Consumer-lending discrimination in the FinTech Era

Robert Bartlett, Adair Morse, Richard Stanton, Nancy Wallace

Journal of Financial Economics, 2021.5

Lv Manni

2021.10.31

Contents

Backgrounds & Motivation

- A literature finds differences between minority and non-minority borrowers in both mortgage-approval probabilities and interest rates paid. However, almost all of this literature looks at mortgages issued prior to the 2008 financial crisis, and focuses on subprime loans.
- Most of this literature also suffers from an omitted-variable problem.
- In the fintech era, the credit approval process increasingly uses algorithms for decision-making. Algorithmic decision-making can reduce face-to-face discrimination in markets, but the use of algorithms can also lead to inadvertent discrimination.

➢ Does consumer-lending discrimination exist? Whether algorithmic decision-making promotes or inhibits impermissible discrimination? How to identify discrimination without omitted-variable concerns?

Research Problem

- Does consumer-lending discrimination exist?
- ➢ Rely on the unique institutional setting that applies to the underwriting of credit risk in the GSE and FHA mortgage markets, we can identify discrimination.
- ➢ Yes. And with geographical variation.
- Whether algorithmic decision-making promotes or inhibits impermissible discrimination?
- ➢ FinTech lenders’ rate disparities were similar to those of non-Fintech lenders for GSE mortgages, but lower for FHA mortgages issued in 2009–2015 and for FHA refi mortgages issued in 2018–2019.

Contribution

- Rely on the unique institutional setting that applies to the underwriting of credit risk in the GSE and FHA mortgage markets, this paper manages to identify discrimination without omitted-variable concerns.
- Use a merged dataset to overcome the weakness of HMDA.
- Fixed the gap in the field of whether fintech algorithms can eliminate lending discrimination.

Model Design

- the function of GSE and FHA mortgage markets
- unique institutional setting that applies to the underwriting of credit risk: When the lender sells a mortgage to the GSE, the lender pays a g-fee to cover the borrower's alleged default and operating costs.

Table 2: All Eligible Mortgages (Excluding MCM): LLPA by Credit Score/LTV

PRODUCT FEATURE	LLPAs by LTV Range								
	≤ 60.00%	60.01 – 70.00%	70.01 – 75.00%	75.01 – 80.00%	80.01 – 85.00%	85.01 – 90.00%	90.01 – 95.00%	95.01 – 97.00%	SFC
Representative Credit Score	Applicable for all mortgages with greater than 15 year terms For whole loans purchased on or before March 31, 2011, or loans delivered into MBS pools with issue dates of March 1, 2011 or earlier								
≥ 740	-0.250%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	N/A
720 – 739	-0.250%	0.000%	0.000%	0.250%	0.000%	0.000%	0.000%	0.000%	N/A
700 – 719	-0.250%	0.500%	0.500%	0.750%	0.500%	0.500%	0.500%	0.500%	N/A
680 – 699	0.000%	0.500%	1.000%	1.500%	1.000%	0.750%	0.750%	0.500%	N/A
660 – 679	0.000%	1.000%	2.000%	2.500%	2.250%	1.750%	1.750%	1.250%	N/A
640 – 659	0.500%	1.250%	2.500%	3.000%	2.750%	2.250%	2.250%	1.750%	N/A
620 – 639	0.500%	1.500%	3.000%	3.000%	3.000%	2.750%	2.750%	2.500%	N/A
< 620 (1)	0.500%	1.500%	3.000%	3.000%	3.000%	3.000%	3.000%	3.000%	N/A

- For two people with the same credit risk in the grid, lenders should impose the same loan price, so any race-related offset to grid reflects discrimination.

Data

- Base sample, HMDA, 2009-2015
- use ATTOM, McDash, Equifax datasets to construct a merged data set of candidate loans with performance information, contract terms, the mortgage lender, and borrower information (accounted for 73.99%)
- focus on candidate loans in each data set that are first- lien, fixed-rate, owner-occupied 30-year single-family residential loans, securitized by the GSEs or insured by the FHA over the period 2009–2015
- exclude manufactured housing, investment properties, condos, duplexes, triplexes, quadraplexes, and loans with outstanding second liens at origination
- impose minimum and maximum LTVs and minimal credit scores

Model Design

- minority applicant: either Latinx or Black
- We augment the HMDA race/ethnicity indicator variable with additional race/ethnicity data obtained from processing the borrower-name field from ATTOM data, using a race and ethnic-name categorization algorithm from Kerr and Lincoln (2010) and Kerr (2008) .

