 2.1 and Appendix A). We argue that pricing appears to be a major factor behind non-banks' success in the GSE market.

The rest of the paper is organized as follows. We review the literature and form our hypotheses in Section 2. We explain data construction and describe the data in Section 3. We explore pricing among non-bank mortgages in Section 4, and briefly conclude in Section 5.

II. Literature review and our contribution

2.1 Literature review

The neoclassical economic theories predict that profits fall with rising levels of market competition; in a perfectly competitive market, firms can only earn zero economic profits. Unfortunately, this prediction does not seem to fit the competitive paradigm in the consumer lending market. For example, Ausubel (1991) shows that with over 4,000 lenders and no regulatory barriers to entry, the profits in the unsecured loan market are still at least three percentage points above the competitive level. This view is also concurred by Philippon (2015).

The literature has proposed a few factors possibly explaining the positive profits in the consumer lending market, such as information asymmetry (Stiglitz and Weiss (1981)) or search costs (Brito and Hartley (1995))). A common feature among these theoretical models is that all these models predict profits to fall as the number of lenders grows. Consequently, it is generally believed that new entrants in the consumer lending market should be able to yield cheaper ways of providing financial services to consumers. This prediction should be particularly true for fintech firms, because fintech lenders adopt new and less costly technologies (for example, moving lending online with potentially more efficient screening techniques) (see for example, Iyer et al. (2015), Croux et al. (2020), and Thakor (2021)).

Nevertheless, collaborative evidence for such a prediction has been elusive. For example, De Roure, Pelizzon, and Thakor (2021) have shown that fintech loans are not necessarily cheaper. Tang (2019) concludes that fintech firms compete primarily on loan availability and loan size rather than on pricing. Di Maggio and Yao (2020) find that fintech firms generally charge higher interest rates than banks do in the unsecured personal loan market.

Specifically in the residential mortgage market, Buchak et al (2020a) find that non-banks, particularly fintech firms, charge slightly higher interest rates on the GSE mortgages they originate. We confirm that, after controlling for the mortgage origination variables publicly available, GSE mortgages originated by fintech lenders have higher interest rates than those of bank mortgages for each year from 2011-2018. Such evidence is also seemingly contradictory to the predictions in theoretical literature. Although Buchak et al (2020a) document that non-fintech firms charge lower interest rates than banks do, we will show later that such results are driven by mortgages originated before 2013. For each year post 2013 and after controlling for the mortgage origination variables publicly available, non-fintech lenders charge higher interest rates than banks do.

In summary, the empirical literature in the consumer lending market has yet to produce evidence rendering support to the theoretical prediction that increased market competition should lead to lower borrowing costs for consumers. We aim to move this literature forward by employing a different methodology.

Our paper is also related to another stream of literature examining non-banks' expansion in the consumer lending market. Studies in this area have explored the questions of why traditional banking has been declining in this segment, what unique contributions non-banks have made, and how non-banks differ from banks.[2]

Non-banks' penetration in the residential mortgage market is particularly notable. The explanations for the rise of non-banks in this market include non-banks' expansion into the riskier residential mortgage market (for example, Buchak et al (2020a) and Lux and Greene (2015)), technological innovations and speedier loan approval process by fintech mortgage lenders (e.g., Buchak et al (2020a), and Fuster et al. (2018)), etc. Nonetheless, none of the reasons can well account for the mounting market share captured by non-banks in the GSE mortgage market, which constitutes the largest fraction of the entire residential mortgage market,[3] and it is a market where loans are originated in compliance with agencies underwriting standards with little regulatory niche for non-banks to take advantage of (see the detailed discussions in Appendix A). In addition,

[2] For example, Danisewicz and Elard (2018), De Roure, Pelizzon, and Thakor (2018), Balyuk (2019), Chava, Paradkar and Zhang (2019), Tang (2019), Vallee and Zeng (2018), Hertzberg, Liberman, and Paravisini (2018), Maggio and Yao (2020), Jiang et al. (2020).

[3]The GSE share is roughly 40-60% of the total residential mortgage market during the past two decades. For more information, see https://www.urban.org/sites/default/files/publication/103539/housing-finance-at-a-glance-a-monthly-chartbook-january-2021_1.pdf.

the literature has not documented pricing as a major factor behind the growth of non-banks in the GSE mortgage market, especially after 2013 and among fintech firms.

2.2 Our hypotheses

Pricing analysis in the current literature (for example, Bartlett et al. (2021) and Buchak et al (2020a)) typically uses loan origination variables on the right-hand-side (RHS), such as FICO scores, LTVs, and DTIs; note that FICO and LTV are the variables in the GSE lending grid. However, the actual interest rates on mortgages can deviate from the lending grid offered by the GSEs.[4]

GSE loans are normally underwritten through the automated underwriter system (Desktop Underwriter for Fannie Mae; Loan Prospector for Freddie Mac). Lender pricing has three components: (1) market interest rate; (2) the guarantee fee (or g-fee) by GSEs to cover borrower default and administration cost, called the Loan Level Price Adjustments (LLPAs) by Fannie Mae and Credit Fee in Price by Freddie Mac;[5] and (3) lenders' discretion in quoting rates, which may reflect lenders' operational costs and strategic volume positioning or monopoly rent-taking (e.g., Bartlett et al. (2021)). As mortgage credit and prepayment risks affect lenders' operational costs, loans with the same guarantee fees according to the GSE pricing grid might have different interest rates if they have different credit and prepayment risks.

Mortgage credit and prepayment risks are not entirely determined by the observed origination variables made publicly available by the GSEs, such as FICOs, LTV, and DTIs, and other variables that are not readily available, such as job duration and income stability, can also

[4] As an example of the lending grid, please see Table 2 of Bartlett et al. (2021).
[5] The g-fee is marginally lower for large sellers:55 basis points for the large (L) seller group, and 56 basis points for the medium (M) and small (S) seller groups in 2019 (source: Fannie Mae and Freddie Mac Single-Family Guarantee Fees in 2019 | Federal Housing Finance Agency (fhfa.gov)).

affect both credit and prepayment risks. Note that, during the loan origination process, lenders acquire information beyond the origination variables publicly available, and such information can be incorporated in loan pricing.[6] Further, a borrower might self-select a higher interest rate loan based on her own financial conditions, which can implicitly reveal her credit and/or prepayment risks beyond the observed variables published by GSEs.[7] As a result, loan pricing may deviate from the GSE pricing grid to account for the additional credit and prepayment risks.

If certain loans consistently have higher credit and/or prepayment risks than other loans with similar variables publicly available, participants in the mortgage market may demand higher interest rates in exchange for bearing the higher risks and higher costs. For instance, mortgage servicers earn servicing fees based on the outstanding mortgage balance each month, and their earning will suffer if the loans are prepaid.[8] Further, even though mortgage servicers do not bear credit risks on GSE mortgages, their earnings are affected by delinquent mortgages. This is because mortgage servicers earn interest on borrowers' escrow account balance until payments are made to tax and insurance organizations, and servicers earn less on loans of borrowers who become delinquent more often or for longer. Moreover, when mortgages are delinquent, servicers still need to make payments to investors of these delinquent loans. Although the servicers can eventually recover such payments when the delinquencies or defaults are resolved later, there are time lags and servicers bear funding costs before they recover the payments. The higher servicing costs from

[6] GSEs have additional information internally but they do not disclose such information in their public data.

[7] A good example can be seen from this website: https://better.com/content/difference-between-mortgage-apr-and-interest-rate/?utm_source=google&utm_medium=search&utm_campaign=Search_G_NonBrand_DSA_App&utm_content=agn:DSA_App_All___m:b_d:c_n:g_tid:dsa-130412282889&utm_term=&gclid=Cj0KCQjws4aKBhDPARIsAIWH0JVRZ2NpEM8VnrL43efKgb04BzWv6awQidD5v53UdKK7BYgecBg5XmQaAt6lEALw_wcB, where a financially constrained borrower may choose the higher cost loan A if she does not have cash to pay the $3,750 fee upfront or if she plans to prepay the loan well ahead of the maturity date.

[8] We can see from the agency data that large lenders service their own loans.

no-payments can be clearly seen from the Call Reports for banks' mortgage portfolio in 2020, when the mortgage servicing income of banks took a significant hit following the rise in mortgage forbearance induced by COVID-19. Lastly, servicers may suffer from the extra servicing costs, such as additional phone calls and operation costs for delinquent or defaulted loans (see Bartlett et la. (2021)). Consequently, mortgage servicers would demand higher interest rates on loans with higher prepayment or delinquency or default rates to service these loans.[9]

We therefore argue that we cannot assess pricing of non-bank loans merely conditioning on the loan origination variables made publicly available by GSEs, as there are other factors that influence loan credit and prepayment risks and thus affect interest rates on mortgages. We examine in this paper whether interest rates offered by non-banks are comparable to banks after we account for the *ex-post* mortgage credit and prepayment risks. We can see from Table 7 in Buchak et al (2020a) that non-banks, including both fintech and non-fintech non-banks have higher delinquency and prepayment risks, and it is possible that non-banks charge lower interest rates given the same level of credit and prepayment risks.

Banks have access to lower cost deposits. However, a well-developed securitization market for GSE mortgages and the GSE guarantees (which enable non-banks to quickly sell the mortgages they originated) might have reduced non-banks' funding costs to a level comparable to banks. At the same time, the additional regulation burdens banks face (such as compliance with regulatory examination requirements, Basel and DFAST capital requirements, or additional surcharges upon being designated as the systematically important financial institutions) may increase banks'

[9] From the agency public data, we find that large mortgage lenders (both banks and non-banks) typically service their own mortgages, and smaller lenders tend to have third parties service their loans. Some non-banks have banks service their loans.

operating costs relative to non-banks. We thus have the following hypothesis based on the theoretical predictions:

H1: Accounting for *ex-post* credit and prepayment risks, non-banks charge lower interest rates to mortgage borrowers than banks.

In addition, the literature (such as Thakor (2021)) has generally expected a fall in borrowing costs with fintech firms' entry in the consumer lending market, since fintech firms have been moving lending online, and use more advanced screening techniques, such as leveraging on nontraditional data (see, Iyer et al. (2015) and Croux et al. (2020)). Such technology innovations can reduce operating costs, and thus fintech mortgage lenders should be able to charge lower interest rates than both banks and non-fintech non-banks. We therefore have the following hypothesis:

H2: Fintech firms charge the lowest interest rates among all mortgage lenders given the same level of credit and prepayment risks.

Lastly, the regulation burden is the highest among large banks, and the literature has found that the small banks have been faring better in the residential mortgage market than large banks (see for example, Begley and Srinivasan (2021)). Because of the much higher regulation burdens large banks face, the operating cost should be the highest among large banks, which feeds into the interest rate they charge to borrowers. We thus have the following hypothesis:

H3: The pricing differentials between banks and non-banks are larger among large banks than among small banks.

III. Data description

3.1 Data

We use in this study the Fannie and Freddie public data.[10] We can identify the originators of each loan from the public data, as long as the originator name is not listed as "other lender." We check whether each lender is a bank or subsidiary of a bank. If a lender does not have a RSSD and a bank charter, nor does it have a bank parent, it is considered a non-bank. Classification of fintech lenders is more judgmental, and we follow the approach adopted by earlier studies, such as Buchak et al (2020a) and Fuster (2019). We check the mortgage lending process of the non-banks. If the lending process up till preapproval can be completed entirely on-line, and the lender's business is generated predominantly online, the lender is classified as a fintech lender. The 20 largest mortgage lenders in term of origination volume for Fannie and Freddie are listed in Appendix C. In total, there are 41 banks, and 71 non-banks from 2010 to 2019. 17% of the non-bank mortgage lenders are fintech firms. The public data report the metropolitan statistical area (MSA) level geographic federal information processing standard (FIPS) code for each loan, and we map the loan to MSA level economic variables.[11] If the property is not in a designated MSA, we use state level economic variables instead.

We include in this study 30-year FRMs originated from January 2010 to March 2018 only. We do not include loans originated post March 2018 because we would like to have 24-months loan performance, and loan performance post March 2020 are to a large extent contaminated by

[10] The Fannie public data can be accessed here: https://capitalmarkets.fanniemae.com/credit-risk-transfer/single-family-credit-risk-transfer/fannie-mae-single-family-loan-performance-data, and the Freddie public data can be accessed here: http://www.freddiemac.com/research/datasets/sf_loanlevel_dataset.page.

