The New HMDA: Fair Lending, Econometrics, and Beyond

ABA Business Law Spring Meetings 2017

April 2017

Consumer Financial Services Committee

Edgeworth Economics

Legal Disclaimer

The opinions expressed herein do not necessarily represent the views of Edgeworth Economics nor any other Edgeworth consultant. The information contained herein is provided for informational purposes only, and should not be construed as legal advice on any subject matter.

Dr. Xiaoling (Ling Ling) Ang

Dr. Xiaoling Ang is an expert in consumer financial services, antitrust, and labor economics. Her experience spans multiple industries with a concentration in banking, insurance, and finance. She focuses on applying rigorous economic and econometric methods to these areas, including through class certification and damages analysis, policy evaluation, cost-benefit analysis, and fair lending analysis.

Prior to joining Edgeworth in November 2015, Dr. Ang was one of the original economists at the Consumer Financial Protection Bureau (CFPB). At the CFPB, Dr. Ang frequently served as the Lead Economist on Bureau initiatives and rulemakings, including TILA/RESPA, a deposit products rulemaking; interagency appraisal rulemakings (Higher Risk Mortgage Rulemaking and ECOA Amendment); larger participant rulemakings in student loan servicing and international money transfers; randomized control trials; disclosure testing; and on a Congressional report on private student loans. Before joining the CFPB, Dr. Ang was a Financial Economist at the FDIC.

Dr. Ang's work has been published in law and economic journals and she has presented her research at over 20 academic and industry conferences. Her academic research includes work on student and mortgage lending, related to behavioral economics, pass through, and bankruptcy. She has also served as a member of the technical panel of the Department of Education's Postsecondary Surveys and a reviewer for the Journal of Financial Services Research and the Review of Economics and Statistics. She is also the recipient of various national fellowships, including the National Science Foundation Graduate Research Fellowship and the Statistics Canada Tom Symons Research Fellowship.

Dr. Ang earned her PhD and MA in economics from Princeton University, and her MS and BS in mathematics from Loyola University Chicago.

Dr. Stephen Bronars

Dr. Stephen Bronars is an expert in the economics of discrimination, wage and hours matters, and the use of criminal background checks in hiring decisions, bringing his expertise in the rigorous analysis of complex datasets to the economic issues considered in these matters. Dr. Bronars testifies on issues related to class certification, liability, and economic damages in matters involving screens used in hiring decisions; wrongful terminations; and age, race, and gender discrimination at the federal and state levels.

He has worked on behalf of litigation and consulting clients in a variety of industries including financial services, federal and state government, food, healthcare, retail, and technology. Dr. Bronars is a sophisticated researcher, with extensive experience in academia and public policy. He previously led the Department of Economics at the University of Texas and has held faculty positions at Georgetown University, the University of Pennsylvania, the University of California Santa Barbara, Yale University, and Texas A&M University.

Dr. Bronars publishes widely on the topics of labor economics, econometrics, and microeconomics and has been featured on CNN's *The Situation Room* and *Good Morning America*. His commentary has been included in the *Wall Street Journal*, *Forbes*, *Bloomberg*, among others. He has also participated in Senate committee meetings, working to draft new legislation regarding immigration matters.

Dr. Bronars earned his PhD and MA in economics from the University of Chicago and his BA in economics from the University of Illinois.

Sanford Shatz

Sanford Shatz (sshatz@mcglinchey.com) is Of Counsel to McGlinchey Stafford in Irvine, CA, and is a member of the firm’s Commercial Litigation and Consumer Financial Services sections. He has been a licensed attorney in California for more than 29 years, during which time he has actively litigated cases in the areas of commercial law, real estate, and consumer financial services, specializing in mortgage-related issues. In 1998, he joined Countrywide Home Loans where he organized and established that company’s California In-House Litigation Group. Mr. Shatz focused on all aspects of mortgage-related litigation, and has tried numerous cases to verdict. In 2008, after Bank of America acquired Countrywide, he managed outside counsel on a pool of several hundred litigation cases, and helped to develop case-resolution strategies. In 2010, he returned to private practice, and currently practices at McGlinchey Stafford, where he works on litigation and regulatory issues, and appeals. Mr. Shatz is active in the American Bar Association’s Consumer Financial Services Committee where he organizes and moderates a monthly call-in program on current issues in consumer financial services. He has published journal articles and papers, and organized and presented seminars on various aspects of current events in the consumer financial services and mortgage world.

