

# Peer-to-peer FinTech Lending, Non-traditional Information, and Racial Discrimination

Oct 2022

Kiridaran (Giri) Kanagaretnam

Schulich School of Business, York University

Xiaoran (Jason) Jia

Lazaridis School of Business and Economics, Wilfrid Laurier University

# Background and Motivation

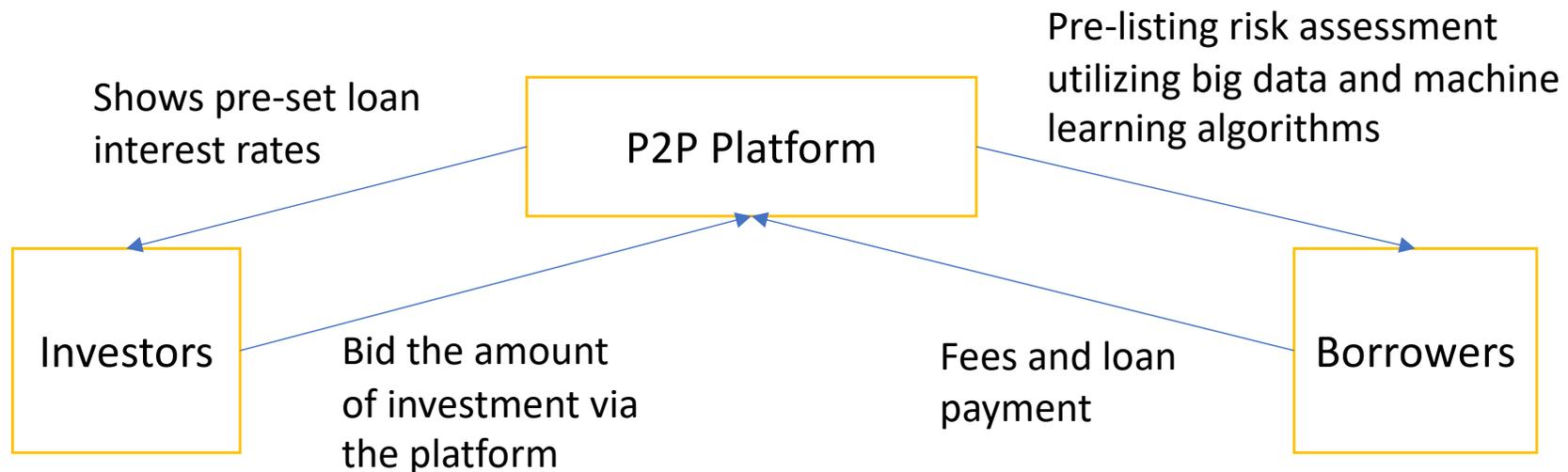
---

- Despite the several regulations in pursuit for equal access to credit, evidence of discrimination against minorities are still prevalent.
- For example, there has been empirical evidence of racial discrimination in traditional bank lending:
  - Several studies find that risk-adjusted loan rates and the probability of loan denial are higher for minority borrowers (e.g., Munnell et al. 1996; Ghent et al. 2014; Reid et al. 2017; Bayer et al. 2018; Butler et al. 2022).

# Background and Motivation

---

- However, much less attention has been paid to racial discrimination in Peer-to-peer (P2P) lending of unsecured personal loans.
- What is P2P lending?



# Background and Motivation

---

- Our study focuses on racial discrimination in P2P lending for two reasons:
  1. Few empirical evidence, especially regarding unsecured personal loans, which is the dominant form of P2P lending.
  2. P2P platforms utilize a broader scope of information when assessing borrower risks, which provides us the opportunity to test the **role of alternative information in lending discrimination**.

# Theoretical Framework

---

- According to the ‘statistical’ discrimination theory, when the observable credit information is noisier for minorities, lenders:
  - may put less weight on the observable information of minority borrowers
  - may find it less costly to use group characteristics, such as race, as a proxy for creditworthiness (e.g., Carr and Megbolugbe 1993; Ladd 1998)
- Thus, increasing the credit information quality (or precision) of minorities should decrease racial discrimination in lending(e.g., Philippon 2019)

# Hypotheses 1 – Existence of Racial Discrimination

---

**H1a:** County-level loan interest rate and loan denial rate are positively related to the county-level proportion of the minority population.

**H1b:** Loan-level interest rate and the probability of loan denial are higher for loans applied in counties with higher proportions of the minority population.