[11] The MSA level unemployment are obtained from the U.S. Bureau of Labor Statistics, and the MSA level real per capita personal income are obtained from the Bureau of Economic Analysis.

the forbearance program during COVID-19 pandemic period.[12] We exclude loans with LTVs outside agencies lending grid as those high LTV loans are non-standard loans.[13]

3.2 Summary statistics on loan characteristics

Table 1 reports the summary statistics for loans originated from January 2010 to March 2018 and Appendix D lists variable definitions. Given that the means and standard deviations may miss important aspects of the distributions, we also depict the kernel density plots of OCLTV, DTI, and FICO scores presented in Figure 2. We can see from both Table 1 and Panel B of Figure 2 that fintech loans have lower LTVs, possibly driven by the fact that the fintech loans are mostly refinance loans with lower LTVs than those of purchase loans. Non-bank borrowers tend to have slightly lower FICO scores and higher DTI than bank borrowers, and fintech borrowers have the lowest FICOs. We can further see from Figure 2 that banks are more likely to grant loans to borrowers with FICO scores around 800 than non-banks, while non-banks are more likely to lend to borrowers with FICO around 700 than banks. Non-bank loans have higher origination balances, and fintech firms are particularly focused on originating refinancing loans including both cash out refinance and rate refinance. Non-banks are more likely to originate their loans via the retail and broker channels than banks do, and fintech firms rely heavily on the retail channel for loan origination.[14] Further, non-banks tend to lend to borrower from MSAs with lower unemployment rate and higher average personal income.

---

[12] Since March 2020, millions of homeowners have received forbearance under the CARES Act, allowing them to temporarily pause or reduce their mortgage payments.

[13] The loan eligibility matrix can be accessed at https://singlefamily.fanniemae.com/media/20786/display (Fannie) ; http://www.freddiemac.com/singlefamily/factsheets/sell/ltv_tltv.htm (Freddie)

[14] Loans can be possibly underwritten via the following three channels: (a) Retail channel: A mortgage loan, for which the lender (the same as the seller) takes the mortgage loan application and then processes, underwrites, funds, and delivers the mortgage loan to GSEs; (b) Broker channel: A mortgage loan that is originated under circumstances where a person or firm other than a mortgage loan seller or lender correspondent is acting as a "broker" and receives a commission for bringing together a borrower and a lender. (c) Correspondent channel: A mortgage loan that is

Despite the differences in the observable loan and borrower characteristics shown in Figure 2 and Panel A of Table 1, we can see that the differences are small, except for the origination channel. Therefore, we conclude that the loan and borrower characteristics are largely comparable between banks and non-banks.[15] These results suggest that the banks and non-banks generally engage in direct competition in the GSE mortgage market.

Panel B of Table 1 reports delinquency and prepayment rates. By various delinquency measures and over different horizons, we can see that non-bank mortgages tend to have higher delinquency rates and higher prepayment rates than bank mortgages, and fintech mortgages have substantially higher prepayment rates than banks and non-fintech non-banks.

The numbers in Panel B do not control for loan and consumer characteristics or credit and prepayment risks, and we next resort to regression analysis to examine the difference in interest rates, delinquency, and prepayment rates among bank, non-fintech non-bank, and fintech mortgages.

IV. Pricing analysis

4.1 Loan level regression of interest rates, delinquency and prepayment rates conditional on origination variables

We report interest rate, 24-month 90 DPD and prepayment rate regressions in the three panels of Table 2, and these results are presented for the purpose of comparison with the existing literature such as Buchak et al (2020). We report results using 24-month 90 DPD rates in Panel B

---

originated by a party other than a mortgage loan seller and is then sold to a mortgage loan seller. The mortgage loan seller is responsible for ensuring that any mortgages that are originated and processed by third parties and sold to GSE meet GSE's eligibility criteria and are originated in a sound manner. See https://singlefamily.fanniemae.com/job-aid/loan-delivery/topic/overview_of_third-party_loans.htm for more information.

[15] Such results are similar to those in Buchak et al. (2020a).

of Table 2, and results using other delinquency measures are qualitatively similar to those in this panel and are available upon request. The control variables in all three panels are observed variables at loan origination, including origination loan amount, number of borrowers, MSA level unemployment rate, dummy variables for cash out refinancing, rate refinancing, first-time home buyer, mortgage insurance, missing information for mortgage insurance, GSE indicator, correspondence channel, retail channel, step functions for FICO (cut-off points at 640, 660, 680, 700, 720, 740, 760, and 780), OCLTV (cut-off points at 60, 70, 75, 80, 85, 90, and 95) and DTI (cut-off points at 20, 30, and 40. Additionally, the spread at origination (SATO) is controlled for in delinquency and prepayment regressions. It signifies the maximum difference between the origination interest rate of a loan and the PMMS30 rate at the time of origination, observed over the 24 months following origination. To save space, we only report results on the key variables in Table 2, and results on the control variables are available upon request.

We report results for all non-banks in the first two columns of Table 2 and separately by fintech and non-fintech non-banks in the last two columns of Table 2. The coefficient estimates of the non-bank dummy, the fintech dummy, and the non-fintech dummy are all significantly positive in all three panels of Table 2, indicating that, after controlling for the origination variables publicly available, non-bank (both fintech and non-fintech) mortgages have significantly higher interest rates, higher 90 DPD rates, and higher prepayment rates by 24 months post loan origination. The interest rate differential between banks and non-banks is around 3 bps. Even though this magnitude is not economically significant, this result does not render support to the predictions of the neoclassical economic theories on falling borrowing costs with new entrants in the market.

Note that the non-fintech coefficients in Columns (3) and (4) of Panel A, Table 2 do not align with those reported in Table 6 of Buchak et al (2020a). To investigate whether the differences are due to different sample periods, we run the regressions in Panel A of Table 2 by each year, and we plot the yearly coefficient estimates of Columns (2) and (4) in Panel A of Figure 3. We can see that, consistent with the results in Buchak et al (2020a), the coefficient estimate of the non-fintech dummy is significantly negative during 2010-2012 and nearly zero in 2013. However, starting from 2014, the coefficient of the non-fintech dummy turns significantly positive, suggesting that, controlling for the mortgage origination variables publicly available, non-fintech non-banks charge higher interest rates than banks do, a pattern different from that before 2013.

Panel A of Figure 3 further shows that the coefficient estimate of the fintech dummy has been positive each year from 2010-2018. Therefore, conditional on the origination variables publicly available, fintech mortgage lenders charge higher interest rates than banks, a finding contradictory to the conventional wisdom.

In Panels B and C of Figure 3, we plot the yearly coefficient estimates of the non-bank, fintech, and non-fintech dummies corresponding to Columns (2) and (4) of Panels B and C of Table 2. We can see that the delinquency rates are higher among non-bank mortgages originated during each year from 2010-2018 except for 2011. Further, Panel B of Figure 3 show that starting from 2011, the non-fintech line and non-bank line have been on a general ascending trend, suggesting that, controlling for the origination variables publicly observable, the credit risks of non-bank and non-fintech mortgages relative to bank mortgages have been climbing from 2011 to 2016. The fintech line in Panel B of Figure 3 has been close to zero or below zero until the 2014 origination, but this line turns significantly positive post 2015, suggesting that the credit risk beyond that reflected in the origination variables publicly available has gone up after 2015 for fintech loans.

Panel C of Figure 3 shows no discernible trend in the three lines, but the coefficients from each year are all significantly positive. Such results suggest that, controlling for mortgage origination variables, non-bank mortgages have higher prepayment rates than bank mortgages, and the prepayment rates are especially high among fintech mortgages.

As a result, the higher interest rates of non-bank mortgages post 2013 in Panel A of Figure 3 might thus be needed to account for their higher credit and prepayment rates than bank mortgages. We accordingly need to consider *ex-post* delinquency and prepayment risks in the pricing analysis, which we turn to in the next section. Further, since the non-bank dummy is significantly positive only after 2013, post 2013 pricing analysis is more interesting and the rest of the analysis reported in the paper are based on mortgages originated after 2013.

4.2 Interest rate regression analysis at the portfolio level accounting for ex post credit and prepayment risks

4.2.1 All banks

We run regressions using interest rate-sorted portfolios rather than loan level regressions, because it is hard to accurately measure delinquency and prepayment risks at the loan level. Loan-level modeling is subject to the censoring issue. Although we could use hazard models to address the censoring problem, we would have to address the added modeling errors, and we will face at least two additional challenges. First, although delinquency and prepayment risks are reflected to a large extent by the observed origination variables such as FICO, LTV, and DTI, there are other unobserved variables contributing to delinquency risks and prepayment risks. Therefore, any modeling at the loan level using the observed origination variables is subject to the omitted variable problem. Second, loan-level delinquency or prepayment risks cannot be easily constructed even if we have well-specified delinquency and prepayment models at the loan-month or loan-quarter level that is free of the omitted variable problem. This is because we still face the difficulty of aggregating the predicted monthly or quarterly delinquency and prepayment rates to a life-time delinquency and prepayment rates for the loan. For example, how much more does 90 DPD in the first year since loan origination add to the lifetime delinquency risk than 90 DPD that happens during Month 25 – 36 since loan origination? Any aggregation to lifetime rates would have to use assumptions, which in turn would have to rely on some rules of arbitration and can be easily challenged. By contrast, the *ex-post* delinquency and prepayment rates at the portfolio level are largely free of the above obstacles, and portfolio level rates can also smooth out the noise at the loan level.

We use the interest rate-sorted portfolios since such portfolios should be the most homogeneous in terms of risk characteristics among all other alternatives. We construct the interest rate-sorted portfolios monthly by grouping all loans originated in the same month into ten interest rate deciles, and then calculate the average interest rates, delinquency rates and prepayment rates, and other loan characteristics for each portfolio-month, and by bank and non-banks separately. We have a total of 63 months from January 2013 to March 2018, as we require two years of performance window. These 1260 bank/non-bank interest rate decile-month portfolios are used for regression analysis. We construct similar decile-month portfolios for non-fintech and fintech loans.

We first present some summary statistics in Figure 4 before reporting regression results. Specifically, we calculate the differences in interest rates, FICO, DTI, OCLTV, delinquency rates and prepayment rates between non-banks and banks by each interest rate decile for each origination month and take the averages of the time-series difference numbers. We take the differences between non-bank and bank loans in the same origination month to address the potential problem of changing interest rates over time.

To better understand the meaning of the lines, let us take the interest rate graph in Figure 4 as an example. The black line represents the difference between non-bank and bank interest rates. If the black line is above the x-axis, the non-bank interest rates are higher than bank interest rates for those interest rate deciles, and vice versa. We can see that black line is very close to the x-axis in Graph A, which means that the interest rates are very close between bank mortgages and non-bank mortgages for each interest rate decile. Although the difference seems to be noticeable for the 10th decile (the decile with the highest interest rates), the difference is economically non-significant, at only 5 bps, and we can see that the non-bank mortgages have lower interest rates than bank mortgages for this decile (the lines falling below the x-axis). Origination FICOs and DTIs are rather comparable in each interest rate-deciles between banks and non-banks, and fintech mortgages have lower OCLTV for each of the interest rate decile.

The non-bank lines in Graphs E-H of Figure 4 are all above the x-axis, suggesting the that non-banks have higher 90 DPD rates and substantially higher prepayment rates than banks for each interest rate decile. For the 10th decile where the interest rates are noticeably lower among non-banks (as shown in Graph A of Figure 4), both the 90 DPD rates and prepayment rates are clearly higher among non-banks. Therefore, the evidence in Figure 4 provides the first piece of evidence that, for the same level of interest rates, non-banks tend to have higher delinquency and prepayment rates. Or reversely, given the same level of delinquency rates and prepayment rates, non-banks charge lower interest rates than banks do.

We next turn to regression analysis, and the regression specification is as follows:

$$Interest\ Rate_{decile,t} = \beta_1 NonBank_{dummy} + \beta_2 PMMS30_t + \beta_3 Delinq\_Rate_{decile,t} + \beta_4 Prepay\_Rate_{decile,t} \qquad (1)$$

where $PMMS30_t$ is the primary mortgage market survey rate for 30-year FRM for month t obtained from Freddie Mac. We add this variable to control for the changing interest rate over our sample period. Since the origination variables are just proxies of borrower credit and prepayment risks, we use in the regression the realized delinquency and prepayment risks instead of the proxies.

Table 3 presents results on the bank and non-bank sample using 24-month 90 DPD rate as the delinquency measure. Column 1 reports results across all interest rate deciles, and Columns 2 and 3 show results for the lower (five low interest rate deciles) and upper (five high interest rate deciles) interest rate deciles, separately. Columns 4 and 5 report results by interest rate portfolios formed among borrowers with credit bureau scores <720 and >=720, respectively.

