Algorithmic Underwriting in High Risk Mortgage Markets

Janel Gao, Hanyi Livia Yi, and David Zhang

Philadelphia Federal Reserve Bank

Mortgage Market Research Conference

May 16, 2024

Discussion: Nancy Wallace

Haas School of Business

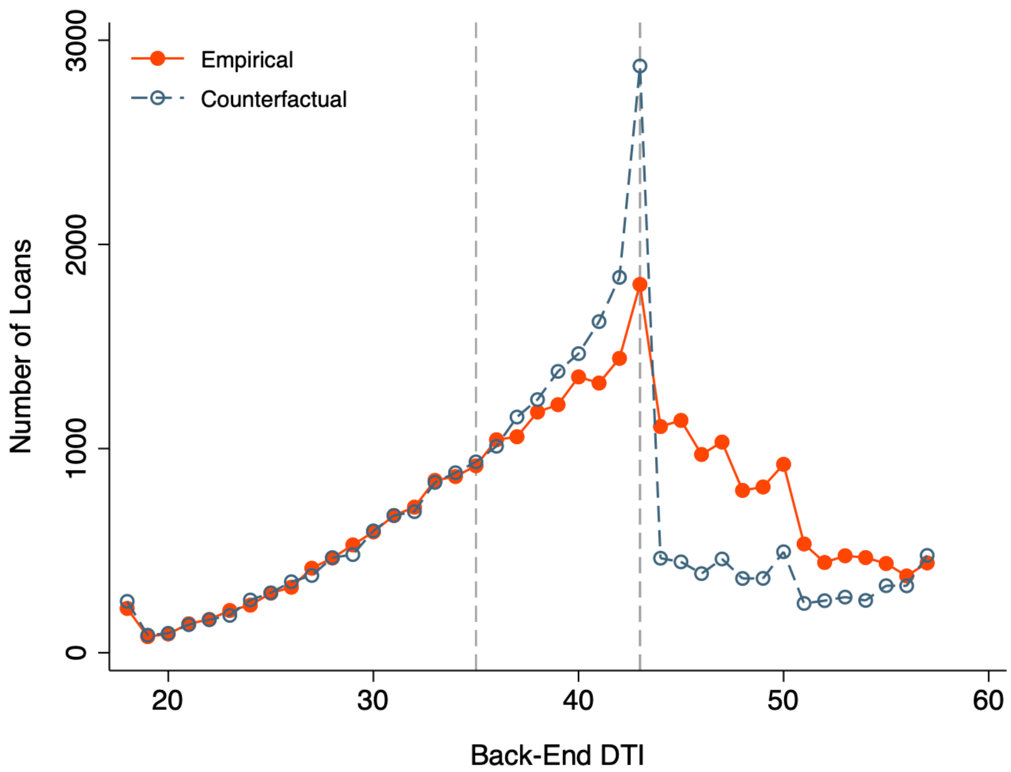
Purpose of the Paper

- Analysis of the 2016 FHA introduction of algorithmic underwriting (AU) for GNMA originations when:
  - Credit scores < 620, Back ended DTI's > 43%.
- **Reduced form analysis of AU policy change on:**
  - **Lender risk management:** efficiency without performance losses, enhanced capacity, reduction in regulatory interventions,
  - **Borrower financial inclusion:** disparate impact, MTO, pricing.
- **Structural demand estimation for borrower welfare measurements.**
  - **Headline result;** 1% change in borrowers' change in DTI target equivalent to 59 bps decrease in their contract rate.

Reduced form strategy

- Counterfactual estimation where placebo is the high-credit-score borrowers who are unaffected by the DTI policy change (Defusco, Johnson, Mondragon, 2020).
- **Efficiency channel:** test of policy effect on performance:
  - DID to treated (low credit score) and control (high credit score) borrowers for loans above and below the DTI cutoff (43%).
  - Triple difference framework to compare delinquency rates of low-credit score, high DTI loans relative to all other groups.
  - Repeat tests for fragile borrowers in areas with highest unemployment rate increases.
  - **Conclusion:** Default risk is not positively correlated with more lax DTI cutoffs and AU policy change, hence policy enhances efficiency.

Figure 4: Effect of the Policy Change on Loan Quantity



Note: This figure plots empirical and counterfactual number of FHA single-family, non-manufactured housing new purchase mortgages in our Ginnie Mae-Endorsements-HMDA sample 12 months after the policy based on the methodology described in Section 4.2. DTI is winsorized at the 1st and 99th percentiles and rounded down to the nearest integer. Dashed lines are drawn at DTI equals 43, above which the policy takes into affect, and at DTI equals 35, at or below which we assume is unaffected by the policy for our baseline bunching analysis. We show in this figure that this assumption along with a parallel trends assumption fits the data well for DTI≤35.