## Hypotheses 2 – The role of traditional and alternative information

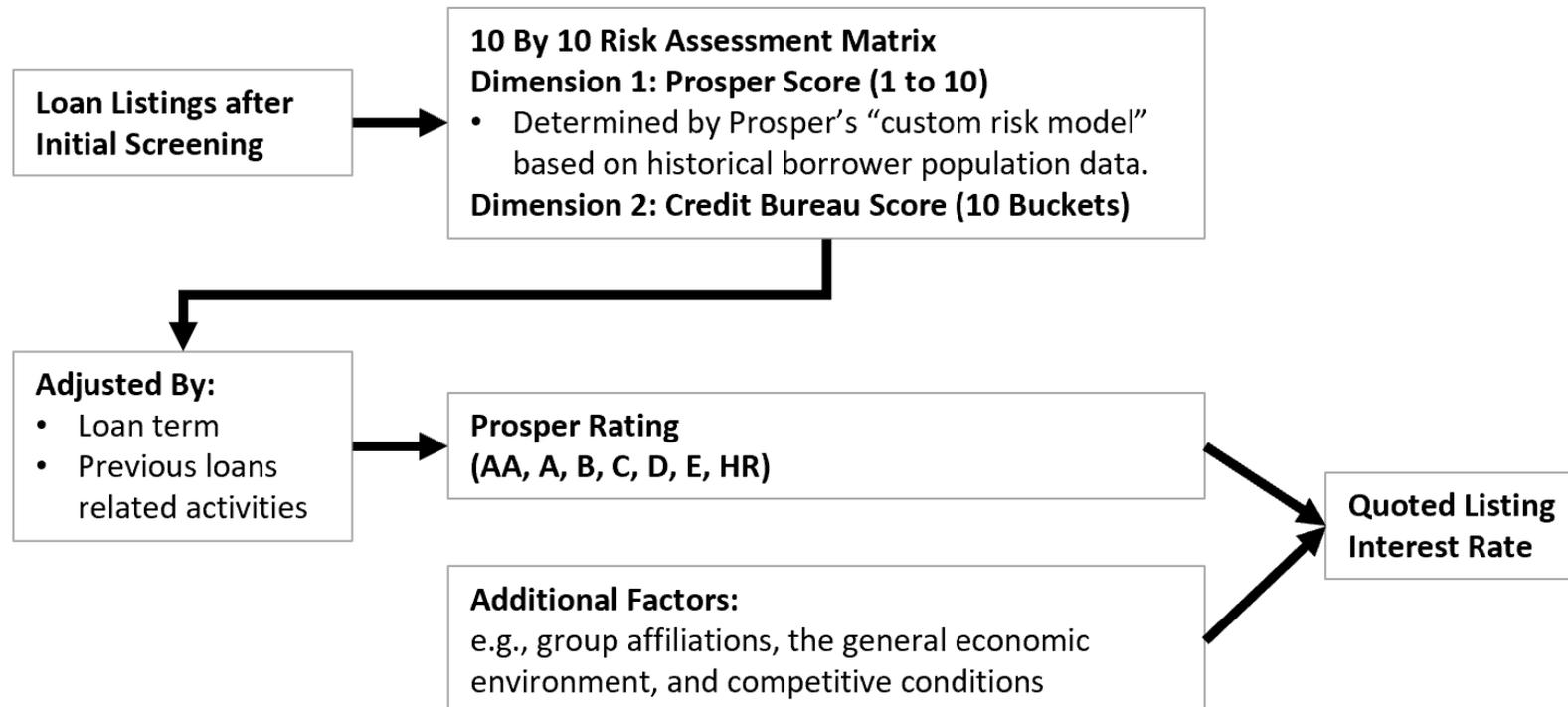
---

**H2a:** The racial discrimination in P2P lending is less pronounced when the availability of non-traditional data is greater.

**H2b:** The racial discrimination in P2P lending is less pronounced when borrowers 1) have prior loans on the P2P platform, 2) have more prior on-time payments on the P2P platform, and 3) have higher credit bureau scores.

