Algorithmic Underwriting in High Risk Mortgage Markets

Janet Gao[1] Livia Yi[2] David Zhang[3]

[1]Georgetown University

[2]Boston College

[3]Rice University

The Mortgage Market Research Conference
May 16, 2024

Introduction

- **Underwriting** is a screening procedure in which lenders collect documents from loan applicants, verify background and financial information, and assess credit risk
- Traditionally performed by humans, underwriting has become **increasingly automated**. Almost all lenders use automated underwriting systems (AUS) in some aspects of lending.
- **Research Question:** How does an increasing reliance on algorithmic underwriting affect the trade-off between risk management and financial inclusion?
- **Setting**: U.S. FHA policy that transitioned from pure human underwriting to increased reliance on algorithmic underwriting (via AUS) in August 2016. Affected the "high-risk" group: Credit Score < 620 & Debt-to-Income (**DTI**) > 43.

Introduction

- **Underwriting** is a screening procedure in which lenders collect documents from loan applicants, verify background and financial information, and assess credit risk
- Traditionally performed by humans, underwriting has become **increasingly automated**. Almost all lenders use automated underwriting systems (AUS) in some aspects of lending.
- **Research Question:** How does an increasing reliance on algorithmic underwriting affect the trade-off between risk management and financial inclusion?
- **Setting**: U.S. FHA policy that transitioned from pure human underwriting to increased reliance on algorithmic underwriting (via AUS) in August 2016. Affected the "high-risk" group: Credit Score < 620 & Debt-to-Income (**DTI**) > 43.

Introduction

- **Underwriting** is a screening procedure in which lenders collect documents from loan applicants, verify background and financial information, and assess credit risk
- Traditionally performed by humans, underwriting has become **increasingly automated**. Almost all lenders use automated underwriting systems (AUS) in some aspects of lending.
- **Research Question:** How does an increasing reliance on algorithmic underwriting affect the trade-off between risk management and financial inclusion?
- **Setting**: U.S. FHA policy that transitioned from pure human underwriting to increased reliance on algorithmic underwriting (via AUS) in August 2016. Affected the "high-risk" group: Credit Score $<$ 620 & Debt-to-Income (**DTI**) $>$ 43.

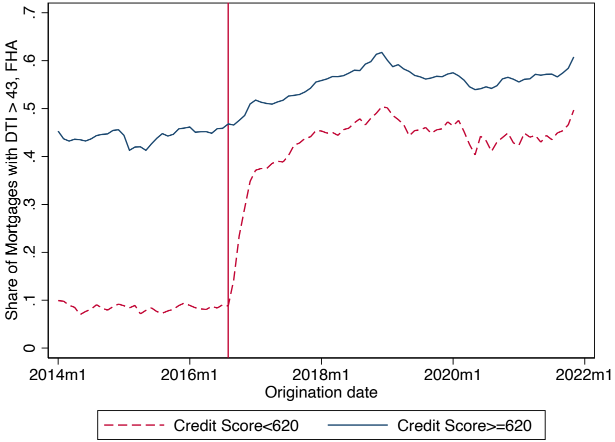
Introduction

- **Underwriting** is a screening procedure in which lenders collect documents from loan applicants, verify background and financial information, and assess credit risk
- Traditionally performed by humans, underwriting has become **increasingly automated**. Almost all lenders use automated underwriting systems (AUS) in some aspects of lending.
- **Research Question:** How does an increasing reliance on algorithmic underwriting affect the trade-off between risk management and financial inclusion?
- **Setting**: U.S. FHA policy that transitioned from pure human underwriting to increased reliance on algorithmic underwriting (via AUS) in August 2016. Affected the "high-risk" group: Credit Score < 620 & Debt-to-Income (**DTI**) > 43.

Data

1. **Ginnie Mae**: Near-universe of FHA mortgages, including interest rates, delinquency status, DTI, credit score, and other underwriting variables, from 2014 onwards,
2. **HMDA**: Race/ethnicity and income demographics, merged to Ginnie Mae data, 2014–2017
3. **Experian**: Consumer credit panel at an annual level, 2015–2017.
4. **CoreLogic LLMA**: Data from a selection of mortgage servicers including for non-FHA loans, used to estimate interest rate elasticities and for certain robustness checks, 2014–2017.