Roadmap of Presentation

- A Primer on Regression and Averages

- Consideration of potential fair lending model implications of specific data fields in the updated HMDA data
    - Fields discussed:
        - Points and Fees
        - Race and Ethnicity
        - Product Choice

A Primer on Regression

Regression of Loan Amount With No Controls

$$loan\ amount_i = \beta_0 + \varepsilon_i$$

Regressions Pick Up Means

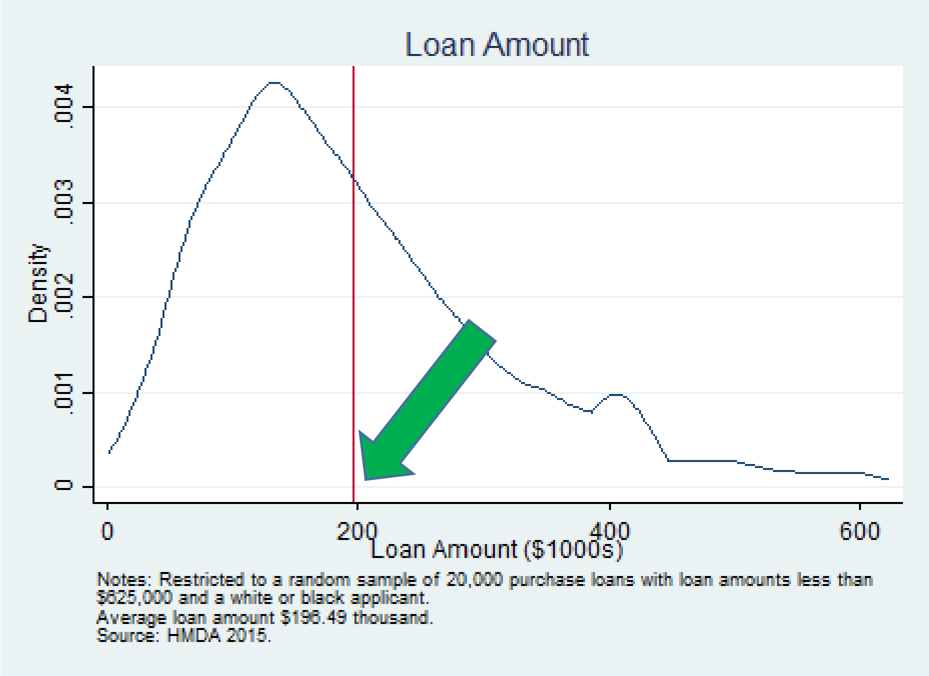
LOAN AMOUNT
NON-JUMBO PURCHASE LOANS[1]

2015

	No Controls
[a]	**[b]**
Constant	196.489 ***
	(0.828)
N	20,000
R^2	0.000

Note:
[1] Restricted to a random sample of 20,000 purchase loans with loan amounts less than $625,000 and a white or black applicant.

Source:
HMDA 2015.