- decompose a borrower's interest rate into (a) a base mortgage rate (captured by time fixed effects), (b) credit risk (captured by a borrower's LTV and credit score), and (c) a residual that reflects a lender's strategic pricing.

$$\begin{aligned} \text{interest rate}_{it} = & \ \alpha I(\text{Latinx or Black})_i \\ & + \mu_{\text{Cash-out} \times \text{GSE-grid} \times \text{year/month}} \\ & + \mu_{\text{Lender} \times \text{year/month}} + \mu_{\text{Amount decile}} + \epsilon_{it} \end{aligned}$$

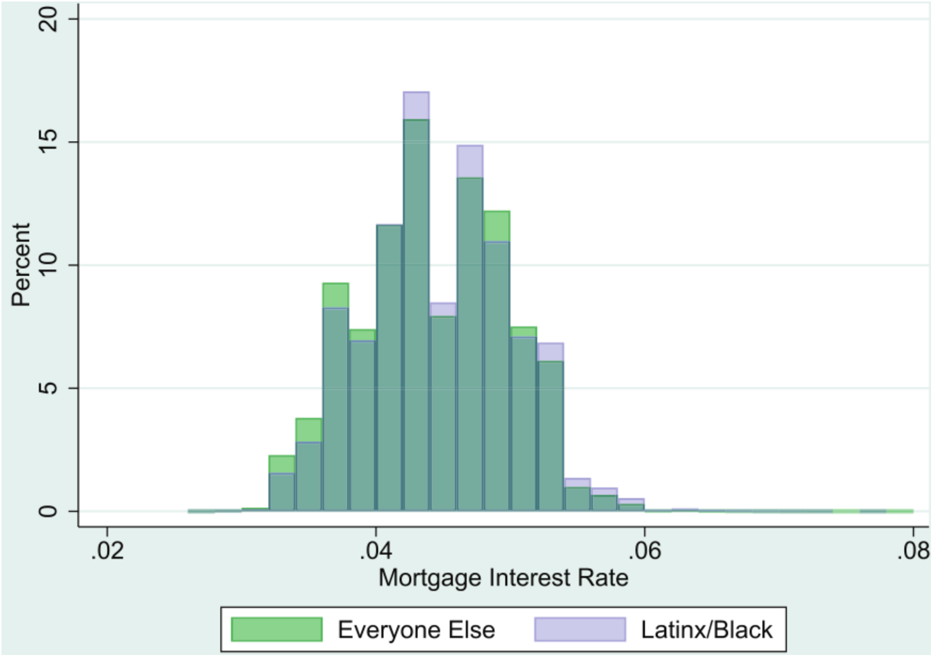
Summary statistics

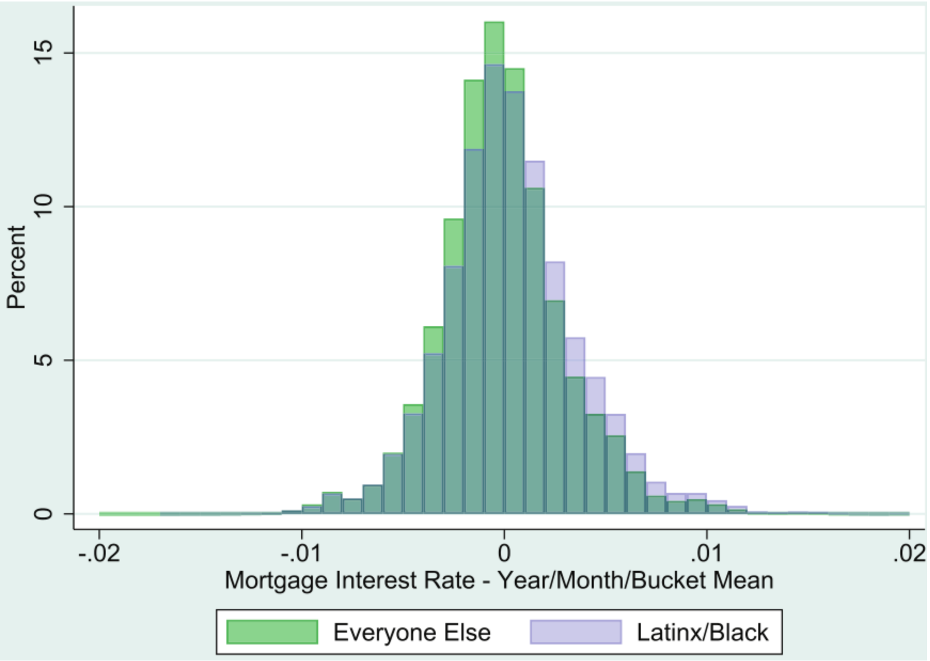
	count	mean	sd	min	max
		(a)GSE loans			
Cash-out refinance	3,376,600	0.0822052	0.2746771	0	1
CRA census tract	3,375,949	0.0926474	0.2899378	0	1
Credit score	2,950,931	757.6442	43.06677	620	850
FinTech	3,376,600	0.0312394	0.1739641	0	1
Income	3,252,686	101.9811	81.34873	20	9755
Loan amount	3,376,600	239.9792	121.9767	40	729
Loan interest rate	3,376,600	0.0442447	0.0054665	0.0275	0.07875
Loan-to-value ratio	3,376,600	0.7397208	0.1488221	0.3	0.95
Minority borrower	3,376,600	0.1001579	0.3002104	0	1
Refinance	3,376,600	0.5309895	0.4990388	0	1
Top-25 lender	3,376,600	0.4095857	0.4917574	0	1
N	3,376,600				