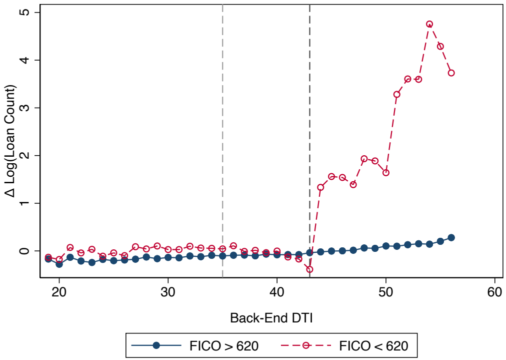
Policy window: 12 months before and after August 2016.



Source: Ginnie Mae data from January 2014 to January 2022

Effects of Policy on Credit Quantity: ΔLog(Loan Count) By DTI

- Descriptive evidence: Changes in log(#loans) from [Aug 2015, July 2016] to [Sep 2016, Aug 2017]



We quantify policy impact across DTI by estimating a counterfactual DTI distribution

- **Methodology**: DeFusco, Johnson, and Mondragon (2020)
- **Assumptions:**
  1. **High credit score (> 620) borrowers unaffected**
     - Already implemented algorithmic underwriting; no impact from the policy
  2. **Very low-DTI ($< \bar{d}$) borrowers unaffected**
     - Provides normalization to adjust for size differences between low- and high-credit-score markets
  3. **Growth of loans in a given DTI bin d for affected borrowers (credit score< 620) would have been the same as that of unaffected borrowers (credit score> 620) absent of the shock**

$$\frac{\hat{n}_{ld}^{post}}{N_{l\bar{d}}^{post}} = \underbrace{\frac{n_{ld}^{pre}}{N_{l\bar{d}}^{pre}}}_{\text{pre-event distribution}} + \underbrace{\left(\frac{n_{hd}^{post}}{N_{h\bar{d}}^{post}} - \frac{n_{hd}^{pre}}{N_{h\bar{d}}^{pre}}\right)}_{\text{change in distribution for high FICO borrowers}} \tag{1}$$

     - "Parallel trends" type assumption. Validate by showing that it generates accurate counterfactual distributions in placebo years.

We quantify policy impact across DTI by estimating a counterfactual DTI distribution

- **Methodology**: DeFusco, Johnson, and Mondragon (2020)
- **Assumptions:**
  1. **High credit score (> 620) borrowers unaffected**
     - Already implemented algorithmic underwriting; no impact from the policy
  2. **Very low-DTI ($< \bar{d}$) borrowers unaffected**
     - Provides normalization to adjust for size differences between low- and high-credit-score markets
  3. **Growth of loans in a given DTI bin d for affected borrowers (credit score< 620) would have been the same as that of unaffected borrowers (credit score> 620) absent of the shock**

$$\frac{\hat{n}_{ld}^{post}}{N_{l\bar{d}}^{post}} = \underbrace{\frac{n_{ld}^{pre}}{N_{l\bar{d}}^{pre}}}_{\text{pre-event distribution}} + \underbrace{\left(\frac{n_{hd}^{post}}{N_{h\bar{d}}^{post}} - \frac{n_{hd}^{pre}}{N_{h\bar{d}}^{pre}}\right)}_{\text{change in distribution for high FICO borrowers}} \tag{1}$$

     - "Parallel trends" type assumption. Validate by showing that it generates accurate counterfactual distributions in placebo years.

We quantify policy impact across DTI by estimating a counterfactual DTI distribution

- **Methodology**: DeFusco, Johnson, and Mondragon (2020)
- **Assumptions:**
  1. **High credit score (> 620) borrowers unaffected**
     - Already implemented algorithmic underwriting; no impact from the policy
  2. **Very low-DTI ($< \bar{d}$) borrowers unaffected**
     - Provides normalization to adjust for size differences between low- and high-credit-score markets
  3. **Growth of loans in a given DTI bin d for affected borrowers (credit score< 620) would have been the same as that of unaffected borrowers (credit score> 620) absent of the shock**

$$\frac{\hat{n}_{ld}^{post}}{N_{l\bar{d}}^{post}} = \underbrace{\frac{n_{ld}^{pre}}{N_{l\bar{d}}^{pre}}}_{\text{pre-event distribution}} + \underbrace{\left(\frac{n_{hd}^{post}}{N_{h\bar{d}}^{post}} - \frac{n_{hd}^{pre}}{N_{h\bar{d}}^{pre}}\right)}_{\text{change in distribution for high FICO borrowers}} \qquad (1)$$

     - "Parallel trends" type assumption. Validate by showing that it generates accurate counterfactual distributions in placebo years.