Reduced form strategies: Capacity and regulatory mitigation tests

- **Capacity channel:** tests on the magnitude of credit expansion across racial and income groups; effects for congested markets.
  - 12% increase for high income; 10% for White borrowers, only 3% for low-income (Black) borrowers.
  - Greater credit expansion in less congested markets.
  - **Conclusion:** Not consistent with simple capacity channel.
- **Regulatory channel:** tests on lenders with differential regulatory risks (i.e. bank and nonbank comparison).
  - At extensive margin, FHA mortgages originations increase by 13% for banks, 9.1% for nonbanks.
  - **Conclusion:** Consistent with regulatory mitigation channel.

Structural Demand Estimation: How does increased use of AU affect welfare?

- Borrowers optimization problem is to smooth a large expenditure in a single period $t = 0$, over a large number of future periods;
- Assume borrower's utility in the initial period is concave.
  - Thus, would borrow until the marginal utility of consumption in the first period is equal to the sum of the discounted future utilities, multiplied by the probability of surviving to each period, times the payment fraction $\phi(r)$.
- Can characterize the borrower surplus: (ρ is the interest rate where borrowing stops and $\bar{V}$ is the no-borrowing utility)

$$V(r) = \bar{V} + \underbrace{\left[\sum_{t=1}^{T} \beta^t (1-\delta)^t u'(w_t)\right]}_{\text{Utility weight}} \underbrace{\left[\int_r^{\rho} L^*(\hat{r}) \frac{d\phi}{dr} d\hat{r}\right]}_{\text{Borrower surplus triangle}},$$

Structural Demand Estimation

- Fit the model to key empirical moments:
  1. Empirical DTI distribution in pre and post-periods.
  2. Extensive margin response to policy change.
  3. Borrowers' extensive margin elasticity of demand to interest rate prior to the policy change.
- Borrower *i*'s utility from taking out a loan, *L*, is a linear function of DTIs and the interest rate

$$v_i^o(L,r) = -\psi|d_{i,r_0}^* - d_{i,r_0}(L)| - \gamma r + \xi^o + \epsilon_i^o$$

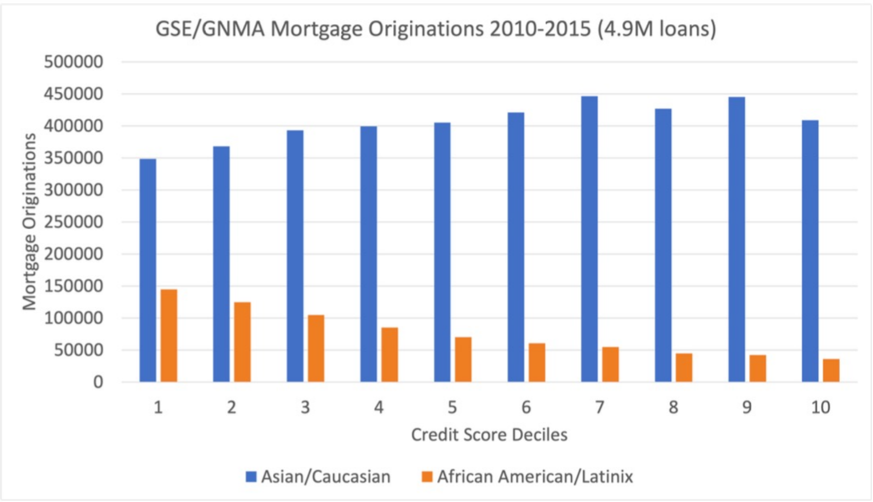
- **Results: Borrower's disutility from higher interest rates, $\gamma = 45$; borrower disutility estimate from not achieving their target interest rate, $\psi = .270$;**

Challenge of disentangling "TOTAL Scorecard Version 3" (2012) from policy change on "Manual Underwriting" (2016)

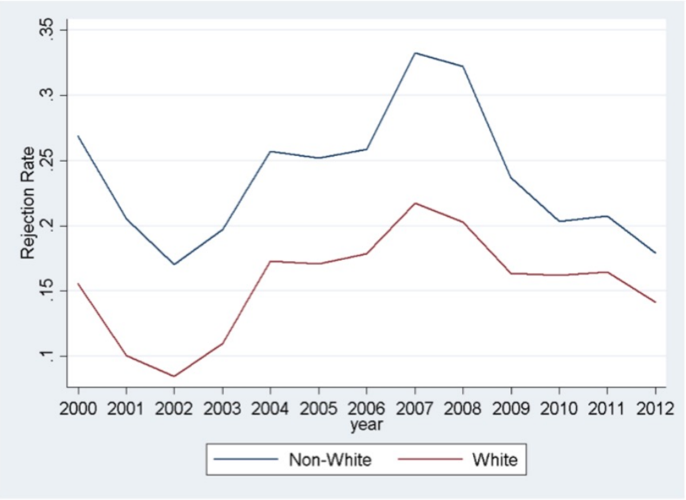
- Missing features from large endogenous contracting space
  - Loan amount, contract rate, LTV, FICO score, DTI, **points, origination fees**, MIP, CRA status, originator type (e.g. broker, correspondent, bank/nonbank)
  - Challenges with data availability (GNMA-HMDA, Experian, GreatSchools).
- TOTAL SCORECARD is a "machine learning application," developed in 2004
  - What do we know about the mathematical/statistical structure of this algorithm?
  - How should we think about the estimation errors of the algorithm (false positives, false negatives, individual vs. group model fairness).
  - What data are used to fit the algorithm?