	count	mean	sd	min	max
		(b)FHA loans			
Cash-out refinance	2,273,444	0.0212822	0.1443237	0	1
CRA census tract	2,273,365	0.1781905	0.3826731	0	1
Credit score	1,994,340	697.3241	49.53322	580	850
FinTech	2,273,444	0.0160492	0.1256648	0	1
Income	1,999,753	66.4546	45.83809	20	7424
Loan amount	2,273,444	176.7100	91.61529	40	729
Loan interest rate	2,273,444	0.0440192	0.0068095	0.0275	0.0775
Loan-to-value ratio	2,273,444	0.9358196	0.0674404	0.3	0.9825
Minority borrower	2,273,444	0.2461666	0.4307768	0	1
Refinance	2,273,444	0.2619699	0.4397065	0	1
Top-25 lender	2,273,444	0.2854264	0.4516174	0	1
N	2,273,444				

Summary statistics - Intuitive cognition

- GSE loans





Empirical Results - Does discrimination exist?

$$\text{interest rate}_{it} = \alpha I(\text{Latinx or Black})_i + \mu_{\text{Cash-out} \times \text{GSE-grid} \times \text{year/month}} + \mu_{\text{Lender} \times \text{year/month}} + \mu_{\text{Amount decile}} + \epsilon_{it}$$

	GSE Loans		FHA Loans	
	Purchase	Refinance	Purchase	Refinance
	(1)	(2)	(3)	(4)
VARIABLES	Interest rate	Interest rate	Interest rate	Interest rate
Minority borrower	4.674***	1.632***	4.866***	1.527***
	(0.255)	(0.227)	(0.333)	(0.253)
Observations	1,371,629	1,540,939	1,533,532	436,420
R-squared	0.803	0.769	0.854	0.869
Lender x year/month FE	Y	Y	Y	Y
Cash-out x bucket x year/month FE	Y	Y	Y	Y
Amount decile FE	Y	Y	Y	Y

Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

- 2 basis points corresponds to 20% of total average profit

Empirical Results - Algorithmic inhibits discrimination?

- Interest-rate differentials: FinTech vs. non-FinTech lenders:

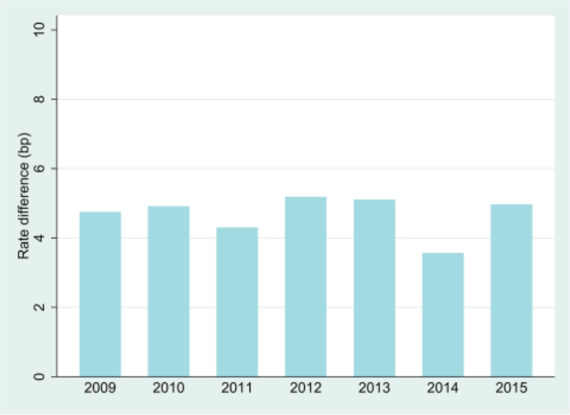
	GSE Loans		FHA Loans	
	Purchase	Refinance	Purchase	Refinance
	(1)	(2)	(3)	(4)
VARIABLES	Interest rate	Interest rate	Interest rate	Interest rate
Non-FinTech × Minority	4.666***	1.631***	4.877***	1.548***
	(0.256)	(0.238)	(0.336)	(0.262)
FinTech × Minority	5.081***	1.565***	3.550***	0.969**
	(0.124)	(0.271)	(0.373)	(0.385)
Observations	1,371,629	1,540,939	1,533,532	436,420
R-squared	0.803	0.769	0.854	0.869
p-value for test of equality	0.1172	0.8570	0.0084	0.2204
Lender x year/month FE	Y	Y	Y	Y
Cash-out x bucket x year/month FE	Y	Y	Y	Y
Amount decile FE	Y	Y	Y	Y
FinTech FE	Y	Y	Y	Y