Regression of Loan Amount on Protected Class

$$loan\ amount_i = \beta_0 + \beta_1 protected\ class_i + \varepsilon_i$$

Regression Picks Up Difference in Means

LOAN AMOUNT

NON-JUMBO PURCHASE LOANS[1]

2015

	Protected Class
[a]	**[b]**
Protected Class[2]	-20.640 ***
	(3.048)
Constant	198.143 ***
	(0.863)
N	20,000
R^2	0.002

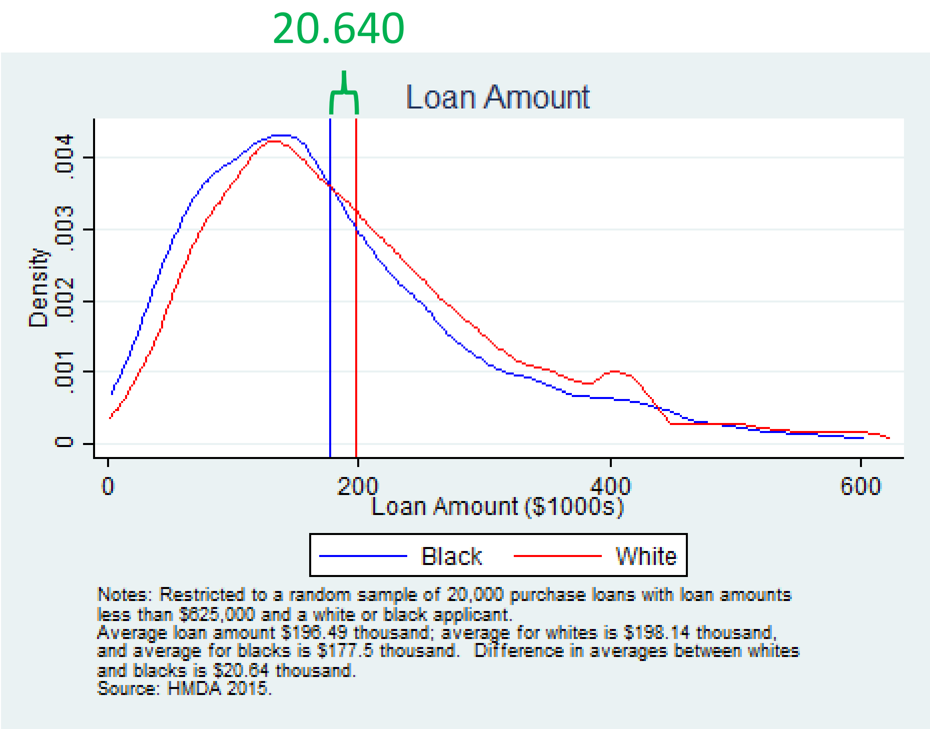
Notes:

1 Restricted to a random sample of 20,000 purchase loans with loan amounts less than $625,000 and a white or black applicant.

2 Note that protected class refers to black applicants.

Source:

HMDA 2015.



Coefficients May Differ Between Groups

LOAN AMOUNT
NON-JUMBO PURCHASE LOANS[1]

2015

[a]	Protected Class and Owner Occupied [b]		Owner Occupied Only [c]		Non-Owner Occupied Only [d]	
Protected Class[2]	-21.610	***	-20.169	***	-51.748	***
	(3.047)		(3.109)		(14.732)	
Owner Occupied	21.945	***				
	(2.976)					
Constant	178.126	***	199.951	***	179.377	***
	(2.848)		(0.899)		(3.001)	
N	20,000		18,313		1,687	
R^2	0.005		0.002		0.007	

Notes:

[1] Restricted to a random sample of 20,000 purchase loans with loan amounts less than $625,000 and a white or black applicant.

[2] Note that protected class refers to black applicants.

Source:
HMDA 2015.

Consideration in Fair Lending Econometric Models

- What are the underlying economics driving how these data are generated?
- Is the analysis grounded in the facts, theory, and acceptable practice?
- Is the model identified—does it measure the relationships you think you are capturing?
  - Do the assumptions in the model align with the data generating process?
  - Are the variables used measuring what we think they measure? Are they measured with error?
  - Are the appropriate variables being used? What about constructed variables (e.g., price per unit)?
  - Is there a factor or process that might cause observations to be related to each other?
- To what extent is a particular concern relevant to the situation?