Challenges of Class-imbalanced Credit Scoring Data Sets (FHA/GSE loans 2010-2015)

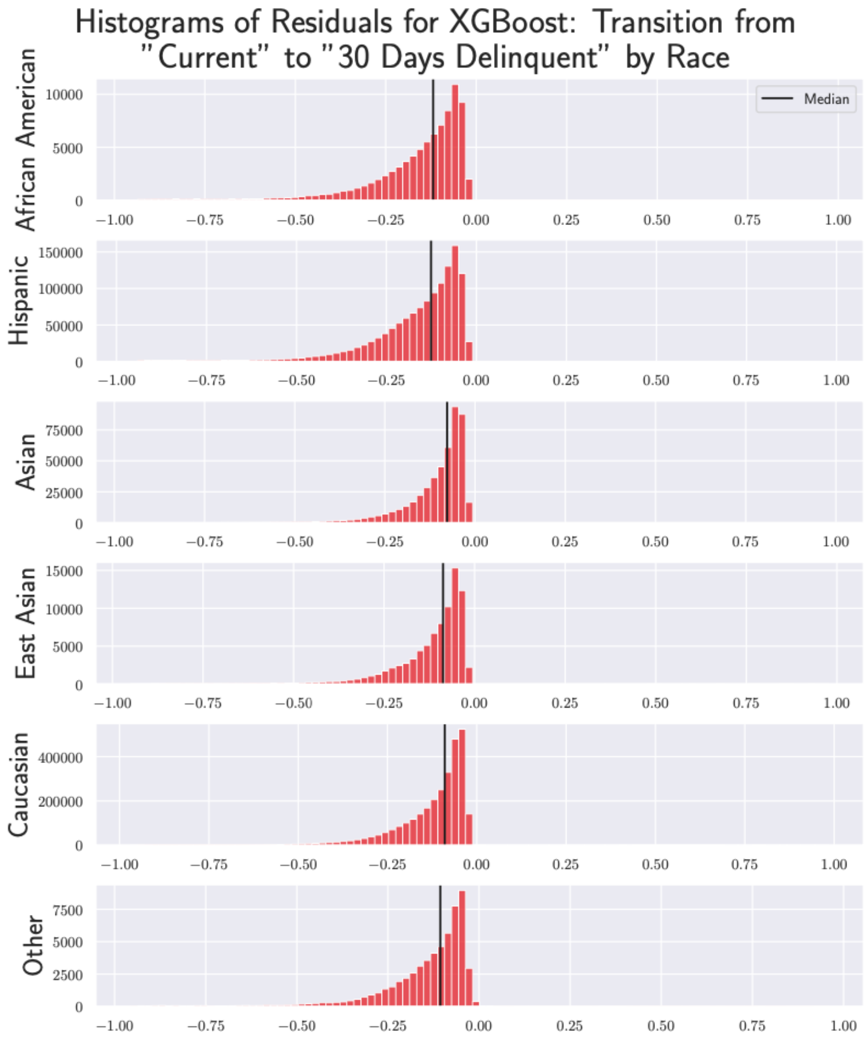
Credit score deciles by race/ethnicity (McDash/Equifax/HMDA/ATTOM merged data, 4.9 M loans)



Mortgage Applicant Rejection rates by race/ethnicity (HMDA data)

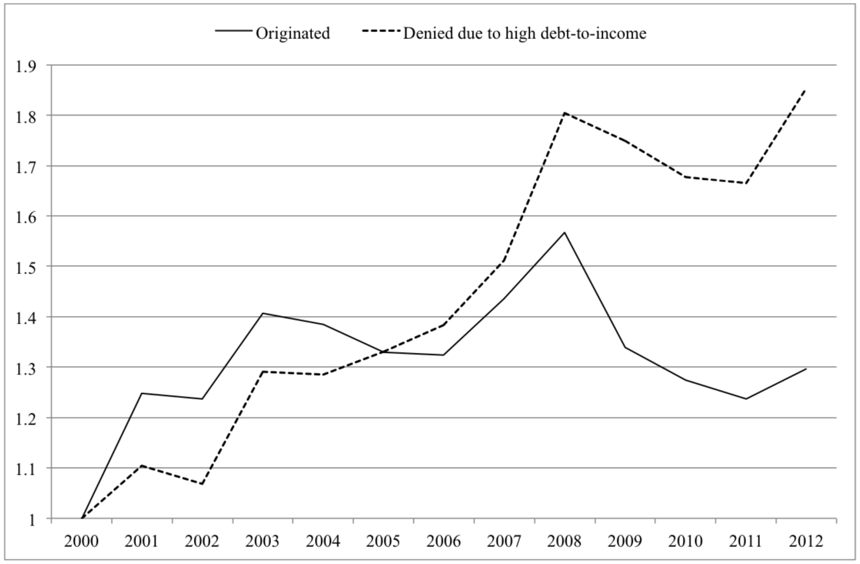


Do scoring models just reconstruct minority status imbalances? Residual Histograms



