sociological science

# Getting a Foot in the Door: A Meta-Analysis of U.S. Audit Studies of Gender Bias in Hiring

So Yun Park, Eunsil Oh

University of Wisconsin–Madison

**Abstract:** For the past three decades, scholars have conducted field experiments to examine gender-based hiring discrimination in the United States. However, these studies have produced mixed results. To further interpret these findings, we performed a meta-analysis of 37 audit studies conducted between 1990 and 2022. Using an aggregated sample of 243,202 fictitious job applications, the study finds no evidence of statistically significant gender discrimination at the study level. However, a series of more focused meta-analyses reveal important variations in the extent of discrimination by occupation type and applicant race. First, the gender composition of an occupation predicts gender bias in hiring. Second, the intersection of gender and race is critical—in female-dominated jobs, White female applicants receive more callbacks than their male counterparts, but Black female applicants experience no such benefit. The study contributes to the literature on labor market and gender (in)equality by synthesizing the findings of field experiments.

**Keywords:** gender; hiring discrimination; field experiments; meta-analysis

**Reproducibility Package:** Data and code are available at the Open Science Framework via https://osf.io/kp5df/

Since the 1960s, women's employment patterns in the United States have significantly changed. During this period, employment rates among women increased dramatically (Cotter, Hermsen, and England 2008), women began to participate in previously male-dominated occupations (Cotter, Hermsen, and Vanneman 2004), and the federal government passed laws banning gender-based discrimination in the workforce (Hirsh 2009). Scholars have argued that despite these changes, the gender revolution in the paid labor market has stalled since the 2000s (England, Levine, and Mishel 2020), and in particular, gender-based wage and hiring gaps persist in the U.S. workforce (Budig 2002; Budig and Hodges 2010; Charles and Grusky 2004; Eaton et al. 2020; Huffman 2004; Killewald and Bearak 2014).

Sociologists, economists, and gender scholars have investigated hiring discrimination based on gender via various research designs. One of the leading methods in this area is field experiments, which entail some of the controlled elements of traditional lab experiments but occur in natural, real-world settings. Specifically, audit studies or correspondence studies are one of the most common forms of field experiments. In audit studies, researchers use fictitious job applications with randomly assigned characteristics (e.g., gender, race/ethnicity, and criminal record) to audit employment outcomes and thereby assess discrimination in hiring based on manipulated characteristics (Gaddis 2018). Unlike surveys, which focus on individuals who have already been hired, audit studies offer a rare opportunity to assess discrimination in the initial stage of hiring. The experimental research design

**Citation:** Park, So Yun, Eunsil Oh. 2025. "Getting a Foot in the Door: A Meta-Analysis of U.S. Audit Studies of Gender Bias in Hiring" Sociological Science 12: 26-50.

**Received:** September 28, 2024

**Accepted:** November 14, 2024

**Published:** January 9, 2025

**Editor(s):** Ari Adut, Vida Maralani

**DOI:** 10.15195/v12.a2

allows researchers to explore decisions made by real employers, thus offering a high internal validity level. Given these strengths, the use of field experiments has expanded; numerous studies have leveraged this methodology to explore hiring discrimination based on social locations and demographic backgrounds, including race and ethnicity (Bertrand and Mullainathan 2004), educational credentials (Deterding and Pedulla 2016), criminal background (Pager 2003), and union membership (Kreisberg and Wilmers 2022).

The first experiment examining patterns of gender discrimination in the United States was conducted by Neumark, Bank, and Van Nort (1996) and explored gender discrimination in hiring practices at restaurants in Philadelphia. In the experiment, male and female research assistants submitted 130 job applications in person. The results showed that in high-priced restaurants, female job applicants had lower probabilities of receiving both interviews and job offers than their male counterparts. However, in-person experiments are rare, mainly because it is difficult to keep the nature of the experiment confidential. In most studies, research confederates apply for jobs via phone or online to avoid in-person encounters between applicants and employers. Nearly all field experiments of hiring patterns use identical human capital and skill sets across different gender groups of confederate applicants, ensuring that the fictitious resumes include the same levels of education and job experiences. The results are often measured by the gender gap in callback rates.

Notably, field experiments that explore gender-based discrimination have produced mixed results. Some studies have found evidence of gender discrimination (Deming et al. 2016; Pedulla 2018), whereas others have observed discrimination only when gender is interacted with additional demographic characteristics (Correll, Benard and Paik 2007; Quadlin 2018, Rivera and Tilcsik, 2016; Thomas 2018); and yet others have found no evidence of gender discrimination (Botelho and Chang, 2022; Mihut 2022; Nunley et al. 2015; Weisshaar 2018).

The central aim of this study is to unpack the mixed results found in prior field experiments to better understand the patterns of gender-based discrimination in the U.S. labor market. The analysis both extends the existing research that employs meta-analyses and synthesizes the results from field experiments in hiring (Flage 2019; Gaddis et al. 2021; Lippens, Vermeiren, and Baert 2023; Quillian et al. 2017).

In particular, this study is the first to perform a meta-analysis that solely situates gender discrimination within the U.S. labor market and policy context. Recent meta-analyses on gender-based discrimination have explored why field experiments puzzlingly find both disproportionately high and disproportionately low callback rates among female job applicants (Galos and Coppock 2023; Lippens et al. 2023; Schaerer et al. 2023). Such a line of existing work provides insight into how occupational characteristics (i.e., gender composition) predict gender bias in hiring (Galos and Coppock 2023). However, these studies pool multiple countries with divergent labor market structures, workplace conditions, and policies. This—using international settings and heterogeneous macro-level contexts—makes it challenging for scholars to adequately sort out the impacts of occupation and group-based differences. Thus, the current study focuses on field experiments conducted within one country, considering the fact that gender and race are context-specific social constructs.

