Discussion of "New Results on the Disparities Between Same-Sex and Different-Sex Couples in the Home Mortgage Market" by Nir Eilam and Yeonjoon Lee

Jinho Kim

University of California, Davis

WEAI Annual Conference
July 3, 2026

Summary of the paper

Question & data

- Disparities in mortgage outcomes by couple **sex-composition** (m–f, f–m, m–m, f–f).
- Linked confidential **HMDA–McDash 2018–21**, **AUS**, **SHED** survey.

Main results

- m–m applicants are **27.6%** more likely to be denied, and pay $+0.73\%$ on rate (conditional on approval) compared to m–f.
- Estimated gaps are about **half** the public-data estimates $\Rightarrow$ prior work likely overstated them (OVB).
- m–m default is $+53.9\%$ in the COVID cohort, with no gap pre-COVID.
- Loan officers overturn AUS *approvals* more often for m–m (**6.7%** vs. 3.8%).

A careful, well-executed paper on a question that matters — and it brings new data to it.

Comment 1: Measure $\rightarrow$ decompose

Comment 1: The field moved from measuring gaps to decomposing mechanisms

1. **Cross-sectional rejection regressions —** Munnell et al. (1996).
   Limit: omitted-variable bias — cannot control for everything lenders see.
2. **Outcome / default tests —** Berkovec et al. (1994, 1998).
   Limit: average default $\neq$ marginal default; default $\neq$ profitability.
3. **Audit / correspondence studies —** Ross et al. (2008), Hanson et al. (2016).
   Causal treatment evidence, but still hard to separate taste-based from statistical discrimination.
4. **Modern mechanism papers —** Bhutta–Hizmo–Ringo (2025), Bartlett et al. (2022), Frame et al. (2025). Use AUS, FinTech, performance, or loan-officer variation to decompose risk, treatment, hard / soft information.

$\Rightarrow$ **The confidential data seem well suited for a richer mechanism analysis.**

Comment 1: From "is there a residual gap?" to "what is in the residual?"

The AUS analysis comes close to the modern toolkit but could be extended.

- A descriptive override gap, on its own, may not separate **bias** from **noise** from **legitimate soft-information risk** (e.g. correlated occupational risk the AUS may not see).
- Whether human discretion adds risk information or mostly adds noise is itself an open question in this literature. (Jansen–Nguyen–Shams, 2025; Gao–Yi–Zhang, 2024)

Proposal: test human overrides using marginal outcomes

- Use the McDash match to ask whether human overrides *predict ex-post performance*
- The key is to avoid the **infra-marginality problem**: compare performance at the **approval margin** (Huang–Mayer–Miller, 2024).
  - **Feasible margin:** compare positive overrides (AUS negative $\rightarrow$ human positive)
  - **Stronger design:** use loan-officer / lender leniency IV to identify applicants approved at the margin.
  - Outcome should ideally be **profitability**; default or delinquency is only a proxy.
- This would turn the "human bias" result into a sharper Becker-style outcome test.

Comment 2: The gap may reflect correlated occupational risk

Comment 2: The gap may reflect correlated occupational risk

Paper's evidence and interpretation

(1) Main disparities are concentrated among m–m applicants.
(2) Larger gaps in the Midwest & South $\Rightarrow$ consistent with lower LGBTQ+ acceptance.
(3) AUS vs. final decisions $\Rightarrow$ suggestive evidence of human bias.

Alternative explanation: Correlated occupational risk rather than discrimination

- Male–male couples are more likely to have **both earners employed in cyclical industries**.
  - Albanesi-Şahin (2018): *"Gender differences in industry composition are the main source of the cyclicality of the unemployment gap."*
  - Hoynes–Miller–Schaller (2012): *"Construction and manufacturing are more cyclical industries while services and government are less cyclical."*
- Consequently, m–m couples may exhibit **greater household-level exposure to macroeconomic shocks.**
  - Consistent with this possibility, m–m default rates are significantly higher.