In Panel A of Table 3, we investigate all non-banks as one group without differentiating between fintech lenders from non-fintech lenders. The coefficients on prepayment rates and 90 DPD rates are all positive, and the coefficient of the 90 DPD rates are more than 10 times those of prepayments. These results suggest that credit risk is the primary driver of interest rates, while prepayment risk is of secondary concern in pricing but still important.

The coefficients of the non-bank dummy variable are significantly negative in all columns of Panel A of Table 3, indicating that non-banks charge lower interest rates than banks do after controlling for delinquency and prepayment risks and the prevailing market interest rates. The coefficient estimate is -0.22 in the first column, suggesting that interest rates offered by non-banks are lower than those by banks by roughly 22 bps on average. This number is close to the average percentage of roughly 25 bps for each point a mortgage borrower can buy at closing, and thus can be economically meaningful. The coefficient estimates range from -0.16 to -0.32 in the remaining four columns of Panel A of Table 3. Therefore, the interest rate differential holds not only in the high FICO (low interest rate) segment but also in the low FICO (high interest rate) segment. Results in Panel A of Table 3 agree with H1.

Panel B of Table 3 reports similar regression results when we construct the portfolio time series by banks, fintech non-banks, and non-fintech non-banks. We can see that the coefficients of both the fintech and non-fintech dummy variable are always significantly negative, ranging from -0.21 to -0.38 for the fintech dummy and -0.08 to -0.18 for the non-fintech dummy. These results suggest that, after controlling for the *ex-post* credit and prepayment risks and the prevailing mortgage interest rates, both fintech and non-fintech mortgage lenders charge lower interest rates than banks do. Upon further testing (unreported here to save space), the differences in the coefficients between the fintech and the non-fintech dummies in Panel B of Table 3 turn out to be all statistically significant. Therefore, the difference in interest rates between fintech lenders and banks is larger than that between non-fintech lenders and banks, and consequently, fintech lenders charge lower interest rates than both non-fintech lenders and banks, a finding consistent with H2.

We next gauge the dollar impact of the lower mortgage prices offered by non-bank. For an average-size 30-year non-bank loan ($255K from Table 1) to have interest rates decline from 4.21% to 3.99% by 22 basis points, the interest savings in lifetime will be $11,715, including $559.77 by the end of the first year, $1,669.74 by the end of the third year, and $2,763.09 by the end of the fifth year. This size of interest savings is non-trivial compared with US personal income. For example, the median weekly personal income in the U.S. is $865 for all full-time workers in

2017, and the annual real median personal income in 2019 is $35,977.[16] Such cost savings could tip some people to borrow from non-banks.

We further estimate the models in Panels A and B of Table 3 by the sub-periods 2013-2014, 2015-2016 and 2017-2018, and report in Table 4 the results based on the overall sample (corresponding to the first column of Table 3). From Panel A of Table 4, we can see that the non-bank dummy carries negative coefficients for each of the three subperiods. The coefficient estimate is most negative during 2015-2016, and least negative in 2017-2018. However, even for loans origination during the period 2017-2018, the coefficient of the non-bank dummy is economically significant at 15.5 (bps). Therefore, the finding from Table 4 lends further credence to H1.

Panel B of Table 4 shows that the coefficients of the fintech and non-fintech dummies are all statistically significant. Further, coefficients of fintech dummy are more negative than those of the non-fintech dummy for each of subperiods, and the difference in the coefficient estimates is large than 10 (bps) in each of the three columns in Panel B of Table 4. Unreported tests again confirm that the coefficients of the fintech dummy are significantly lower than those of the non-fintech dummy in this panel. Therefore, results in Panel B of Table 4 renders additional support to H2.

4.2.2 Large banks versus small banks

We next investigate the pricing differences between fintech, non-fintech, and large, and small banks separately in Tables 5. We classify as large banks using the list of banks subject to Federal Reserve's annual Comprehensive Capital Analysis and Review (CCAR),[17] and the non-

[16] https://en.m.wikipedia.org/wiki/Personal_income_in_the_United_States.

[17] https://www.federalreserve.gov/publications/comprehensive-capital-analysis-and-review-summary-instructions-2020.htm.

CCAR banks are classified small banks. Results in Table 5 are from a regression with four sets of times series for large bank, small bank, fintech firms, and non-fintech firms by origination month and decile. The regression specification is as follows:

$$\begin{aligned} Interest\ Rate_{decile,t} &= \beta_1 LargeBank_{dummy} + \beta_2 Fintech_{dummy} \\ &+ \beta_3 NonFintech_{dummy} + \beta_4 PMMS30_t \\ &+ \beta_5 Delinq_{Rate_{decile,t}} + \beta_6 Prepay\_Rate_{decile,t} \end{aligned} \quad (2)$$

We use small banks as the reference in this regression. We can see that the large bank dummy carries a positive coefficient in each column of this panel, with the value ranging from 7 bps to 14 bps. Accordingly, on average for loans with similar credit and prepayment risks, the interest rates charge by large banks are higher than those by small banks by roughly 10 bps, consistent with H3. The coefficient estimates are all negative for the fintech dummy, ranging from 14 to 22 basis points. Therefore, the interest rates charged by fintech lenders are lower than those by small banks by roughly 20 basis points. Finally, the coefficient estimate for the non-fintech dummy ranges from non-significant to 5 bps in Table 5. So, in the overall sample, non-fintech lenders seem to charge lower interest rates than small banks, but this evidence does not hold in the low interest rate segment. We have tested the difference in the coefficient estimates of the fintech and non-fintech dummies in each column of Table 5, and they are all statistically significant, suggesting that the interest rates charged by fintech lenders are significantly lower than those by non-fintech lenders, in agreement with H2. All in all, Table 5 suggests that among all mortgage lenders, fintech firm charge the lowest interest rates, large banks charge the highest interest rates, while non-fintech non-banks and small banks fall in the middle.

We report in Table 6 the sub-period results corresponding to the first column of Table 5, and we report the subperiod, FICO group results in Table 7. We can see that the results in Tables 6 and 7 are not qualitatively different from those in Table 5, that is, the large bank coefficients are consistently positive, the fintech coefficients are consistently negative, while the non-fintech coefficients are negatively only in the period 2013-2014. These results again provide support to all our hypotheses.

4.3 Additional investigations

We have conducted many robustness tests. First, we have tried different delinquency measures (such as 30 DPD and 60 DPD) and different horizons post loan origination in all analyses in the paper (12 months to 60 months post loan origination). Second, since purchase loans and refinance loans are quite different, we have also conducted all the analyses in the paper by purchase and refinance loans separately. Third, we have conducted all the analyses by 1) Fannie and Freddie loans separately and 2) by different origination channels separately. Fourth, we have conducted all the additional analyses by loans originated from each of the three subperiods, 2013-2014, 2015-2016 and 2017-2018, and we have also conducted the same analysis for mortgages originated during 2011-2012. Lastly, we have conducted all these additional analysis by risk groups as well, i.e, high and low FICO and interest rate groups. All results reported in Tables 2-7 and Figure 4 hold qualitatively among the additional tests, all pointing toward lower pricing by non-banks, particularly fintech mortgage lenders. Further, the interest rate differentials are always larger between non-banks and large banks than between non-banks and smaller banks. These additional results from robustness tests are not reported here because of space limitations but they are available upon request.

Finally, the interest rates used in our analysis do not reflect closing costs. Is it possible that the closing costs are significantly higher among non-banks so that the pricing differences documented above among non-banks relative to banks are substantially over-stated? Note that annual percentage rates (APRs) reflect both the interest rate and closing costs. We thus examine the difference between interest rates and APRs using the post-2018 Home Mortgage Disclosure Act (HMDA) data.[18] The HMDA data reports rate spread, which is the difference between the APR and the average prime offer rate (APOR). APOR is publicly available at (https://ffiec.cfpb.gov/tools/rate-spread), and we can thus back out APR for each loan. We then calculate the difference between APR and interest rate for each loan. We find that mean difference is 0.24% among non-banks (both fintech and non-fintech) and 0.25% among banks. These results suggest that the closing costs are comparable between banks and non-banks, and the pricing differential we uncover in this study is not contaminated by different closing costs.

4.4 Changing mortgage interest rates over-time

In this subsection, we examine the over-time change in the interest rate for a hypothetical portfolio. For each month, we first assemble four portfolios using historical loans originated by small banks, large banks, fintech, and non-fintech firms. We then run a regression using the following specification:

$$\begin{aligned} Interest\ Rate_{lender\ type,t} &= \beta_1 LargeBank_{dummy} + \beta_2 Fintech_{dummy} \\ &\quad + \beta_3 NonFintech_{dummy} + \beta_4 PMMS30_t \\ &\quad + \beta_5 Delinq_{Rate_{lender,t}} + \beta_6 Prepay_{Rate_{lender,t}} \end{aligned}$$

[18] The pre-2018 HMDA data do not have interest rate information. This analysis is based on the non-public HMDA data available to OCC.

$$+\beta_7 LoanCharacteristic_{lender,t} \quad (3)$$

We simulate four hypothetical portfolios using the coefficient estimates from equation (3). In the simulation, loan characteristics, default rates, and prepayment rates do not exhibit time variation by lender type; they are the historical average values from 2010 to 2019. The PMM30 rates in simulation correspond to the historical monthly values used in the regression.

We finally use the corresponding historical monthly market share of the four lender types to generate a market hypothetical portfolio, and the weighted average interest rates of the market hypothetical portfolio over time are depicted in Figure 5.

From Figure 5, we can see that the weighted average interest rates have been falling nearly monotonically from 4.34% in 2010 to 4.27% 2019. Since we set the loan characteristics, default rates, and prepayment rates to fixed numbers across lender types and over time, the decline in interest rate is entirely due to the changing market share of the different lender types. In other words, Figure 5 suggests that the entrance of non-bank lenders, and in particular, fintech lenders, decrease the GSE mortgage rates overall by roughly 7 bps from 2010 to 2019.

V. Conclusion

Non-banks have become major players in the GSE residential mortgage market in the past decades. Although the neoclassical economic theories predict that higher levels of competition bring about lower costs for borrowers, this prediction has not found collaborative evidence in the empirical literature in the consumer lending market.

Consistent with the prior literature, we find that, given the same variables publicly available at the time of loan origination, non-banks, especially fintech firms seem to charge slightly higher interest rates than banks. At the same time however, non-bank loans have higher delinquency and prepayment risks after controlling for the observed origination variables, and the higher interest rates might thus be necessary to account for the higher delinquency and prepayment risks for non-bank loans.

When we control for *ex-post* delinquency and prepayment risks, we find that non-banks, and in particular fintech non-banks, charge lower interest rates to mortgage borrowers than banks do. These findings render support to the theoretical predictions on competitive pricing in the consumer lending market when the number of lenders increases (for instance, Stiglitz and Weiss (1981), Brito and Hartley (1995), or Ausubel (1991)).

This finding is also related to the literature examining the factors behind non-banks' growth in the residential mortgage market. Although the existing literature has proposed several reasons for the rise of non-banks in the mortgage market, these explanations can hardly explain why the market share of banks has been declining in the GSE mortgage market, relative to both non-fintech and fintech non-banks. We argue that pricing may play a role behind non-banks' rise in the GSE mortgage market. The pricing differential between banks and non-banks is largely concentrated among large banks, the banks with the steepest decline in the GSE market shares during the past decade. We cannot find conclusive evidence on the difference in mortgage pricing schemes between non-fintech lenders and small banks. Such a finding is consistent with the heavy regulatory burden large banks faces, such as Basel and DFAST capital requirements, which adds to their operating costs. By contrast, smaller banks face less regulatory burdens, and they should be nimbler than large banks. Finally, with the entrance of non-banks, the GSE mortgage interest rates decrease by roughly 7 bps from 2010 to 2019.

References:

Ausubel, L.M., 1991, The failure of competition in the credit card market, American Economic Review, 81(1), 50-81.

Balyuk, T., 2019, Financial innovation and borrowers: Evidence from peer-to-peer lending, Rotman School of Management Working paper No. 2802220.

Bartlett, Robert, Adair Morse, Richard Stanton, and Nancy Wallace, 2021, Consumer-lending discrimination in the fintech era, Journal of Financial Economics, forthcoming.

Begley, Taylor A., and Kandarp Srinivasan, 2021, Small bank lending in the era of fintech and shadow banks: a sideshow? Working paper, Washington University in St. Louis.

Brito, Dagobert L. and Peter R. Harley, 1995, Consumer rationality and credit cards, Journal of Political Economy 103(2), 400-433.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2020a, Fintech, regulatory arbitrage, and the rise of shadow banks, Journal of Financial Economics, forthcoming.