**Table 1:** Theories of status predicting gender-based hiring discrimination.

| Conceptual Framework | Measurements | Predictions for Callback Rates of Female Applicants |
|---|---|---|
| Gender Status Beliefs | Applicants' gender | $-$ |
| Intersectionality | Interaction between diverse demographic characteristics and gender | $+/-$ Depending on the interaction between background measures |
| Compositional Status (tokenism and statistical discrimination) | Gender composition (the average proportion of female workers) of a job or occupation | $+$ For occupations with a *higher* proportion of women<br>$-$ For occupations with a *lower* proportion of women |

To explain the mixed results found in field experiments in the U.S. labor market, we compiled audit studies of gender-based hiring discrimination that were administered in the United States from the first experiment in 1996 to the most recent in 2022 ($N = 37$). Then, we conducted meta-analyses using an aggregate of the focal study samples, including 243,202 fictitious job applications.

## Theoretical Frameworks

The theoretical foundations of this study are demand-side theoretical models that predict and explain the patterns of gender-based hiring bias in multiple contexts. Field experiments test whether employers make discriminatory decisions. Thus, demand-side explanations of hiring discrimination, which focus on employers' decisions to hire, promote, and evaluate employees, are particularly useful in explaining the results of prior field experiments (in contrast, supply-side theories focus on factors related to the employees themselves).

Table 1 presents three key demand-side theories and their predictions for gender-based hiring discrimination. The first possible explanation of hiring discrimination is that employers make decisions that are rooted in status beliefs—specific beliefs about which groups are more capable and competent (Ridgeway 2014, 2019). Status beliefs may derive from stereotypes or purely discriminatory beliefs about the superiority of a particular group. The top row in Table 1 refers to a framework that considers discrimination based on an applicant's gender, which results in fewer callbacks and worse hiring outcomes for women.

A second conceptual framework concerning status beliefs is the intersectional perspective, which emphasizes that status beliefs are not based solely on the applicant's gender but rather represent compounded beliefs about multiple groups (Choo and Ferree 2010; Crenshaw 1989). Thus, the intersection of race, ethnicity, citizenship, and human capital likely influences the magnitude and direction of gender bias. For example, this approach acknowledges the persistence of racial inequality in the U.S. labor market (Quillian et al. 2017) and stresses that racial differences among female applicants in field experiments will impact the results. Based on the intersectional perspective, we predict that there is heterogeneity in results across audit studies that hinges on interaction effects between gender and other variables such as education credentials, race/ethnicity, and parental status.

A third key perspective is a composition-based status perspective, which posits that the proportion of women in an organization or institution (i.e., occupation) is an important factor associated with employers' hiring bias. Both economists and organizational scholars have argued that the compositional positioning of women measured by the proportion of female employees is a salient factor for explaining gender-based discrimination in hiring. Economists have introduced the idea of statistical discrimination, which posits that employers make hiring decisions based on their prior knowledge about average productivity levels across different groups (e.g., women and other identifiable classes of job applicants) (Bielby and Baron 1986; McPherson, Smith-Lovin, and Cook 2001). Statistical discrimination theory argues that hiring discrimination against women in male-dominant jobs is the result of employers' rational choices, given that they seek to overcome uncertainty around applicant performance (Arrow 1973).

Organizational scholars also emphasize the composition of the workforce as a factor driving gender discrimination, arguing that when women hold a proportional minority of positions within an organization, it induces gender inequality (Kanter 1977). Kanter (1977) proposed the idea of tokenism to explain the disadvantages faced by either male or female workers when they constitute less than 15 percent of the population of an organization. In this framework, those of the "token" gender face discrimination in promotions, evaluations, and hiring (Campero and Fernandez 2019; Kanter 1977).

### *Additional Predictions: Occupational-Level Status*

Given that field experiments have been conducted in a wide range of occupations (Galos and Coppock 2023), this study further predicts that employers' hiring patterns may vary across occupations. Specifically, we explore how the different occupational-level statuses shape hiring decisions to further analyze the impact of hiring discrimination on producing inequality. Although we hold an exploratory stance in predicting the directionality of the relationship between occupational-level status and hiring discrimination, we predict that different forms of occupational-level status—including symbolic status (prestige), median wage, and skill sets—are likely to shape hiring patterns.

First, occupational prestige is a measure of the social status and symbolic status of an occupation. Higher-prestige jobs often enhance individuals' career choices and opportunities for advancement (Lamont et al. 2016; Valentino 2020). In addition, median wage level is used as an indicator of the economic status of an occupation. Finally, skill level is a salient characteristic of an occupation often correlated with prestige and wages. Thus, we explore how these status measures of an occupation relate to different levels of gender inequality in hiring.

## Data

To conduct the meta-analysis, we first identified and compiled relevant field experiments and then constructed a novel data set. We employed data collection approaches similar to those used in prior meta-analysis studies (Flage 2019; Gaddis et al. 2021; Quillian et al. 2017).

## Study Identification and Inclusion Criteria

We conducted the data collection in two steps. First, we defined a set of eligibility criteria based on the research problem specification. To be included in the analysis, a study had to meet three criteria: (1) the field experiment (correspondence audit) was conducted in the U.S. labor market, (2) job applicant gender was either the focal variable in the within-subject (matching) design or randomized within and/or across jobs in the application process, and (3) the outcome of interest could be transformed into a binary hiring decision (e.g., callback) for or against job applicants.