Model Considerations: Points and Fees

Rate Regressions

$$rate_i = \beta_0 + \color{red}{\beta_1} protected\ class_i + \gamma X_i + \varepsilon_i$$

Other Controls

Points and Fees Schedule and Takeup Rates

HYPOTHETICAL
POINTS AND FEES SCHEDULE
AND TAKEUP RATES

		Proportion Paying Points	
Points	**Rate Discount**	**Whites**	**Protected Class**
[a]	**[b]**	**[c]**	**[d]**
0	0	50%	33%
1	0.5	0%	33%
2	1.75	50%	33%

Rate Regressions

$$rate_i = \beta_0 + \beta_1 protected\ class_i + \gamma X_i + \varepsilon_i$$

Other Controls

$$rate_i = \beta_0 + \beta_1 protected\ class_i + \underbrace{\gamma_1 Jumbo_i + \gamma_2 GSE_i + \gamma_3 Owner\ Occupied_i}_{\text{Other Controls (Take 1)}} + \varepsilon_i$$

Fair Lending Model Not Controlling for Points

[a]	No Controls [b]	Jumbo, GSE, and Owner Occupied Controls [c]
Protected Class	0.105 ***	0.102 ***
	(0.024)	(0.024)
Jumbo		-0.276 ***
		(0.050)
GSE Conforming Loan		-0.142 ***
		(0.021)
Owner Occupied		-0.244 ***
		(0.029)
Points		
Purchased 1 Point		
Purchased 2 Points		
Constant	4.877 ***	5.137 ***
	(0.009)	(0.029)
N	10,000	10,000
R^2	0.002	0.014

Notes:

Restricted to a random sample of 10,000 purchase loans with race reported from HMDA 2015

Points and rates simulated.

Sources:

HMDA 2015, author's simulations.

Rate Regressions

$$rate_i = \beta_0 + \beta_1 protected\ class_i + \underbrace{\gamma_1 Jumbo_i + \gamma_2 GSE_i + \gamma_3 Owner\ Occupied_i}_{\text{Other Controls}} + \varepsilon_i$$

$$rate_i = \beta_0 + \beta_1 protected\ class_i + \underbrace{\gamma_1 Jumbo_i + \gamma_2 GSE_i + \gamma_3 Owner\ Occupied_i}_{\text{Other Controls}} + \gamma_4 points_i + \varepsilon_i$$

Linear Control for Points

Fair Lending Model for Rate, Controlling for Points Linearly

	No Controls	Jumbo, GSE, and Owner Occupied Controls	Jumbo, GSE, Owner Occupied, and Points Controls
[a]	[b]	[c]	[d]
Protected Class	0.105 ***	0.102 ***	0.129 ***
	(0.024)	(0.024)	(0.003)
Jumbo		-0.276 ***	-0.191 ***
		(0.050)	(0.007)
GSE Conforming Loan		-0.142 ***	-0.100 ***
		(0.021)	(0.003)
Owner Occupied		-0.244 ***	-0.249 ***
		(0.029)	(0.004)
Points			-0.874 ***
			(0.001)
Purchased 1 Point			
Purchased 2 Points			
Constant	4.877 ***	5.137 ***	5.998 ***
	(0.009)	(0.029)	(0.004)
N	10,000	10,000	10,000
R^2	0.002	0.014	0.980

Notes:

Restricted to a random sample of 10,000 purchase loans with race reported from HMDA 2015.

Points and rates simulated.

Sources:

HMDA 2015, author's simulations.