Challenge of disentangling the effects of policy uncertainty from other macro factors (Lucas critique, 1976)

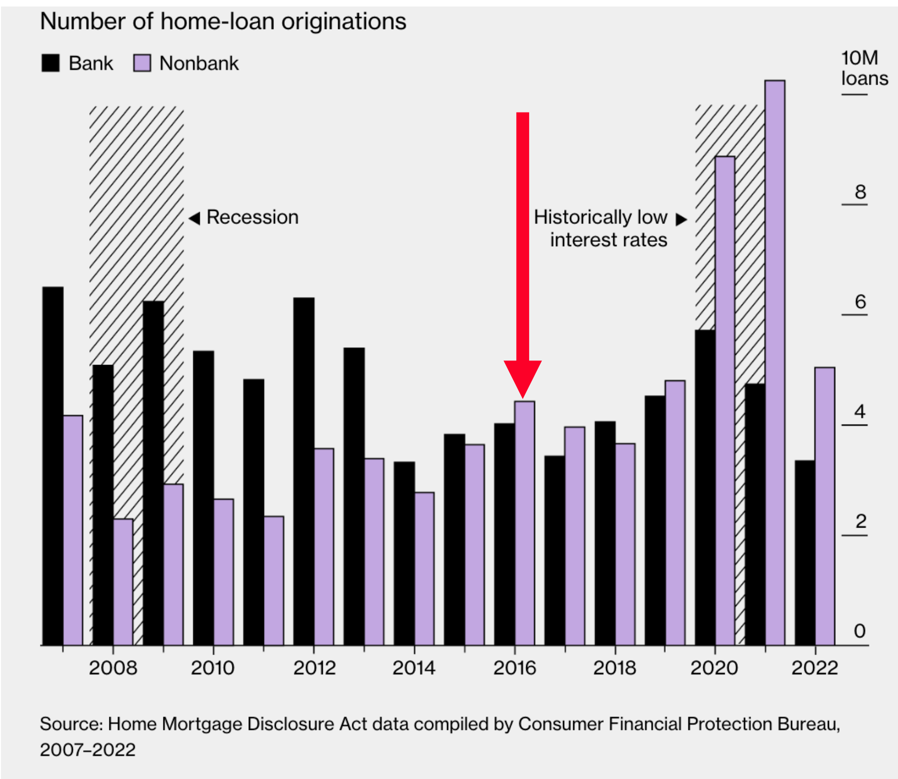
- **Effects of policy uncertainty on the supply of mortgage credit**
  - Especially during the rule making process for the TILA amendments
    - » "qualified mortgages" (QM)
    - » "ability-to-pay" (ATR).
  - Focus on debt-to-income (DTI) ratio – no clear DTI guidelines until January 2013.
  - TILA requirements for portfolio loans and FHA/GSE securitized loans were different.



HMDA Data: constructed DTI

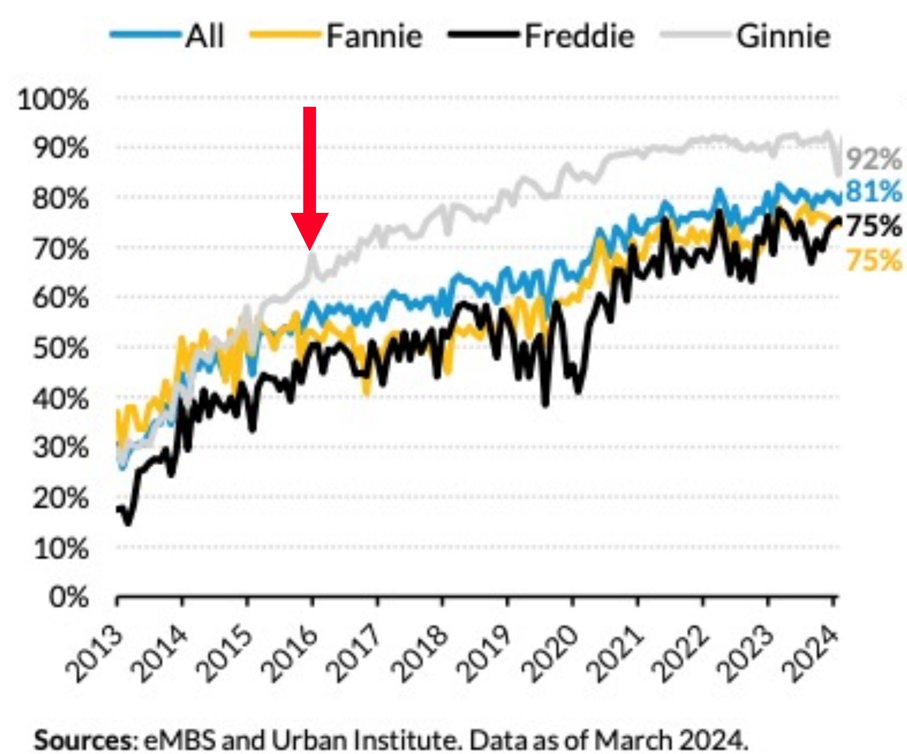
What about the post 2016 policy effects on nonbank dominance?