Chava, S., N. Paradkar, and Y. Zhang, 2019, Winners and losers of marketplace lending: evidence from borrower credit dynamics, Georgia Tech Scheller Colleage of Business Research Paper, 18-16.

Croux, C., Jagtiani, J., Korivi, T. and Vulanovic, M., 2020. Important factors determining Fintech loan default: Evidence from a lendingclub consumer platform. Journal of Economic Behavior & Organization, 173, pp.270-296

Danisewicz, Piotr and Ilaf Elard, 2018, The real effects of financial technology: Marketplace lending and personal bankruptcy. Available at SSRN: https://ssrn.com/abstract=3208908.

Davidson, Andrew, and Alexander Levin, 2014, Mortgage Valuation Models: Embedded Options, Risk, and Uncertainty Oxford University Press.

De Roure, Calebe, Loriana Pelizzon, and Anjan V. Thakor, 2021, P2P lenders versus banks: Cream skimming or bottom fishing? Working paper, Leibniz Information Center for Economics.

Di Maggio, M. and Yao, V., 2021. Fintech borrowers: lax screening or cream-skimming?. The Review of Financial Studies, 34(10), pp.4565-4618.

Fuster, Andreas, Matthew Plosser, Schnabl Philipp, and James Vickery, 2018, The role of technology in mortgage lending, Federal Reserve Bank of New York Staff Reports, no.836, February 2018.

Hertzberg, A., A. Liberman, and D. Paravisini, 2018, Screening on loan terms: Evidence from maturity choice in consumer credit, The Review of Financial Studies 3532-3567.

Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo F.P. Luttmer, and Kelly Shue, 2015, Screening peers softly: Inferring the quality of small borrowers, Management Science 62(6), 1554-1577

Jiang, Erica, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2020, Banking without deposits: Evidence from shadow bank call reports, NBER Working paper.

Lux, Marshall, and Robert Greene, 2015, What's behind the non-bank mortgage boom? M-RCBG Associate Working Paper Series No. 42.

Maggio, Marco Di, and Vincent Yao, 2020, Fintech borrowers: Lax-screening or cream-skimming? Working paper, Harvard Business School.

Philippon, Thomas, 2015, Has the US finance industry become less efficient? On the theory and measurement of financial intermediation, The American Economic Review 105, 1408-1438.

Stiglitz, Joseph E., and Andrew Weiss, 1981, Credit rationing in markets with imperfect information, American Economic Review, 71(3), 393-410.

Tang, H., 2019, Peer-to-peer lenders versus banks: Substitutes or complements? The Review of Financial Studies 32(5), 1900-1938.

Thakor, Anjan, 2021, Fintech and banking: What do we know? Journal of Financial Intermediation, forthcoming.

Vallee, B. and Y. Zeng, 2018, Marketplace lending: A new banking paradigm? The Review of Financial Studies, forthcoming.

Figure 1. Origination Volume by Lender Types (in $MM)

This figure shows the total volumes of GSE residential mortgages originations by different lender types between 2010 and 2019 as reported by Freddie Mac and Fannie Mae in their public data. The GSEs group all small lenders into the "other lenders" category without revealing lender names, and as a result, we cannot differentiate banks from non-banks in the "other" category and these loans are excluded from the analyses in the rest of the paper. We can identify the originators of each loan from the public data as long as the originator name is not listed as 'other.' We check whether each lender is a bank or subsidiary of a bank. If a lender does not have a RSSD and a bank charter nor does it have a bank parent, it is considered a non-bank. Classification of fintech shadow banks is more judgmental, and we follow the approach adopted by earlier studies, such as Buchak et al (2020a) and Fuster (2019). We check the mortgage lending process of the non-banks. If the lending process up till preapproval can be completed entirely on-line, and the lender's business is generated predominantly online, the lender is deemed a fintech lender. The 20 largest mortgage lenders in term of origination volume in 2012 to 2019 for Fannie and Freddie are listed in Appendix C.

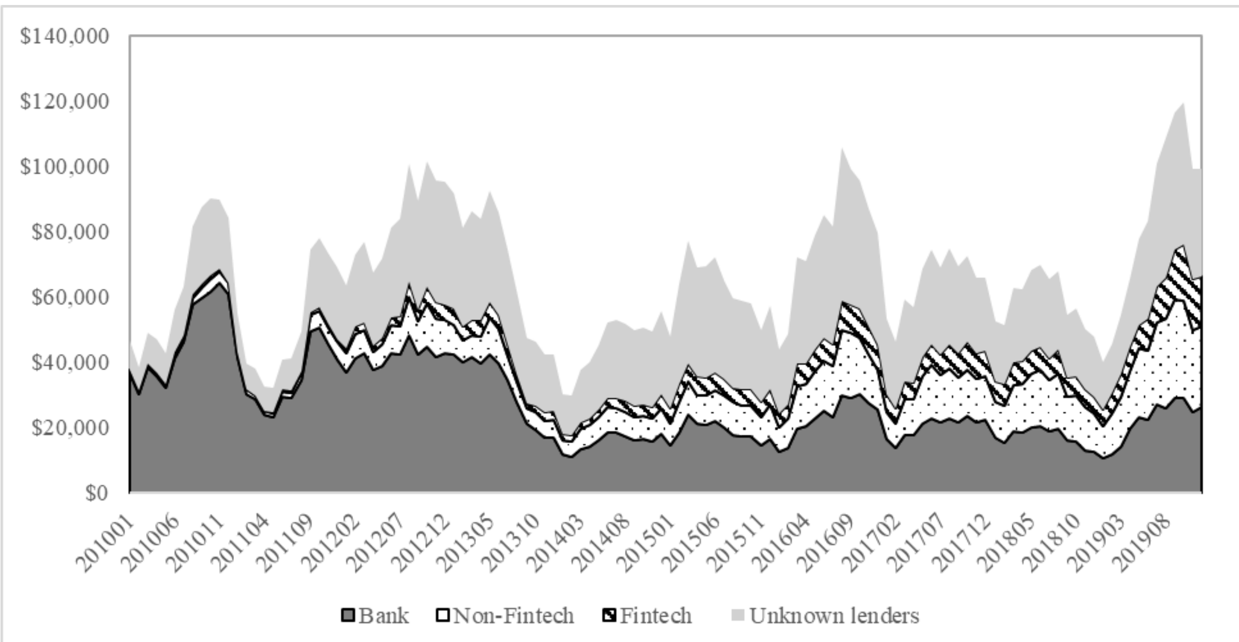
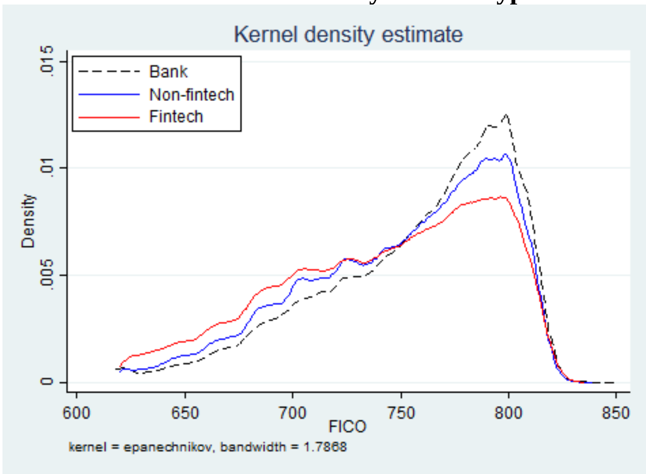


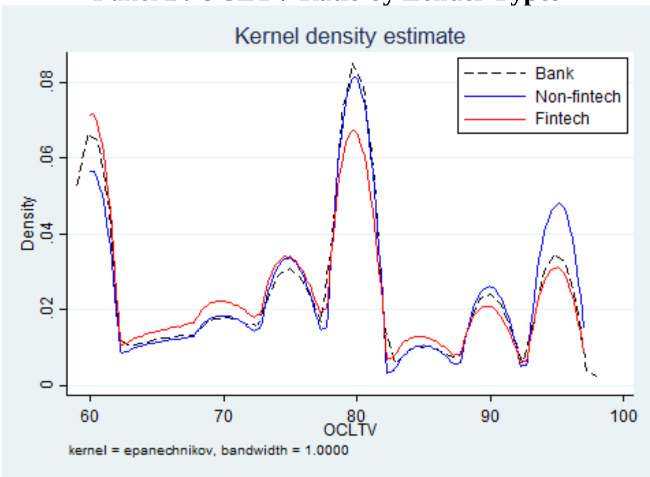
Figure 2. Kernel density plots of FICO, OCLTV and DTI at origination by lender types

These graphs show the kernel density plots of FICO score, OCLTV, and DTI for GSE 30-year FRMs originated from January2010 to March 2018.

Panel A: FICO Score by Lender Types



Panel B: OCLTV Ratio by Lender Types



Panel C: DTI Ratio by Lender Types

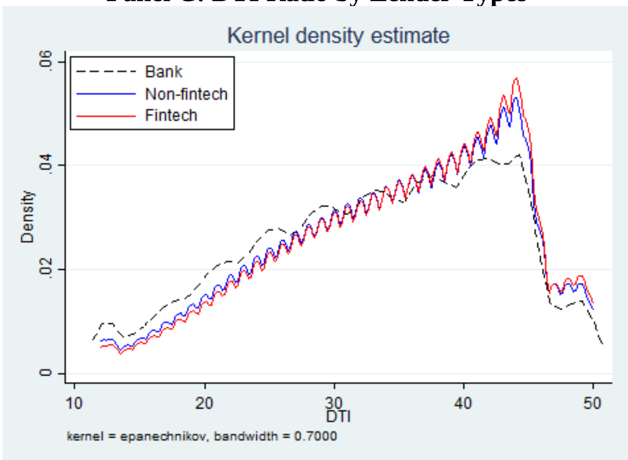
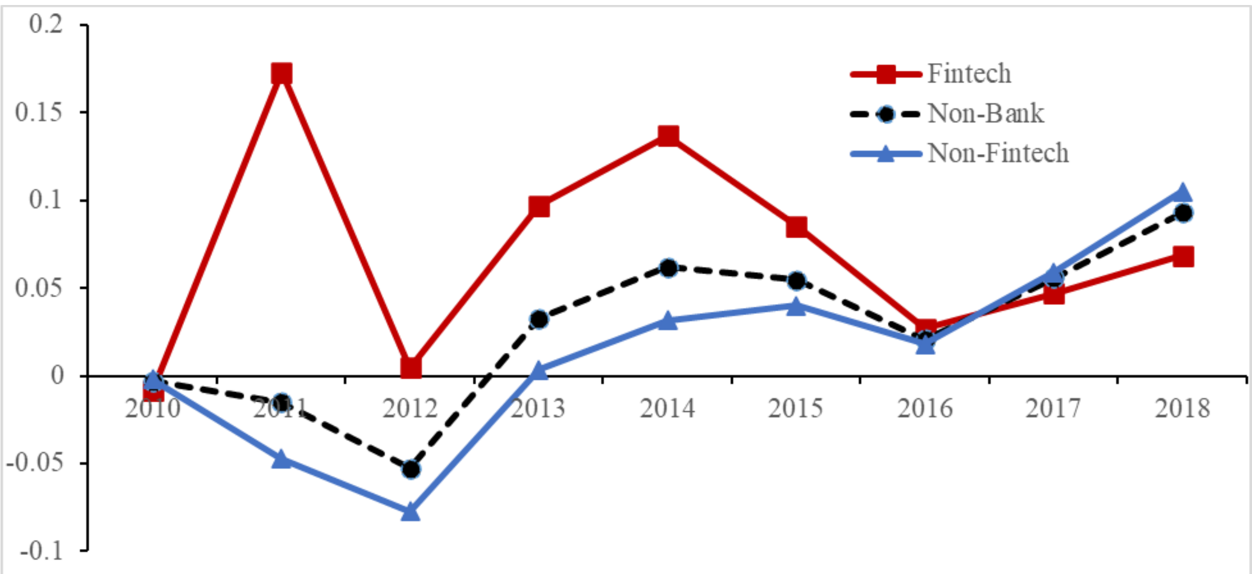