# Institutional Background and Data

- We use P2P loan applications data from Prosper Marketplace Inc.
- Prosper's credit pricing system:



- Summarized from Prosper's prospectus report in 2010

# Institutional Background and Data

---

- Data can be downloaded for free for registered members of Prosper at:
  - <https://www.prosper.com/investor/marketplace#/download>
- We retrieve all loan applications from 2013 to 2019 – the raw sample contains 1,538,451 loan listings.
  - Since Prosper launched the home equity line of credit at the end of 2019 and credit card in 2022, we cutoff our sample by the end of 2019 so that it only contains one product – securitized personal loans.
- Key variables:
  - Listing creation date, funding date, loan amount, loan term, FICO credit scores, self-disclosed income, self-disclosed employment length, city and state of residence, information on prior interactions with the platform.

# Data for the proxy of Non-traditional information

---

- Data from National Telecommunications and Information Administration (NTIA) Internet Use Survey
  - <https://www.ntia.doc.gov/data/explorer#sel=socialNetworkUser&disp=map>
- Considered as an important source for policymakers and researchers on internet use and digital divide, the survey includes 50 questions to 50,000 households in the U.S. at the state level (NTIA 2020).
- Variables of interest (state-year level):
  - the percentage of 15+ persons doing online shopping (**ECOM%**),
  - the percentage of 15+ persons using online financial services such as banking, investing, and paying bills (**OFS%**), and
  - the percentage of 15+ persons using online social networks (**SNW%**).

# Other Data Source

---

- Race (county-year level)
  - American Community Survey 5 Years (ACS5) database from the U.S. Census Bureau (pulled using the Census API).
  - Minority percentage (*MP*) is measured as county-year-level percentage of African American and/or Latinx population
- Education, Gender, Marital Status (county-year level)
  - American Community Survey 5 Years (ACS5) database from the U.S. Census Bureau (pulled using the Census API).
- Personal income per capita (county-year level)
  - U.S. Bureau of Economic Analysis (BEA)
- Unemployment rate (county-year level)
  - U.S. Bureau of Labor Statistics
- Housing Price Index
  - Federal Housing Finance Agency
- GDP per capita
  - U.S. Bureau of Economic Analysis (BEA)

# Sample Construction

Original loan listings data from 2013 to 2019	1,538,451
After matching the original data to U.S. counties and deleting unmatched observations (possibly due to city name entry errors)	1,317,474
After matching with county level race, personal income, GDP, HPI, unemployment rate, marital status, and education variables, and deleting unmatched observations (due to some missing values in these variables)	1,300,886
After deleting missing values of loan-level characteristic variables	1,093,797
<ul style="list-style-type: none"> <li>Final sample with successfully funded loans</li> </ul>	825,975
<ul style="list-style-type: none"> <li>Final sample with loan listings that are cancelled or expired</li> </ul>	259,731
<ul style="list-style-type: none"> <li>Final sample with loan listings that are under other status</li> </ul>	8,091
After collapsing into county-level data	13,867
After deleting county-level missing values of loan amount, loan term, loan status, loan rate, FICO scores, self-reported income range, and self-reported employment length	12,600

# Baseline Empirical Design

---

We estimate the following cross-sectional regressions:

$$DECTY \text{ or } RATECTY = \alpha_0 + \alpha_1 MP + \alpha_2 VC + \alpha_3 W + Year\_FE + State\_FE + \varepsilon \quad (1)$$

$$DE \text{ or } RATE = \beta_0 + \beta_1 MP + \beta_2 VL + \beta_3 W + Year\_FE + State\_FE \text{ or } County\_FE + \varepsilon \quad (2)$$

<i>DECTY</i>	County-level loan denial rate	<i>VC</i>	Vector of loan characteristics at the county-year level
<i>RATECTY</i>	County-level loan interest rate	<i>W</i>	Vector of county-level demographic and economic characteristics
<i>DE = 1</i>	Loan level denial indicator	<i>VL</i>	Vector of individual loan-level loan characteristics
<i>RATE</i>	Loan level interest rate	<i>Year,</i> <i>State or</i> <i>County_FE</i>	Year, State, or County Fixed Effects
<i>MP</i>	County-level minority percentage		