Fair Lending Model for Rate, Offsetting Whites' Rates by "Disparity"

[a]	No Controls [b]	Jumbo, GSE, and Owner Occupied Controls [c]	Jumbo, GSE, Owner Occupied, and Points Controls [d]	12.9 bps to Whites' Rates [e]
Protected Class	0.105 ***	0.102 ***	0.129 ***	0.000
	(0.024)	(0.024)	(0.003)	(0.003)
Jumbo		-0.276 ***	-0.191 ***	-0.191 ***
		(0.050)	(0.007)	(0.007)
GSE Conforming Loan		-0.142 ***	-0.100 ***	-0.100 ***
		(0.021)	(0.003)	(0.003)
Owner Occupied		-0.244 ***	-0.249 ***	-0.249 ***
		(0.029)	(0.004)	(0.004)
Points			-0.874 ***	-0.874 ***
			(0.001)	(0.001)
Purchased 1 Point				
Purchased 2 Points				
Constant	4.877 ***	5.137 ***	5.998 ***	6.127 ***
	(0.009)	(0.029)	(0.004)	(0.004)
N	10,000	10,000	10,000	10,000
R^2	0.002	0.014	0.980	0.980

Notes:

Restricted to a random sample of 10,000 purchase loans with race reported from HMDA 2015.

Points and rates simulated.

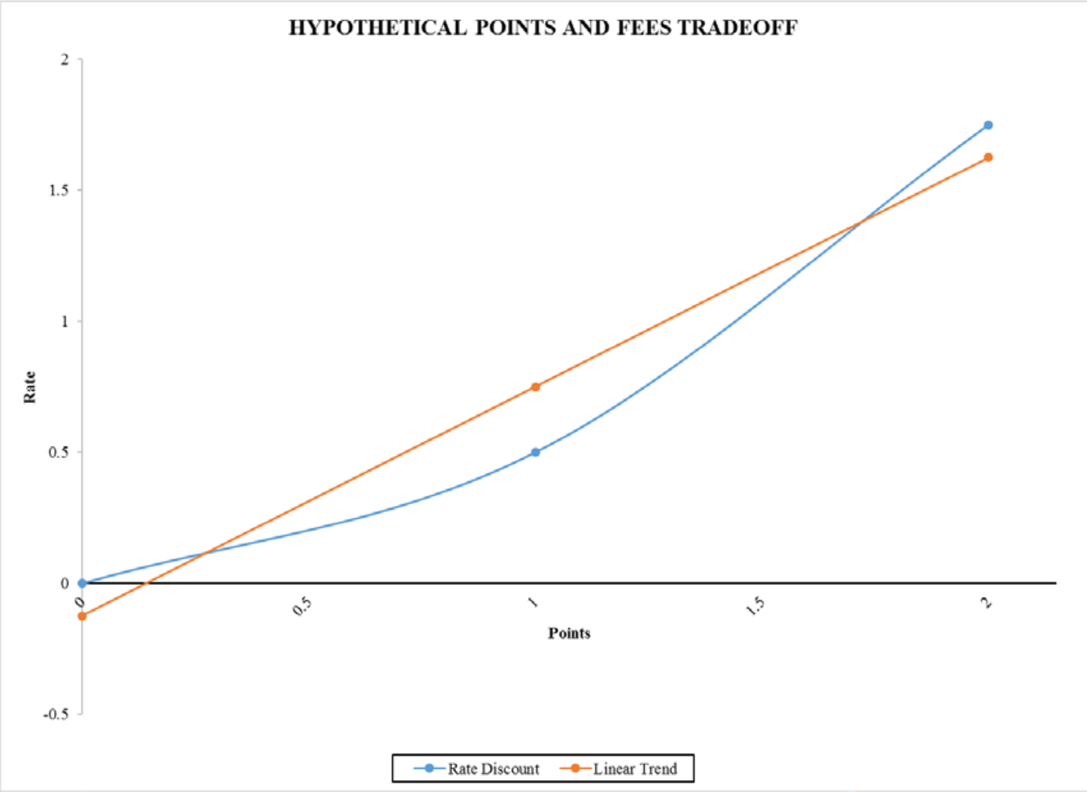
Sources:

HMDA 2015, author's simulations.

What's Going on With Points?

