Hiring Discrimination Against Transgender Job Applicants in the US Labor Market

Pre-analysis plan

Emily Beam and Ivy Stanton
April 17, 2024

1 Introduction

Transgender individuals throughout the United States face discrimination in a range of sectors, harming their well-being and exacerbating inequalities with the general population (Grant et al., 2011; Fumarco et al., 2023; Abbate et al., 2024). Additionally, they report lower incomes and higher rates of poverty (Shannon, 2022). Access to employment opportunities is a critical determinant of lifetime earnings and other outcomes, and surveys of transgender individuals have documented frequent employment discrimination (Kattari et al., 2016) However, quantitative work that details the nature and extent of this discrimination remains underexplored (McFadden, 2020). Recent work finds evidence of substantial hiring discrimination against transgender applicants based on correspondence studies and list experiments (Granberg et al., 2020; Aksoy et al., 2022; Eames, 2024). Notably absent are analyses of intersectionality with gender and race, as well as how employment discrimination varies geographically—especially in the face of shifting state and local anti-trans policies (Bardales, 2013; Rainey et al., 2015).

This study aims to understand the intersectionality of race and gender identity to fully understand the range of potential impacts that transgender people might face in the labor market. Two, to explore heterogeneity in discrimination across the United States and explore the relationships with local political climates toward transgender individuals.

Specifically, we set out to answer the following research questions:

RQ1 How does employment discrimination against transgender people vary across the United States, and to what extent does it vary with local legislation and political climates?

RQ2 How does gender identity impact callback rates?

RQ3 How does race impact callback rates for transgender people?

To answer these questions, we conduct a correspondence study in which we submit 5,000 applications to 2,500 postings in early 2024, randomly varying race and gender identity in labor markets across the United States. We also include data from an earlier phase, which included 1,000 applications to 250 postings in four major US markets in early 2023 (Stanton, 2023)

Specifically, we vary gender identity by using gender-specific names for cisgendered men and women and then indicate transgender identity by using those same names alongside a phrase that

indicates a different-gender legal name. Conditional on gender, we use traditionally white names and traditionally black names to signal race (Bertrand and Mullainathan, 2004), which enables us to examine not only the impact of being transgender overall, but also the impact of being a transgender man versus a transgender woman relative to cisgender men and cisgender women, respectively. It also enables us to measure whether there are differential penalties for black versus white transgender applicants.

We will analyze our results in aggregate and separately by local political climate towards transgender people. We also propose a secondary analysis to examine the role discrimination against non-binary individuals versus transgender men and transgender women, by use of they/them pronouns to indicate non-binary status. We use cis-gender individuals with and without indicated pronouns as a benchmark.

This study sheds light on the extent of discrimination against transgender people, overall and separately by specific gender identities and by race, across the majority of the United States in a range of entry-level markets. Additionally, it speaks to how labor market experiences vary by political climate.

Resume submission began prior to registering this PAP. However, we have submitted it before analyzing any data on callback rates.

2 Experimental Design

2.1 Market selection

We include 49 markets across 27 U.S. states, listed in Table A.2. We identified markets and states via the following procedure:

1. To identify sufficiently active markets, we counted the number of job postings with unique employers over a two day-period (August 21 and 22, 2023) within food-service positions, one of our two main industries, for each of 417 Craigslist markets in the United States. See Table A.1 for the distribution of postings by market.

2. We excluded any market with fewer than 5 postings during that period, leaving 110 markets across 40 states. We automatically included the four Phase 1 states: Arizona, California, New York, and Texas.

3. We merged the state-level Republican vote share from the 2020 presidential election (MIT Election Data and Science Lab, 2017) for the remaining 36 states. We randomly selected two-thirds of states (24), stratifying by Republican vote-share quartile, for a total of 28 states with 90 markets.

4. We select the 4 Phase 1 markets and then up to two markets per state, with the state-specific probability of selection proportional to the number of job postings in that market. This yielded a total of 48 markets in 27 states.

5. We revised the list of markets to (1) ensure we did not include overlapping markets and (2) exclude markets that we later found had insufficient postings to be viable. We then revised selection to include up to three markets per state to compensate for the additional loss of markets.

This led to a final sample of 49 markets in 27 states. Revision details are described below.