# Descriptive Statistics – Selected Variables

## Selected Variable Descriptive Statistics at the County Level

Variable	N	Mean	Q1	p50	Q3	Std dev
<i>MP</i>	12,600	0.191	0.050	0.123	0.294	0.179
<i>AA</i>	12,600	0.101	0.010	0.035	0.131	0.143
<i>DECTY</i>	12,600	0.283	0.167	0.250	0.346	0.165
<i>RATECTY</i>	12,600	15.680	14.081	15.327	17.222	2.561
<i>AMTCTY</i>	12,600	12.664	11.157	12.746	14.174	2.693
<i>TERMCTY</i>	12,600	43.564	41.684	43.368	45.231	3.851
<i>SCORECTY</i>	12,600	706.128	697.756	704.955	713.187	15.191
<i>DPRIORCTY</i>	12,600	0.226	0.111	0.204	0.313	0.141
<i>ECOM%</i>	7,670	0.652	0.614	0.662	0.716	0.086
<i>OFS%</i>	7,670	0.620	0.592	0.642	0.674	0.085
<i>SNW%</i>	6,091	0.736	0.712	0.732	0.762	0.033

## Selected Variable Descriptive Statistics at the Loan Level

Variable	N	Mean	Q1	p50	Q3	Std dev
<i>DE</i>	1,093,797	0.237	0.000	0.000	0.000	0.426
<i>RATE</i>	1,093,797	14.935	10.280	13.540	18.550	6.237
<i>AMT</i>	1,093,797	13.630	7.500	12.000	19.000	8.132
<i>TERM</i>	1,093,797	43.206	36.000	36.000	60.000	11.001
<i>SCORE</i>	1,093,797	704.42	669.50	689.50	729.50	39.65
<i>INCSELF</i>	1,093,797	4.315	3.000	4.000	5.000	1.161
<i>EMPLN</i>	1,093,797	109.55	28.00	76.00	161.00	106.36
<i>PRIOR</i>	1,093,797	0.290	0.000	0.000	0.000	1.077

# Result of Tests of Existence of Racial Discrimination in P2P lending – County Level

The proportion of county-level minority population is positively related to both county-level loan denial rate and county-level loan interest rate.