We quantify policy impact across DTI by estimating a counterfactual DTI distribution

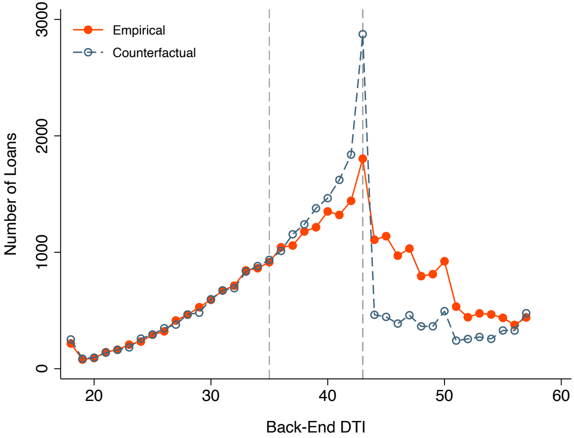
- **Methodology**: DeFusco, Johnson, and Mondragon (2020)
- **Assumptions:**
  1. **High credit score (> 620) borrowers unaffected**
     - Already implemented algorithmic underwriting; no impact from the policy
  2. **Very low-DTI ($< \bar{d}$) borrowers unaffected**
     - Provides normalization to adjust for size differences between low- and high-credit-score markets
  3. **Growth of loans in a given DTI bin d for affected borrowers (credit score< 620) would have been the same as that of unaffected borrowers (credit score> 620) absent of the shock**

$$\frac{\hat{n}_{ld}^{post}}{N_{l\bar{d}}^{post}} = \underbrace{\frac{n_{ld}^{pre}}{N_{l\bar{d}}^{pre}}}_{\text{pre-event distribution}} + \underbrace{\left(\frac{n_{hd}^{post}}{N_{h\bar{d}}^{post}} - \frac{n_{hd}^{pre}}{N_{h\bar{d}}^{pre}}\right)}_{\text{change in distribution for high FICO borrowers}} \tag{1}$$

     - "Parallel trends" type assumption. Validate by showing that it generates accurate counterfactual distributions in placebo years.

Policy impact by DTI

- **Extensive**: 10.3% total loan growth
- **Intensive**: shifting distribution
  - *Less bunching* left of threshold (9% "missing mass")
  - ΔAverage DTI $= 1.3$
- Event window: 12M before and after



Distributional impact across income and race/ethnicity

	Income		Race/Ethnicity		
	Below Median	Above Median	Non-Hispanic White	Black	Hispanic
Δ*Loans Originated*	0.038	0.136***	0.108***	0.014	0.109**
	(0.025)	(0.019)	(0.018)	(0.040)	(0.043)
Observations	324,061	324,058	428,086	83,120	112,658

- Sample: Ginnie Mae-HMDA merged, low FICO (<620) borrowers
- Weaker effects for lower-income and Black borrowers, highlighting the difficulty of increasing financial inclusion for those borrower groups.
- Later: use structural model to estimate the share of difference explained by supply and demand factors.

Delinquency Risk

- Little change in delinquency rates conditional on FICO and DTI category

Sample	High DTI (> 43)			Low DTI (≤ 43)		
Dep. Var.: *Delinquency Rate*	(1)	(2)	(3)	(4)	(5)	(6)
Treated (FICO<620) $\times$ *Post*	-0.00651	-0.00648	-0.00323	-0.0000618	-0.000317	0.00143
	(0.0116)	(0.0120)	(0.0123)	(0.00709)	(0.00740)	(0.00624)
Controls		Yes	Yes		Yes	Yes
Month FE	Yes			Yes		
FICO FE	Yes			Yes		
FICO-DTI FE		Yes	Yes		Yes	Yes
Month-DTI FE		Yes	Yes		Yes	Yes
County FE			Yes			Yes
Lender FE			Yes			Yes

Increase in dollar volume and delinquency rates in the FHA market

Dep. Var: *Volume ($ mil)*	(1) With Policy	(2) No Policy	(3) Difference (1)-(2)	(4) % Difference ((1)-(2))/(2)*100
Treated (FICO $<$ 620)	5,990***	5,189***	802***	15.5***
	(37)	(69)	(66)	(1.49)
Full Sample	73,411***	72,609***	802***	1.10***
	(103)	(121)	(66)	(0.09)

