

Systemic Discrimination: Theory and Measurement

Aislinn Bohren
UPenn

Peter Hull
Brown

Alex Imas
Chicago Booth

NBER Labor Studies, March 2022

Introduction

- Disparities in treatment and outcomes (by, e.g., race/gender) are widely documented, but not always easily interpreted
 - In labor markets, housing, criminal justice, education, healthcare...

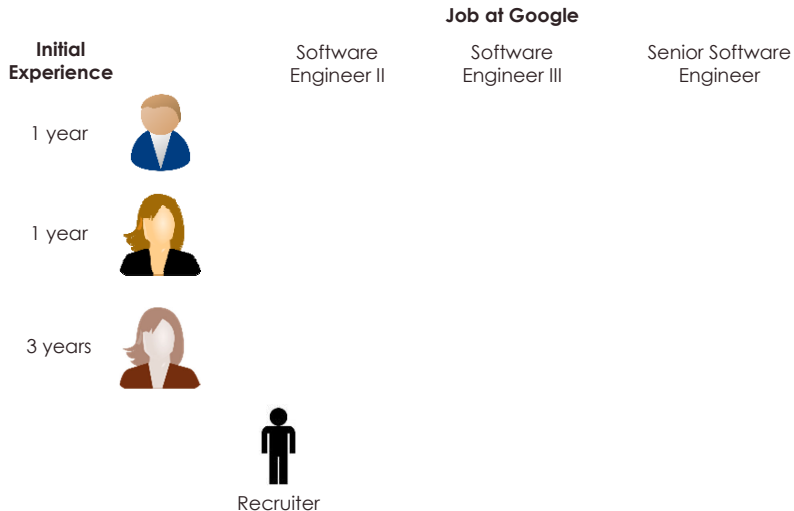
Introduction

- Disparities in treatment and outcomes (by, e.g., race/gender) are widely documented, but not always easily interpreted
 - In labor markets, housing, criminal justice, education, healthcare...
- In economics, disparity analyses tend to focus on *direct discrimination*: differential treatment on the basis of a protected characteristic
 - Theoretical models of how race/gender affect treatment through a decision-maker's preferences/beliefs (e.g. Becker 1957; Phelps 1972)
 - Empirical studies of the causal effects of perceived race/gender, holding other characteristics fixed (e.g. Bertrand & Mullainathan 2004)

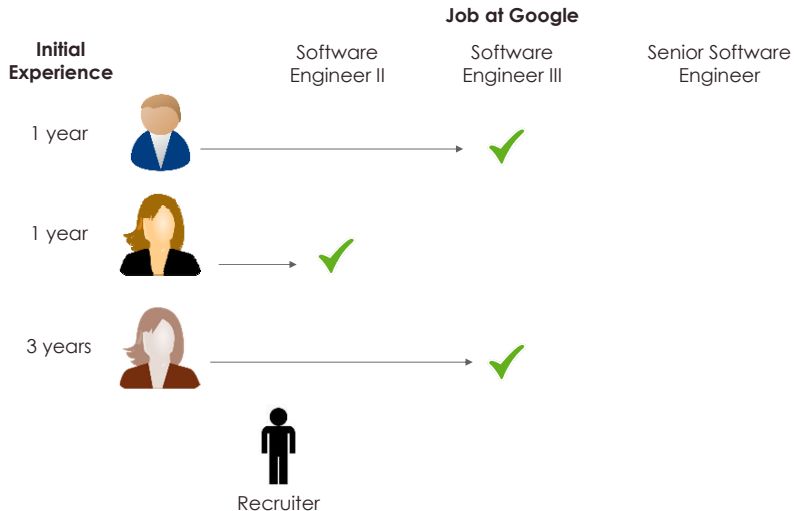
Introduction

- Disparities in treatment and outcomes (by, e.g., race/gender) are widely documented, but not always easily interpreted
 - In labor markets, housing, criminal justice, education, healthcare...
- In economics, disparity analyses tend to focus on *direct discrimination*: differential treatment on the basis of a protected characteristic
 - Theoretical models of how race/gender affect treatment through a decision-maker's preferences/beliefs (e.g. Becker 1957; Phelps 1972)
 - Empirical studies of the causal effects of perceived race/gender, holding other characteristics fixed (e.g. Bertrand & Mullainathan 2004)
- A large body of work in other fields suggests this view is incomplete
 - Sociological analyses often take a systems-based approach, with discrimination arising *indirectly* through non-group characteristics
 - This *systemic* perspective is echoed in some legal/economic analyses, and most recently in computer science (i.e. algorithmic discrimination)