Comment 2: Men's unemployment is more cyclical

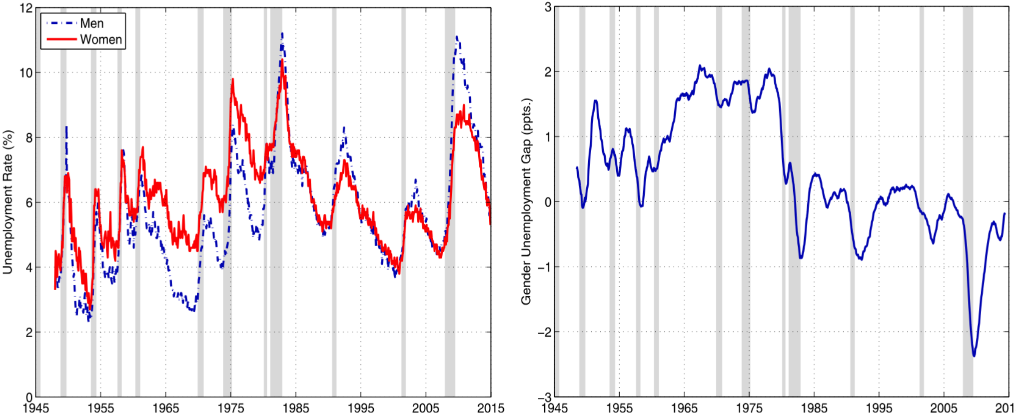
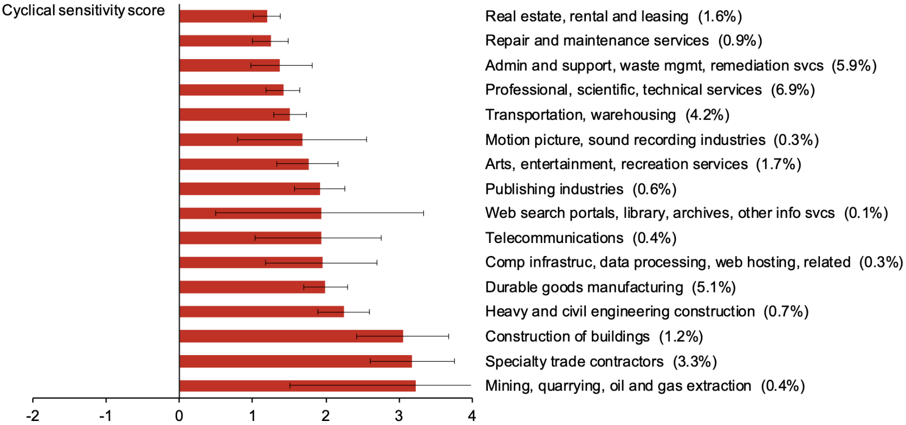


Fig. 1. Left panel: unemployment rate by gender (monthly). Right panel: the gender unemployment gap defined as the difference between the female and male unemployment rates (12-month centered moving average). Source: Current Population Survey.

Source: Albanesi-Şahin (2018), "The Gender Unemployment Gap," *Review of Economic Dynamics*

Comment 2: Men are more concentrated in cyclical industries

Chart 2
16 most cyclical sectors account for 34 percent of total employment



NOTES: Estimated score and confidence intervals are shown. Numbers in parentheses are percentages of total employment.
SOURCES: Bureau of Labor Statistics; Haver Analytics; authors' calculations.

Federal Reserve Bank of Dallas

Comment 2: Regional patterns are consistent with occupational risk

Regional heterogeneity is consistent with the mechanism, but cannot identify taste

- **Prediction of the mechanism:** The gap should be largest where male cyclical-sector concentration is highest → Midwest/South. The rejection pattern is consistent with this.
- **But taste is not identified**: at the region level, low LGBTQ+ sentiment and blue-collar occupational composition coincide. Taste and correlated-risk predict the same pattern.

Table: Cyclical Sector Employment Share by Region

Metric	**Northeast**	**West**	**Midwest**	**South**
Share of Emp. in Cyclical Sectors (%)				
Men	22.1	22.8	30.1	26.3
Women	6.3	7.0	9.9	7.4
Rejection Coefficient (vs. Male–Female)				
Male–Male	1.33	-0.18	5.41*	3.03**
Female–Female	-0.48	0.11	-0.10	1.28

Cyclical sectors: mining/oil & gas, construction, and manufacturing.

Comment 2: AUS may miss joint income risk

Why standard underwriting may miss this risk

- **AUS/FICO do not measure joint income-risk exposure.**
  - DU/LPA do not appear to use detailed occupation/industry codes; FICO reflects past credit behavior, not income volatility.
  - Neither directly models two earner's *joint* exposure to macro employment shock.
- **Human underwriting may incorporate soft information and loan-performance incentives.**
  - Lenders can review employer, job title, and roughly two years of income history, but usually do not observe a long earnings panel.
  - Selling to GSEs does not fully remove lender exposure to loan performance (Bosshardt–Kakhbod–Kermani, 2025).

Interpretation: the residual could reflect (1) *risk management*, (2a) *statistical discrimination*, or (2b) *taste-based discrimination*. → claiming (2b) requires stronger evidence.