Robust standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

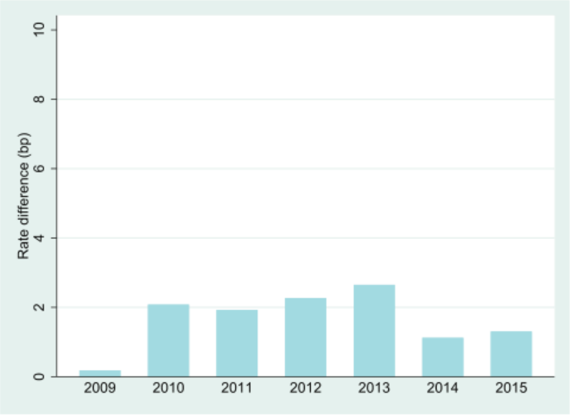
Empirical Results - Time pattern

- Interest-rate differentials by year

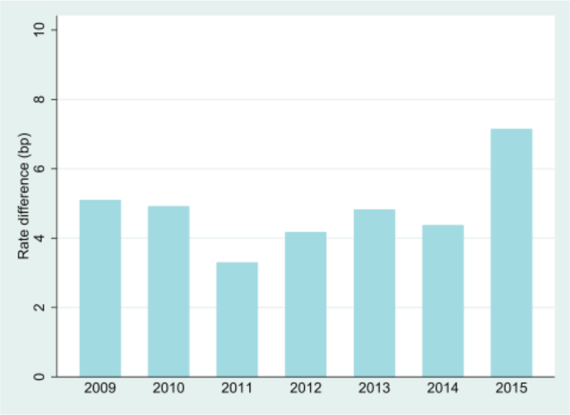
$$\text{interest rate}_{it} = \alpha I(\text{Latinx or Black})_i + \mu_{\text{Cash-out} \times \text{GSE-grid} \times \text{year/month}} + \mu_{\text{Lender} \times \text{year/month}} + \mu_{\text{Amount decile}} + \epsilon_{it}$$