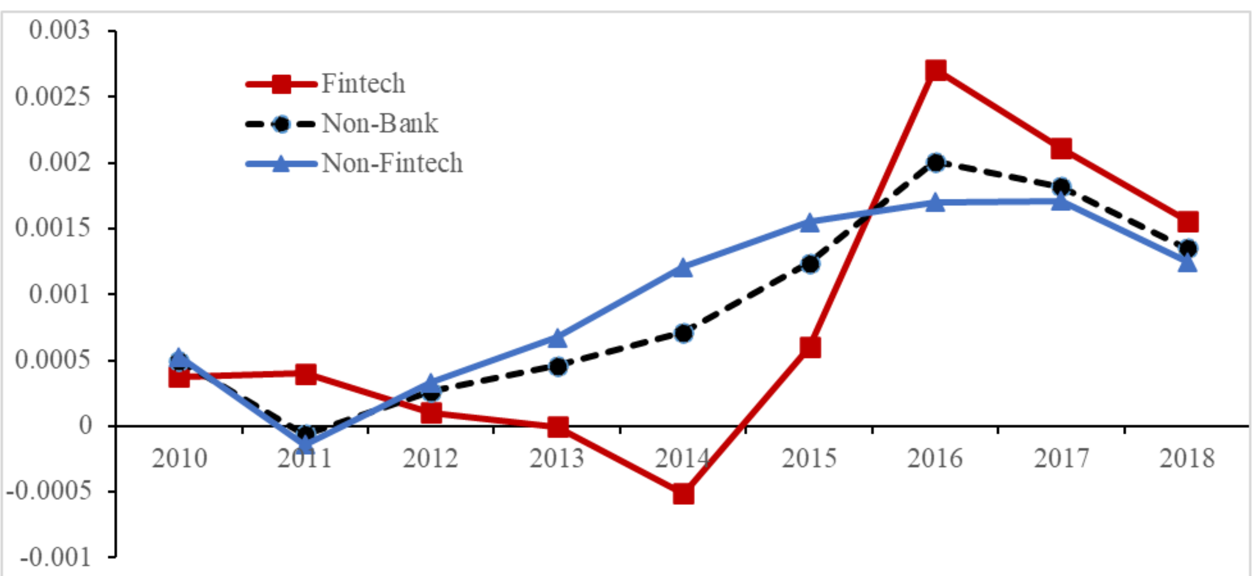
Figure 3. Yearly Coefficients of Loan Level Regression

These graphs report the yearly regression coefficients on non-bank, fintech, and non-fintech dummies from the models specified in Columns 2 and 4 of Table 2. Panels A, B, and C report the regression coefficients with the origination interest rate, delinquency rate, and prepayment rate as the LHS variable (corresponding to Panels A-C in Table 2), respectively. Delinquency is identified if the loan becomes delinquent (90 DPD) within 24 months after loan origination, and prepayment is identified if the loan is prepaid in 24 months post loan origination.

Panel A: Yearly Coefficients for the Origination Interest Rate Regression



Panel B: Yearly Coefficients for the Delinquency Regression



Panel C: Yearly Coefficients for the Prepayment Regression

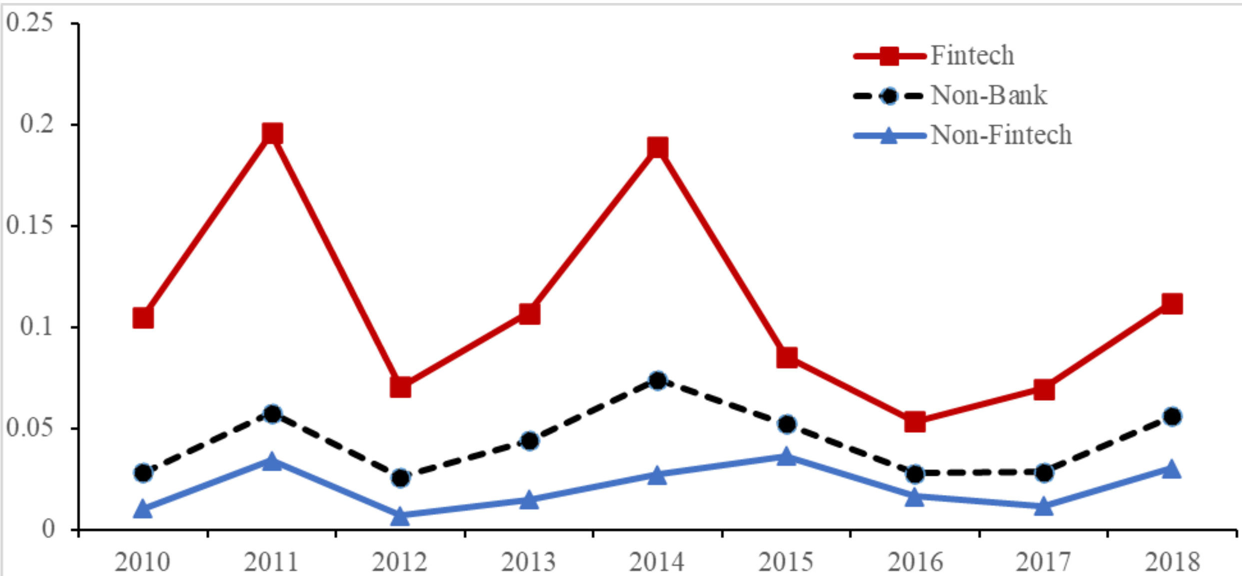


Figure 4. Loan Characteristics and Performance by Interest Rate Deciles

The following graphs display the differences in interest rates, FICO, DTI, OCLTV, delinquency rates and prepayment rates between non-banks and banks by each interest rate decile (i.e., banks minus non-banks). Specifically, we calculate the differences in interest rate (Panel A), FICO (Panel B), DTI (Panel C), OCLTV (Panel D), delinquency rates (Panel E and F) and prepayment rates (Panel G and H) between non-banks and banks by each interest rate decile for each origination month and then take the average of the time-series difference numbers. Please refer to section 4.2.1 for details on the portfolio construction.

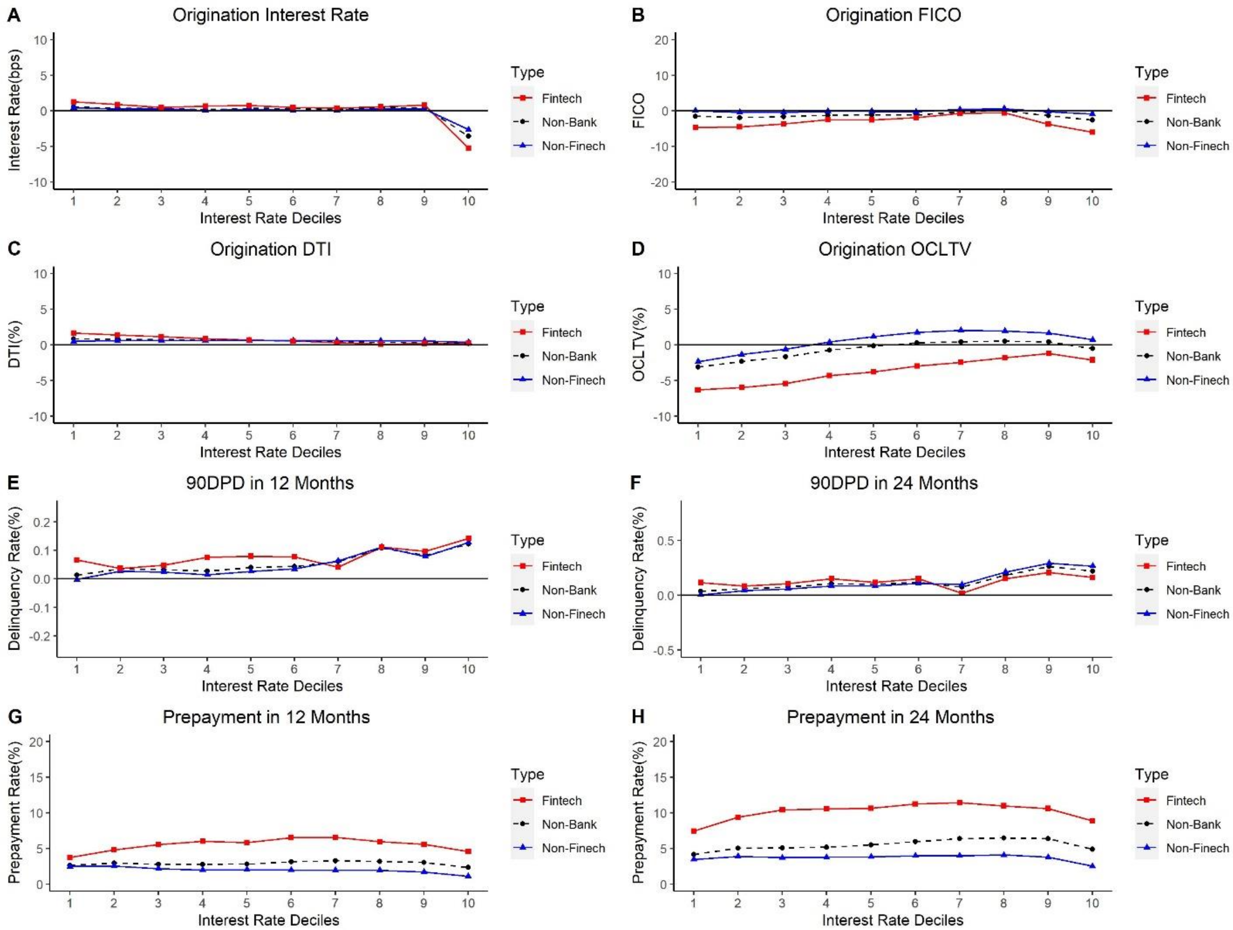


Figure 5. Simulated Over-Time Weighted Average Interest Rates of a Hypothetical Portfolio of Loans

We first construct four historical portfolios of GSE loans, corresponding to the four lender types, and estimate a regression model as expressed in equation 3. We then simulate monthly interest rates for the four hypothetical portfolios for large bank, small bank, fintech, and non-fintech, non-banks using average historical values of the overall portfolio characteristics and default/prepayment rates from 2010 to 2019. We then calculate the weighted average interest rate for each month for the hypothetical market portfolio using the over-time market share of the four lender types as the weights. The variation in the simulated weighted average interest rates of the hypothetical market portfolio depicted below is driven by the changes in market share of the four lender types over time, as the portfolio characteristics and default/prepayment rates are the same across the lender types and over time.

Table 1. Summary of Statistic by Lender Types

This table reports the summary statistics of 30-year GSE fixed-rate mortgages originated from Jan. 2010 to Mar. 2018.

Panel A: Loan Origination Characteristic

	Banks		Non-Banks		Non-Fintech		Fintech	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Origination Rate	4.33	0.53	4.21	0.46	4.19	0.47	4.26	0.44
Origination Balance ($ Thousand)	239.63	130.70	255.05	128.84	260.80	127.63	241.80	130.63
Original Loan-to-Value Ratio (OLTV)	72.73	16.29	74.35	15.82	75.15	15.74	72.51	15.87
Original Combined LTV(OCLTV)	73.74	16.04	74.91	15.66	75.76	15.55	72.97	15.76
Debt-to-Income Ratio	33.32	9.38	34.80	9.01	34.67	9.05	35.11	8.91
FICO Score	757.48	43.39	748.78	46.33	751.24	45.09	743.12	48.62
Refinance	0.56	0.50	0.57	0.50	0.50	0.50	0.74	0.44
Cash Out Refinance	0.21	0.41	0.26	0.44	0.22	0.41	0.35	0.48
Non-Cash Out Refinance	0.35	0.48	0.32	0.47	0.28	0.45	0.39	0.49
Purchase	0.44	0.50	0.43	0.50	0.50	0.50	0.26	0.44
First Time Home Buyer	0.16	0.37	0.17	0.37	0.20	0.40	0.11	0.31
Number of Borrowers	1.55	0.51	1.50	0.51	1.51	0.51	1.48	0.51
Has Mortgage Insurance	0.20	0.40	0.26	0.44	0.28	0.45	0.23	0.42
Mortgage Insurance Unknown	0.38	0.49	0.45	0.50	0.43	0.50	0.49	0.50
Primary Residence	0.87	0.33	0.88	0.33	0.87	0.34	0.89	0.31
Investment or 2nd Property	0.13	0.33	0.12	0.33	0.13	0.34	0.11	0.31
MSA Unemployment Rate	7.31	2.58	5.80	2.18	5.86	2.24	5.65	2.01
MSA Real Personal Income ($ Thousand)	46.25	6.84	47.88	7.18	47.76	7.10	48.17	7.35
CHG MSA UNEMP RTE	-0.49	0.87	-0.73	0.59	-0.73	0.60	-0.73	0.56
CHG MSA HPI	0.01	0.06	0.047	0.046	0.047	0.048	0.048	0.042
Correspondent Channel	0.55	0.50	0.34	0.47	0.47	0.50	0.03	0.17
Retail Channel	0.40	0.49	0.43	0.50	0.24	0.43	0.86	0.34
Broker Channel	0.05	0.21	0.24	0.43	0.30	0.46	0.11	0.31
Number of Loans	5,199,954		2,126,803		1,482,348		644,455	