U.S. Bank/Nonbank Originations

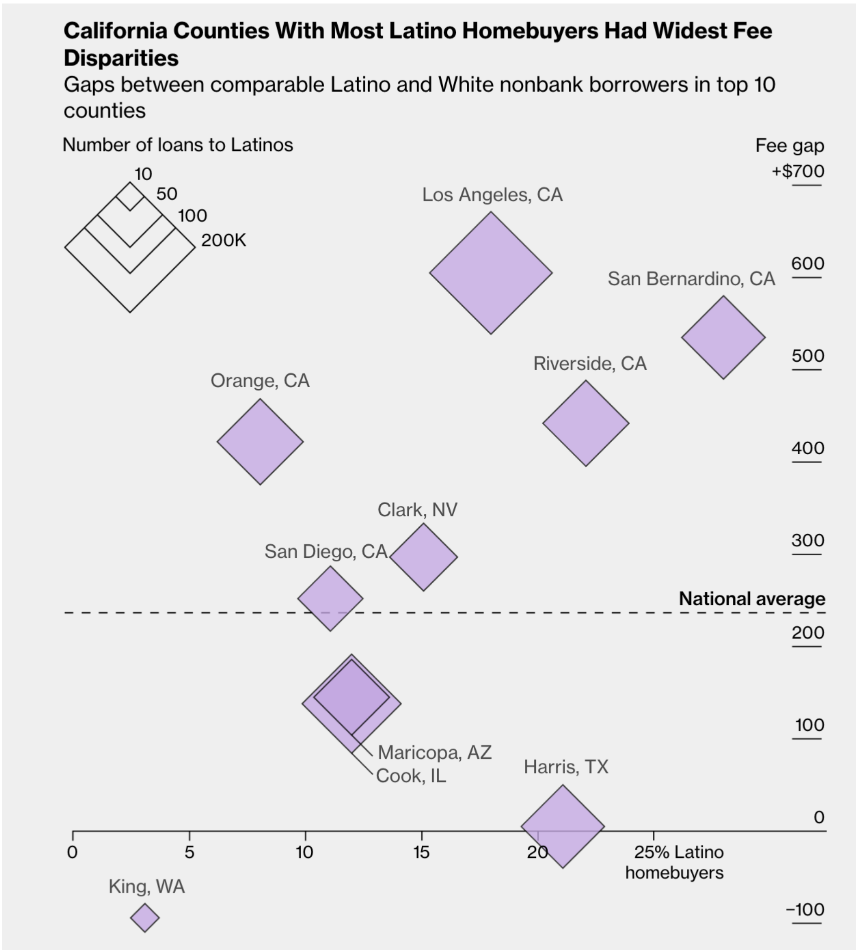


Source: Bloomberg

U.S. Nonbank origination share: Purchase loans



Challenge of disentangling disparate impact of policies (AU, market structure, pricing, other)?



Source: Home Mortgage Disclosure Act data compiled by Consumer Financial Protection Bureau, 2018–2022

Differential fees for Latinos compared to whites in high Latino homeowner markets.

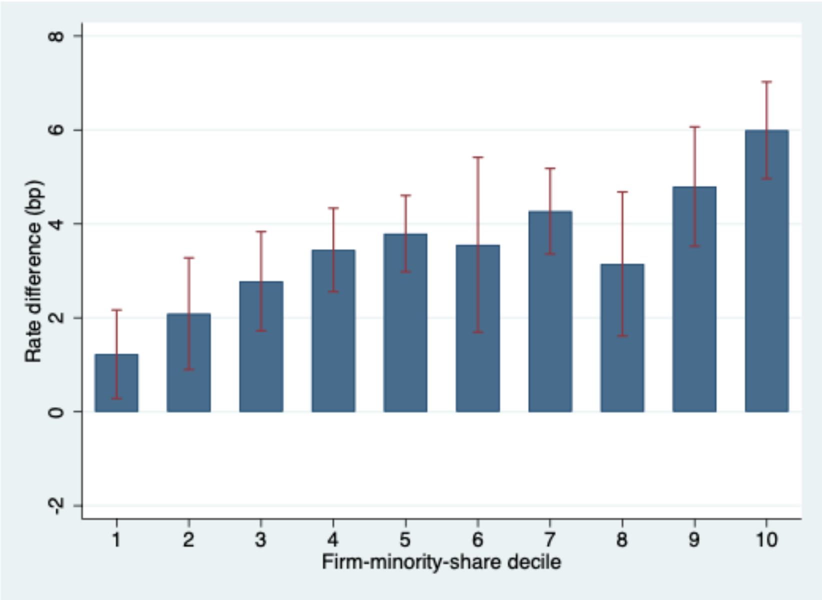
- Broker originators,
- Lack of shopping options.
- Language barriers,
- Other.