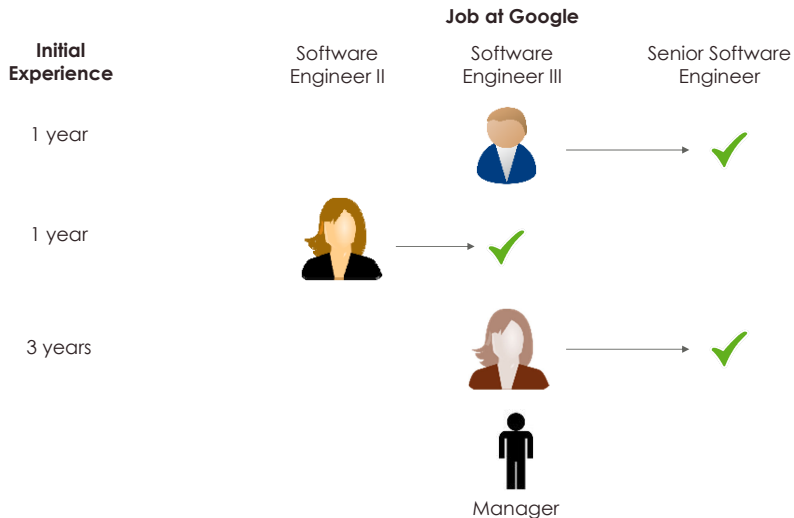
A Simple Labor Market Example



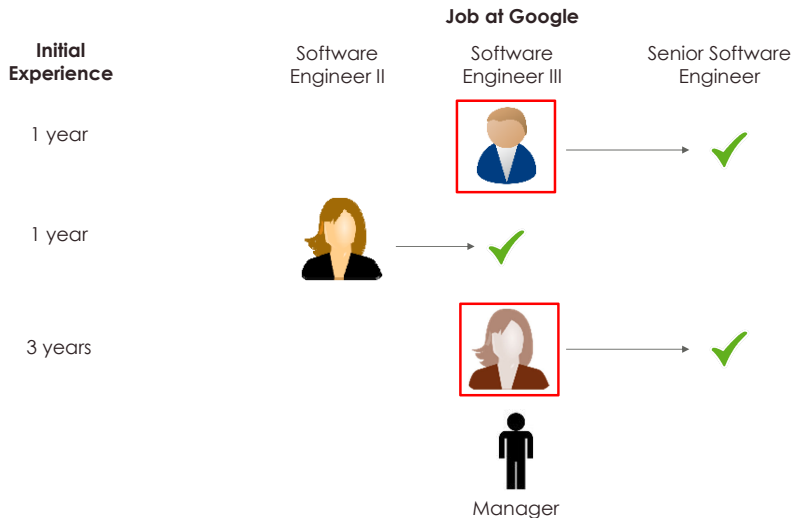
Direct Discrimination by Recruiters



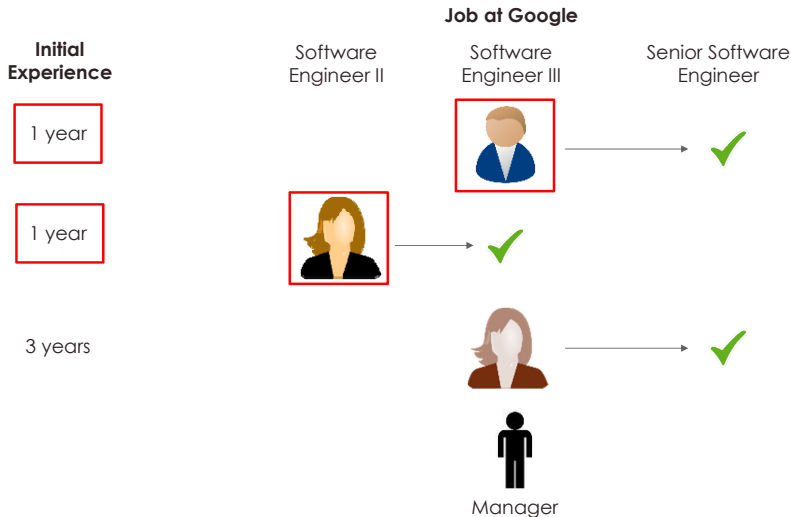
Manager Promotion Decisions



No Direct Discrimination by Managers



Yet Equally Qualified Workers Face Unequal Promotions



Griggs v. Duke Power Co. (1970)

- A landmark Supreme Court case on the interpretation of the Civil Rights Act Title VII, which set the disparate impact standard
 - A policy requiring a high school diploma for job transfers was found to be racially discriminatory, since high school diplomas were found irrelevant to an individual's qualification for different jobs
- Importantly, racial discrimination was found at Duke Power despite their transfer policy being facially “race-blind”
 - White/Black employees with the same educational background had the same ability to transfer
 - But equally-qualified workers had different high school diploma rates, in part due to discriminatory policies in secondary education

Griggs v. Duke Power Co. (1970)

- A landmark Supreme Court case on the interpretation of the Civil Rights Act Title VII, which set the disparate impact standard
 - A policy requiring a high school diploma for job transfers was found to be racially discriminatory, since high school diplomas were found irrelevant to an individual's qualification for different jobs
- Importantly, racial discrimination was found at Duke Power despite their transfer policy being facially “race-blind”
 - White/Black employees with the same educational background had the same ability to transfer
 - But equally-qualified workers had different high school diploma rates, in part due to discriminatory policies in secondary education
- Such discrimination would not have been found in a correspondence study which randomizes distinctively white/Black names
 - Also not naturally modeled as taste-based/statistical discrimination

This Paper

- We propose new tools to model and measure such discrimination
- We first develop a general theoretical framework to distinguish *direct* and *systemic* discrimination
 - Shifts focus from causal effects in a given decision to disparities conditional on a reference qualification measure (*total* discrimination)
 - Discrimination can be individual or *institutional* (e.g. firm-level)

This Paper

- We propose new tools to model and measure such discrimination
- We first develop a general theoretical framework to distinguish *direct* and *systemic* discrimination
 - Shifts focus from causal effects in a given decision to disparities conditional on a reference qualification measure (*total* discrimination)
 - Discrimination can be individual or *institutional* (e.g. firm-level)
- We discuss and illustrate the drivers of both forms of discrimination
 - Systemic discrimination can arise from biases in interactions over time or across different domains in the same period
 - *Informational* systemic discrimination arises from disparities in a signaling process: e.g. signal inflation or screening
 - *Technological* systemic discrimination arises from endogenous disparities in productivity, e.g. opportunities for skill development

This Paper (Cont.)

- We then develop and apply a new measure of systemic discrimination
 - Based on a decomposition of total discrimination into direct and systemic components (Kitagawa-Oaxaca-Blinder style)

This Paper (Cont.)

- We then develop and apply a new measure of systemic discrimination
 - Based on a decomposition of total discrimination into direct and systemic components (Kitagawa-Oaxaca-Blinder style)
- This empirical framework can guide data collection and identification strategies in observational and (quasi-)experimental analyses
 - Direct discrimination is identified by correspondence studies
 - Total discrimination is identified by the qualification distribution, maybe leveraging IV methods (e.g. Arnold et al. 2020)
 - Systemic discrimination is given by the decomposition residual

This Paper (Cont.)

- We then develop and apply a new measure of systemic discrimination
 - Based on a decomposition of total discrimination into direct and systemic components (Kitagawa-Oaxaca-Blinder style)
- This empirical framework can guide data collection and identification strategies in observational and (quasi-)experimental analyses
 - Direct discrimination is identified by correspondence studies
 - Total discrimination is identified by the qualification distribution, maybe leveraging IV methods (e.g. Arnold et al. 2020)
 - Systemic discrimination is given by the decomposition residual
- We demonstrate these tools in two hiring experiments
 - Recruiters propose lower wages / hiring rates for female Workers than male Workers with identical productivity signals (direct discrimination)
 - Signal inflation / screening lead Hiring Managers to indirectly discriminate against female Workers
 - Standard measures fail to capture this systemic discrimination

Some Literature Connections

- **Large literatures on systemic/structural discrimination** (Tiny sample: Feagin and Feagin 1978; Hill 1988; Pincus 1996; Gynter 2003; Powell 2007; De Plevitz 2007; Pager and Shepherd 2008; Feagin 2013; Allard and Small 2013; Rothstein 2017)
- **More recent literatures on algorithmic discrimination/unfairness** (Angwin, Larson, Mattu, and Kirchner. 2016; Hard, Prize, and Srebro 2016; Zafar, Valera, Gomez Rodriguez, and Gummadi 2017; Berk, Heidari, Jabbari, Kearns, and Roth 2018)
- **Recent theoretical departures from classical taste-based/statistical discrimination** (Darity 2005; Bordalo, Coffman, Gennaioli, and Schleifer 2019; Bohren, Imas, and Rosenberg 2019; Bohren, Haggag, Imas, and Pope 2020; Barron, Dittmann, Gehrig, and Schweighofer-Kodritsch 2020; Hübert and Little 2020); Rose (2022)
- **Recent empirical advances in measuring disparate impact/racial bias** (Arnold, Dobbie, and Yang 2018; Arnold, Dobbie, and Hull 2020, 2021; Grau and Vergara 2021; Canay, Mogstad, and Mountjoy 2020; Hull 2021; Gelbach 2021)