In the second step, we conducted several rounds of searches to find all possible qualifying studies to include in the data set. The initial list of studies was created through systematic searches using EBSCO SocINDEX, NBER working papers, ProQuest Dissertations and Theses, ProQuest Sociological Abstracts, and Web of Science SSCI. We then conducted both forward and backward citation searches using the selected studies. The backward search involved a careful review of the bibliography of each study in the initial list. For the forward citation search, we used Google Scholar to identify all later works citing studies from the initial list.

To avoid publication bias (bias resulting from the exclusion of studies that were not published due to a lack of findings demonstrating discrimination or other factors), we emailed the authors of unpublished manuscripts and sought out nonpublished field experiments—dissertations, pre-prints, working papers, and conference papers. These studies are included in the main sample and the results reflect both published and non-published studies. After compiling the study list, the two researchers cross-coded the data by reading each article to extract and code the variables necessary for the analysis.

## Study-Level Data Set

In total, we included 37 individual studies that analyzed 243,202 job applications in the study-level data set. This data set is composed of 37 studies that met our criteria. Despite searching for studies published from the year 1990, our data span field experiments that were conducted between 2002 and 2021 and published between the years 2004 and 2022—Figure 1 graphs the number of publications per year as well as when the studies conducted their experiments. Our data have been restricted to correspondence audit studies that do not entail an in-person component. This meant that earlier publications conducted in person, including Neumark et al. (1996), were excluded. Approximately 84 percent of the studies in our data found job postings online, mainly using online job boards and job search websites. The remaining 16 percent of the studies used job advertisements in newspapers and journals, directories, and directly emailed institutions.

In addition, an important characteristic of our study-level data set is that several studies have more than one manipulation in their experiment. In other words, the fictitious applications submitted by researchers varied in more than one feature, allowing researchers to examine hiring discrimination across several variables.

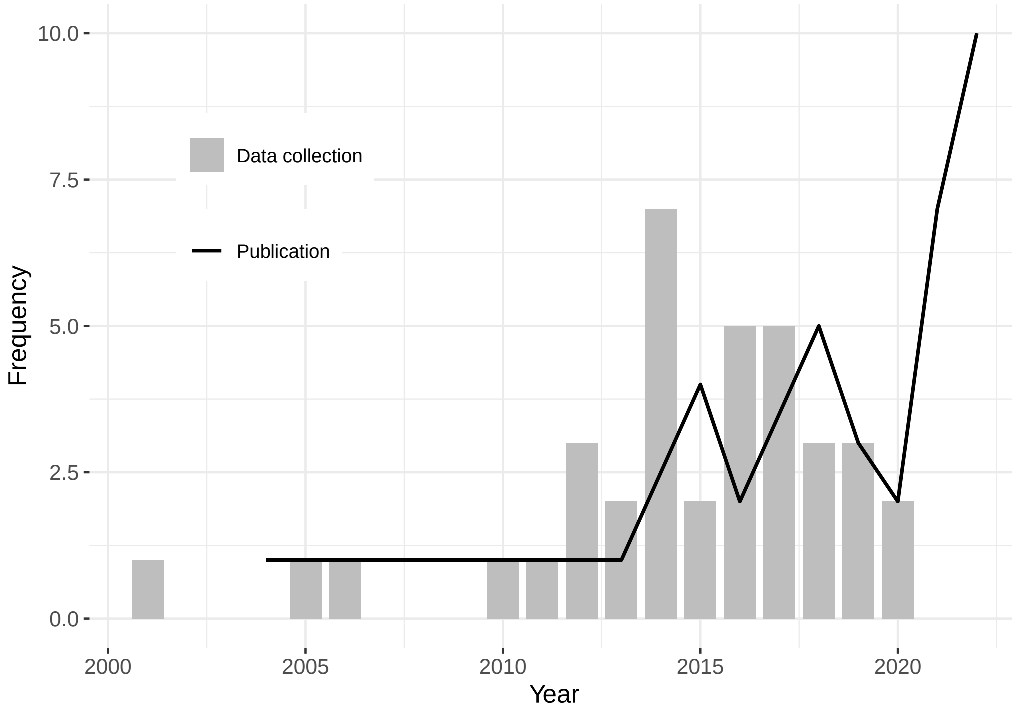



**Figure 1:** Number of U.S. audit studies by year of data collection and publication. *Note:* For studies that collected data across multiple years, we used the earliest year of data collection. There were two studies that did not indicate when the data were collected. For these two cases, we followed an approach used in prior meta-analysis studies (Quillian et al. 2017) and used publication year minus 2 years.

Figure 2 summarizes the central manipulations in the fictitious resumes used in the 37 selected studies. The three most frequent manipulations are gender (76 percent[1] of studies), educational credentials (41 percent), and race/ethnicity (35 percent). Among all the selected studies, 35 percent focus primarily on gender, whereas 65 percent include gender as part of the manipulation.

## *Occupational-Level Data Set*

To adequately analyze the occupational variations across and within the studies, we constructed a separate data set at the occupation level. Approximately 84 percent of the studies conducted their experiments across more than one occupation group. The remaining 16 percent of the studies focused on one specific occupation group (e.g., lawyer and teacher). In the end, we have constructed a sub-data set that comprises 85 individual occupation-level data points[2].