Source: Bloomberg, 2024

FHA Purchase Loan Rate Differentials

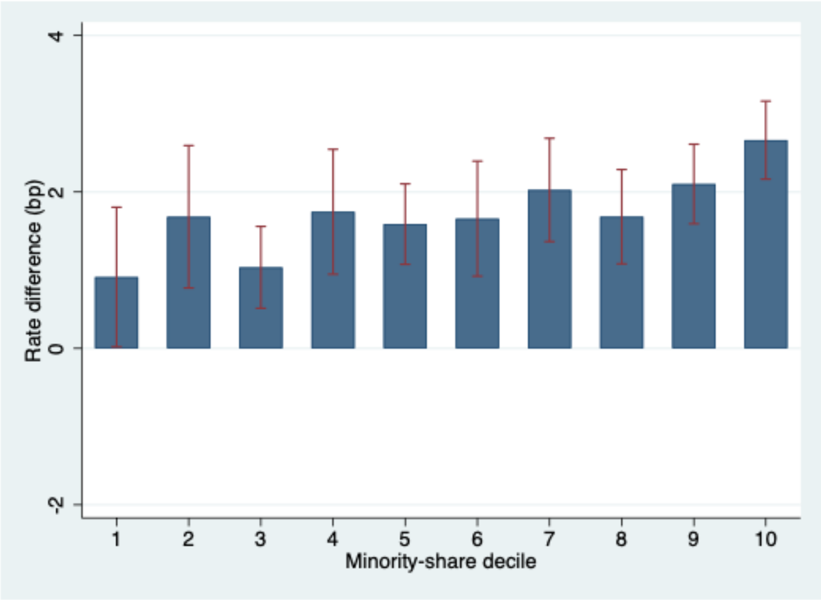
(2009- 2015 with controls for LLPA grid – FICO & LTV)