Outline

1. Theory
2. Measurement
3. Demonstration

Setup

- Consider a set of hiring managers j evaluating a set of workers i
 - Workers have a group $G_i \in \{m, f\}$, productivity Y_i^* , and a signal S_i
 - Managers take action A_{ij} after observing (G_i, S_i)
 - Reduced-form action rule $A_j(g, s)$, derived from beliefs/preferences
 - Firm action rule: $\alpha(g, s) = \sum_j \pi_j A_j(g, s)$ for manager shares π_j

Setup

- Consider a set of hiring managers j evaluating a set of workers i
 - Workers have a group $G_i \in \{m, f\}$, productivity Y_i^* , and a signal S_i
 - Managers take action A_{ij} after observing (G_i, S_i)
 - Reduced-form action rule $A_j(g, s)$, derived from beliefs/preferences
 - Firm action rule: $\alpha(g, s) = \sum_j \pi_j A_j(g, s)$ for manager shares π_j
- The hiring task is embedded in a larger economy, of other managers and firms (potentially in other sectors / periods / etc)
 - Workers enter the economy with qualification Y_i^0
 - Allow $Y_i^0 = Y_i^*$ or for Y_i^* to be generated endogenously from Y_i^0 and the actions of other managers/firms

[Equivalent Setups](#)

Three Types of (Individual) Discrimination

- **Def. 1:** manager j 's actions exhibit *direct discrimination* if $A_j(m, s) \neq A_j(f, s)$ for some s
 - Discrimination arising from group membership itself (i.e. causally), holding relevant non-group characteristics fixed
- **Def. 2:** manager j 's actions exhibit *systemic discrimination* if $A_j(g, S_i) \not\perp G_i \mid Y_i^0$ for some g
 - Discrimination arising indirectly from non-group characteristics, among equally qualified workers

Three Types of (Individual) Discrimination

- **Def. 1:** manager j 's actions exhibit *direct discrimination* if $A_j(m, s) \neq A_j(f, s)$ for some s
 - Discrimination arising from group membership itself (i.e. causally), holding relevant non-group characteristics fixed
- **Def. 2:** manager j 's actions exhibit *systemic discrimination* if $A_j(g, S_i) \not\perp G_i \mid Y_i^0$ for some g
 - Discrimination arising indirectly from non-group characteristics, among equally qualified workers
- **Def. 3:** manager j 's actions exhibit (*total*) *discrimination* if $A_j(G_i, S_i) \not\perp G_i \mid Y_i^0$
 - Discrimination arising from either direct or systemic channels, holding worker qualification fixed

[In DAG Form](#) | [Institutional Analogs](#)

The Choice of Y_i^0

- Varying the qualification measure Y_i^0 brings focus to different systemic forces. Two extremes:
 - $Y_i^0 = S_i$ makes direct and total discrimination coincide; no role for systemic discrimination (implicit in most economic analyses)
 - $Y_i^0 = 0$ interprets any unconditional disparities as discrimination: no “inherent” qualification differences across groups

The Choice of Y_i^0

- Varying the qualification measure Y_i^0 brings focus to different systemic forces. Two extremes:
 - $Y_i^0 = S_i$ makes direct and total discrimination coincide; no role for systemic discrimination (implicit in most economic analyses)
 - $Y_i^0 = 0$ interprets any unconditional disparities as discrimination: no “inherent” qualification differences across groups
- Setting $Y_i^0 = Y_i^*$ defines discrimination as disparities among equally productive workers (i.e. “disparate impact,” Arnold et al. (2020))
 - Also seen in CS analyses of algorithmic unfairness (“equalized odds”)
- Other choices of Y_i^0 allow for discrimination in the objective itself
 - E.g. setting Y_i^0 to i ’s initial labor market qualifications allows for systemic discrimination in hiring/promotion practices

Sources of Discrimination

- Canonical sources of direct discrimination focus on manager preferences and beliefs
 - Accurate statistical discrimination (e.g. Aigner & Cain 1977)
 - Taste-based discrimination (e.g. Becker 1957)
 - Inaccurate beliefs (e.g. Bohren et al. 2020)
- We formalize two analogous sources of systemic discrimination:
 - *Informational*: differences in the signaling process (i.e. $S_i \not\perp G_i \mid Y_i^*, Y_i^0$)
 - *Technological*: differences in skill accumulation (i.e. $Y_i^* \not\perp G_i \mid Y_i^0$)

Theoretical Applications

- Signal Inflation with Unaware / Aware Evaluators

- Shows how direct discrimination “reversals” over time (e.g. Bohren et al. 2019) may not imply total discrimination is mitigated or reversed

- Screening Workers

- Shows how disparities in signal quality lead to both direct and systemic discrimination, heterogeneously by worker productivity

- Signaling Across Markets

- Shows how discrimination in one market can lead to discrimination in another, through endogenous investments (e.g. Bursztyn et al. 2017)

- Accurate Statistical Discrimination with Social Learning

- Shows how accurate statistical discrimination can lead to persistent systemic discrimination when information is social

- Managerial Composition and Institutional Discrimination

- Shows how the share of managers from different groups can impact institutional direct/systemic discrimination arising from in-group biases

Outline

1. Theory✓
2. Measurement
3. Demonstration

Correspondence Studies

- Direct discrimination is a causal concept, so it can be measured by an experimental manipulation of (perceived) group
 - Both at the individual and institutional level (e.g. Kline et al. 2022)
 - Note: abstracting away some conceptual issues with randomizing group perceptions (Fryer and Levitt 2004; Kohler-Hausmann 2019; Rose 2022)

Correspondence Studies

- Direct discrimination is a causal concept, so it can be measured by an experimental manipulation of (perceived) group
 - Both at the individual and institutional level (e.g. Kline et al. 2022)
 - Note: abstracting away some conceptual issues with randomizing group perceptions (Fryer and Levitt 2004; Kohler-Hausmann 2019; Rose 2022)
- Systemic discrimination is *not* a causal concept: randomizing group breaks any dependence with S_i
 - We propose an alternative approach which backs it out of a decomposition of total discrimination into direct and systemic factors
 - Correspondence studies identify the direct component...

