Generous to Men or Harsh to Women?

Experimentally Unpacking Gender Bias in Lending

Youngjoo Jung
(Ph.D. Candidate, UIUC)

October 2025

- Small businesses worldwide struggle with credit constraints, especially in low-income economies where financing options are limited.

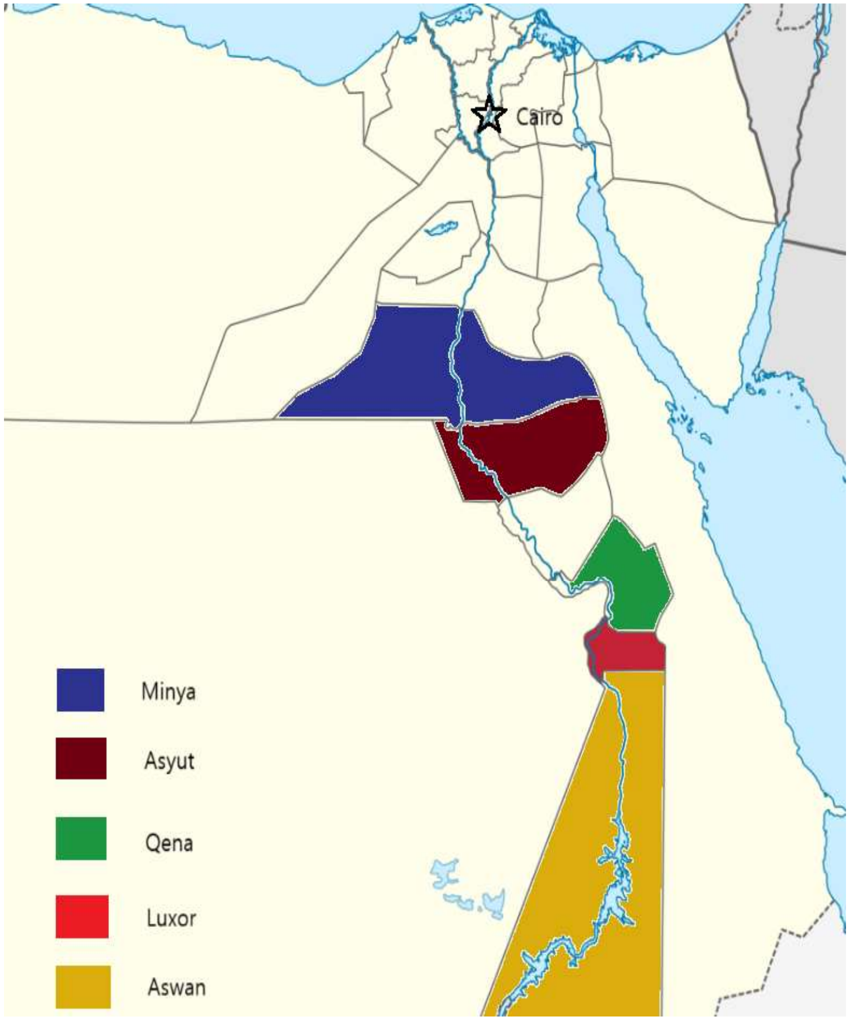
- As gatekeepers to capital, lenders' biases not only limits credit access, but can also contribute to capital misallocation and productivity loss.
    - Rigol and Roth 2021, Dobbie et al. 2021, Butler et al. 2023, Frame et al. 2025, Hsieh and Klenow 2009, Moll 2014, Hsieh et al. 2019

- The persistent gender gap in business lending suggests potential gender bias, but previous work has shown limited evidence.
    - Business loan borrowing experience in developing countries: men 12.3%, women 8.9% (The Global Findex 2025)
    - Demirguc-Kunt et al. 2018, Brock and De Haas 2023, Ubfal 2024

- Bias decomposition into generosity and harshness was proposed in psychology (Glick and Fiske, 1996), but rarely used in economics.
- If a gender gap favoring men exist, it could stem from generosity to men, harshness to women, or both.
  1. Generous to men: approving low-quality loans from male applicants.
  2. Harsh to women: rejecting high-quality loans from female applicants.
- Disentangling the nature of bias is empirically challenging but crucial.