Table 1: Distribution of 2020 Republican vote shares

	Mean	p50	Min	Max
0	0.323	0.34	0.054	0.404
1	0.437	0.436	0.406	0.477
2	0.505	0.499	0.478	0.533
3	0.595	0.585	0.551	0.654
Total	0.462	0.478	0.054	0.654
Phase 1	0.432	0.433	0.343	0.521

Source: MIT Election Data and Science Lab (2017).

2.1.1 Market selection revisions

We made the following revisions to our initial market selection

- Dropped Miami and Long Island because Miami is equivalent to South Florida and Long Island is nested inside NYC.
- Upon second review in late November 2023, we identified 9 selected markets with 0–5 usable postings (excluding postings for door dash drivers, for example) in the food service sector for the entire month of November. We removed these markets and replaced them with the next selected market if one was available.
- Revised selection to include the top three randomly selected markets, excluding those dropped above.

2.1.2 Additional market notes

Four selected markets (Hawaii, New Hampshire, Rhode Island, and Vermont) cover job postings for their entire respective states. For each of these markets, we select the city with the highest population for the applicants' backgrounds (location of high schools and work experience) and where they apply. The cities with the highest populations are Honolulu, HI; Manchester, NH; Providence, RI; and Burlington, VT.

We also note that two pairs of selected markets are close to one another (about a one-hour drive). These pairs are Annapolis and Baltimore, MD, and Madison and Milwaukee, WI. If one market in a pair has a shortage of job postings, it could draw upon postings from the other nearby market, so we made sure to check that the location of the postings is in the actual market and check for duplicate postings across markets.

2.2 Resume generation

Each resume is tailored to the food or retail industry, and each applicant will have two entry-level job experiences in each industry (one customer-facing, one non-customer-facing). Job descriptions were generated using ChatGPT and then edited appropriately.

The jobs and companies are standardized. Food industry positions include server and line cook, and the retail jobs are variants of retail associate and inventory assistant. The companies are selected from the highest-grossing chain restaurants and retailers in the United States. The chain restaurants must have a sit-down option so that job applicants could have been a server at said restaurants. Then, the locations of these jobs are selected from the companies' locations nearest the applicants' cities of residence.

Each resume will have a high school degree from the high school located within their city of residence, and we target high schools that are ranked as "average" (5–6) by GreatSchools, extending the range to 4–7 when necessary. We include both public and charter high schools, excluding magnet schools.

2.3 Posting identification

We apply for entry-level food service and retail positions. Specifically, we target all postings listed on the food service and retail pages of the Craigslist job ads, excluding those that are (1) not entry-level (ex: manager positions that require more experience); (2) require the applicant to call, text, complete an application, write a cover letter, or go in-person to apply (3) and/or mass positions for gig positions through firms like DoorDash and InstaCart, which require applications via an external website.

Each week, we use a Python script to scrape food and retail job postings from all selected markets on Craigslist. The script removes ineligible postings, and we manually remove any ineligible postings that the code did not remove. In addition, we remove any posts from employers to whom we previously submitted job applications to avoid suspicion of the authenticity of the applications. If employers post multiple positions, we randomly select one position to apply to. Then, we divide up the job postings into groups by market and send them to the research assistants.

2.4 Name selection

We signal applicant gender and race by selecting names that are particularly common to white women, black women, white men, and black men, using the SSA list of first names given to newborns

in 2003 and 2000 U.S. Census data on surnames.

We identify gender-specific first names as those that are given to only girls or only boys at least 90% of the time. We identify white-signaling names as those given to white babies at least 90% of the time, and we identify black-signaling names as those given to black babies at least 50% of the time. Because many white names met these criteria, we restricted to names that were at least in the top 50 most popular names. Within the pool of potential names, we randomly select two for each gender-race category.

We select surnames that are at least 90 percent white or at least 40 percent black. While many first and last names are at least 90 percent white, few are above 40 percent black, which is why the cutoff percentage is lower for black first and last names.

We signal transgender identity by including a statement near the top of each resume stating that the applicant's preferred and legal names differ. For example, "My preferred name is John, but my legal name is Mary." One name is unambiguously masculine, while the other name is unambiguously feminine to establish a contrast in gender between the applicant's legal name and preferred name that signals transgender identity. We use the same set of names for legal names and preferred names.

2.5 Randomization

Stratifying by market, we randomize half of all applications to be from white applicants and half to be from black applicants. We cross-randomize this with gender identity, such that applications are equally likely to be assigned to one of four gender identities: Cis, no pronouns; Cis, pronouns; Trans, binary; and Trans, non-binary. The first 1000 submitted resumes are also stratified by employer, but the remainder is stratified only by market, reflecting recommendations from Phillips (2019).