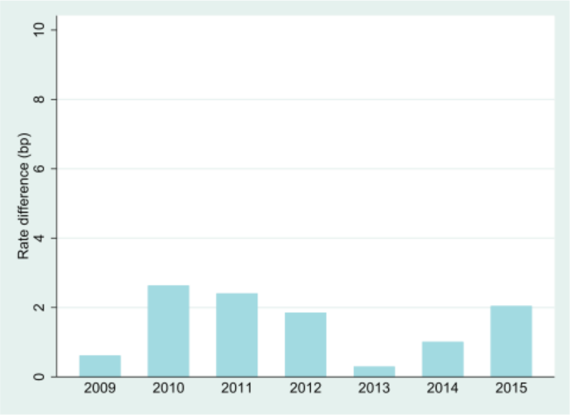
(a) GSE purchase loans



(b) GSE refi loans



(c) FHA purchase loans

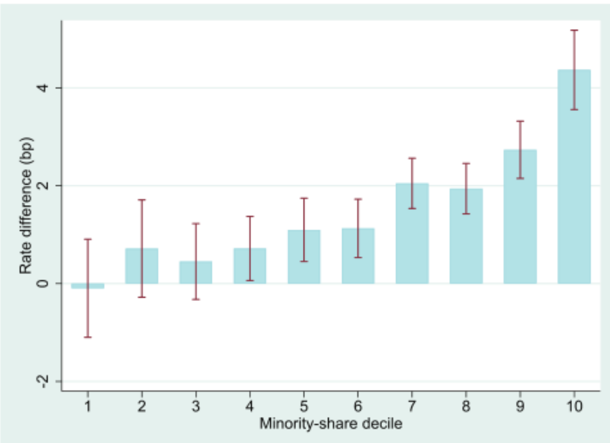


(d) FHA refinance loans

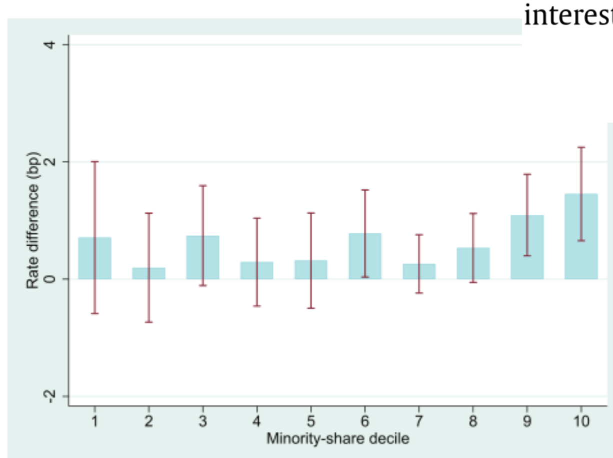
Empirical Results - Geographical variation

- Interest-rate differentials with census-tract controls by minority-share decile

$$
\begin{aligned}
\text{interest rate}_{it} = {} & \alpha I(\text{Latinx or Black})_i \\
& + \mu_{\text{Cash-out}\times\text{GSE-grid}\times\text{year/month}} \\
& + \mu_{\text{Census tract}\times\text{year FE}} + \mu_{\text{Amount decile}} + \epsilon_{it}
\end{aligned}
$$

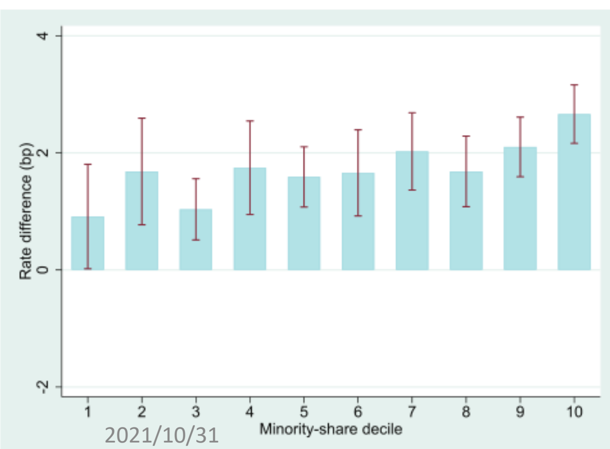


(a) GSE purchase loans

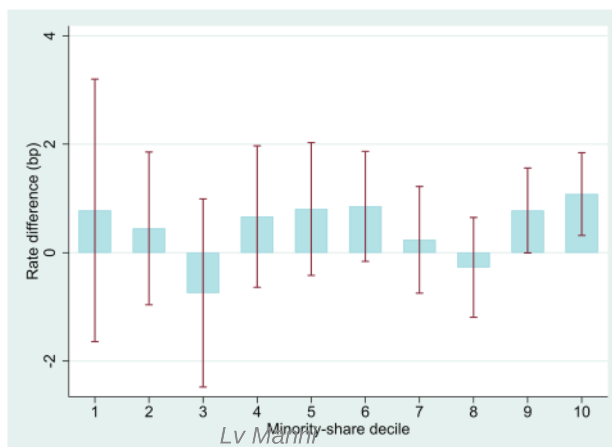


(b) GSE refi loans

In those same census tracts, minority borrowers also pay higher rates than non-minority borrowers.