A Kitagawa-Oaxaca-Blinder Decomposition

$$\underbrace{\Delta(y)}_{\text{Total discrimination}} = \underbrace{E[\tau(S_i) \mid G_i = m, Y_i^0 = y]}_{\text{Average direct discrimination}} + \underbrace{\delta(f, y)}_{\text{Systemic discrimination}}$$

where

- $\Delta(y) = E[\alpha(G_i, S_i) \mid G_i = m, Y_i^0 = y] - E[\alpha(G_i, S_i) \mid G_i = f, Y_i^0 = y]$
- $\tau(s) = \alpha(m, s) - \alpha(f, s)$, identified by correspondence studies
- $\delta(g, y) = E[\alpha(g, S_i) \mid G_i = m, Y_i^0 = y] - E[\alpha(g, S_i) \mid G_i = f, Y_i^0 = y]$

[Alternative Decompositions](#)

Last Piece: Measuring Total Discrimination

- When qualification Y_i^0 is observed, $\Delta(y)$ is directly measurable
 - E.g. $Y_i^0 = 0$ or some upstream characteristic (like education)
 - Easy addition to correspondence studies, potentially for a range of Y_i^0

Last Piece: Measuring Total Discrimination

- When qualification Y_i^0 is observed, $\Delta(y)$ is directly measurable
 - E.g. $Y_i^0 = 0$ or some upstream characteristic (like education)
 - Easy addition to correspondence studies, potentially for a range of Y_i^0
- In some cases Y_i^0 is selectively observed, given manager actions
 - E.g. $Y_i^0 = Y_i^*$ and we only observe realized output $Y_{ij} = A_{ij}Y_i^*$ given worker productivity and manager hiring decision A_{ij}
 - Similarly, pretrial misconduct potential Y_i^* may only be observed among released ($A_{ij} = 1$) defendants (Lakkaraju et al. 2017)

Last Piece: Measuring Total Discrimination

- When qualification Y_i^0 is observed, $\Delta(y)$ is directly measurable
 - E.g. $Y_i^0 = 0$ or some upstream characteristic (like education)
 - Easy addition to correspondence studies, potentially for a range of Y_i^0
- In some cases Y_i^0 is selectively observed, given manager actions
 - E.g. $Y_i^0 = Y_i^*$ and we only observe realized output $Y_{ij} = A_{ij}Y_i^*$ given worker productivity and manager hiring decision A_{ij}
 - Similarly, pretrial misconduct potential Y_i^* may only be observed among released ($A_{ij} = 1$) defendants (Lakkaraju et al. 2017)
- Arnold et al. (2020, 2021) show how quasi-experimental assignment of managers/firms can solve this selection challenge
 - Core idea: use assignment as an IV to “selection correct” key moments
 - Applies to $Y_i^0 = Y_i^*$ as well as when Y_i^0 is “upstream” qualification

Last Piece: Measuring Total Discrimination

- When qualification Y_i^0 is observed, $\Delta(y)$ is directly measurable
 - E.g. $Y_i^0 = 0$ or some upstream characteristic (like education)
 - Easy addition to correspondence studies, potentially for a range of Y_i^0
- In some cases Y_i^0 is selectively observed, given manager actions
 - E.g. $Y_i^0 = Y_i^*$ and we only observe realized output $Y_{ij} = A_{ij}Y_i^*$ given worker productivity and manager hiring decision A_{ij}
 - Similarly, pretrial misconduct potential Y_i^* may only be observed among released ($A_{ij} = 1$) defendants (Lakkaraju et al. 2017)
- Arnold et al. (2020, 2021) show how quasi-experimental assignment of managers/firms can solve this selection challenge
 - Core idea: use assignment as an IV to “selection correct” key moments
 - Applies to $Y_i^0 = Y_i^*$ as well as when Y_i^0 is “upstream” qualification
- Framework may guide data collection/(quasi-)experimental design

Outline

1. Theory✓
2. Measurement✓
3. Demonstration

Experiment 1: Signal Inflation

- We first illustrate these tools in a setting like the motivating example
 - Three types of participants recruited from *Prolific*: Workers ($N = 100$), Recruiters ($N = 200$), and Hiring Managers ($N = 500$)
- Workers complete a test of basic math/business/history knowledge
 - Two parts (A and B), each with 10 randomly-selected questions
 - No difference in test performance or part correlation by Worker gender

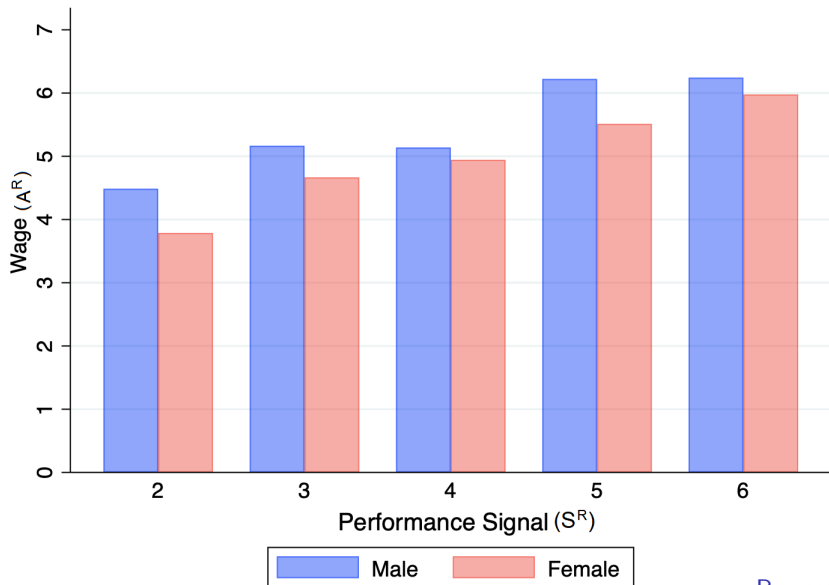
Experiment 1: Signal Inflation

- We first illustrate these tools in a setting like the motivating example
 - Three types of participants recruited from *Prolific*: Workers ($N = 100$), Recruiters ($N = 200$), and Hiring Managers ($N = 500$)
- Workers complete a test of basic math/business/history knowledge
 - Two parts (A and B), each with 10 randomly-selected questions
 - No difference in test performance or part correlation by Worker gender
- Recruiters observe part-A performance S_i^R and submit a wage A_{ij}^R
 - Then paid on the basis of part-B performance Y_i^{R*} and the submitted wage, via the Becker-DeGroot-Marschak mechanism
 - Gender disparities reflect direct discrimination (specifically, bias)

Experiment 1: Signal Inflation

- We first illustrate these tools in a setting like the motivating example
 - Three types of participants recruited from *Prolific*: Workers ($N = 100$), Recruiters ($N = 200$), and Hiring Managers ($N = 500$)
- Workers complete a test of basic math/business/history knowledge
 - Two parts (A and B), each with 10 randomly-selected questions
 - No difference in test performance or part correlation by Worker gender
- Recruiters observe part-A performance S_i^R and submit a wage A_{ij}^R
 - Then paid on the basis of part-B performance Y_i^{R*} and the submitted wage, via the Becker-DeGroot-Marschak mechanism
 - Gender disparities reflect direct discrimination (specifically, bias)
- Managers observe recruiter wage offers S_i^H and submit a wage A_{ij}^H
 - Then paid on the basis of part-A performance Y_i^{H*} and the submitted wage, via the Becker-DeGroot-Marschak mechanism
 - Gender disparities from both direct and systemic discrimination

Direct Discrimination by Recruiters



Total Discrimination by Managers

	(1)
Gender (1=Male, 0=Female)	0.92*** (0.19)
Signal S_i^H	
Constant	5.18*** (0.14)

Notes: This table reports coefficients from regressing Hiring Manager wage offers on Worker gender and Recruiter wage offers. The sample includes 506 Hiring Managers, each evaluating one Worker ($N = 506$). Robust standard errors are reported in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Robustness

Total Discrimination by Managers

	(1)	(2)
Gender (1=Male, 0=Female)	0.92*** (0.19)	-0.09 (0.13)
Signal S_i^H		0.72*** (0.03)
Constant	5.18*** (0.14)	1.78*** (0.15)

Notes: This table reports coefficients from regressing Hiring Manager wage offers on Worker gender and Recruiter wage offers. The sample includes 506 Hiring Managers, each evaluating one Worker ($N = 506$). Robust standard errors are reported in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Robustness

Systemic Discrimination Drives Total Discrimination

	(1)	(2)	(3)	(4)	(5)
	Worker Performance Level Y_i^{H0}				
	2	3	4	5	6
Total Discrimination	1.00	1.39	0.47	2.01	0.33
Average Direct Discrimination	-0.10	0.30	0.11	0.51	0.10
Systemic Discrimination	1.10	1.09	0.36	1.50	0.23

Notes: This table reports estimates of total discrimination, average direct discrimination, and systemic discrimination in Hiring Manager wage offers across different levels of Worker productivity. The sample includes 506 Hiring Managers, each evaluating one Worker ($N = 506$).