Based on race, each applicant was randomly assigned to receive one of four race-specific names, two for men and two for women. Among the white trans, non-binary applicants, half were randomized to receive a non-binary name. We did not generate applications for black trans non-binary applicants because popular names that were distinctly Black were also distinctly gendered. For trans binary applicants, a legal name was chosen that reflected the same race but a different gender identity, so trans men have a recognizably female legal name. Table 2 includes the set of names used.

For each job category and market, we generate two different resume versions, which we randomly assigned so that each employer received one application from each type. These versions were designed to be of equal quality.

2.6 Resume submissions

After randomizing the resumes and filling in the appropriate information, research assistants submit applications to Craigslist. First, they copy a unique email address for the job posting on Craigslist and paste it into the "to" field on the appropriate Yahoo email account. Second, they copy and

Table 2: Applicant names

	White	**Black**
Female	Katelyn Walsh	Ebony Dorsey
	Mary Hansen	Kenya Mosley
Male	Justin Carlson	Demetrius Banks
	Hunter Weber	Prince Booker
Gender ambiguous	Skyler Jensen	
	Peyton Meyer	
Legal names (transgender applicants)		
Female	Brooke	Aisha
	Nicole	Ayanna
Male	Thomas	Jermaine
	Gavin	Reginald

paste the email subject line (which varies to indicate gender identity internally) and add the job title. Third, they copy and paste the email template and edit it to include the job title in the body of the email and the applicant's name in the email signature. Fourth, they attach the resume to the email and send it to the employer.

2.7 Outcomes

The key outcome variable will be whether a submitted resume receives a callback from an employer. We define a callback as any e-mail, text message, or phone call to invite the applicant for an interview or to request more information. We will consider all responses within two weeks to constitute a callback. Neutral interactions, such as an employer sending a message to confirm receipt, are not considered a callback.

2.8 Power calculations

We calculate minimum detectable effect sizes based on 80% power and a 5% significance level. We use a baseline callback rate of 19% and assume that covariates and market fixed effects explain 10% of the variation in callback rates, reflecting Stanton (2023).

With 1,500 observations per arm in Equation 1, we have 80% power to detect aa MDE of 3.8 percentage points among non-binary applicants or pooled transgender men and transgender women. We can detect an MDE of 4.8 percentage points among transgender men and women, respectively, based on pair-wise comparisons with cisgender men. Based on simulation, for interaction effects by race, as shown in Equation 2, we have 80% power to detect an interaction effect of roughly 5.5

percentage points.

2.9 Ethical approvals

This study is approved under protocol STUDY00002218 by the University of Vermont Institutional Review Board.

3 Empirical specifications

Our primary specification will measure the differential impact of gender identity relative to cisgender men:

$$\begin{aligned} Callback_{ijm} = \alpha + \beta_1 TransBinary_{ijm} + \beta_2 TransWoman_{ijm} + \beta_3 NonBinary_{ijm} + \\ \beta_4 CisWoman_{ijm} + \gamma_1 CisProunouns_{ijm} + \gamma_2 CustFacing_{jm} + \phi_m + + X'\psi + \epsilon_{ijm} \end{aligned} \tag{1}$$

where $Callback_{ij}$ equals 1 if employer j in market m called back resume i for an interview. The indicators $TransMan$, $TransWoman$, $NonBinary$, and $CisWoman$ are equal to 1 based on the signaled gender identity in the submitted resume. $CisPronouns$ is equal to one if a cisgender applicant includes pronounces (she/her or he/him). ϕ_m represents a set of market-level fixed effects, which form our stratification cells.

We will improve the precision of our estimates by including day-of-week (MTWThFSSu), week-of-submission (1–8) FE, and submitter FE, captured in the vector X. If these meaningfully affect the results, we will also report a version without these covariates. We will report both unadjusted p-values and p-values adjusted to account for multiple hypothesis testing.