Dep. Var: *Delinquency Rate*	(1) With Policy	(2) No Policy	(3) Difference (1)-(2)	(4) % Difference ((1)-(2))/(2)*100
Treated (FICO $<$ 620)	12.92***	12.45***	0.47***	3.75***
	(0.60)	(0.48)	(0.19)	(1.47)
Full Sample	5.85***	5.76***	0.09***	1.61***
	(0.06)	(0.05)	(0.02)	(0.34)

Migration

- Given that low-credit-score households have improved access to credit from the FHA policy, do they become more mobile and migrate to neighborhoods with higher school quality?
- **Two-stage approach** to connect the effects to the FHA policy

$$\textit{New FHA Mortgage}_{i,t} = \beta_1 Treated_i \times Post_t + X_{i,t} + \alpha_{fico} + \tau_{z,t} + \phi_{g,t} + \eta_{a,t} + \epsilon_{i,t}$$

$$\textit{d(School Rating)}_{i,t} = \gamma_1 \widehat{\textit{New FHA Mortgage}}_{i,t} + X_{i,t} + \alpha_{fico} + \tau_{z,t} + \phi_{g,t} + \eta_{a,t} + \nu_{i,t},$$

  - **Data**: 1% **credit bureau panel** that tracks debt and location at annual freq.
  - *Treated* $= 1$ for individuals with FICO< 620 in 2015

Policy impact on neighborhood choice

Panel A. First Stage, Y = Obtaining FHA Mortgage

Post × Treat (2015)	0.0019***	0.0018***	0.0018***
	(0.0001)	(0.0001)	(0.0001)
F-statistic	380.40	313.03	319.34

Panel B. Second Stage, Y= Changes in School Ratings

New Purchase FHA	1.9332***	1.1625**	1.8315***
	(0.5196)	(0.5414)	(0.5302)
Individual Char	Yes	Yes	Yes
Year FE	Yes		Yes
FICO FE	Yes	Yes	Yes
Zipcode FE	Yes		
Zipcode-Year FE		Yes	
Gender-Zipcode FE			Yes
Age Group-Zipcode FE			Yes
Married-Zipcode FE			Yes
Observations	10,698,445	10,690,370	10,698,445

Structural mode, introduction

- **Next**: use a structural model to further separate supply vs demand and quantify welfare
- Intuition: assuming the target DTI distribution is smooth, kinds in the empirical distribution identifies supply restriction.
- Changes in bunching identifies % supply expansion, with the remainder to be explained by demand.

Changes in high-DTI loan eligibility

Panel B: % Changes in High-DTI Eligibility Rates

Full Sample	99.430***
	[92.656, 105.788]

Race/Ethnicity:	Non-Hispanic White	Black	Hispanic
	111.704***	63.729***	94.218***
	[103.696, 120.710]	[56.765, 71.157]	[78.483, 111.205]

Income:	Below Median	Above Median
	49.763***	152.373***
	[44.826, 55.145]	[143.491, 161.917]

- Large credit supply expansion, and some differential expansion by borrower race/ethnicity and income.
- Residual differences in demand still present: relaxing DTI constraint is insufficient for financial inclusion by race, likely due to other reasons such as down payment and information constraints.

Conclusion

1. Increased reliance on algorithmic underwriting can help **increase financial inclusion while controlling risk conditional on observables**, leading to **sizable gains in consumer welfare**
   - For **society**, trade-offs are not obvious, but the FHA's stated position is that making loans at these risk levels are of net social benefit (McFarlane, 2010).

2. The increase in financial inclusion was **not equally distributed**, but are concentrated on white and high-income borrowers
   - Highlights both demand and supply factors in limiting financial inclusion for these subgroups.

Conclusion

1. Increased reliance on algorithmic underwriting can help **increase financial inclusion while controlling risk conditional on observables**, leading to **sizable gains in consumer welfare**
   - For **society**, trade-offs are not obvious, but the FHA's stated position is that making loans at these risk levels are of net social benefit (McFarlane, 2010).
2. The increase in financial inclusion was **not equally distributed**, but are concentrated on white and high-income borrowers
   - Highlights both demand and supply factors in limiting financial inclusion for these subgroups.