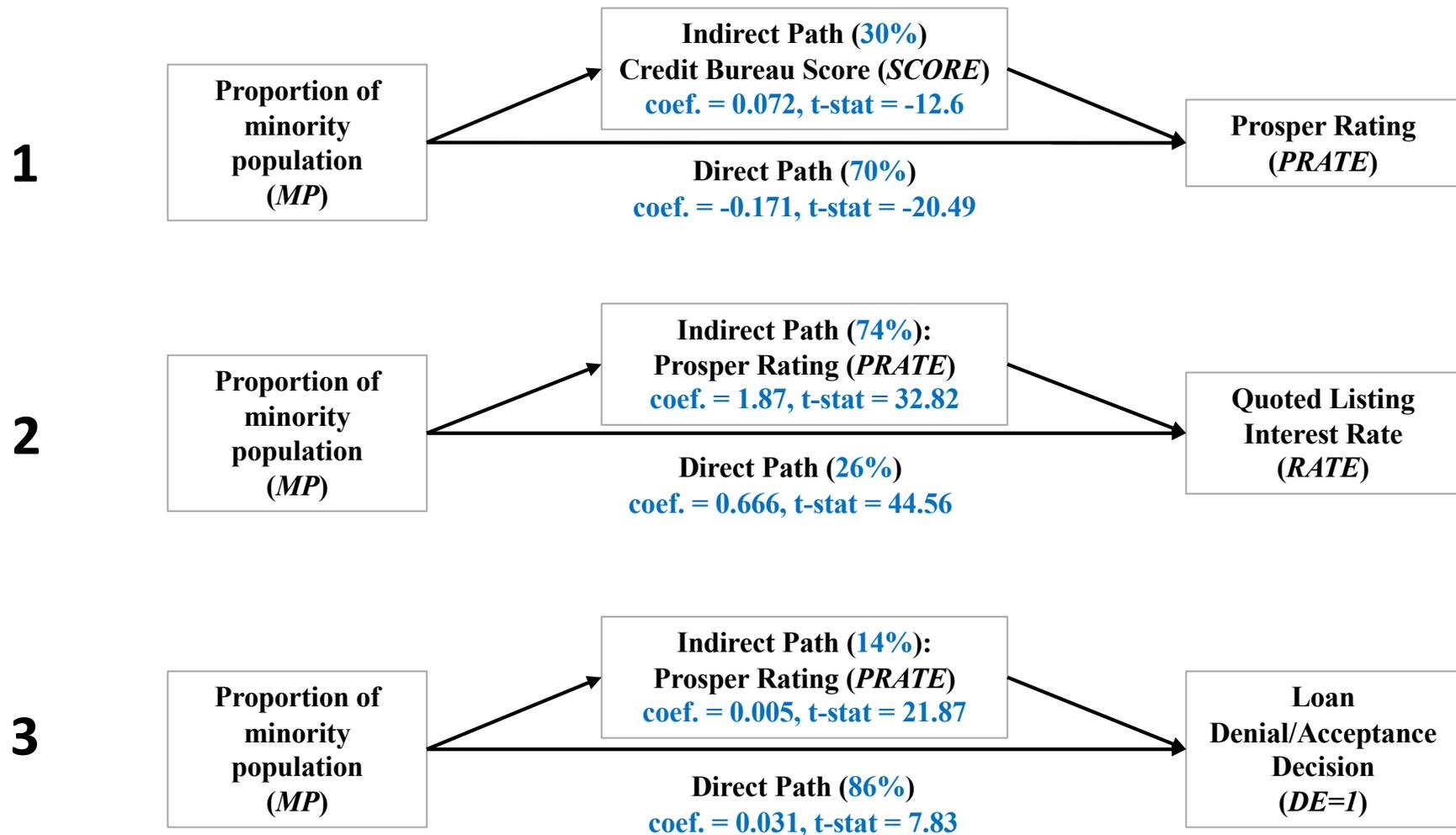
	(1) <i>DECTY</i>	(2) <i>DECTY</i>	(3) <i>RATECTY</i>	(4) <i>RATECTY</i>
<b><i>MP</i></b>	<b>0.052***</b>	<b>(7.44)</b>	<b>0.041***</b>	<b>(4.62)</b>
<i>AMTCTY</i>	0.006***	(8.25)	0.006***	(8.69)
<i>TERMCTY</i>	-0.001*	(-2.06)	-0.001	(-1.41)
<i>SCORECTY</i>	-0.000	(-0.06)	0.000	(0.03)
<i>INCSELFCTY</i>	-0.016***	(-4.67)	-0.014***	(-2.86)
<i>EMPLENCTY</i>	-0.000	(-0.77)	-0.000	(-0.55)
<i>PRIORCTY</i>	-0.014	(-1.42)	-0.017*	(-1.73)
<i>PRIORACTY</i>	-0.029*	(-2.19)	-0.028	(-1.49)
<i>PRIORPCTY</i>	0.002**	(3.09)	0.002***	(2.72)
<i>PRIOROCTY</i>	-0.002*	(-2.04)	-0.001**	(-2.57)
<i>PRIORLCTY</i>	0.016	(1.69)	0.016**	(2.05)
<i>INCOME</i>	0.000	(0.93)	0.000***	(3.23)
<i>EDU</i>	-0.165**	(-3.22)	-0.169***	(-5.35)
<i>MALE</i>	0.157**	(2.71)	0.068	(1.03)
<i>MARRIED</i>	0.262***	(15.78)	0.232***	(8.25)
<i>PEN</i>	-0.012***	(-6.07)	-0.014***	(-21.58)
<i>UNEMP</i>	0.001	(1.44)	0.003***	(3.56)
<i>HPI</i>	0.000***	(4.92)	0.000***	(5.39)
<i>GDP</i>	-0.001	(-0.20)	-0.000	(-0.13)
<i>Constant</i>	0.179	(1.30)	0.187**	(2.04)
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	No	Yes	No	Yes
SE Cluster	Year, County	Year, County	Year, County	Year, County
R-squared	0.46	0.59	0.60	0.62
N	12,600	12,600	12,600	12,600

# Result of Tests of Existence of Racial Discrimination in P2P lending – Loan Level

The proportion of county-level minority population is positively related to both loan-level probability of denial and loan-level quoted interest rate.