Based on the data at the occupational level, we examined four occupational characteristics: the overall skill levels required gender ratio, the associated level of prestige, and median hourly wage level. All occupation-level measures were extracted from external data. Thus, the occupation categories used in the studies were re-categorized into the Standard Occupational Classification System and the Census occupation codes. The two researchers found the next best matching occupation

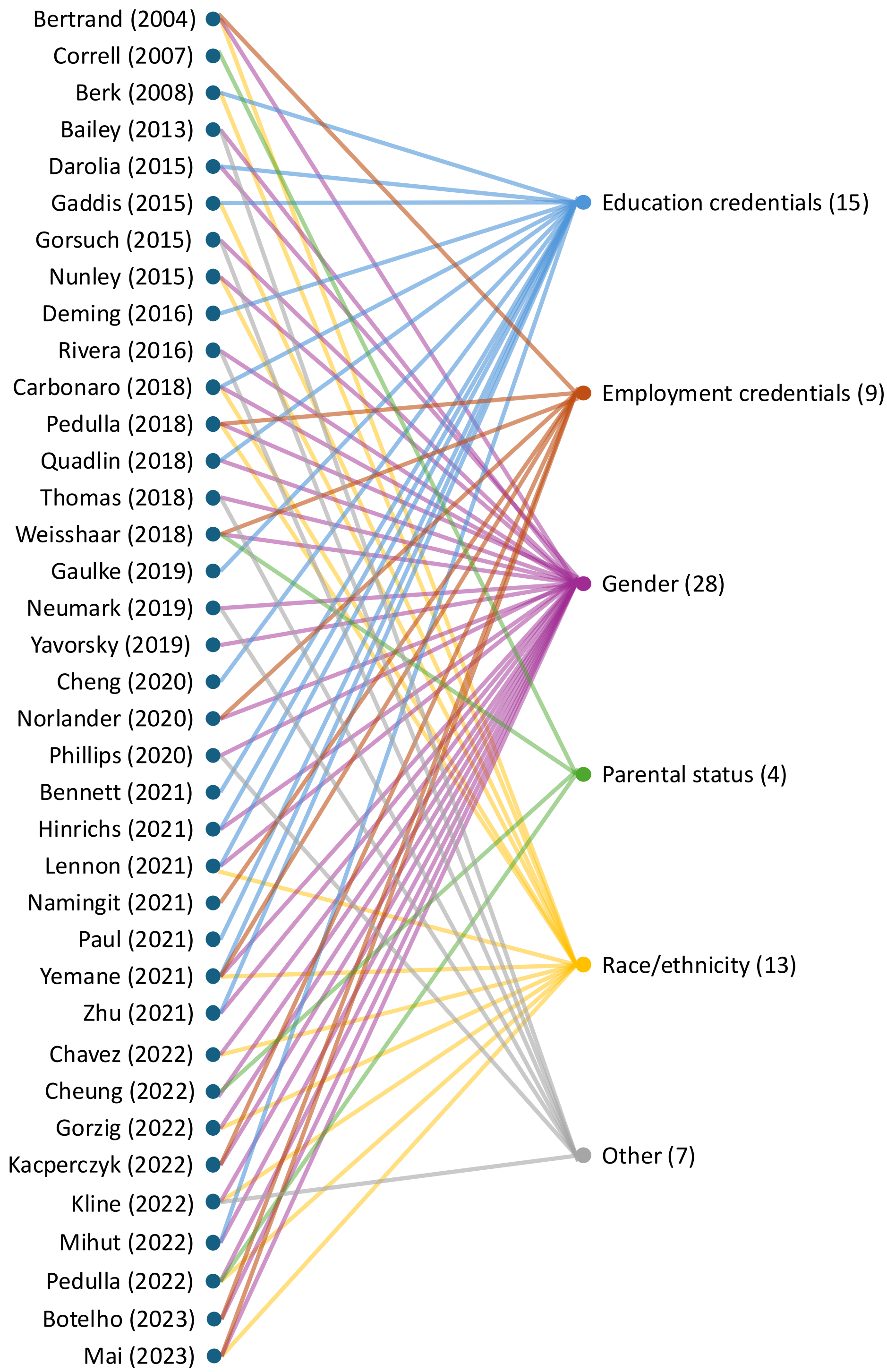



**Figure 2:** Mapping resume manipulations. *Note:* The first author's name is used for co-authored studies. The "other" category includes manipulations such as age, class, geographical area, political leaning, and sexual orientation. Our study eligibility criteria cover the period from January 1990 to June 2022. For Botelho (2023) and Mai (2023), we used earlier versions of the article but included the published articles in Figure 2 and the list of studies used in the meta-analysis in the online supplement for reference.

**Table 2:** Descriptive statistics for occupational-level data points ($N = 85$) from 37 studies.

| | Frequency | Percentage |
|---|---|---|
| Race | | |
| White only | 15 | 18 |
| Mixed | 62 | 73 |
| Unknown | 8 | 9 |
| Parenthood status | | |
| Yes | 7 | 8 |
| No | 78 | 92 |
| | Median | Std. Dev. |
| Year of data collection | 2014 | 3.83 |
| Education (years) | 16 | 1.82 |
| Percent female | 55.7 | 23.59 |
| Median hourly wage (USD) | 32.57 | 17.16 |
| Prestige score | 47 | 11.41 |
| Required skill level | 67.2 | 20.47 |
| ***N*** | 85 | |

category for categories that did not match exact descriptions. All occupation categories were a one-to-one match with data[3].

We provide descriptive statistics of the occupation-level data set in Table 2. The demographic characteristics of the study sample include the applicants' race, which we coded as whether the experiment was limited to White job applicants (18 percent), whether applicants of varying race/ethnicity were examined (73 percent), or not specified (9 percent), the average education level coded in years of education completed (median = 16 years), and a binary variable indicating whether a portion of the applicants signaled statuses of parenthood or pregnancy (8 percent of the studies included parenthood status signals). The median percent of females is 55.7 percent, indicating a slightly higher proportion of female workers. The median prestige score, a relative scale measuring the perceived prestige of an occupation, is 47. The median required skill level, a relative scale created using the percentage of individuals in an occupation group who have completed postsecondary education, is 67.2. We also coded the year of data collection as a research design factor. The median year of data collection is 2014.