[Alternative Decompositions](#)

Experiment 2: Screening

- Our second experiment shows how disparities in signal observations can yield *heterogeneous* systemic discrimination
 - Same setup as before, except now Recruiters decide whether or not to “hire” Workers after observing G_i and S_i^R
 - Hiring Managers see S_i^R *only if* the worker is hired (always see G_i)
- Theory predicts Hiring Manager actions will exhibit the most systemic discrimination against high-performing females
 - Recruiter discrimination hurts them the most, since female workers with high signals are hired at a higher rate than workers w/ no signal

Screening Yields Heterogeneous Systemic Discrimination

	(1)	(2)	(3)	(4)
	Task A Performance Y_i^{H0}			
	2	3	4	5
Total Discrimination	0.05	0.09	0.16	0.27
Average Direct Discrimination	0.06	0.09	0.14	0.20
Systemic Discrimination	-0.01	0.00	0.02	0.07

Notes: This table reports estimates of total discrimination, average direct discrimination, and systemic discrimination in Hiring Manager hiring rates across different levels of Worker performance on Task A. The sample includes 501 Hiring Managers, each evaluating one Worker ($N = 501$).

[Alternative Decompositions](#)

Conclusion

- We develop new tools to model and measure systemic discrimination
 - Moving beyond direct discrimination involves a choice of the reference qualification level Y_i^0
 - Systemic discrimination can stem from both informational and technological channels
 - Measurement requires going beyond audit/correspondence studies
- We show how systemic discrimination can arise and persist, in both theory and practice
 - Signal inflation in one period or domain can drive persistent systemic discrimination in another
 - Screening problems can yield important heterogeneity in discrimination across qualification levels
- The framework suggests high returns to developing new models and measures of systemic discrimination
 - The U.S. EEOC launched nearly 600 investigations into systemic discrimination in 2020, highlighting *facially neutral* hiring practices

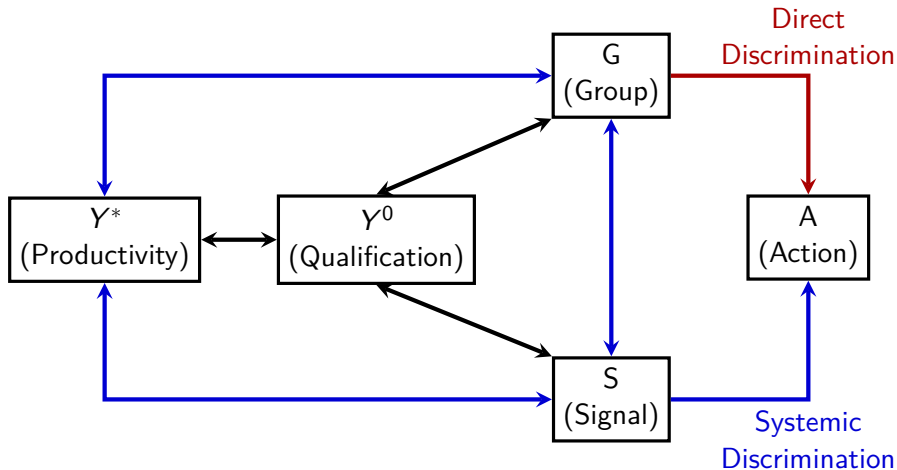
Thank you!

Appendix Slides

Other Setups

- **Lending:** Loan officers at a bank decide to lend to borrowers
 - Y_i^* : borrower i 's potential ability-to-repay
 - Y_i^0 : initial lending qualifications (may interact with employment)
 - S_i : credit scores, income, zip code...
- **Education:** Admissions officers at a school decide to admit students
 - Y_i^* : student i 's potential academic performance
 - Y_i^0 : initial education (may interact with familial obligations)
 - S_i : test scores, recommendation letters...
- **Healthcare:** Doctors at a hospital decide to test patients for a disease
 - Y_i^* : patient i 's latent disease state
 - Y_i^0 : underlying health (may interact with prior healthcare access)
 - S_i : blood pressure, BMI, previous tests...
- **Bail:** Judges in a district decide to release defendants before trial
 - Y_i^* : defendant i 's potential for pretrial misconduct
 - Y_i^0 : underlying "propensity for crime" (may interact with many things)
 - S_i : criminal record, face tattoos, demeanor in court...

In DAG Form



Three Types of Institutional Discrimination

- **Def. 1:** the firm's actions exhibit *institutional direct discrimination* if $\alpha(m, s) \neq \alpha(f, s)$ for some s
- **Def. 2:** the firm's actions exhibit *institutional systemic discrimination* if $\alpha(g, S_i) \not\perp G_i \mid Y_i^0$ for some g
- **Def. 3:** the firm's actions exhibit *institutional systemic discrimination* if $\alpha(G_i, S_i) \not\perp G_i \mid Y_i^0$

[Back](#)

Signal Inflation with Unaware / Aware Evaluators

- A recruiter observes G_i and a signal S_i^R with $E[Y_i^* | S_i^R = s] = s$.
Assume $(Y_i^*, S_i^R) \perp G_i$ and set $Y_i^0 = Y_i^*$
 - The recruiter submits a forecast A_i^R with direct discrimination (bias):
 $A^R(f, s) = s$ but $A^R(m, s) = s + 1$
 - A hiring manager observes (A_i^R, G_i) and makes a wage offer A_i^H

Signal Inflation with Unaware / Aware Evaluators

- A recruiter observes G_i and a signal S_i^R with $E[Y_i^* | S_i^R = s] = s$.
Assume $(Y_i^*, S_i^R) \perp G_i$ and set $Y_i^0 = Y_i^*$
 - The recruiter submits a forecast A_i^R with direct discrimination (bias):
 $A^R(f, s) = s$ but $A^R(m, s) = s + 1$
 - A hiring manager observes (A_i^R, G_i) and makes a wage offer A_i^H
- Suppose managers take signals at face value: $A^H(g, s) = s$
 - No direct discrimination: $A^H(m, s) - A^H(f, s) = 0$
 - Positive systemic discrimination: $E[A^H(g, S_i^H) | G_i = m, S_i^R = s] = s + 1$
vs. $E[A^H(g, S_i^H) | G_i = f, S_i^R = s] = s$
 - Positive total discrimination:
 $E[A^H(G_i, S_i^H) | G_i = m, S_i^R = s] - E[A^H(G_i, S_i^H) | G_i = f, S_i^R = s] = 1$