	(1)		(2)		(3)		(4)	
	<i>RATE</i>		<i>RATE</i>		<i>DE = 1</i>		<i>DE = 1</i>	
<b><i>MP</i></b>	<b>0.550***</b>	<b>(6.74)</b>	<b>0.486***</b>	<b>(4.72)</b>	<b>0.210***</b>	<b>(6.00)</b>	<b>0.231***</b>	<b>(4.06)</b>
<i>AMT</i>	-0.081***	(-25.04)	-0.067***	(-67.06)	0.017***	(22.93)	0.017***	(22.96)
<i>TERM</i>	0.130***	(61.64)	0.129***	(61.64)	-0.001***	(-4.17)	-0.001***	(-4.35)
<i>SCORE</i>	-0.076***	(-93.58)	-0.082***	(-283.6)	0.002***	(7.63)	0.002***	(7.64)
<i>INCSELF</i>	-0.331***	(-25.60)	-0.380***	(-25.75)	-0.043***	(-7.02)	-0.039***	(-6.10)
<i>EMPLEN</i>	-0.000	(-0.21)	-0.000	(-0.65)	-0.000***	(-7.72)	-0.000***	(-8.29)
<i>PRIOR</i>	-0.493***	(-9.61)	-0.491***	(-9.60)	-0.633***	(-23.93)	-0.631***	(-23.87)
<i>PRIORA</i>	0.044	(0.56)	0.043	(0.55)	0.073**	(2.29)	0.073**	(2.27)
<i>PRIORP</i>	0.068***	(12.55)	0.068***	(12.54)	-0.039***	(-13.20)	-0.039***	(-13.21)
<i>PRIORO</i>	-0.018***	(-7.51)	-0.018***	(-7.51)	-0.010***	(-5.47)	-0.010***	(-5.45)
<i>PRIORL</i>	0.030*	(1.85)	0.030*	(1.84)	0.030***	(5.43)	0.030***	(5.33)
<i>INCOME</i>			-0.0007	(-0.57)			0.002***	(3.20)
<i>EDU</i>			-1.3178***	(-5.39)			-1.307***	(-12.75)
<i>MALE</i>			-0.2310	(-0.19)			0.310	(0.61)
<i>MARRIED</i>			0.1454	(0.45)			0.736***	(4.17)
<i>UNEMP</i>			-0.0132	(-1.33)			0.008*	(1.78)
<i>HPI</i>			0.0014	(1.64)			0.002***	(6.86)
<i>GDP</i>			0.0000	(0.36)			0.000	(0.16)
Constant	69.416***	(60.37)	69.779***	(53.19)	-0.555	(-1.38)	-1.209**	(-2.38)
Year FE	Yes		Yes		Yes		Yes	
State FE	Yes		Yes		Yes		Yes	
County FE	No		No		No		No	
SE Cluster	Year, County		Year, County		Year, County		Year, County	
Observations	1,093,797		1,093,797		1,093,797		1,093,797	
<i>R</i> <sup>2</sup>	0.36		0.36		n/a		n/a	
ROC	n/a		n/a		0.66		0.67	

# Path Analyses – How is racial information transmitted through the P2P platform?



## Cross-sectional Tests – Role of Alternative Information Availability

Our proxy of alternative information (*MP*) negatively moderates the relation between minority proportion and loan denial rates (loan interest rates)

### County-Level Regressions:

	(1)	(2)
	<i>DECTY</i>	<i>RATECTY</i>
<i>MP</i>	0.087***	1.020***
<b><i>INFO × MP</i></b>	<b>-0.055***</b>	<b>-0.703***</b>
<i>INFO</i>	0.017***	0.188
<i>AMTCTY</i>	0.006***	-0.128***
<i>TERMCTY</i>	-0.001	0.099***
<i>SCORECTY</i>	0.000	-0.076***
<i>INCSELFCTY</i>	-0.013***	-0.147**
<i>EMPLENCTY</i>	-0.000	-0.000
<i>PRIORCTY</i>	-0.017*	-0.169
<i>PRIORACTY</i>	-0.027	0.692**
<i>PRIORPCTY</i>	0.002***	-0.028
<i>PRIOROCTY</i>	-0.001**	-0.002
<i>PRIORLCTY</i>	0.016**	-0.106
<i>Constant</i>	0.171*	72.202***
<i>Macro Controls</i>	Yes	Yes
Year Fixed Effects	Yes	Yes
State Fixed Effects	Yes	Yes
County Fixed Effects	No	No
SE Cluster	Year, State	Year, State
R-squared	0.47	0.62
N	12,600	12,600