# Methods

## *Dependent Variable*

The dependent variable for the meta-analysis is the logged discrimination ratio. The original studies assessed discrimination based on whether the applications submitted to jobs received "callbacks," which is a measure of an employer's request for either additional information or an interview that is common in hiring field experiments. The discrimination ratio is the proportion of positive callbacks for men's applications relative to positive callbacks for women's applications.

Calculating the ratio requires three data points for each study: sample size, callback rate for women, and callback rate for men. Following an approach used in prior literature (Schaerer et al. 2023), we changed rates of 0 to 0.5 to avoid dropping cases; this methodological choice also reduces bias in the estimator and prevents division by zero. The adjusted data points totaled four cases, which include two data points in the occupation-level White applicant data set and two data points in the occupation-level Black applicant data set

$$Discrimination\ Ratio_i = \frac{\frac{callback_{male_i}}{applications_{male_i}}}{\frac{callback_{female_i}}{applications_{female_i}}}. \tag{1}$$

The variance of the logged discrimination ratio was calculated using the following equation:

$$\begin{aligned} Var(\ln{(DiscriminationRatio_i)}) = & \frac{1}{callback_{female_i}} - \frac{1}{applications_{female_i}} \\ & + \frac{1}{callback_{male_i}} - \frac{1}{applications_{male_i}}. \end{aligned} \tag{2}$$

## Independent Variables

The independent variables in the meta-regression models fall into two broad categories: study sample demographic characteristics and study sample occupational characteristics. Study sample demographic characteristics include applicants' education, parenthood status, and race (whether the experiment was limited to White job applicants, included applicants of various races/ethnicities, or did not clearly state applicants' race/ethnicity). Study sample occupational characteristics include occupational skill level (a measure of the overall skill level required by an occupation), the average ratio of female-to-male workers in the occupation, the median hourly wage in the occupation, and occupational prestige.

We extracted the data points for the four variables from multiple sources, including an index of occupational skill level developed by Hauser and Warren (1997) and updated by Frederick (2010)[4]. We have also used the Bureau of Labor Statistics' *Employed Persons by Detailed Occupation in 2022* and *May 2021 National Occupational Employment and Wage Estimates*. Finally, we utilized Treiman's Standard International Occupational Prestige Scale (Ganzeboom and Treiman 2019). Furthermore, merging the occupational characteristics data involved matching occupation categories between the selected studies and the Bureau of Labor Statistics report. The authors independently cross-coded the occupations and went through multiple rounds of calibration. All four measures are continuous variables. All scales range from 0 to 100 apart from the median hourly wage variable measured in U.S. dollars.

In addition, the analysis included only White and Black applicants because the number of applications from candidates of other races was negligible. The full data set included 66,145 White applicants and 34,328 Black applicants. The education levels used in each study and occupation group of applications were comparable, with an average of 14.82 years of education for White applicants and 14.75 years of education for Black applicants. Occupations were also comparable across Black and White applicants. All occupation groups were included except domestic workers, human resources, lawyers, manufacturing, and teachers (these five groups either had only White job candidates or race was not introduced into the study). We excluded Hispanic and Asian applicants because there were so few. For example, for the 15 occupation-level data points that included information on Hispanic job candidates and callback results, half had less than 100 observations per gender.

### Analytical Model

We conducted a series of meta-regressions using a logged version of the calculated discrimination ratio as the dependent variable. More specifically, we used a random-effect model to account for potential variability across studies that may result from two factors: individual studies targeting different geographic areas and occupation categories and variation across studies in the characteristics of job applicants. The random-effect model assumes that there is not a single true effect size but rather a distribution of true effect sizes. Therefore, the model's goal is to estimate the mean distribution of true effects. Because we anticipated considerable inter-study heterogeneity, we used a random-effect model to pool the effect sizes. The meta-regression model equation is

$$\ln\ (y_i) = x_i\beta + u_i + \varepsilon_i, \text{ where } u_i \sim N\left(0,\ \tau^2\right) \text{ and } \varepsilon_i \sim N(0,\ \sigma_i^2). \tag{3}$$

In the first step, we conducted a random-effect meta-analysis regression for the pooled effects of the 37 studies in the data set. We then conducted a random-effect meta-analysis regression of the 85 individual occupation-level data points. Our second prediction was that there might be a high level of heterogeneity in the jobs selected for field experiments, which might increase the likelihood of obtaining inconclusive results. To address this possibility, we introduced the four occupation-level variables into the model.

Finally, to assess the generalizability of the meta-regression results, we divided the sample by the racial composition of the applicants, and following previous audit studies conducted in the United States, which showed race-based discrimination in initial hiring stages (Quillian et al. 2017), we conducted additional analyses considering race-based variations in callback rates for women. We ran two separate models: one exclusively for White applicants and the other for Black applicants (racial background was indicated by fictitious names) to tease out the divergent patterns of gender discrimination. The synthesis of all the field experiments in the data set also revealed evidence of racial differences for the Hispanic population. However, we did not include that analysis due to the small sample size.