	Inefficiencies	Policy Implications
Generous to men	lender loss (esp. in negative shocks)	male approval $\downarrow$
Harsh to women	borrower productivity + lender loss	female approval $\uparrow$

Experiment Context



- I traveled to Egypt and ran a randomized experiment with 720 loan officers in a context with potential gender bias.
- Implemented from Nov 2024 to May 2025.
- Randomized applicant names to test whether identical applications were approved differently by gender.

Research Questions

1. Is there gender bias in small business lending?
2. If so, is it mainly driven by generosity to men or harshness to women?
3. How can we reduce gender bias in lending?
4. Could AI (e.g., ChatGPT) help officers make more accurate decisions? But do officers use AI in a gender-neutral way?

Debiasing Strategies

Extrinsic and intrinsic approaches can be used to counter bias:

1. Higher performance pay (extrinsic): raising the cost of biased behavior.
   - Heywood and O'Halloran 2005; Heywood and Parent 2012; Marinescu et al. 2018; Mejia and Parker 2021
2. A nudge using Implicit Association Test (IAT) feedback (intrinsic): realizing one's unconscious bias prompts self-correction.
   - Greenwald et al. 1998; Alesina et al. 2024

→ How do debiasing policies affect gender bias and capital allocation?

What Do I Do?

- Officers approve or reject past loan applications with randomized gender.
- Randomization (4 groups):
  - control
  - gender-career IAT feedback
  - higher performance pay
  - combined interventions
- High vs. low-quality loans classification: ex-post outcome, late repayment history, and credit score.

Preview of Results

1. The same application is approved more with male names.
2. The gap is concentrated in low-quality loans, implying generosity toward men rather than harshness toward women.
3. A nudge using IAT Feedback does not appear to be effective, whereas higher performance pay closes the gender gap.

Contributions

1. The decomposition of bias into generosity and harshness has not typically been addressed in the broader bias literature.

→ I contribute to the literature on biased decision-making by being the first to experimentally separate generosity and harshness.

- Gender: Bertrand and Mullainathan 2004; Barasinska and Schäfer 2010; Bellucci et al. 2010; Agier and Szafarz 2013; Alesina et al. 2013; Stefani and Vacca 2013; Sarsons 2017; Beck et al. 2018; Andreeva and Matuszyk 2019; Hebert 2023; Brock and De Haas 2023; Ayalew et al. 2023; Montoya et al. 2024; Bartŏs et al. 2024
- Race: Berkovec et al. 1998; Altonji and Blank 1999; Blanchflower et al. 2003; Charles and Hurst 2002; Han 2004; Charles et al. 2008; Ross et al. 2008; Cohen-Cole 2011; Pope and Sydnor 2011; Deku et al. 2016; Hanson et al. 2016; Bayer et al. 2018; Bartlett et al. 2022; Butler et al. 2023; Frame et al. 2025
- Others: Fisman et al. 2017; Haselmann et al. 2018; Dobbie et al. 2021

2. While many observational studies report gender bias in lending, there are only a few experimental studies which show mixed results.

→ I contribute to the literature on loan officers' gender bias by providing the first experimental evidence of bias in small business loan approvals.

- Alibhai et al. 2019; Ayalew et al. 2023, Brock and De Haas 2023, Bartŏs et al. 2024, Montoya et al. 2024

3. Most debiasing work focuses on either extrinsic or intrinsic methods.

→ I contribute to the debiasing literature by comparing intrinsic, extrinsic, and their combination at the same time.