### Loan-Level Regressions:

	(1)	(2)	(3)	(4)
	<i>RATE</i>	<i>RATE</i>	<i>RATE</i>	<i>RATE</i>
<i>MP</i>	0.9429***	0.6752***	0.7494***	0.8470***
<b><i>INFO × MP</i></b>	<b>-0.5653***</b>			
<b><i>ECOM × MP</i></b>		<b>-0.2619**</b>		
<b><i>OFS × MP</i></b>			<b>-0.3574**</b>	
<b><i>SNW × MP</i></b>				<b>-0.4131*</b>
<i>INFO</i>	0.1006			
<i>ECOM</i>		-0.0895		
<i>OFS</i>			0.2155**	
<i>SNW</i>				0.0604
<i>AMT</i>	-0.0805***	-0.0805***	-0.0805***	-0.0805***
<i>TERM</i>	0.1296***	0.1296***	0.1296***	0.1296***
<i>SCORE</i>	-0.0760***	-0.0760***	-0.0760***	-0.0760***
<i>INCSELF</i>	-0.3253***	-0.3255***	-0.3255***	-0.3256***
<i>EMPLEN</i>	-0.0000	-0.0000	-0.0000	-0.0000
<i>Constant</i>	69.6371***	69.8145***	69.5936***	69.6845***
<i>Other Loan Controls</i>	Yes	Yes	Yes	Yes
<i>Macro Controls</i>	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
SE Cluster	Year, State	Year, State	Year, State	Year, State
<i>R</i> <sup>2</sup>	0.375	0.375	0.375	0.375
Observations	1,093,797	1,093,797	1,093,797	1,093,797

## Cross-sectional Tests – Role of Traditional Information

Our proxy of alternative information (*MP*) negatively moderates the relation between minority proportion and loan denial rates (loan interest rates)

	(1) <i>RATE</i>	(2) <i>RATE</i>	(3) <i>RATE</i>
<i>MP</i>	0.5443***	0.5174***	0.7159***
<b><i>PRIOR</i> × <i>MP</i></b>	<b>-0.3581***</b>		
<i>PRIORO_High</i> × <i>MP</i>		<b>-0.4056***</b>	
<i>SCORE_High</i> × <i>MP</i>			<b>-0.5324**</b>
<i>PRIOR</i>	-0.2925***	0.0838***	-0.2577***
<i>PRIORO_High</i>		-0.6388***	
<i>SCORE_High</i>			0.0060
<i>AMT_High</i>			
<i>TERM_Long</i>			
<i>INCOME_High</i>			
Year FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
SE Cluster	Year, State	Year, State	Year, State
<i>R</i> <sup>2</sup>	0.360	0.360	0.377
Observations	1,093,797	1,093,797	1,093,797

# Quantile Regressions

Model	Dependent Variable	Explanatory Variable of Interest	Quantile Regression Estimates			OLS Estimates	
			0.25	0.5	0.75	Coef.	R <sup>2</sup>
(1)	<i>RATECTY</i>	<i>MP</i>	0.938*** (11.86)	0.814*** (11.50)	0.781*** (8.15)	0.954*** (4.90)	0.60
(2)	<i>RATECTY</i>	<i>INFO × MP</i>	<b>-0.248*</b> <b>(-1.67)</b>	<b>-0.685***</b> <b>(-4.54)</b>	<b>-0.966***</b> <b>(-5.40)</b>	<b>-0.703***</b> <b>(-2.85)</b>	<b>0.62</b>
(3)	<i>DECTY</i>	<i>MP</i>	0.050*** (10.44)	0.042*** (8.17)	0.055*** (6.21)	0.052*** (6.39)	0.46
(4)	<i>DECTY</i>	<i>INFO × MP</i>	<b>-0.018**</b> <b>(-2.55)</b>	<b>-0.044***</b> <b>(-5.89)</b>	<b>-0.071***</b> <b>(-5.18)</b>	<b>-0.055***</b> <b>(-3.76)</b>	<b>0.47</b>
	Year Fixed Effects		Yes	Yes	Yes	Yes	
	State Fixed Effects		No	No	No	No	
	Control Variables		Yes	Yes	Yes	Yes	
	Observations		12,600	12,600	12,600	12,600	

# Additional Tests

---

- Tests using alternative measure of the availability of alternative information
- Tests using alternative measure of minority population percentages

# Contributions and Conclusion

---

- We provide empirical evidence showing racial discrimination in P2P lending.
- We document the effects of the precision of different credit quality signals on P2P lending racial discrimination
- We examine how racial discrimination is transmitted through the entire credit pricing and funding process of a P2P platform
- *Policy implications:* policy makers can promote the use of big data for the entire banking sector to reduce racial and economic discrimination.