## Findings

### Patterns at the Study Level

Similar to prior studies that examined the overall trend of gender discrimination in hiring (Galos and Coppock 2023), the meta-analysis of the full set of sample studies

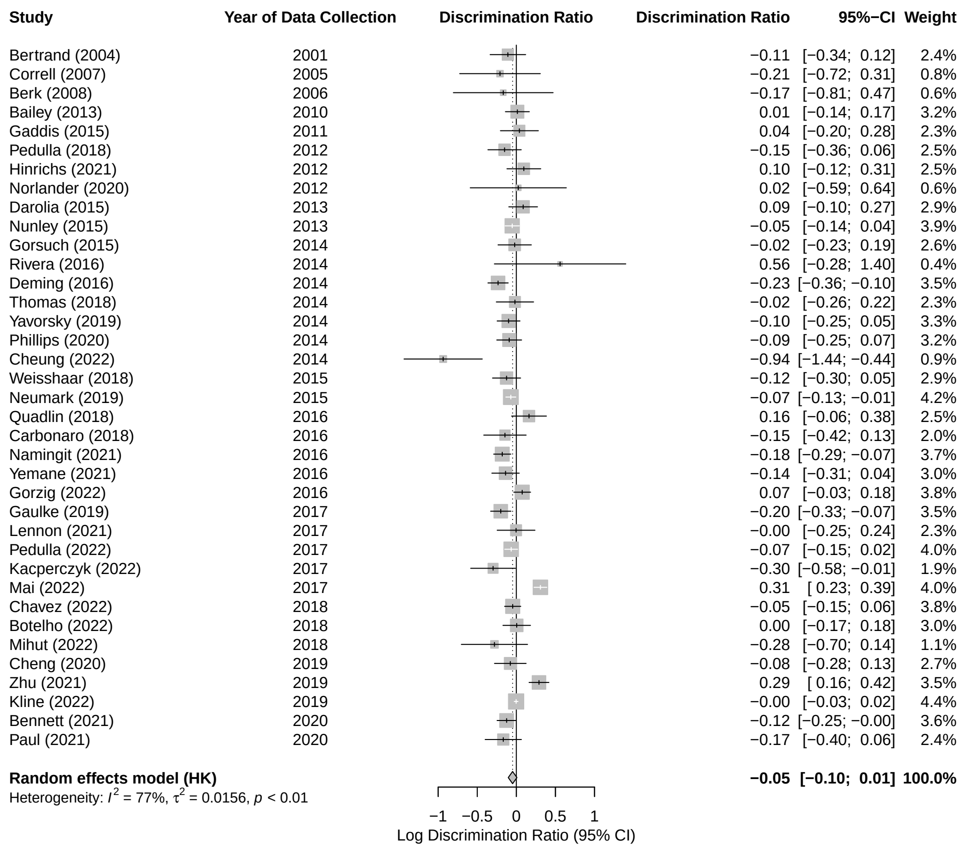


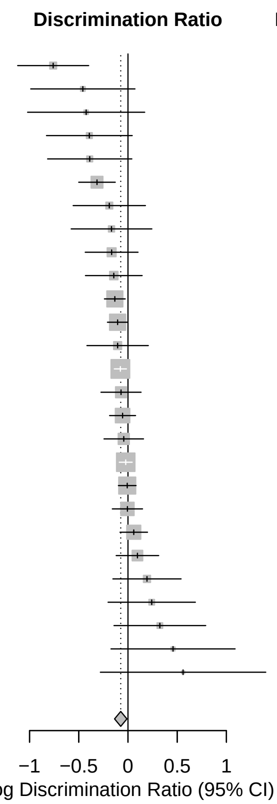

**Figure 3:** Gender bias at the study level. *Note:* The first author's name is used for co-authored studies. Negative ratios reflect a preference for female applicants and positive ratios reflect a preference for male applicants. The weights are assigned based on the inverse of the variance of each study, which incorporates both within-study sampling error and between-study heterogeneity. Studies with higher precision are given a higher weight in the analysis (Borenstein et al. 2010).

shows mixed results. At the individual study level, some studies identified gender discrimination, some did not, and others observed gender bias only in conjunction with variables such as race/ethnicity, credentials, or parenthood status. This pattern mirrors the lack of consensus in the literature on the directionality of gender bias due to varied institutional contexts and individual attributes.

Figure 3 shows the meta-analysis results at the aggregate study level. When analyzing the overall level of gender discrimination observed across all studies, the weighted average male/female discrimination ratio (transformed into a log risk ratio) is –0.05. The negative ratio indicates a preference for female applicants. However, the results are not statistically significant when studies are the unit of analysis (95 percent confidence interval: –0.10 and 0.01). In sum, the results suggest a mixed pattern of gender-biased hiring decisions, which we posit may stem from the diversity in institutional contexts and occupational characteristics across studies.

In addition, as expected, we found that because most audit studies were conducted after 2010, there are not sufficient empirical findings to assess trends or

| Occupation | Discrimination Ratio | 95%−CI | Weight |
|---|---|---|---|
| Nurse | −0.76 | [−1.12; −0.40] | 1.5% |
| Human Resources | −0.46 | [−0.99; 0.07] | 0.7% |
| Bookkeeping | −0.42 | [−1.02; 0.17] | 0.6% |
| Supervisor | −0.39 | [−0.83; 0.04] | 1.1% |
| Domestic Worker | −0.39 | [−0.82; 0.04] | 1.1% |
| Administrative | −0.31 | [−0.50; −0.13] | 4.4% |
| Clerk | −0.19 | [−0.56; 0.18] | 1.5% |
| Banking | −0.17 | [−0.58; 0.24] | 1.2% |
| Server | −0.17 | [−0.43; 0.10] | 2.5% |
| Finance | −0.14 | [−0.43; 0.14] | 2.2% |
| Customer Service | −0.13 | [−0.24; −0.03] | 8.1% |
| Marketing | −0.11 | [−0.21; −0.00] | 8.2% |
| Medical | −0.11 | [−0.42; 0.21] | 2.0% |
| Retail | −0.08 | [−0.14; −0.02] | 11.0% |
| Insurance | −0.07 | [−0.27; 0.13] | 3.8% |
| Information Technology | −0.05 | [−0.19; 0.08] | 6.6% |
| Analyst | −0.04 | [−0.24; 0.16] | 3.9% |
| Sales | −0.02 | [−0.09; 0.04] | 10.7% |
| Accounting | −0.01 | [−0.10; 0.08] | 9.1% |
| Management | −0.01 | [−0.16; 0.15] | 5.6% |
| Cook | 0.06 | [−0.08; 0.20] | 6.2% |
| Teacher | 0.10 | [−0.12; 0.31] | 3.5% |
| Maintenance | 0.19 | [−0.15; 0.54] | 1.6% |
| Parking Attendants | 0.24 | [−0.20; 0.68] | 1.0% |
| Manufacturing | 0.32 | [−0.14; 0.79] | 1.0% |
| Engineer | 0.46 | [−0.17; 1.09] | 0.5% |
| Lawyer | 0.56 | [−0.28; 1.40] | 0.3% |
| **Random effects model (HK)** | **−0.07** | **[−0.14; −0.01]** | **100.0%** |