- Heywood and O'Halloran 2005, Beaman et al. 2009, Heywood and Parent 2012, Marinescu et al. 2018, Bohren et al. 2019, Mejia and Parker 2021, Montoya et al. 2024, Alesina et al. 2024

Outline

Experiment Design context

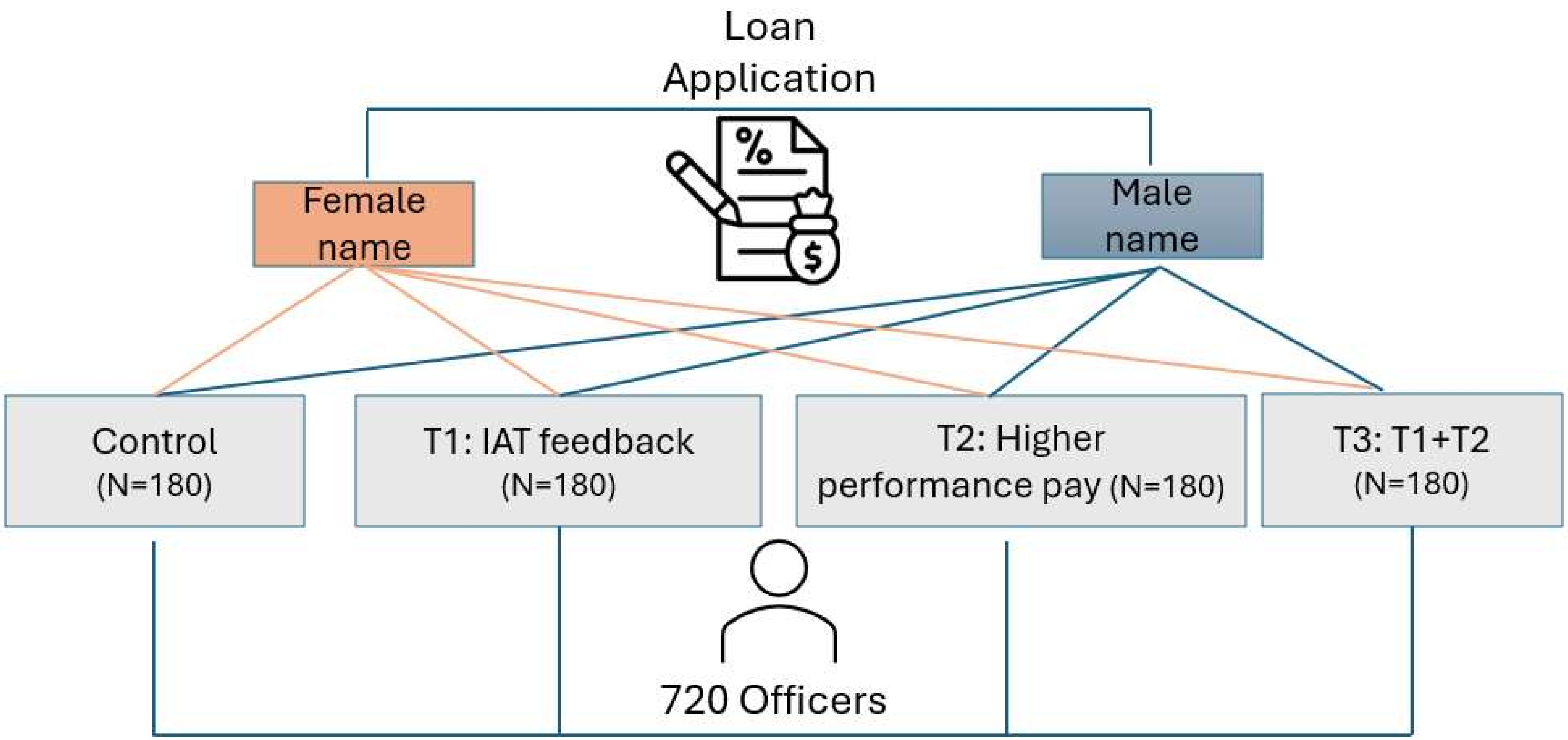
Loan officers: 50% female, avg. age 35, 16 yrs schooling, 6 yrs experience.