To detect differences by race, will estimate a single, fully interacted model to test whether coefficients on β_1 through β_4 are each equal between black and white applicants:

$$\begin{aligned} Callback_{ijm} = \alpha + \beta_1 Trans_{ijm} + \beta_2 Black_{ijm} + \beta_3 TransXBlack_{ijm} + \\ \gamma_1 CisProunouns_{ijm} + \gamma_2 CustFacing_{jm} + \phi_m + + X'_{ij}\psi + \epsilon_{ijm} \end{aligned} \tag{2}$$

We will also estimate a version of Equation 2 that disaggregates the group transgender applicants, $Trans$, into transgender men, transgender women, and non-binary applicants, noting that there will not be a $NonBinary$ interaction term for Black applicants

To examine differences in impact by political climate, we will estimate a model similar to Equation 2, but using a political climate variable as the interaction term, which will vary at the market or state level.

$$\begin{aligned} Callback_{ijm} = \alpha + \beta_1 Trans_{ijm} + \beta_2 Pol.Climate_m + \beta_3 TransXPol.Climate_{ijm} + \\ \gamma_1 CisProunouns_{ijm} + \gamma_2 CustFacing_{jm} + Y'_m\kappa + X'_{ij}\psi + \epsilon_{ij} \end{aligned} \tag{3}$$

For this analysis, we omit market-level fixed effects to avoid collinearity. Instead, we will estimate a parsimonious specification with no other covariates and then we will examine whether these estimates are robust to controlling for demographics (share White, Black, LGBT (as available)), economic conditions, and minimum wage, represented by the vector Y.

We will consider several potential measures of political climate toward transgender people, including the following:

- State and/or local-level Republican vote share in the 2020 election as a proxy for the initial political climate
- State-level measures of anti-trans legislation passed in 2022 and 2023
- Local-level measures of anti-trans or trans-affirming legislation passed in 2022–2023

4 Hypotheses

H1: Discrimination against women and transgender applicants means that cisgender men will have the highest callback rates and transgender women will have the lowest, with rates for cisgender women, transgender men, and non-binary applicants falling in between. From Equation 1 β_1, β_2, β_3 will be jointly less than zero, with $\beta_2 < \beta_1$ and $\beta_2 < \beta_3$

H2: Discrimination by gender identity will be greater among black than white applicants, with γ_2, γ_3, and γ_4 from Equation 2 jointly less than zero.

H3: The callback differential between cisgender and transgender applicants will be greater in areas with more anti-trans political climates. In the model in which we use a measure of anti-trans political climate, reflecting Equation 3, we predict that γ_2, γ_3, and γ_4 will be jointly less than zero.

H4: If discrimination is driven by employer perceptions of customers' preferences, then discrimination will be greater for positions that are customer-facing. Using a model equivalent to Equation 2 but interacting whether positions are customer-facing with gender identity, then γ_2, γ_3, and γ_4 from Equation 2 will be jointly less than zero.

4.1 Secondary Analysis

We will explore the following dimensions in our secondary analysis:

- Differential impact of race among gender identities, disaggregating transgender applicants into transgender men, transgender women, and non-binary applicants
- Impact of pronoun inclusion among cis applicants, which will allow us to isolate the impact of being non-binary vs. using pronouns

- Impact of feminine vs. masculine vs. gender-neutral names on call-back rates among non-binary applicants

References

Abbate, N., I. Berniell, J. Coleff, L. Laguinge, M. Machelett, M. Marchionni, J. Pedrazzi, and M. F. Pinto (2024, April). Discrimination against gay and transgender people in Latin America: A correspondence study in the rental housing market. *Labour Economics 87*, 102486.

Aksoy, B., C. S. Carpenter, and D. Sansone (2022, September). Understanding labor market discrimination against transgender people: Evidence from a double list experiment and a survey.

Bardales, N. (2013). Finding a job in "a beard and a dress"': Evaluating the effectiveness of transgender anti-discrimination laws. *Unpublished manuscript*.

Bertrand, M. and S. Mullainathan (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review 94*(4), 23.

Eames, T. (2024). Taryn versus Taryn (she/her) versus Taryn (they/them): A Field Experiment on Pronoun Disclosure and Nonbinary Hiring Discrimination. Technical report.

Fumarco, L., B. Harrell, P. Button, D. Schwegman, and E. Dils (2023). Gender identity, race, and ethnicity discrimination in access to mental health care: Evidence from a multi-wave audit field experiment. Technical report, National Bureau of Economic Research.

Granberg, M., P. A. Andersson, and A. Ahmed (2020, August). Hiring Discrimination Against Transgender People: Evidence from a Field Experiment. *Labour Economics 65*, 101860.

Grant, J. M., L. A. Mottet, J. J. Tanis, and D. Min (2011). Transgender discrimination survey. *National Center for Transgender Equality and National Gay and Lesbian Task Force: Washington, DC, USA*.