Firm minority share deciles



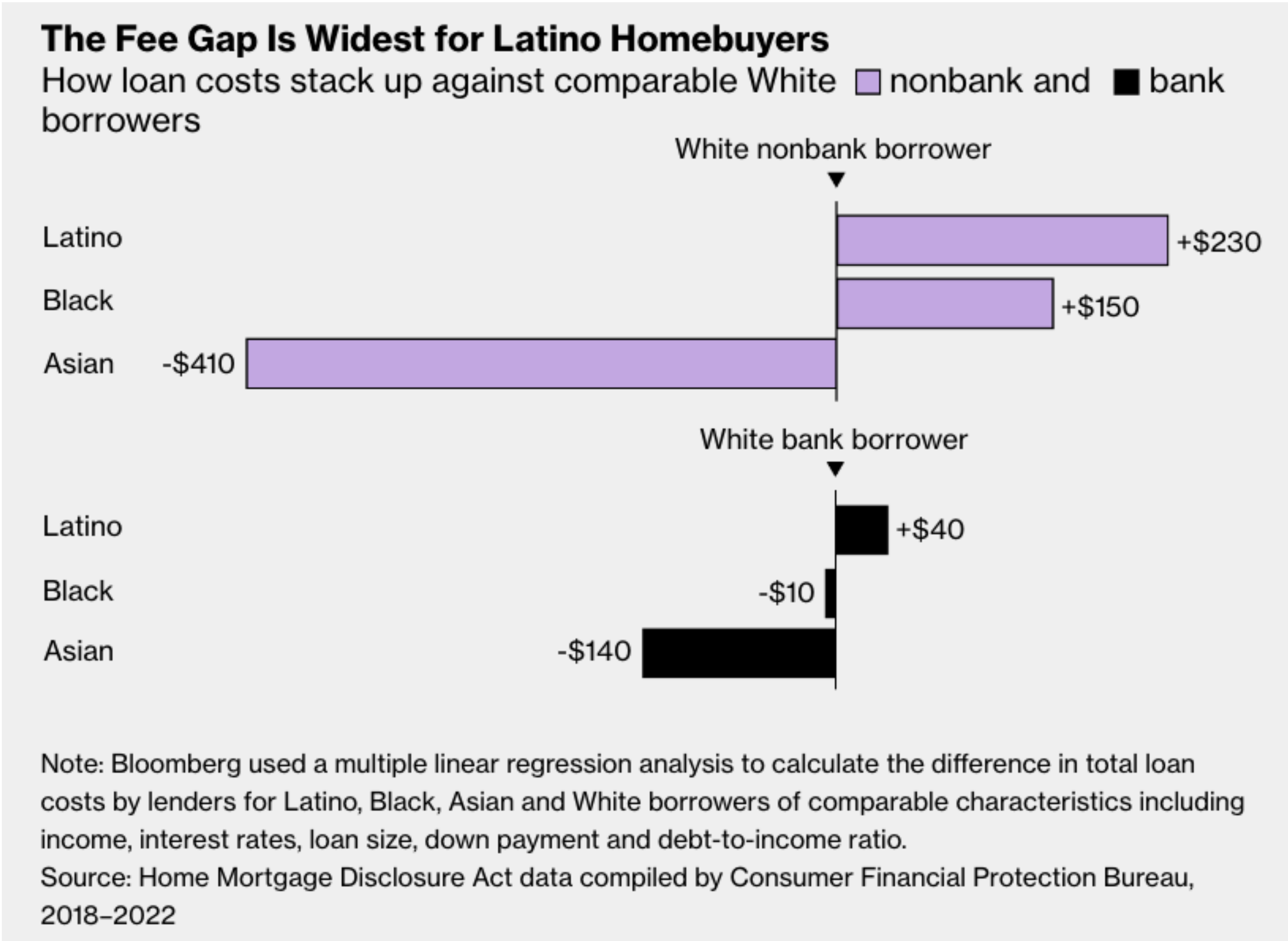
(c) FHA purchase loans

Census tract minority share deciles



(c) FHA purchase loans

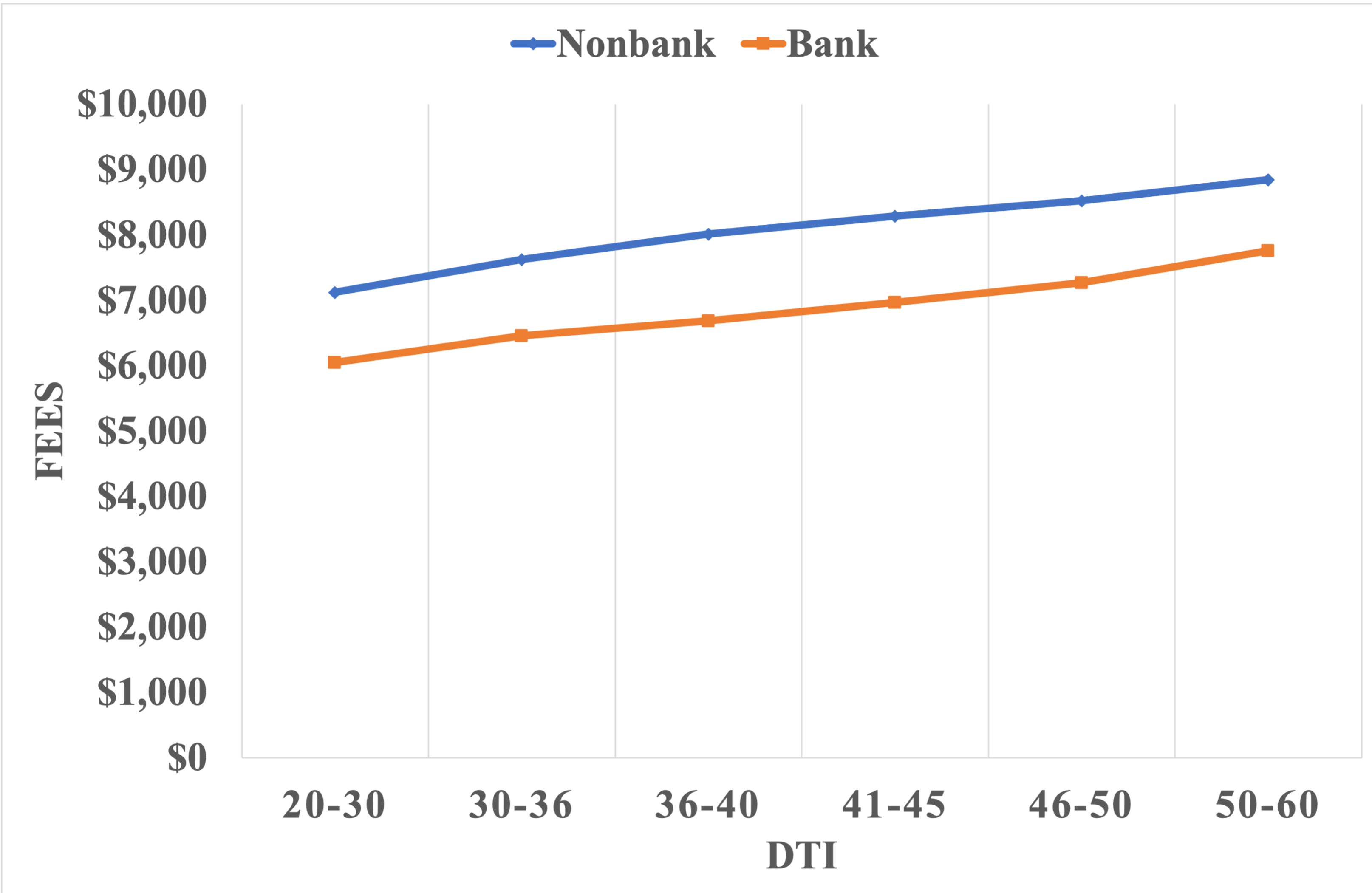
Nonbanks Charge Latino Borrowers Higher Fees Relative to White Borrowers