1. **Task**: Approve/reject 10 past approved loan applications. Example
   - Out of 10 loans, 6 are repaid and 4 defaulted. Comparison
   - 50% female applicants; repayment rates equal by gender.
2. **Incentives**: Incentives/penalties by ex-post outcome (Cole et al., 2015).
   - Basic payment: 400 EGP ($\sim$ \$8)
   - Approval: 10 EGP if repaid, -15 EGP if defaulted.
   - Rejection: -15 EGP if repaid, 10 EGP if defaulted.
3. **IAT**: Officers took the Gender–Career IAT before loan decisions. IAT

Taste vs. Statistical Discrimination

- Even with identical applications, if male businesses face more favorable environments, officers may be more likely to approve men.
- Fieldwork evidence suggests that both industries and officers' priors are gender-segregated: officers rarely visit other-gender businesses.
- To mitigate inaccurate statistical discrimination (Bohren et al., 2023):
  1. Using only industries with balanced gender composition.
     - ★ Agriculture, supermarkets, grocery stores, clothing shops, fish vendors, poultry, and livestock.
  2. Officers are told repayment outcomes are equal by gender on average.

Randomization: Applications x Officers Balance Check



1. **IAT feedback**: Receive feedback prior to loan decisions. Feedback
2. **Higher performance pay**: Doubled incentives/penalties. Incentive
3. **Combined**: IAT feedback and higher performance pay together
   $\rightarrow$ possible complexity (multitasking theory, Holmstrom and Milgrom 1991)

Outline

Q: Do loan officers approve same applications more with male names?

$$Y_{ip} = \alpha + \beta Male_p + \mu_p + \delta_{LO\ gender} + \epsilon_{ip}$$

($i =$ loan officer, $p =$ loan application)

- Y_{ip}: 1 for approval, but 0 for rejection.
- $Male_p$: 1 for male, but 0 for female application, μ_p: loan application FE.
- $\delta_{LO\ gender}$: loan officer gender FE, ϵ_{ip}: error term.
- Standard errors are clustered at loan officer level.

Strong Evidence of Gender Bias in Loan Approval

	Approval Rate
Male Names	0.044**
	(0.019)
Female Applicants Approval Mean	0.708
Number of Applications	1800
Number of Loan Officers	180
Loan Application FE	✓
Loan Officer Gender FE	✓