Signal Inflation with Unaware / Aware Evaluators

- A recruiter observes G_i and a signal S_i^R with $E[Y_i^* | S_i^R = s] = s$.
Assume $(Y_i^*, S_i^R) \perp G_i$ and set $Y_i^0 = Y_i^*$
 - The recruiter submits a forecast A_i^R with direct discrimination (bias):
 $A^R(f, s) = s$ but $A^R(m, s) = s + 1$
 - A hiring manager observes (A_i^R, G_i) and makes a wage offer A_i^H
- Suppose managers take signals at face value: $A^H(g, s) = s$
 - No direct discrimination: $A^H(m, s) - A^H(f, s) = 0$
 - Positive systemic discrimination: $E[A^H(g, S_i^H) | G_i = m, S_i^R = s] = s + 1$
vs. $E[A^H(g, S_i^H) | G_i = f, S_i^R = s] = s$
 - Positive total discrimination:
 $E[A^H(G_i, S_i^H) | G_i = m, S_i^R = s] - E[A^H(G_i, S_i^H) | G_i = f, S_i^R = s] = 1$
- Now suppose managers are aware: $A^H(f, s) = s$ and $A^H(m, s) = s - 1$
 - Same systemic discrimination as before, but now offset by negative manager direct discrimination $\rightarrow$ no total discrimination
 - Gives lens for interpreting direct discrimination “reversals” documented in recent work on discrimination dynamics (Bohren et al. 2019)

Screening Workers

- A hiring manager observes a signal $S_i = Y_i^* + \varepsilon_i$ of productivity $Y_i^* \mid G_i \sim N(0, 1)$, with $\varepsilon_i \mid Y_i^*, G_i \sim N(0, 1/\eta_{G_i})$
 - Hires using a cutoff rule: $A(g, s) = \mathbf{1}[E[Y_i^* \mid G_i = g, S_i = s] \geq t]$
 - $E[Y_i^* \mid G_i = g, S_i = s] = s \frac{\eta_g}{1 + \eta_g}$, so workers with $S_i \geq t \frac{1 + \eta_{G_i}}{\eta_{G_i}}$ are hired
 - Suppose $\eta_m > \eta_f$: noisier signals for group- f workers

Screening Workers

- A hiring manager observes a signal $S_i = Y_i^* + \varepsilon_i$ of productivity $Y_i^* \mid G_i \sim N(0, 1)$, with $\varepsilon_i \mid Y_i^*, G_i \sim N(0, 1/\eta_{G_i})$
 - Hires using a cutoff rule: $A(g, s) = \mathbf{1}[E[Y_i^* \mid G_i = g, S_i = s] \geq t]$
 - $E[Y_i^* \mid G_i = g, S_i = s] = s \frac{\eta_g}{1 + \eta_g}$, so workers with $S_i \geq t \frac{1 + \eta_{G_i}}{\eta_{G_i}}$ are hired
 - Suppose $\eta_m > \eta_f$: noisier signals for group- f workers
- Direct discrimination (accurate statistical) through signal thresholds
 - Workers with $S_i \in (t \frac{1 + \eta_m}{\eta_m}, t \frac{1 + \eta_f}{\eta_f})$ are hired iff $G_i = m$

Screening Workers

- A hiring manager observes a signal $S_i = Y_i^* + \varepsilon_i$ of productivity $Y_i^* \mid G_i \sim N(0, 1)$, with $\varepsilon_i \mid Y_i^*, G_i \sim N(0, 1/\eta_{G_i})$
 - Hires using a cutoff rule: $A(g, s) = \mathbf{1}[E[Y_i^* \mid G_i = g, S_i = s] \geq t]$
 - $E[Y_i^* \mid G_i = g, S_i = s] = s \frac{\eta_g}{1 + \eta_g}$, so workers with $S_i \geq t \frac{1 + \eta_{G_i}}{\eta_{G_i}}$ are hired
 - Suppose $\eta_m > \eta_f$: noisier signals for group- f workers
- Direct discrimination (accurate statistical) through signal thresholds
 - Workers with $S_i \in (t \frac{1 + \eta_m}{\eta_m}, t \frac{1 + \eta_f}{\eta_f})$ are hired iff $G_i = m$
- Systemic discrimination through the differential signal process itself:

$$\begin{aligned} & E[A(g, S_i) \mid Y_i^* = y, G_i = m] - E[A(g, S_i) \mid Y_i^* = y, G_i = f] \\ &= \Phi\left(\eta_f \left(t \frac{1 + \eta_g}{\eta_g} - y\right)\right) - \Phi\left(\eta_m \left(t \frac{1 + \eta_g}{\eta_g} - y\right)\right) \neq 0 \end{aligned}$$

- Systemic discrimination is heterogeneous: hurts high-performers more

Signaling Across Markets

- Workers participate in the job and marriage markets, which differentially value productivity Y_i^*
 - Workers choose S_i , paying $(S_i - Y_i^*)^2$ to deviate from endowed Y_i^*
 - Employers are unaware, and have a uniform action rule of $A_1(g, s) = s$
 - Partners are unaware & discriminate: $A_2(m, s) = -s$ but $A_2(f, s) = -s$
 - Workers max. utility $U_i = \gamma A_1(G_i, S_i) + (1 - \gamma) A_2(G_i, S_i) - (S_i - Y_i^*)^2$

Signaling Across Markets

- Workers participate in the job and marriage markets, which differentially value productivity Y_i^*
 - Workers choose S_i , paying $(S_i - Y_i^*)^2$ to deviate from endowed Y_i^*
 - Employers are unaware, and have a uniform action rule of $A_1(g, s) = s$
 - Partners are unaware & discriminate: $A_2(m, s) = -s$ but $A_2(f, s) = -s$
 - Workers max. utility $U_i = \gamma A_1(G_i, S_i) + (1 - \gamma) A_2(G_i, S_i) - (S_i - Y_i^*)^2$
- Differential signal inflation by G_i leads to systemic discrimination
 - Group- m workers always shade up: $S_i = Y_i^* + \frac{1}{2}$
 - Group- f signal depends on marriage market weight: $S_i = Y_i^* + \gamma - \frac{1}{2}$
 - So $E[A_1(g, S_i) \mid Y_i^0 = y, G_i = m] - E[A_1(g, S_i) \mid Y_i^0 = y, G_i = f] = 1 - \gamma$