Panel B: Loan Performance

	Banks		Non-Banks		Non-Fintech		Fintech	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
30 DPD in 12 Months	2.20%	14.50%	3.20%	17.50%	3.50%	18.30%	2.50%	15.50%
30 DPD in 24 Months	4.00%	19.60%	5.90%	23.50%	6.40%	24.50%	4.60%	21.00%
30 DPD in 36 Months	5.20%	22.20%	7.30%	26.00%	8.00%	27.20%	5.60%	22.90%
60 DPD in 12 Months	0.30%	5.40%	0.50%	7.00%	0.50%	6.90%	0.50%	7.10%
60 DPD in 24 Months	0.80%	8.80%	1.20%	11.00%	1.20%	10.90%	1.30%	11.20%
60 DPD in 36 Months	1.20%	10.90%	1.70%	12.80%	1.60%	12.70%	1.70%	12.80%
90 DPD in 12 Months	0.10%	3.70%	0.20%	4.90%	0.20%	4.70%	0.30%	5.40%
90 DPD in 24 Months	0.40%	6.60%	0.70%	8.30%	0.60%	8.00%	0.80%	8.90%
90 DPD in 36 Months	0.70%	8.50%	1.00%	9.90%	0.90%	9.60%	1.10%	10.50%
Prepaid in 12 Months	6.60%	24.80%	7.90%	27.00%	7.30%	26.00%	9.30%	29.00%
Prepaid in 24 Months	17.80%	38.20%	19.40%	39.60%	18.00%	38.40%	22.70%	41.90%
Prepaid in 36 Months	28.80%	45.30%	30.00%	45.80%	28.30%	45.00%	34.00%	47.40%

Table 2: Loan Level Regression of Origination Interest Rate, Delinquency, and Prepayment

This table reports the loan level regression results with the dependent variables being the origination interest rates (Panel A), delinquency event (Panel B), and prepayment event (Panel C). The control variables in all three panels are observed variables at loan origination, including origination loan amount, number of borrowers, MSA level unemployment rate, dummy variables for cash out refinancing, rate refinancing, first-time home buyer, mortgage insurance, missing information for mortgage insurance, GSE indicator, correspondence channel, retail channel, step functions for FICO (cut-off points at 640, 660, 680, 700, 720, 740, 760, and 780), OCLTV (cut-off points at 60, 70, 75, 80, 85, 90, and 95) and DTI (cut-off points at 20, 30, and 40). Additionally, the spread at origination (SATO) is controlled for delinquency and prepayment regressions. It signifies the maximum difference between the origination interest rate of a loan and the PMMS30 rate at the time of origination, observed over the 24 months following origination. Standard errors are reported in the parenthesis. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Interest Rate Regression

	(1)	(2)	(3)	(4)
	Interest rate (%)			
Non-Bank	0.0287***	0.0271***		
	(0.00101)	(0.000994)		
Non-Fintech			0.0216***	0.0189***
			(0.00110)	(0.00108)
Fintech			0.0454***	0.0464***
			(0.00122)	(0.00119)
Borrower and Loan Controls	Yes	Yes	Yes	Yes
Zip x Quarter FE	No	Yes	No	Yes
Quarter FE	Yes	No	Yes	No
Adjust R-sq	0.667	0.681	0.667	0.681
N	7,270,207	7,270,207	7,270,207	7,270,207

Panel B: Delinquency Regression

	(1)	(2)	(3)	(4)
	90+ DPD within 24 months since origination			
Non-Bank	0.000950***	0.00100***		
	(0.0000891)	(0.0000860)		
Non-Fintech			0.000959***	0.00110***
			(0.0000915)	(0.0000870)
Fintech			0.000929***	0.000774***
			(0.000155)	(0.000150)
Borrower and Loan Controls	Yes	Yes	Yes	Yes
Zip x Quarter FE	No	Yes	No	Yes
Quarter FE	Yes	No	Yes	No
Adjust R-sq	0.0111	0.0191	0.0111	0.0191
N	7270189	7270189	7270189	7270189

Panel C: Prepayment Regression

	(1)	(2)	(3)	(4)
	Prepaid within 24 months since origination			
Non-Bank	0.0394***	0.0390***		
	(0.000773)	(0.000552)		
Non-Fintech			0.0186***	0.0168***
			(0.000800)	(0.000581)
Fintech			0.0887***	0.0909***
			(0.00112)	(0.000935)
Borrower and Loan Controls	Yes	Yes	Yes	Yes
Zip x Quarter FE	No	Yes	No	Yes

Quarter FE	Yes	No	Yes	No
Adjust R-sq	0.0790	0.113	0.0808	0.115
N	7270189	7270189	7270189	7270189

Table 3. Pricing Analysis Using Interest Rate Sorted Portfolios (Banks vs Non-banks)

This table reports the regression results by the interest rate sorted portfolios. To construct such portfolios, for each month, we sort all loans originated in the month into ten deciles based on origination interest rates. Panel A constructs portfolios by banks and non-banks, and Panel B is based on the three sets of time series of portfolios of banks, non-fintech and fintech loans by origination month and decile. The dependent variable is portfolio interest rates. Column 1 reports results across all interest rate deciles, and Columns 2 and 3 show results for the lower (five low interest rate deciles) and upper (five high interest rate deciles) interest rate deciles, separately. Columns 4 and 5 report results by interest rate portfolios formed among borrowers with credit bureau scores <720 and >720, respectively. Portfolio delinquency rate is the percentage of delinquent loans (90 DPD) within 24 months post loan origination, and portfolio prepayment rate is the proportion of loan prepaid in 24 months after loan origination. PMMS30 is the Primary Mortgage Market Survey for 30-Year Fixed-Rate from Freddie Mac. Standard errors are reported in the parenthesis. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Unreported test statistics show that the coefficient estimates of the fintech dummy are significantly different from the coefficient estimates of the non-fintech dummy in all columns of Panel B of Table 3.

Panel A: Non-Banks v.s. Banks

	(1)	(2)	(3)	(4)	(5)
	All Loans	Low Interest Rates (Bottom Five Interest Rate Deciles)	High Interest Rates (Top Five Interest Rate Deciles)	FICO <720	FICO>=720
Non-Bank	-0.223***	-0.157***	-0.244***	-0.318***	-0.226***
	(0.00918)	(0.0113)	(0.0139)	(0.0136)	(0.0117)
PMMS30	0.361***	0.405***	0.414***	0.332***	0.304***
	(0.0176)	(0.0222)	(0.0264)	(0.0228)	(0.0228)
Prepayment Rate	3.014***	2.574***	2.888***	4.167***	3.297***
	(0.0777)	(0.0917)	(0.131)	(0.122)	(0.0910)
Delinquency Rate	45.34***	42.28***	37.27***	22.29***	76.65***
	(0.875)	(2.835)	(1.120)	(0.746)	(2.700)
Adjust R-sq	0.885	0.845	0.855	0.786	0.784
N	1,260	630	630	1,260	1,260

Panel B: Fintech and Non-Fintech v.s. Banks

	(1)	(2)	(3)	(4)	(5)
	All Loans	Low Interest Rates (Bottom Five Interest Rate Deciles)	High Interest Rates (Top Five Interest Rate Deciles)	FICO <720	FICO>=720
Fintech	-0.298***	-0.209***	-0.285***	-0.379***	-0.287***
	(0.0114)	(0.0132)	(0.0171)	(0.0172)	(0.0144)
Non-Fintech	-0.143***	-0.0835***	-0.150***	-0.181***	-0.119***
	(0.0100)	(0.0115)	(0.0139)	(0.0141)	(0.0128)
PMMS30	0.377***	0.446***	0.463***	0.379***	0.360***
	(0.0170)	(0.0200)	(0.0248)	(0.0221)	(0.0220)
Prepayment Rate	2.384***	1.880***	2.177***	3.156***	2.375***
	(0.0585)	(0.0633)	(0.100)	(0.0930)	(0.0695)
Delinquency Rate	43.03***	29.38***	33.41***	19.20***	57.60***
	(0.751)	(1.871)	(0.966)	(0.596)	(2.022)
Adjust R-sq	0.836	0.808	0.802	0.694	0.697
N	1,890	945	945	1,890	1,890

Table 4. Pricing Analysis Using Interest Rate Sorted Portfolio – Subperiods

This table reports the sub-period results corresponding to the first column of Table 3 – interest rate decile regression. To construct such portfolios, for each month, we sort all loans originated in the month into ten deciles based on origination interest rates. Panel A reports the results of bank loans versus non-bank loans, while Panel B presents the results of banks vs fintech and non-fintech non-banks loans. The dependent variable is portfolio interest rates. Column 1, 2, 3 represent the subperiod of 2013-2014, 2015-2016, and 2017-2018, respectively. Portfolio delinquency rate is the percentage of delinquent loans (90 DPD) within 24 months. Portfolio prepayment rate is the proportion of loan prepaid in 24 months. PMMS30 is the Primary Mortgage Market Survey for 30-Year Fixed-Rate from Freddie Mac. Standard errors are reported in the parenthesis. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Unreported test statistics show that the coefficient estimates of the fintech dummy are significantly different from the coefficient estimate of the non-fintech dummy in all columns of Panel B of Table 4.

Panel A: Banks vs Non-Banks by Subperiods

	(1) 2013-2014	(2) 2015-2016	(3) 2017-2018
Non-bank	-0.228***	-0.286***	-0.155***
	(0.0155)	(0.0117)	(0.0142)
PMMS30	0.315***	0.160***	0.142***
	(0.0283)	(0.0288)	(0.0501)
Prepayment Rate	2.608***	4.617***	2.277***
	(0.123)	(0.127)	(0.188)
Delinquency Rate	65.93***	33.73***	44.99***
	(2.016)	(1.253)	(1.495)
Adjust R-sq	0.911	0.907	0.905
N	480	480	300

Panel B: Banks vs Fintech and Non-Fintech Non-Banks by Subperiods

	(1) 2013-2014	(2) 2015-2016	(3) 2017-2018
Fintech	-0.271***	-0.356***	-0.308***
	(0.0225)	(0.0135)	(0.0176)
Non-Fintech	-0.143***	-0.225***	-0.0880***
	(0.0173)	(0.0123)	(0.0153)
PMMS30	0.370***	0.146***	0.0489
	(0.0297)	(0.0268)	(0.0453)
Prepayment Rate	1.899***	4.358***	2.699***
	(0.0953)	(0.106)	(0.159)
Delinquency Rate	59.39***	30.29***	39.60***
	(1.846)	(1.027)	(1.245)
Adjust R-sq	0.844	0.877	0.876
N	720	720	450

Table 5. Pricing Analysis Using Interest Rate Sorted Portfolios of Large and Small Banks and Fintech and Non-Fintech Lenders

We pool four sets of time series of interest rate sorted portfolios in this regression. To construct such portfolios, for each month, we sort all loans originated in the month into ten deciles, and we then construct large bank, small bank, fintech, and non-fintech time series by decile-origination month. We divided the banks loans into two subsamples – large (CCAR) banks and small (non-CCAR) banks. The dependent variable is portfolio interest rates. Portfolio delinquency rate is the percentage of delinquent loans (90 DPD) within 24 months. Portfolio prepayment rate is the proportion of loan prepaid in 24 months. PMMS30 is the Primary Mortgage Market Survey for 30-Year Fixed-Rate from Freddie Mac. Standard errors are reported in the parenthesis. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Unreported test statistics show that the coefficient estimates of the fintech dummy are significantly different from the coefficient estimates of the non-fintech dummy in all columns of Table 6.