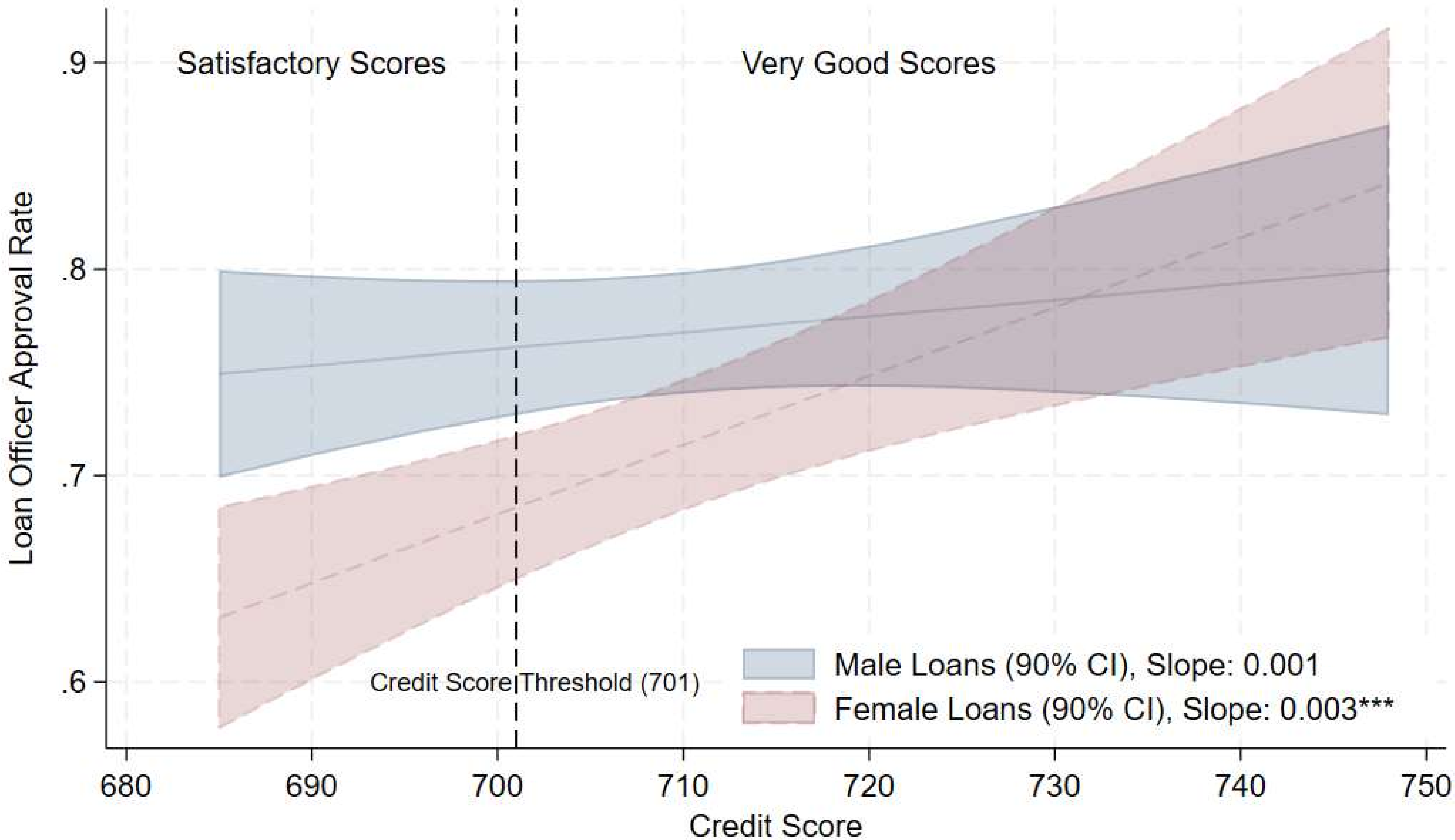
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Nature of Gender Bias