Signaling Across Markets

- Workers participate in the job and marriage markets, which differentially value productivity Y_i^*
 - Workers choose S_i , paying $(S_i - Y_i^*)^2$ to deviate from endowed Y_i^*
 - Employers are unaware, and have a uniform action rule of $A_1(g, s) = s$
 - Partners are unaware & discriminate: $A_2(m, s) = -s$ but $A_2(f, s) = -s$
 - Workers max. utility $U_i = \gamma A_1(G_i, S_i) + (1 - \gamma) A_2(G_i, S_i) - (S_i - Y_i^*)^2$
- Differential signal inflation by G_i leads to systemic discrimination
 - Group- m workers always shade up: $S_i = Y_i^* + \frac{1}{2}$
 - Group- f signal depends on marriage market weight: $S_i = Y_i^* + \gamma - \frac{1}{2}$
 - So $E[A_1(g, S_i) \mid Y_i^0 = y, G_i = m] - E[A_1(g, S_i) \mid Y_i^0 = y, G_i = f] = 1 - \gamma$
- Highlights that systemic discrimination may not be dynamic
 - Related to “acting wife” findings in Bursztyn et al. (2017), and notion of “side-effect” discrimination in Feagin and Feagin (1978)

[Back](#)

Discrimination with Social Information

- A series of evaluators $t = 1, 2, \dots$ predict productivity $Y_i^* \sim N(\mu_g, 1)$
 - Evaluator t observes history of past forecasts $H_{it} = \{A_{i1}, \dots, A_{i,t-1}\}$ and new signal $\tilde{S}_{it} = Y_i^* + \varepsilon_{it}$ with $\varepsilon_{it} \mid H_{it}, G_i \sim N(0, 1)$
 - Forecasts are accurate: $A_{it} = E[Y_i^* \mid G_i, S_{it}]$ where $S_{it} = \{H_{it}, \tilde{S}_{it}\}$

Discrimination with Social Information

- A series of evaluators $t = 1, 2, \dots$ predict productivity $Y_i^* \sim N(\mu_g, 1)$
 - Evaluator t observes history of past forecasts $H_{it} = \{A_{i1}, \dots, A_{i,t-1}\}$ and new signal $\tilde{S}_{it} = Y_i^* + \varepsilon_{it}$ with $\varepsilon_{it} \mid H_{it}, G_i \sim N(0, 1)$
 - Forecasts are accurate: $A_{it} = E[Y_i^* \mid G_i, S_{it}]$ where $S_{it} = \{H_{it}, \tilde{S}_{it}\}$
- Evaluator 1 exhibits direct discrimination (accurate statistical)
 - $A_1(g, S_{i1}) = (\mu_g + \tilde{S}_{i1})/2$, so direct discrimination of $(\mu_m - \mu_f)/2$
 - No systemic discrimination: signal process is identical by group

Discrimination with Social Information

- A series of evaluators $t = 1, 2, \dots$ predict productivity $Y_i^* \sim N(\mu_g, 1)$
 - Evaluator t observes history of past forecasts $H_{it} = \{A_{i1}, \dots, A_{i,t-1}\}$ and new signal $\tilde{S}_{it} = Y_i^* + \varepsilon_{it}$ with $\varepsilon_{it} \mid H_{it}, G_i \sim N(0, 1)$
 - Forecasts are accurate: $A_{it} = E[Y_i^* \mid G_i, S_{it}]$ where $S_{it} = \{H_{it}, \tilde{S}_{it}\}$
- Evaluator 1 exhibits direct discrimination (accurate statistical)
 - $A_1(g, S_{i1}) = (\mu_g + \tilde{S}_{i1})/2$, so direct discrimination of $(\mu_m - \mu_f)/2$
 - No systemic discrimination: signal process is identical by group
- But every subsequent evaluator exhibits systemic discrimination
 - $E[A_t(g, S_{it}) \mid G_i = g', Y_i^*] = (\mu_{g'} + tY_i^*)/(t+1)$ for $g, g' \in \{m, f\}$ so systemic discrimination of $(\mu_m - \mu_f)/(t+1)$
 - No direct discrimination: the worker's forecast history is a sufficient statistic for average group productivity difference

Discrimination with Social Information

- A series of evaluators $t = 1, 2, \dots$ predict productivity $Y_i^* \sim N(\mu_g, 1)$
 - Evaluator t observes history of past forecasts $H_{it} = \{A_{i1}, \dots, A_{i,t-1}\}$ and new signal $\tilde{S}_{it} = Y_i^* + \varepsilon_{it}$ with $\varepsilon_{it} \mid H_{it}, G_i \sim N(0, 1)$
 - Forecasts are accurate: $A_{it} = E[Y_i^* \mid G_i, S_{it}]$ where $S_{it} = \{H_{it}, \tilde{S}_{it}\}$
- Evaluator 1 exhibits direct discrimination (accurate statistical)
 - $A_1(g, S_{i1}) = (\mu_g + \tilde{S}_{i1})/2$, so direct discrimination of $(\mu_m - \mu_f)/2$
 - No systemic discrimination: signal process is identical by group
- But every subsequent evaluator exhibits systemic discrimination
 - $E[A_t(g, S_{it}) \mid G_i = g', Y_i^*] = (\mu_{g'} + tY_i^*)/(t+1)$ for $g, g' \in \{m, f\}$ so systemic discrimination of $(\mu_m - \mu_f)/(t+1)$
 - No direct discrimination: the worker's forecast history is a sufficient statistic for average group productivity difference
- Social learning is the key driver: if signals were directly observed (i.e. $H_{it} = \{\tilde{S}_{i1}, \dots, \tilde{S}_{i,t-1}\}$) there would only be accurate statistical disc.

Institutional Discrimination from In-Group Bias

- In the Example 1 setup, suppose recruiters and managers also have groups $g \in \{m, f\}$, with shares $\pi_m^R, \pi_m^H \in [0, 1]$
 - In-group bias among both recruiters and managers:
 $A^{R,m}(f, s) = A^{R,f}(m, s) = s$ but $A^{R,m}(m, s) = A^{R,f}(f, s) = s + 1$;
 $A^{H,m}(f, s) = A^{H,f}(m, s) = s$ but $A^{H,m}(m, s) = A^{H,f}(f, s) = s + 1$

Institutional Discrimination from In-Group Bias

- In the Example 1 setup, suppose recruiters and managers also have groups $g \in \{m, f\}$, with shares $\pi_m^R, \pi_m^H \in [0, 1]$
 - In-group bias among both recruiters and managers:
 $A^{R,m}(f, s) = A^{R,f}(m, s) = s$ but $A^{R,m}(m, s) = A^{R,f}(f, s) = s + 1$;
 $A^{H,m}(f, s) = A^{H,f}(m, s) = s$ but $A^{H,m}(m, s) = A^{H,f}(f, s) = s + 1$
- Recruiters and hiring managers both exhibit individual direct discrimination against the out-group
 - There is institutional direct discrimination against group f when group m is the dominant type (among recruiters or managers)