HYPOTHETICAL
POINTS AND FEES SCHEDULE

Points [a]	Rate Discount [b]
0	0
1	0.5
2	1.75



Rate Regressions

$$rate_i = \beta_0 + {\color{red}\beta_1} protected\ class_i + \underbrace{\gamma_1 Jumbo_i + \gamma_2 GSE_i + \gamma_3 Owner\ Occupied_i}_{\text{Other Controls}} + \gamma_4 points_i + \varepsilon_i$$

Linear Control for Points

$$rate_i = \beta_0 + {\color{red}\beta_1} protected\ class_i + \underbrace{\gamma_1 Jumbo_i + \gamma_2 GSE_i + \gamma_3 Owner\ Occupied_i}_{\text{Other Controls}} + \underbrace{\gamma_4 onepoint_i + \gamma_5 twopoints_i}_{\text{Non-Linear Control for Points}} + \varepsilon_i$$

Fair Lending Model for Rate, Controlling for Points Nonlinearly

[a]	No Controls [b]	Jumbo, GSE, and Owner Occupied Controls [c]	Jumbo, GSE, Owner Occupied, and Points Controls [d]	12.9 bps to Whites' Rates [e]	Non-Linear Points and Rate Tradeoff [f]
Protected Class	0.105 ***	0.102 ***	0.129 ***	0.000	0.000
	(0.024)	(0.024)	(0.003)	(0.003)	(0.003)
Jumbo		-0.276 ***	-0.191 ***	-0.191 ***	-0.198 ***
		(0.050)	(0.007)	(0.007)	(0.006)
GSE Conforming Loan		-0.142 ***	-0.100 ***	-0.100 ***	-0.101 ***
		(0.021)	(0.003)	(0.003)	(0.002)
Owner Occupied		-0.244 ***	-0.249 ***	-0.249 ***	-0.252 ***
		(0.029)	(0.004)	(0.004)	(0.003)
Points			-0.874 ***	-0.874 ***	
			(0.001)	(0.001)	
Purchased 1 Point					-0.497 ***
					(0.006)
Purchased 2 Points					-1.746 ***
					(0.002)
Constant	4.877 ***	5.137 ***	5.998 ***	6.127 ***	6.000 ***
	(0.009)	(0.029)	(0.004)	(0.004)	(0.004)
N	10,000	10,000	10,000	10,000	10,000
R^2	0.002	0.014	0.980	0.980	0.987

Notes:

Restricted to a random sample of 10,000 purchase loans with race reported from HMDA 2015.

Points and rates simulated.

Sources:

HMDA 2015, author's simulations.

Modeling Considerations: Race and Ethnicity

Levels of Aggregation

- Asians may report disaggregated categories

- Rates of reporting subcategories depend on:
  - Subcategory (Chinese, Filipino, etc.)
  - Unobservable factors correlated with interest rates
    - This can vary by subcategory

Rate Regressions

$$rate_i = \beta_0 + \beta_1 \times protected\ class_i + \gamma X_i + \varepsilon_i$$

Which protected class?
Can you compare protected classes?

Rate Regressions

$$rate_i = \beta_0 + \underbrace{\beta_1 \times Filipino_i + \beta_2 \times Chinese_i + \beta_3 \times Asian_i}_{} + \gamma X_i + \varepsilon_i$$

Which protected class?

Can you compare protected classes?

Reported Asian Subcategories

REPORTED ASIAN SUBCATEGORY[1]

Reported Asian Subcategory [a]	Percent [b]
Filipino	15%
Chinese	34%
Not Specified	51%

Note:
[1] Restricted to Asian respondents.

Sources:
HMDA 2015, author's simulations.