Kattari, S. K., D. L. Whitfield, N. E. Walls, L. Langenderfer-Magruder, and D. Ramos (2016, September). Policing Gender Through Housing and Employment Discrimination: Comparison of Discrimination Experiences of Transgender and Cisgender LGBQ Individuals. *Journal of the Society for Social Work and Research 7*(3), 427–447.

McFadden, C. (2020, April). Hiring discrimination against transgender job applicants – considerations when designing a study. *IJM 41*(6), 731–752.

MIT Election Data and Science Lab (2017). U.S. President 1976–2020.

Phillips, D. C. (2019). Do comparisons of fictional applicants measure discrimination when search externalities are present? Evidence from existing experiments. *The Economic Journal 129*(621), 2240–2264.

Rainey, T., E. Imse, and A. Pomerantz (2015). Qualified and transgender: a report on the results of resume testing for employment discrimination based on gender identity. *Washington, DC: Office of Human Resources.*.

Shannon, M. (2022, August). The labour market outcomes of transgender individuals. *Labour Economics 77*, 102006.

Stanton, I. (2023). Hiring discrimination against transgender individuals in the us labor market. Technical report.

A Appendix Tables

Table A.1: Initial posting distribution, by market

# Postings	Freq.	Percent	Cum.
0	167	40.05	40.05
1	71	17.03	57.07
2	30	7.19	64.27
3	26	6.24	70.50
4	13	3.12	73.62
5	13	3.12	76.74
6	4	0.96	77.70
7	11	2.64	80.34
8	10	2.40	82.73
9	8	1.92	84.65
10	5	1.20	85.85
11	8	1.92	87.77
12	3	0.72	88.49
13	5	1.20	89.69
14	3	0.72	90.41
15	5	1.20	91.61
16	2	0.48	92.09
17	5	1.20	93.29
18	1	0.24	93.53
19	3	0.72	94.24
20	2	0.48	94.72
21	3	0.72	95.44
22	2	0.48	95.92
23	1	0.24	96.16
24	1	0.24	96.40
25	2	0.48	96.88
27	1	0.24	97.12
28	1	0.24	97.36
32	1	0.24	97.6
33	1	0.24	97.84
36	1	0.24	98.08
40	1	0.24	98.32
41	1	0.24	98.56
46	1	0.24	98.80
47	1	0.24	99.04
62	1	0.24	99.28
90	1	0.24	99.52
96	1	0.24	99.76
200	1	0.24	100.00
Total	417	100	100

Table A.2: Selected Markets

Phoenix	AZ	Phase 1
Tucson	AZ	Phase 2
Los Angeles	CA	Phase 1
Orange County	CA	Top-up
San Francisco Bay Area	CA	Phase 2
Boulder	CO	Phase 2
Denver	CO	Phase 2
High Rockies	CO	Top-up
Palm Beach County	FL	Top-up
South Florida	FL	Phase 2
Tampa Bay Area	FL	Phase 2
Atlanta	GA	Phase 2
Hawaii	HI	Phase 2
Iowa City	IA	Phase 2
Boise	ID	Phase 2
Chicago	IL	Phase 2
New Orleans	LA	Phase 2
Annapolis	MD	Phase 2
Baltimore	MD	Phase 2
Kansas City	MO	Phase 2
St Louis	MO	Phase 2
Charlotte	NC	Top-up
Raleigh / Durham / Ch	NC	Phase 2
Wilmington	NC	Phase 2
New Hampshire	NH	Phase 2
Albuquerque	NM	Phase 2
Santa Fe / Taos	NM	Phase 2
Las Vegas	NV	Phase 2
Reno / Tahoe	NV	Phase 2
Glens Falls	NY	Top-up
Hudson Valley	NY	Phase 2
New York City	NY	Phase 1
Cleveland	OH	Phase 2
Bend	OR	Top-up
Medford-Ashland	OR	Phase 2
Salem	OR	Phase 2
Rhode Island	RI	Phase 2
Charleston	SC	Phase 2
Nashville	TN	Phase 2
Dallas / Fort Worth	TX	Top-up
Houston	TX	Phase 1
San Antonio	TX	Phase 2
Richmond	VA	Phase 2
Vermont	VT	Phase 2
Seattle-Tacoma	WA	Phase 2
Spokane / Coeur D'Alene	WA	Phase 2
Wenatchee	WA	Top-up
Madison	WI	Phase 2
Milwaukee	WI	Phase 2
Total states	27	
Total markets	49	