Heterogeneity: $I^2$ = 51%, $\tau^2$ = 0.0043, $p$ < 0.01

**Figure 4:** Discrimination ratios by occupation. *Note:* Negative ratios reflect a preference for female applicants and positive ratios reflect a preference for male applicants. The weights are assigned based on the inverse of the variance of each study, which incorporates both within-study sampling error and between-study heterogeneity. Studies with higher precision are given a higher weight in the analysis (Borenstein et al. 2010).

changes in gender (non)discrimination over time. Using fieldwork years instead of publication years revealed a similar pattern: the meta-regression results for changes in the level of gender bias over fieldwork years are not statistically significant, regardless of whether applicant and occupation characteristics are controlled. Importantly, however, the meta-analysis found two significant patterns of gender bias: an occupational characteristic significantly predicted gender bias in job callback rates and findings varied by race. We explain these findings further in the following sections.

## *Patterns at the Occupation Level*

In our additional analysis, we added occupational characteristics into the model. The aggregate results for the occupation-level data show a statistically significant discrimination ratio of –0.07 ($p < 0.05$), signaling a preference for female over male applicants.

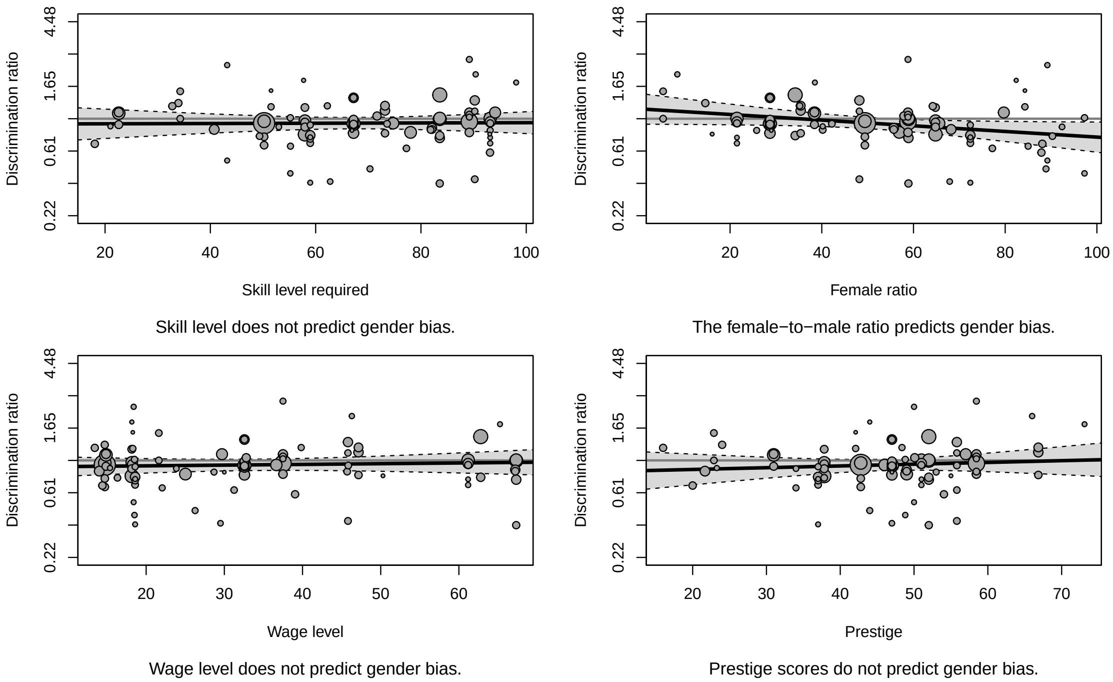



**Figure 5:** Relationships between occupational characteristics and gender bias.

However, looking into the data in more detail, Figure 4, which depicts a forest plot with results aggregated by the target occupations in the selected audit studies, reveals that the discrimination ratio varies across occupation groups. The discrimination ratios at the top of the forest plot are negative, indicating a preference for female applicants. These occupations include nursing, administrative positions, and clerk jobs, which tend to be dominated by women. In contrast, the discrimination ratios at the lower end of the forest plot are positive, indicating a preference for male applicants. These occupations include maintenance work, manufacturing, engineering, and law, which tend to be dominated by men. These results align with a body of research that has found an enduring pattern of gender-based occupational segregation (Breda et al. 2020; Charles and Bradley 2002). Such results suggest gender discrimination occurs across the horizontal sex segregation of occupations in both female- and male-dominant occupations. However, further analysis is needed to explain how gendered occupational segregation creates gender discrimination in hiring across occupations, potentially increasing gender-based inequality, including wage differentials. Thus, we further unpack occupational segregation by examining how the gender discrimination ratio differs by occupational characteristics.