Respondents Are Actually 32% Filipino

Asian Subcategory Regressions

[a]	No Controls [b]	Loan Characteristic Controls [c]	Self-Reported Asian Subcategories[1] With Controls [d]
Filipino			0.042 ***
			(0.005)
Chinese			0.018 ***
			(0.004)
Asian	-0.006	-0.001	-0.013 ***
	(0.004)	(0.002)	(0.003)
Black	0.011 ***	-0.001	-0.001
	(0.003)	(0.002)	(0.002)
Native Hawaiian or Other Pacific Islander	0.009	0.006	0.006
	(0.014)	(0.007)	(0.007)
American Indian	0.007	0.002	0.002
	(0.010)	(0.005)	(0.005)
Constant	5.488 ***	5.999 ***	5.998 ***
	(0.001)	(0.002)	(0.002)
GSE, Jumbo, Owner Occupied Controls	No	Yes	Yes
N	50,000	50,000	50,000
R^2	0.000	0.733	0.733

Notes:

[1] Reporting of Asian subcategories is correlated with interest rate, and the relationship differs by subcategory.

Restricted to a random sample of 50,000 purchase loans with reported race, ethnicity, and applicant income.

Sources:

HMDA 2015, author's simulations.

More Potential Issues

- **Other Ways to Cut Categories**
  - Multiple race borrowers
  - Multiple borrowers with different races
  - Applicant race ordering
  - Co-applicant/applicant ordering

- **Measurement Error (may cause bias):**
  - The extent to which lender-reported race is reported with error may differ:
    - Depending on which borrowers choose to report race
    - Between different loan officers

Modeling Considerations: Product Choice

Potential Selection in Product Choice

- 2 parents seeking to use home equity to finance college costs
- Identical HMDA characteristics
- Both 48 years old with children entering as freshmen at the same school with $26,000 annual cost of attendance

Applicant 1	Applicant 2
Has an "Exceptional" Child	Skeptical about Child's Ability to Complete Degree
May be more willing to borrow full amount up front.	May be less willing to borrow full amount at once.
Closed-End Home Equity Loan	HELOC (open ended) Draw Every Semester

Challenges in Comparing Borrowers Like these Applicants

- **Omitted variables:** Does borrower think his child is “Exceptional”
  - If the threshold for thinking a child is “Exceptional” differs between protected and non-protected classes, then draw rates may differ among those who select the HELOC in the two groups.

- **Product Selection:** How do we compare HELOC and home equity loan terms?
  - Interest rates
  - Amount borrowed: size of line vs. amount drawn
  - Expected total borrowing costs
    - Applicant’s expectations are based on expectations of child
    - Lender’s expectations are based on expectations of portfolio
      - Historic performance and credit risk models
      - Expectations over a distribution of borrower’s children

Simulations

- Assumptions
  - Assessment of Child Abilities

	White	Protected Class
Exceptional	75%	25%
Skeptical	25%	75%

  - Product Choice

	Exceptional	Skeptical
Home Equity Loan	75%	25%
HELOC	25%	75%

Regression Results

[a]	Odds Ratio of HELOC (Logit) [b]	Rate, no HELOC Control [c]	Rate, With HELOC Control [d]
Protected Class	2.609 ***	0.119 ***	0.002
	(0.146)	(0.007)	(0.003)
Unsecured	1.015	0.249 ***	0.248 ***
	(0.043)	(0.005)	(0.002)
HELOC			0.501 ***
			(0.002)
Constant	0.605 ***	6.189 ***	6.001 ***
	(0.017)	(0.004)	(0.002)
N	10,000	10,000	10,000
R^2		0.212	0.885

50 bps HELOC premium

Note:

Restricted to a random sample of 10,000 home improvement loans with populated information for race, applicant income, and ethinicity.

Sources:

HMDA 2015, authors' simulations.

Points to Discuss With Your Economist

- What is being measured?
  - By the field being used?
  - Who reports/when is it reported?
  - Do you record how you record information?
- How do you use this information in the course of business?
- What choices do consumers have?
  - In what order?
  - How is the equivalence of alternatives documented?
- Please let us know if something is unclear!

•

Dr. Xiaoling (Ling Ling) Ang
Principal Consultant
xang@edgewortheconomics.com
202-435-7744

Dr. Stephen Bronars
Partner
sbronars@edgewortheconomics.com
202-580-7777

Sanford Shatz
McGlinchey Stafford
sshatz@mcglincey.com
949-381-5911