Comment 2: Build a joint-risk proxy

Proposal: Add joint-risk×region controls

- Caveat: Same-sex couples may differ in education, income, and location, which can also shape occupation and industry choice (Del Río–Alonso–Villar, 2018)
  - → ACS PUMS identifies co-resident couples (MM, FF, MF) and reports each partner's industry/occupation and region.
- Build a predicted exposure index by region × couple type:

$$Exposure_{r,c} = E[H_{s_1} H_{s_2} \mid region = r,\ couple\ type = c]\,,$$

$$H_s = 1\{s \in \text{cyclical sectors}\}.$$

- Or use a continuous sector-sensitivity measure:

$$\Delta \log E_{s,t} = \alpha_s + \beta_s \Delta \log E_{\text{agg},t} + u_{s,t},$$

$$Exposure^{\beta}_{r,c} = E\left[\hat{\beta}_{s_1} \hat{\beta}_{s_2} \mid region = r,\ couple\ type = c\right].$$

- This gives a region-level predicted joint-risk proxy.

Comment 2: Test risk versus taste

A further regression could help separate statistical from taste-based discrimination

$$Rejected_i = \sum_{g \neq g_0} \beta_g \mathbf{1}\{G_i = g\} + \gamma Exposure_{r,c} + \delta LGBTQ\_Sentiment_r + X_i'\theta + FE + \varepsilon_i.$$

$Exposure_{r,c}$: predicted joint cyclical-labor-risk exposure by region r and couple type c.
$LGBTQ\_Sentiment_r$: regional support for marriage equality, e.g., PRRI American Values Atlas.

- If the m–m gap is concentrated in high-exposure regions $\rightarrow$ supports a risk/statistical-discrimination channel.
- If the m–m gap persists in low-exposure regions, or loads strongly on LGBTQ sentiment $\rightarrow$ supports a taste-based channel.

Potential contribution to the literature

- Prior same-sex mortgage studies have not, to my knowledge, modeled joint income-shock covariance from borrowers' industry/occupation pairings.
- Incorporating this channel would help separate correlated risk from differential treatment.

Comment 3: The right benchmark?

Comment 3: Additional benchmark – single male applicants

- Joint applications involve two institutional forces:
  - **Approval benefit:** two borrowers pool income and debt, potentially lowering DTI.
  - **Pricing cost:** the weaker co-borrower credit profile may bind for mortgage pricing.
- Therefore, applying jointly is itself informative:
  - households likely apply jointly when the expected approval gain is valuable.
- Male–female couples may be a "best-case" joint benchmark:

$$\text{income pooling} + \text{lower correlation in labor-income shocks.}$$

Comment 3: Additional benchmark – single male applicants

Proposal: add single-male applicants as a benchmark.

$$\widehat{r}_g \equiv \text{predicted rejection rate for group } g$$

- $\widehat{r}_{\text{MF}} < \widehat{r}_{\text{MM}} < \widehat{r}_{\text{M}}$: male–male applicants still receive a joint-application benefit, but less than male–female applicants.
- $\widehat{r}_{\text{MF}} < \widehat{r}_{\text{M}} \lesssim \widehat{r}_{\text{MM}}$: : male–male applicants receive little or no joint-application premium; this is more consistent with differential treatment.
- Caveat: This does not by itself prove discrimination.
  - joint and single applicants are selected differently;
  - male–male households may have higher correlated labor-income risk;
  - unobserved underwriting variables may still matter.

  But it is a useful diagnostic beyond the male–female benchmark.

Comment 3: Preliminary Test Using Public HMDA '18-'19

$$Rejected_i = \sum_{g \neq \text{single male}} \beta_g \mathbf{1}\{G_i = g\} + X_i'\gamma + \alpha_{c(i)\times t(i)} + \lambda_{\ell(i)} + \varepsilon_i.$$

$\alpha_{c(i)\times t(i)}$: county $\times$ year FE; $\lambda_{\ell(i)}$: loan-type FE.
X_i: income, loan size, LTV, property value, term, DTI, occupancy, race/ethnicity, age.

$\Rightarrow$ The results partially support the paper.

- Male–male applicants are the only group with a higher rejection rate than single-male applicants.
- However, the regional gap is significant only in the Northeast, which does not align with the paper's taste-based story.

Applicant type	**Rejection gap**	**s.e.**
Single female	-0.58^{***}	(0.09)
Male–male joint	$+0.72^{**}$	(0.31)
Male–female joint	-0.83^{***}	(0.11)
Female–male joint	-0.64^{***}	(0.14)
Female–female joint	-1.35^{***}	(0.32)

Region	**MM – single male**
Northeast	$+1.77^{**}$
Midwest	$+0.95$
South	$+0.43$
West	$+0.42$

Bottom line

1. **Decompose the residual:** move from *measuring* the gap to asking *what is in it* — e.g. test whether human overrides of the AUS predict ex-post performance.
2. **Correlated occupational risk:** a nationwide channel — m–m couples may carry higher joint labor-income risk. The regional pattern is consistent with it but cannot separate it from taste.
3. **Benchmark:** compare m–m against *single* borrowers, not only m–f.

Great data; the paper is well positioned to push from gaps toward mechanisms.