	(1)	(2)	(3)	(4)	(5)
	All Loans	Low Interest Rates (Bottom Five Interest Rate Deciles)	High Interest Rates (Top Five Interest Rate Deciles)	FICO <720	FICO>=720
Large Bank	0.0959***	0.0678***	0.0853***	0.136***	0.0756***
	(0.0113)	(0.0117)	(0.0156)	(0.0157)	(0.0137)
Fintech	-0.204***	-0.139***	-0.175***	-0.219***	-0.197***
	(0.0117)	(0.0123)	(0.0164)	(0.0163)	(0.0143)
Non-Fintech	-0.0490***	-0.0177	-0.0515***	-0.0280*	-0.0398***
	(0.0111)	(0.0116)	(0.0151)	(0.0152)	(0.0136)
PMMS30	0.407***	0.485***	0.515***	0.435***	0.401***
	(0.0165)	(0.0173)	(0.0237)	(0.0208)	(0.0203)
Prepayment Rate	2.396***	1.872***	1.958***	2.874***	2.338***
	(0.0584)	(0.0575)	(0.0964)	(0.0846)	(0.0662)
Delinquency Rate	36.98***	22.65***	26.59***	12.81***	40.99***
	(0.695)	(1.546)	(0.850)	(0.477)	(1.699)
Adjust R-sq	0.791	0.797	0.748	0.616	0.648
N	2,520	1,260	1,260	2,520	2,520

Table 6- Pricing Analysis Using Interest Rate Sorted Portfolios of Large and Small Banks and Fintech and Non-Fintech Lenders – Sub-periods

This table shows subperiod analysis for the first column of Table 5. Portfolio delinquency rate is the percentage of delinquent loans (90 DPD) within 24 months. Portfolio prepayment rate is the proportion of loan prepaid in 24 months. PMMS30 is the Primary Mortgage Market Survey for 30-Year Fixed-Rate from Freddie Mac. Standard errors are reported in the parenthesis. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Unreported test statistics show that the coefficient estimates of the fintech dummy are significantly different from the coefficient estimates of the non-fintech dummy in all columns of Table 6.

	(1) 2013-2014	(2) 2015-2016	(3) 2017-2018
Large Bank	0.0685***	0.208***	0.0980***
	(0.0192)	(0.0151)	(0.0195)
Fintech	-0.230***	-0.145***	-0.225***
	(0.0223)	(0.0145)	(0.0190)
Non-Fintech	-0.0777***	-0.0208	0.00528
	(0.0189)	(0.0144)	(0.0191)
PMMS30	0.398***	0.159***	-0.0257
	(0.0290)	(0.0283)	(0.0469)
Prepayment Rate	2.004***	4.152***	3.213***
	(0.0953)	(0.104)	(0.161)
Delinquency Rate	48.13***	26.50***	31.29***
	(1.608)	(0.982)	(1.198)
Adjust R-sq	0.801	0.818	0.818
N	960	960	600

Table 7. Pricing Analysis Using Interest Rate Sorted Portfolio – FICO Group, Bank Size and Subperiods

This table reports the FICO subgroup analysis corresponding to Table 6. The dependent variable is portfolio interest rates. Portfolio delinquency rate is the percentage of delinquent loans (90 DPD) within 24 months since loan origination, and portfolio prepayment rate is the proportion of loan prepaid in 24 months since loan origination. PMMS30 is the Primary Mortgage Market Survey for 30-Year Fixed-Rate from Freddie Mac. In addition to these reported variables, we control for all other loan origination variables. Standard errors are reported in the parenthesis. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Unreported test statistics show that the coefficient estimates of the fintech dummy are significantly different from the coefficient estimates of the non-fintech dummy in all columns of Table 7.

	2013-2014		2015-2016		2017-2018	
	FICO<720	FICO>=720	FICO<720	FICO>=720	FICO<720	FICO>=720
Large Bank	0.0570**	0.0762***	0.284***	0.186***	0.252***	0.0694***
	(0.0261)	(0.0227)	(0.0229)	(0.0183)	(0.0291)	(0.0249)
Fintech	-0.289***	-0.258***	-0.194***	-0.116***	-0.211***	-0.234***
	(0.0309)	(0.0258)	(0.0219)	(0.0179)	(0.0279)	(0.0260)
Non-Fintech	-0.0881***	-0.0414*	0.0246	-0.0151	0.103***	-0.0174
	(0.0260)	(0.0223)	(0.0214)	(0.0176)	(0.0277)	(0.0251)
PMMS30	0.426***	0.430***	0.171***	0.113***	-0.154**	-0.139**
	(0.0367)	(0.0341)	(0.0405)	(0.0342)	(0.0623)	(0.0607)
Prepayment Rate	2.385***	2.060***	4.515***	4.272***	4.628***	3.884***
	(0.134)	(0.106)	(0.147)	(0.118)	(0.211)	(0.179)
Delinquency Rate	13.06***	35.18***	9.575***	28.26***	11.41***	39.13***
	(0.967)	(3.177)	(0.667)	(2.402)	(0.833)	(2.896)
Adjust R-sq	0.634	0.694	0.624	0.670	0.638	0.620
N	960	960	960	960	600	600

Appendix A Literature review: non-banks' rise in the residential mortgage market

Scholars have explored how non-banks have successfully penetrated the residential mortgage market. The first explanation proposed in the literature is non-banks' expansion into riskier segments. Buchak et al (2020a) finds that non-banks gain their share of the residential mortgage market mainly by targeting "riskier, less creditworthy Federal Housing Administration (FHA) borrowers and areas with larger minority populations." This view is also shared by other studies, such as Lux and Greene (2015). The potential regulations that may drive non-banks' penetration in the riskier market segment that are cited in the literature include the risk-based capital requirements for mortgages held on balance sheet,[19] Qualified Mortgage (QM) Rule,[20] the harsher regulatory treatment for mortgage servicing assets (MSA) under the new Basel III capital rules,[21] mortgage-related lawsuits, etc. The majority of these regulations make banks less willing to lend to riskier borrowers, and not surprisingly, both Lux and Greene (2015) and Buchak et al (2020a) use these regulations to explain non-banks' rise in the risky FHA market.

Nevertheless, the specific rules cited above cannot well explain the growing market share by non-bank mortgage originators in the GSE mortgage market. The risk-based capital requirements affect balance sheet mortgage lending by banks but not mortgages originated to sell. In addition, the QM rule should not apply to GSE mortgages, since GSE mortgages are exempted from the QM rule during our sample period.

---

[19] See https://www.fdic.gov/resources/regulations/federal-register-publications/2007/07basel2dec7.pdf.

[20] The Ability-to-Repay/Qualified Mortgage Rule (ATR/QM Rule) requires a creditor to make a reasonable, good faith determination of a consumer's ability to repay a residential mortgage loan according to its terms. See https://www.consumerfinance.gov/rules-policy/final-rules/ability-to-pay-qualified-mortgage-rule/ for more details.

[21] Under the Basel III framework that was first proposed in 2012 and was approved in 2013 (Final Rule: Basel III, 2014), the risk weights applied to mortgage servicing rights (MSRs) for regulatory capital calculations increased from 100 percent to 250 percent. For summary of the latest updates to the rule, see websites such as https://blogs.claconnect.com/financialinstitutions/new-capital-rule-changes-treatment-of-mortgage-servicing-rights/.

The MSA capital rule affects bank mortgage servicers directly, but not necessarily bank mortgage origination, which is the focus of our investigation. Of course, bank mortgage origination decisions might be affected by the MSA capital rule if banks reduce origination because of their lack of willingness to service these loans. However, banks, and especially large banks, are typically well under the capital rule cutoff for MSAs since 2012 according to a study by federal regulators,[22] which is confirmed by our own calculations. As a result, it is unlikely that the MSA capital rule is the driving factor behind non-banks' growth in the GSE market.

The literature has further suggested that the mortgage lawsuits and legal fears may drive banks to become more selective about borrowers and less willing to lend to riskier borrowers, but this concern should be negligible in the GSE mortgage market. The underwriting standards of these loans are largely set by the agencies and are transparent, and lenders typically underwrite these loans after making sure that these loans meet the underwriting standard and are eligible to be acquired by Fannie or Freddie after origination. In addition, the reps and warrants practices, which generally allow the agencies (the buyers of the loans) to put the loan back to the lender if they find that the loan does not meet the agency standards (Davidson and Levin (2014)), apply to both banks and non-banks. Moreover, if banks were concerned with mortgage lawsuits, they would be more likely to exit the low FICO segments of the GSE mortgage market, and we can test this hypothesis using the empirical data.

We plot the over-time changes in the market sizes of different FICO segments of the GSE mortgage market in Figure A. Since we study 30-year fixed rate agency mortgages in this paper,

[22] https://www.federalreserve.gov/publications/other-reports/files/effect-capital-rules-mortgage-servicing-assets-201606.pdf.

we only plot in Figure A the market share of the 30-year fixed rate mortgages (FRMs) by different types of lenders in various FICO segments.[23]

Panel A of Figure A shows that the origination volume of mortgages in the segment with FICO<680 by banks has been largely stable, and the entrance of non-banks has been responsible for the increase in overall volumes of mortgages in this segment. This panel thus indicates that non-banks are more likely to offer mortgages to risky borrowers who might be under-served before. Panel B of Figure A depicts a slight decline in bank volume over time. We can also clearly see non-banks' expansion to new borrowers in this FICO segment, which nearly doubles the total volume of GSE mortgages in this segment during the period 2010-2019. By contrast, the total size of the FICO >720 market in Panel C of Figure A does not experience much increase, and the market share of banks has been clearly declining, relative to both non-fintech and fintech lenders. Since loans with FICO >720 make up the lion's share of the GSE mortgage market, we can conclude from Figure A that non-banks' penetration in the GSE market is primarily featured by non-banks' rising market share among the high FICO segments, supplemented by non-banks' expansion into the low FICO segments that might be under-served before by banks.

If banks reduce mortgage originations because of lack of willingness to lend to riskier borrowers as argued in the existing literature, they are more likely to reduce origination in the relatively riskier segment, for example, the FICO <680 segment. Nevertheless, the opposite has been happening in the GSE loan market. In the FICO>720 segment, a market segment banks should be least worried about from the perspective of riskiness and legal concerns, the total volume

[23] Results using all mortgages in the Fannie and Freddie public data are very similar to those reported in the three panels of Figure A and are available upon request.

originated by banks has been clearly shrinking, a phenomenon that cannot be explained by the risky-segment expansion and regulation concern story proposed in the literature.

The second factor identified by the literature driving the competitiveness of non-banks in the residential mortgage market is the relatively higher utilization of information technology (IT). However, the findings in Buchak et al (2020a) are mainly driven by technology innovation among the fintech residential mortgage lenders. Similarly, Fuster et al. (2018) finds that fintech lenders can process mortgage applications noticeably faster than other lenders, but the evidence in that paper is again restricted to fintech lenders. We can see from Figures 1 and A that the proportions of GSE mortgages originated by both fintech and non-fintech non-banks have been climbing noticeably post 2010. Further, the market share of non-fintech non-banks overshadows that of fintech lenders substantively in both Figures 1 and A and Appendix B, but neither factor above can explain the rise of non-fintech non-bank market share in the GSE mortgage market.

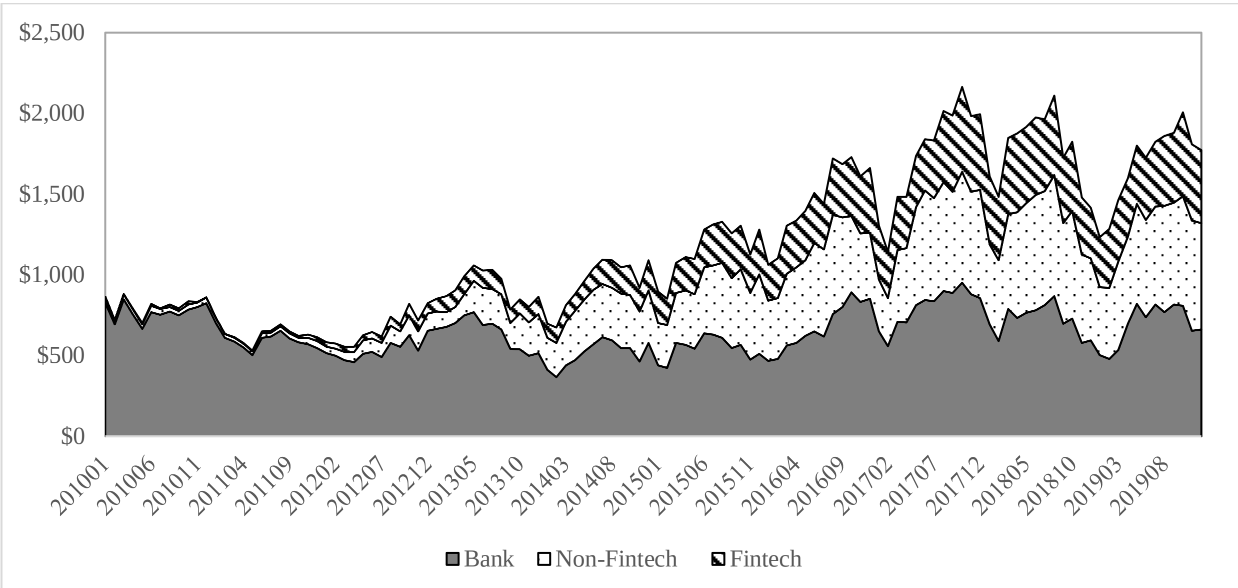
Finally, the literature has studied pricing of the GSE mortgages by banks and non-banks, as the conventional wisdom in the existing literature predicts that new entrants in the consumer lending market should lower the borrowing costs (see, for example, Ausubel (1991) and Philippon (2015), Stiglitz and Weiss (1981),(Brito and Hartley (1995), Ausubel (1991), and Thakor (2021)). Nevertheless, collaborative evidence for such a prediction has been elusive.