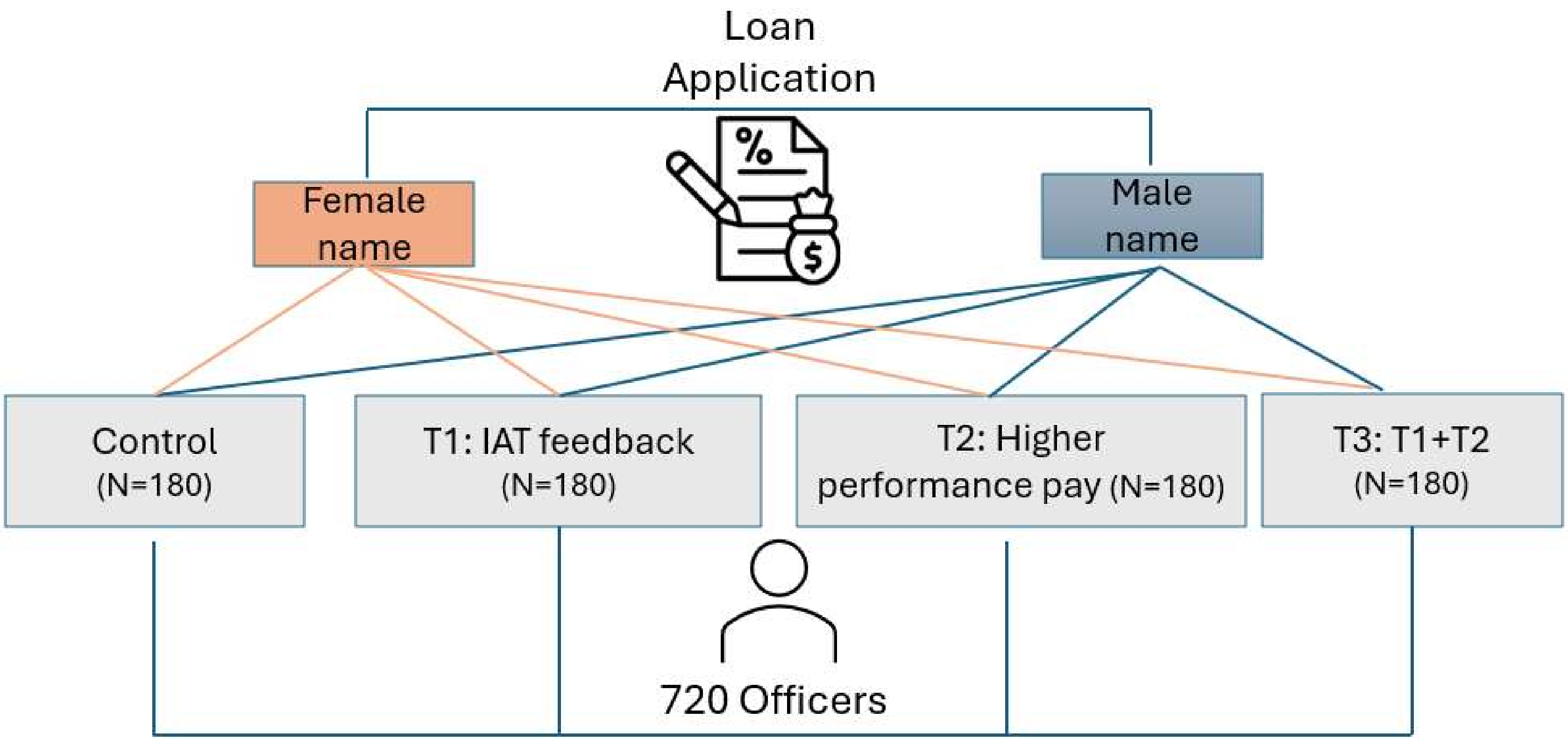
- Where does the gender gap come from—approving low-quality men (generous to men) or rejecting high-quality women (harsh to women)?
- I find the pro-male gender gap is concentrated in low-quality loans, indicating generosity toward men.
    1. Ex-post defaulted ($14.7\%^{***}$) vs. repaid loans (1.2%) Table
    2. With late repayment in prior loan ($21.6\%^{**}$) vs. without (3.4%) Table
    3. Low credit score ($16.8\%^{***}$) vs. high (0.8%) Table

Less responsive to low credit scores for male loans

Scatter Plot | Regression | Fractional Polynomial | Robustness Check



Randomization: Applications x Officers



1. **IAT feedback**: Receive feedback prior loan decisions. Feedback
2. **Higher performance pay**: Doubled incentives/penalties. Incentive
3. **Combined**: IAT feedback and higher performance pay together
   $\rightarrow$ possible complexity (multitasking theory, Holmstrom and Milgrom 1991)

Treatment effects: Debiasing

Q: Does gender gap in approval change under IAT feedback (T1), higher performance pay (T2), or their combination (T3) relative to control?

$$Y_{ip} = \alpha + \sum_{k=1}^{3} \beta_k T_{ik} Male_p + \sum_{k=1}^{3} \gamma_k T_{ik} + \delta Male_p + \mu_p + \delta_{LO\ gender} + \epsilon_{ip}$$

(i = loan officer, p = loan application, k = treatment group)

- Y_{ip}: 1 for approval, but 0 for rejection.
- $T_{ik} = 1$ if officer i in treatment k, 0 otherwise (e.g., $T_{i1} = 1$ for T1, 0 for others).
- All other definitions remain the same as in the above equation.