Source: Bloomberg, 2024

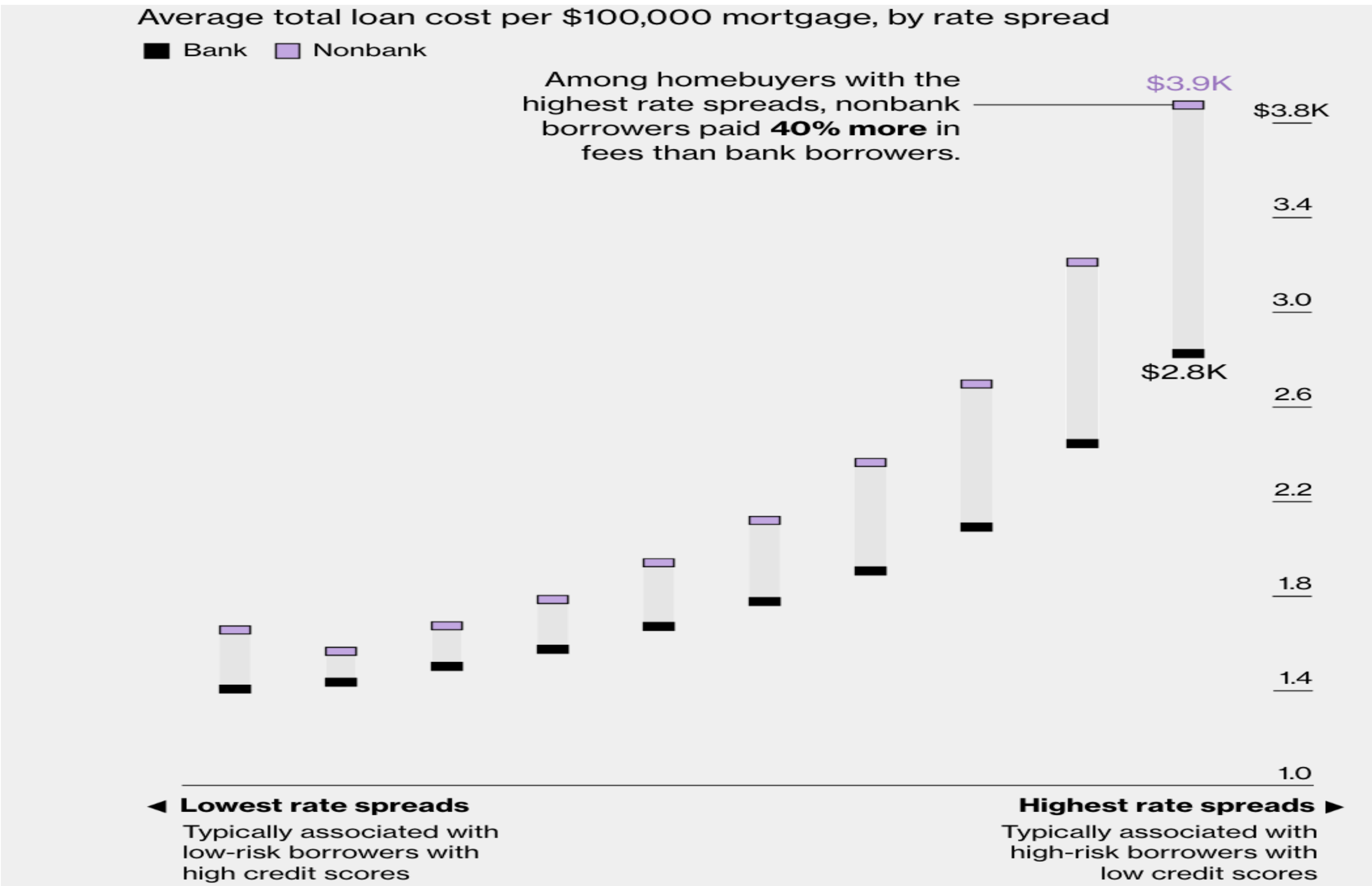
Nonbanks Charge Higher Fees by DTI

(FHA purchase loans -- HMDA 2019)



Nonbanks Charge Higher Fees by Rate Spread

(rate spread = contract rate minus prime mortgage rate)



Source: Home Mortgage Disclosure Act data compiled by Consumer Financial Protection Bureau, 2018–2022

Source: Bloomberg

Conclusions

- Excellent paper with well identified empirical strategies that deliver credible measurement of an important policy change (2016 use of algorithmic underwriting (AU) for "high-risk" GNMA originations decisions)
  - Efficiency and regulatory mitigation channels appear to drive results.
- Structural demand estimation indicates that borrowers are very sensitive to DTI constraints.

Conclusions

- **But,** primary beneficiaries are the White borrowers.
  - **Likely omitted variables bias:** disparate rate setting effects to LLPAs, disparate origination fee effects.
  - **Issues with AU technology:** Compared with prior technology some classes of borrower may now be systematically scored as riskier?
    - ML algorithms, by design, reduce the predictive mean-squared error,
    - ML algorithms will produce predictions with greater variance.