In summary, the literature has not documented pricing as a major factor behind the growth of non-banks in the GSE mortgage market, especially after 2013 and among fintech firms.

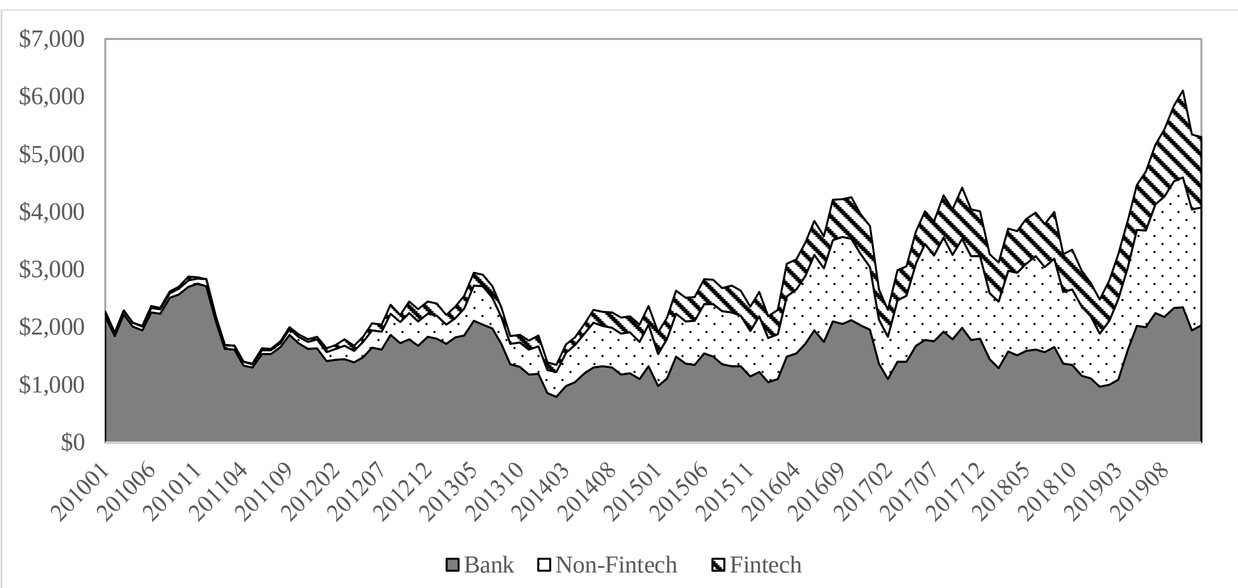
Figure A. Origination Volume by Lender Type for Different FICO Segmentations

The following figures show the origination volume of the 30-year fixed rate mortgages by different types of lenders in various FICO segments based on the public Freddie and Fannie data.

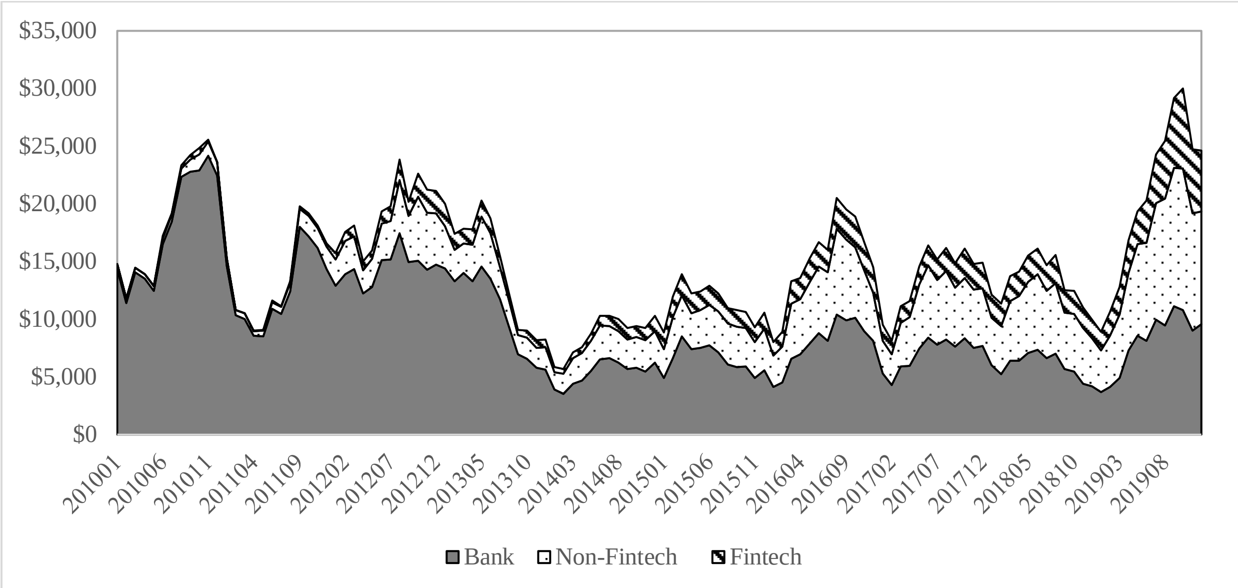
Panel A: Origination Volume by Lender Type for FICO < 680 (in $MM)



Panel B: Origination Volume by Lender Type for FICO 680-720 (in $MM)



Panel C: Origination Volume by Lender Type for FICO above 720 (in $MM)



Appendix B. Market Share by Loan Types

This analysis is based on the public Fannie and Freddie data for loans originated from 2011 to 2019. Mortgages originated by lenders that cannot be identified (labelled as "other" in the datasets) are excluded from the table. We can identify the originators of each loan from the public data, as long as the originator name is not listed as "other lender." We check whether each lender is a bank or subsidiary of a bank. If a lender does not have a RSSD and a bank charter or it is does not have a bank parent, it is considered a non-bank. Classification of fintech lenders is more judgmental, and we follow the approach adopted by earlier studies, such as Buchak et al (2020a) and Fuster (2019). We check the mortgage lending process of the non-banks. If the lending process up till preapproval can be completed entirely on-line, and the lender's business is generated predominantly online, the lender is deemed a fintech lender.

	Bank		**Non-Fintech**		**Fintech**		**All**
Origination Year	Percentage	$Amount (MM)	Percentage	$Amount (MM)	Percentage	$Amount (MM)	$Amount (Total)
2011	94.99%	$ 564,778.55	4.05%	$ 24,097.15	0.95%	$ 5,674.18	$ 594,549.88
2012	91.61%	$ 427,024.01	7.09%	$ 33,047.84	1.30%	$ 6,038.97	$ 466,110.83
2013	77.51%	$ 502,232.61	16.39%	$ 106,233.64	6.10%	$ 39,514.23	$ 647,980.48
2014	74.62%	$ 382,309.77	16.85%	$ 86,310.07	8.54%	$ 43,756.71	$ 512,376.55
2015	61.56%	$ 186,566.46	26.89%	$ 81,479.64	11.55%	$ 35,001.82	$ 303,047.92
2016	56.44%	$ 223,877.24	28.16%	$ 111,705.00	15.40%	$ 61,074.78	$ 396,657.02
2017	52.04%	$ 279,173.29	32.09%	$ 172,162.00	15.87%	$ 85,117.32	$ 536,452.61
2018	51.27%	$ 243,110.01	32.34%	$ 153,345.95	16.39%	$ 77,717.18	$ 474,173.14
2019	45.40%	$ 206,007.86	37.42%	$ 169,773.44	17.18%	$ 77,958.36	$ 453,739.66
Total	68.76%	$ 3,015,079.80	21.39%	$ 938,154.73	9.85%	$ 431,853.55	$ 4,385,088.08

Appendix C. Top 20 Lenders Ranked by Origination Volume

This table report the top 20 lenders in the confirming residential mortgage market based on the public Fannie and Freddie data. We can identify the originators of each loan from the public data, as long as the originator name is not listed as “other lender.” We check whether each lender is a bank or subsidiary of a bank. If a lender does not have a RSSD and a bank charter or it is does not have a bank parent, it is considered a non-bank. Classification of fintech lenders is more judgmental, and we follow the approach adopted by earlier studies, such as Buchak et al (2020a) and Fuster (2019). We check the mortgage lending process of the non-banks. If the lending process up till preapproval can be completed entirely on-line, and the lender’s business is generated predominantly online, the lender is deemed a fintech lender. The ranking is determined by the origination volume of the GSE residential mortgage issued by the lender from 2012 to 2019.

Rank	Lender Name	Lender Type	2012-2019 Origination Volume ($MM)
1	Wells Fargo Bank	Bank	468,627
2	JPMorgan Chase Bank	Bank	184,085
3	Quicken Loans	Fintech Non-Bank	177,101
4	U.S. Bank	Bank	103,964
5	United Shore Financial Services	Non-Fintech Non-Bank	87,689
6	PennyMac Loan Services	Non-Fintech Non-Bank	66,860
7	Caliber Home Loans	Non-Fintech Non-Bank	53,824
8	Flagstar Bank	Bank	51,273
9	AmeriHome Mortgage	Non-Fintech Non-Bank	43,121
10	SunTrust Mortgage	Bank	40,587
11	Franklin American Mortgage	Non-Fintech Non-Bank	38,479
12	Branch Banking and Trust	Bank	37,310
13	Stearns Lending	Non-Fintech Non-Bank	36,770
14	LoanDepot.com LLC	Fintech Non-Bank	34,003
14	Stearns Lending	Non-Fintech Non-Bank	36,770
15	CitiMortgage Inc	Bank	30,663
16	Nationstar Mortgage	Non-Fintech Non-Bank	29,734
17	Bank of America	Bank	28,199
18	Fairway Independent Mortgage	Non-Fintech Non-Bank	25,070
19	Provident Funding Associates	Non-Fintech Non-Bank	23,661
20	Freedom Mortgage	Non-Fintech Non-Bank	16,790

Appendix D. Variable definition

Variable	Definition
Borrower and loan Characteristic	
Origination Rate	The original interest rate on a mortgage loan as identified in the original mortgage note
Origination Balance ($ Thousand)	The dollar amount of the loan as stated on the note at the time the loan was originated
Original Loan-to-Value Ratio (OLTV)	Loan amount divided by the value of property at origination
Original Combined LTV(OCLTV)	The amount of all known outstanding loans (including home equity) at origination divided by the value of property
Debt-to-Income Ratio	Loan amount divided by borrower income at origination
FICO Score	Borrower's FICO score at origination
Refinance	Indicator variables for whether the loan is a home refinancing or not
Cash Out Refinance	Indicator variables for whether the loan is a cash-out refinance or not
Non-Cash Out Refinance	Indicator variables for whether the loan is a non cash-out (rate) refinance or not
Purchase	Indicator variables for whether the loan is a home purchase or not
First Time Home Buyer	An indicator that denotes if the borrower or co-borrower qualifies as a first-time homebuyer
Number of Borrowers	The number of individuals obligated to repay the mortgage loan
Has Mortgage Insurance	Indicator variables for whether the loan has mortgage insurance or not
Mortgage Insurance Unknown	Indicator variables for whether the loan's mortgage insurance status is unknown
Primary Residence	An indicator that denotes whether the property occupancy status is for primary residence or not
Investment or 2nd Property	An indicator that denotes whether the property occupancy status is for secondary home/investment purpose or not
Correspondent Channel	Indicator variables for whether the loan is originated through the correspondent channel or not
Retail Channel	Indicator variables for whether the loan is originated through the retail channel or not
Broker Channel	Indicator variables for whether the loan is originated through the broker channel or not
MSA Macroeconomic Indicators	
MSA Unemployment Rate	Unemployment rate by metropolitan statistical area, seasonally adjusted; obtained from the U.S. Bureau of Labor Statistics
MSA Real Personal Income	Real per capita personal income (Chained 2012 dollars) by metropolitan statistical area; obtained from the Bureau of Economic Analysis
Loan Performance	
30 (60/90) DPD in 12 (24/36) Months	Indicator variables for whether the loan is 30 (60/90) days past due within 12 (24/36) months after origination
Prepaid in 12 (24/36) Months	Indicator variables for whether the loan is prepaid within 12 (24/36) months after origination