Debiasing Effects: Higher performance pay and Combined

IAT Mechanism

Outcome: Approval Binary	Approval Rate
IAT Feedback (T1) x Male Names	-0.039
	(0.026)
Higher Performance Pay (T2) x Male Names	-0.066**
	(0.027)
Combined (T3) x Male Names	-0.048*
	(0.027)
Male Names	0.044**
	(0.019)
Number of Applications	7200
Number of Loan Officers	720

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Screening Time Change

Accuracy Heterogeneity

$$Y_{ip} = \alpha + \sum_{k=1}^{3} \gamma_k T_{ik} + \mu_p + \delta_{LO\ gender} + \epsilon_{ip}$$

	Average Screening Time per Loan (minutes)
IAT Feedback (T1)	-0.033
	(0.045)
Higher Performance Pay (T2)	0.124**
	(0.052)
Combined (T3)	0.006
	(0.044)
Control Mean	1.098
Number of Loan Officers	704

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

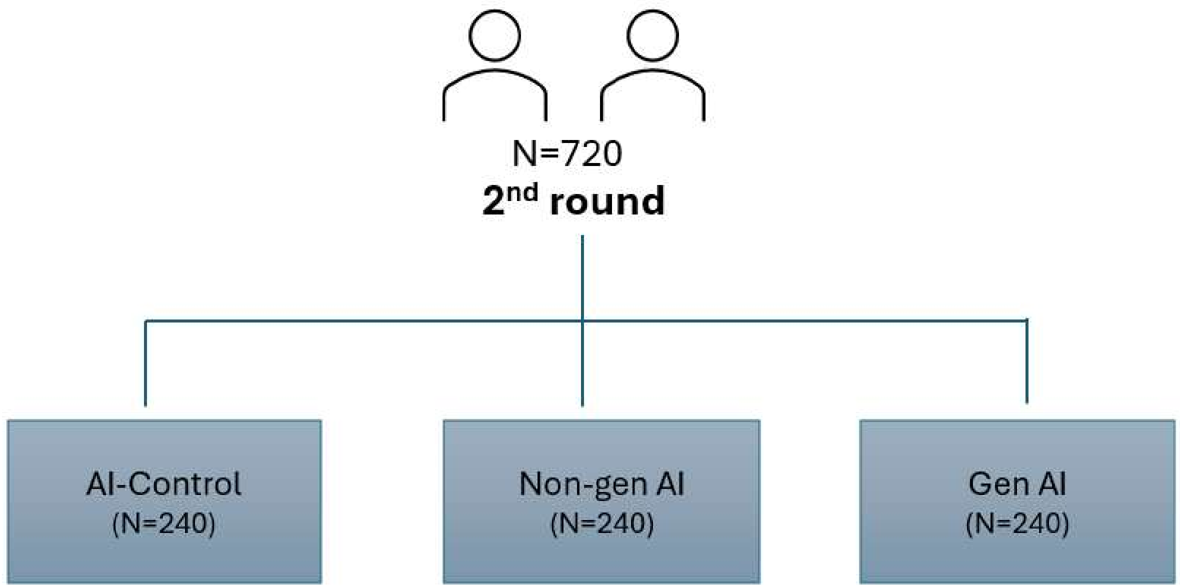
Outline

Beyond Debiasing: Augmenting Officer Ability with AI

- Monetary incentives have limits, as they do not enhance skills.
  - Enhancing screening ability is critical in low-income economies with high default rates and limited resources.
- Could AI (e.g., ChatGPT) help officers make more accurate decisions?
  - But do officers use AI in a gender-neutral way?