Institutional Discrimination from In-Group Bias

- In the Example 1 setup, suppose recruiters and managers also have groups $g \in \{m, f\}$, with shares $\pi_m^R, \pi_m^H \in [0, 1]$
 - In-group bias among both recruiters and managers:
 $A^{R,m}(f, s) = A^{R,f}(m, s) = s$ but $A^{R,m}(m, s) = A^{R,f}(f, s) = s + 1$;
 $A^{H,m}(f, s) = A^{H,f}(m, s) = s$ but $A^{H,m}(m, s) = A^{H,f}(f, s) = s + 1$
- Recruiters and hiring managers both exhibit individual direct discrimination against the out-group
 - There is institutional direct discrimination against group f when group m is the dominant type (among recruiters or managers)
- Hiring managers do not exhibit systemic discrimination iff $\pi_m^R = 1/2$
 - Given π_m^R , institutional systemic/total discrimination depends on π_m^H
 - E.g. if $\pi_m^R = 1$ then there is no total discrimination iff $\pi_m^H = 0$

Institutional Discrimination from In-Group Bias

- In the Example 1 setup, suppose recruiters and managers also have groups $g \in \{m, f\}$, with shares $\pi_m^R, \pi_m^H \in [0, 1]$
 - In-group bias among both recruiters and managers:
 $A^{R,m}(f, s) = A^{R,f}(m, s) = s$ but $A^{R,m}(m, s) = A^{R,f}(f, s) = s + 1$;
 $A^{H,m}(f, s) = A^{H,f}(m, s) = s$ but $A^{H,m}(m, s) = A^{H,f}(f, s) = s + 1$
- Recruiters and hiring managers both exhibit individual direct discrimination against the out-group
 - There is institutional direct discrimination against group f when group m is the dominant type (among recruiters or managers)
- Hiring managers do not exhibit systemic discrimination iff $\pi_m^R = 1/2$
 - Given π_m^R , institutional systemic/total discrimination depends on π_m^H
 - E.g. if $\pi_m^R = 1$ then there is no total discrimination iff $\pi_m^H = 0$
- Highlights how manager composition can differentially drive individual/institutional & direct/systemic discrimination

Other Decompositions

$$\underbrace{\Delta(y)}_{\text{Total discrimination}} = \underbrace{E[\tau(S_i) \mid G_i = m, Y_i^0 = y]}_{\text{Average direct discrimination}} + \underbrace{\delta(f, y)}_{\text{Systemic discrimination}}$$

- As usual, the decomposition “order” can matter. Equivalently:

$$\Delta(y) = E[\tau(S_i) \mid G_i = f, Y_i^0 = y] + \delta(m, y)$$

and $\Delta(y) = \bar{\tau}(y) + \bar{\delta}(y)$, where $\bar{\tau}(y)$ and $\bar{\delta}(y)$ are simple averages

[Back](#)

Experiment 1: Direct Discrimination by Recruiters

TABLE 1. Signal Inflation: Direct Discrimination in Recruiter Wage Offers

	(1)	(2)	(3)
Gender (1=Male, 0=Female)	0.47*** (0.12)	0.47*** (0.12)	0.49*** (0.12)
Signal S_i^R		0.49*** (0.09)	0.52*** (0.09)
Constant	4.99*** (0.14)	3.04*** (0.36)	5.71*** (0.60)
Recruiter Demographic Controls	N	N	Y

Notes: This table reports coefficients from regressing Recruiter wage offers on Worker gender and the Worker's Task A performance. Columns 3 controls for Recruiter characteristics: age, gender, employment status, an indicator for the Recruiter being white, and an indicator for being college-educated. The sample includes 201 Recruiters, each evaluating two Workers ($N = 402$). Standard errors, clustered at the Worker level, are reported in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$;

Experiment 1: Total Discrimination by Managers

TABLE 2. Signal Inflation: Total Discrimination in Hiring Manager Wage Offers

	(1)	(2)	(3)	(4)
Gender (1=Male, 0=Female)	0.92*** (0.19)	-0.09 (0.13)	0.95*** (0.20)	-0.09 (0.13)
Signal S_i^H		0.72*** (0.03)		0.72*** (0.03)
Constant	5.18*** (0.14)	1.78*** (0.15)	5.36*** (0.42)	1.76*** (0.30)
Hiring Manager Demographic Controls	N	N	Y	Y

Notes: This table reports coefficients from regressing Hiring Manager wage offers on self-reported Worker gender and the Worker's Recruiter wage offer. Columns 3 and 4 control for Manager characteristics: age, gender, employment status, an indicator for the Manager being white, and an indicator for being college-educated. The sample includes 506 Hiring Managers, each evaluating one Worker ($N = 506$). Robust standard errors are reported in parentheses. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$;

Experiment 1: Manager Discrimination Decompositions

TABLE 3. Signal Inflation: Total, Direct, and Systemic Discrimination in Manager Wages

	(1)	(2)	(3)	(4)	(5)
	Worker Performance Level Y_i^{H0}				
	2	3	4	5	6
Total Discrimination	1.00	1.39	0.47	2.01	0.33
<i>Equation (4)</i>					
Average Direct Discrimination	-0.10	0.30	0.11	0.51	0.10
Systemic Discrimination	1.10	1.09	0.36	1.50	0.23
<i>Equation (5)</i>					
Average Direct Discrimination	-0.25	-0.17	0.29	-0.08	0.13
Systemic Discrimination	1.25	1.56	0.17	2.08	0.20
<i>Equation (6)</i>					
Average Direct Discrimination	-0.18	0.07	0.20	0.22	0.12
Systemic Discrimination	1.18	1.33	0.27	1.79	0.22

Notes: This table reports estimates of total discrimination, average direct discrimination, and systemic discrimination in Hiring Manager wage offers across different levels of Worker productivity. Total discrimination is measured by the average difference in wage offers among male vs. female Workers with a given Task A score. Average direct and systemic discrimination are measured by equations [Equations \(4\) to \(6\)](#), as described in the text. The sample includes 506 Hiring Managers, each evaluating one Worker ($N = 506$).

Experiment 2: Manager Discrimination Decompositions

TABLE 4. Screening: Total, Direct, and Systemic Discrimination in Hiring Manager Actions

	(1)	(2)	(3)	(4)
	Task A Performance Y_i^{H0}			
	2	3	4	5
Total Discrimination	0.05	0.09	0.16	0.27
<i>Equation (4)</i>				
Average Direct Discrimination	0.06	0.09	0.14	0.20
Systemic Discrimination	-0.01	0.00	0.02	0.07
<i>Equation (5)</i>				
Average Direct Discrimination	0.07	0.10	0.13	0.16
Systemic Discrimination	-0.02	0.00	0.04	0.11
<i>Equation (6)</i>				
Average Direct Discrimination	0.06	0.09	0.13	0.18
Systemic Discrimination	-0.01	0.00	0.03	0.09

Notes: This table reports estimates of total discrimination, average direct discrimination, and systemic discrimination in Hiring Manager hiring rates across different levels of Worker performance on Task A. Total discrimination is measured by the average difference in hiring rates among male vs. female Workers with a given Task A score. Average direct and systemic discrimination are measured by equations [Equations \(4\) to \(6\)](#), as described in the text. The sample includes 501 Hiring Managers, each evaluating one Worker ($N = 501$).