(c) FHA purchase loans

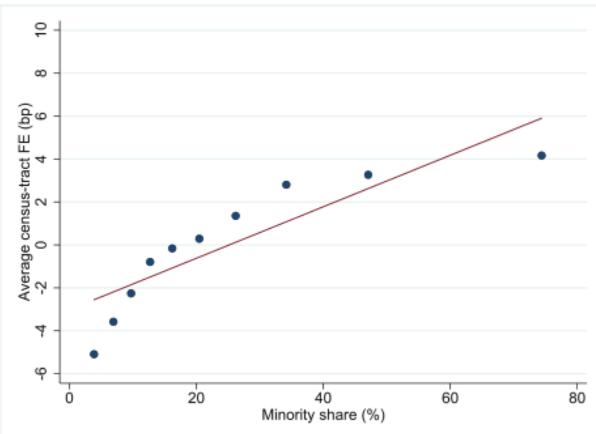


(d) FHA refinance loans

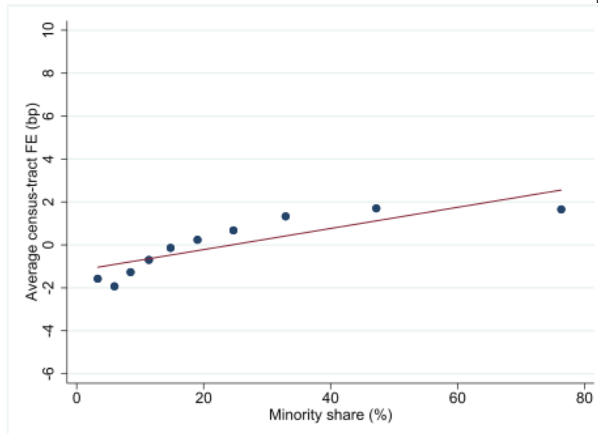
Empirical Results - Geographical variation

- Average interest-rate levels by minority-share decile

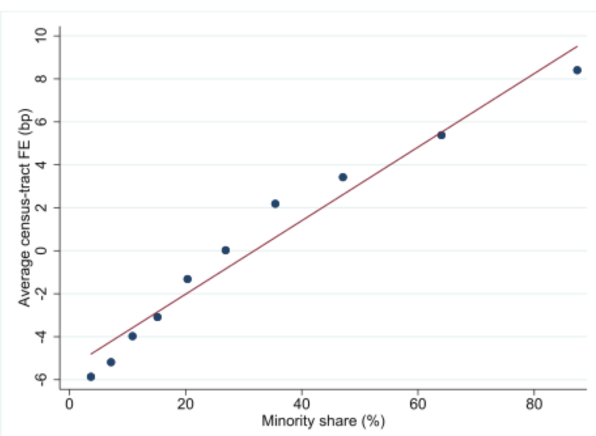
$$\text{interest rate}_{it} = \alpha I(\text{Latinx or Black})_i + \mu_{\text{Cash-out}\times\text{GSE-grid}\times\text{year/month}} + \mu_{\text{Census tract} \times \text{year FE}} + \mu_{\text{Amount decile}} + \epsilon_{it}$$



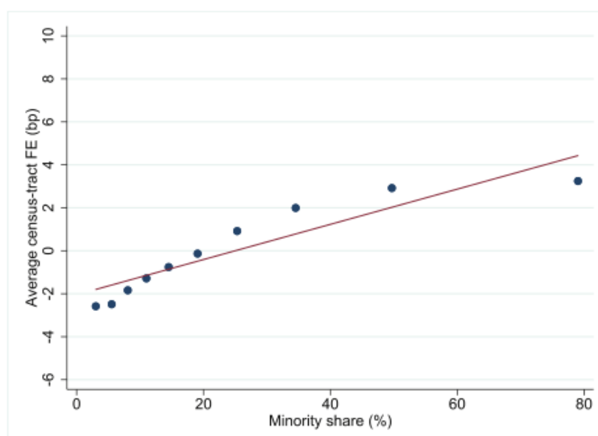
(a) GSE purchase loans



(b) GSE refi loans



(c) FHA purchase loans



(d) FHA refinance loans

The average level of mortgage rates is higher for all borrowers in high-minority-share census tracts.

Reflect different costs in different area, such as differential default risk?

Empirical Results - Robustness Check

- Put-back risk - Loans from 2013 on/ High-quality borrowers/ Banks vs. non banks
- Servicing costs - include as controls three dummy variables for whether each loan subsequently went into foreclosure/REO, 60-days-plus delinquent, or 90-days-plus delinquent
- Measurement of minority status - Using only observations for which race or ethnicity is provided by HMDA / Setting the treatment variable to 1 if either the borrower or the first coborrower is Latinx or Black
- Discount points – use 2018–2019 HMDA data

Empirical Results – Other result

- Accept/reject discrimination: Even though an application might receive a approval in the GSE underwriter system, the lender might still reject it.

- Traditional lenders have a about 6% higher loan application rejection rate for Latino and African American groups than non-minorities in the same situation.
- Fintech lenders have completely without discrimination in deciding whether to accept minority loan applications.

Conclusion

- Using lots of data and clever empirical design, this paper clearly identifies whether lenders have lending discrimination against Latino and African-American borrowers.
- The results suggest that fintech algorithms can reduce discrimination against minority lending by traditional lenders, but are not enough to completely eliminate discrimination in loan pricing.

Limitations

- 少数族裔贷款者是否选择去金融科技贷款机构并不是随机的，也没有完全体现在回归中的控制变量上，所以相关结果准确性值得商榷。