$\rightarrow$ In round 2, officers get assistance from an AI chatbot that I custom-built, suggesting approval or rejection along with supporting reasons. Chatbot

Two-Stage Randomization: First-round Stratified Full



1. Control (no AI): Loan decisions without AI-assistance.
2. Non-generative-AI (algorithm): Officers can view AI chatbot's recommendation, but cannot interact with it. Non-Generative AI
3. Generative-AI (Chat GPT): Officers can view and interact with the AI chatbot's recommendations by asking questions. Generative-AI

AI and Loan Screening Performance (Accuracy)

Q: Do officers using AI make more accurate loan decisions? Any difference between generative AI and non-generative AI?

$$Y_{ip} = \alpha + \gamma_1 \times \textit{Gen } AI_i + \gamma_2 \times \textit{Non-Gen } AI_i + \mu_p + \delta_{LO\ gender} + \epsilon_{ip}$$

(i = loan officer, p = portfolio, k = treatment group)

- Y_{ip}: 1 for correct, but 0 for wrong decisions.
- *Gen* AI_i: 1 for Generative AI users, and 0 for others.
- *Non-Gen* AI_i: 1 for Non-generative AI users, and 0 for others.

Officers with AI have Higher Performance (accuracy)

- The communication feature of generative AI adds no additional benefit over non-generative AI. Robustness Check

Outcome: Accuracy Binary	Ex-post Outcome Accuracy
Generative AI	0.066***
	(0.013)
Non-Generative AI	0.089***
	(0.013)
2nd Control Accuracy Mean	0.553
Number of Portfolios	7200
Number of Loan Officers	720

Follow Rate of AI Suggestions by Applicant Gender Bias

Q: Do officers follow the same AI recommendations more for male vs. female applicants? Does this differ between Gen and non-Gen AI?

$$Y_{ip} = \alpha + \beta \ \textit{Gen AI}_i \cdot Male_p + \gamma \textit{Gen AI}_i + \theta Male_p + \mu_p + \delta_{LO\ gender} + \epsilon_{ip}$$

(i = loan officer, p = portfolio)

- Y_{ip}: 1 for following AI's suggestions, but 0 for not following it.
- *Gen AI*$_i$: 1 for generative AI, but 0 for non-generative AI users.

Follow AI's Risky Loan Approval more for Male Names

Outcome: Follow AI Recommendation Binary	Ex-post Repaid Loans	Ex-post Defaulted Loans	
	(1) AI Decisions : Approve	(2) AI Decisions : Reject	(3) AI Decisions : Approve
Gen AI Binary x Male Names	-0.021	0.025	0.113***
	(0.025)	(0.056)	(0.042)
Gen AI Binary	-0.029	-0.012	-0.063*
	(0.021)	(0.041)	(0.033)
Male Names	0.024	0.013	-0.025
	(0.016)	(0.038)	(0.030)
Number of Portfolios	2880	960	960
Number of Loan Officers	480	480	480

- Gen AI can be a medium where gender bias manifests. Robust NLP
- Relative to non-Gen AI, when Gen AI advises approval for defaulted loans, officers are more likely to approve men and reject women.

Outline

Summary and Policy Implications

1. Evidence of gender-based discrimination in loan approvals.

   $\rightarrow$ Bias driven primarily by generosity toward men

2. Performance pay as a debiasing tool in lending

3. AI-assisted lending as a promising approach, though generosity toward men persists

Policy Implications

- Generosity toward men raises defaulted loans, lowers bank's profits.
  - Back of the envelope: Excessive generosity toward men drives up EGP 3.6 billion (USD 72.9 million) in annual losses for Egypt's banks. detail
- MFIs need to consider doubling down on performance pay. detail
  - Back of the envelope: Doubling performance-based pay can reduce annual bank losses in Egypt by EGP 3.2 billion (USD 64.3 million).
- AI-assisted lending is promising, but we must account for how users interact with it.