The next steps in the analysis explored the associations between four occupational characteristics—skill level, gender ratio, prestige, and median hourly wage—and callback ratios. Figure 5 presents the results for each of these occupational characteristics. Plot points below 1 indicate a higher preference for female applicants, and plot points above 1 indicate a higher preference for male applicants. A table with the results of the meta-regression including occupational traits is provided in the online supplement.

Four noteworthy patterns emerge. First, skill level does not predict gender bias, as demonstrated in the top-left graph of Figure 5. In this graph, the slope of the line is flat across occupational skill levels, indicating that there is no difference in the callback ratio between male and female candidates based on skill level. Both meta-regression analyses controlling and not controlling for other occupational characteristics produced statistically insignificant results for the skill-level difference variable.

Second, as shown in the top-right figure, the female-to-male ratio predicts gender bias. The larger the proportion of female workers in an occupation, on average, the smaller the ratio of male-to-female callbacks. This result indicates that a higher proportion of female workers is associated with a higher likelihood that female job applicants will receive a callback. This variable is statistically significant ($p < 0.05$) in the meta-regression. Such a result aligns with prior meta-analysis findings that demonstrate how the gender composition of an occupation predicts gender bias in hiring at the international level.

Third, wage level does not predict gender bias. The results for wage level and gender discrimination in callbacks show a small upward-trending slope, but these results are not statistically significant; thus, we found no empirical support for an effect of wage levels in the data. Because the data on hourly wages often consist of lower- and higher-end hourly wages associated with each occupation, we conducted a robustness check on this finding by running the same model using the lower-end values of hourly wage estimates provided by the Bureau of Labor Statistics. As in the main results, wage level did not predict gender bias in hiring. In addition, we examined whether the median hourly wage predicted gender discrimination in callbacks when the female ratio of occupations was not controlled for. These variables still did not significantly predict gender bias.

Finally, prestige scores do not predict gender bias. The bottom-right panel illustrates the association between occupational prestige scores and gender discrimination. The line has an upward slope, starting below 1 and approaching 1 for occupations with higher prestige scores. This result means that female candidates are more likely to receive callbacks for lower-prestige jobs, and as job prestige increases, the preference for female applicant declines. However, despite this overall trend, the results of the meta-regression are not statistically significant. Additional analysis that does not control for the ratio of females in the occupation also shows that the prestige level does not predict gender bias.

In sum, based on field experiments conducted in the U.S. labor market, we find that only the ratio of females in the occupation predicts gender bias. Other occupational-level characteristics measured by the average skill level, prestige score, and median hourly wage do not predict gender discrimination in callbacks. Such a result holds for both meta-regressions when the ratios of females in occupations are controlled for and are not controlled for.

## *The Intersection of Race and Gender*

Given prior studies that highlight the importance of adopting an intersectional perspective (Choo and Ferree 2010; Crenshaw 1989), we explored whether patterns of gender discrimination varied across White and Black applicants. We first examined

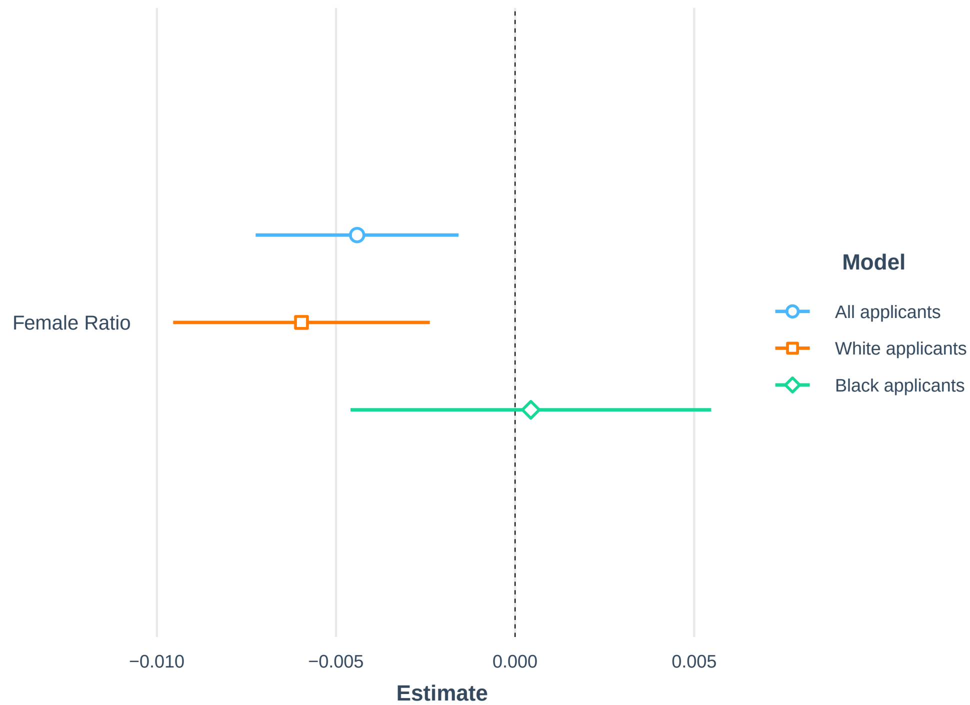



**Figure 6:** Percent female estimates of gender bias by race.

these applicant groups at the study level. The results for both White and Black applicant groups were not statistically significant. The White applicant group sample yielded a male/female discrimination ratio of –0.03 (95 percent confidence interval: –0.10 and 0.03), whereas the Black applicant group sample was 0.01 (95 percent confidence interval: –0.10 and 0.12). As we had seen for the pooled analysis, we expect that this result to be caused by diversity in institutional contexts and occupational characteristics.

Next, we conducted a random-effect meta-analysis at the occupation level to examine whether there were differences in gender discrimination by race. As for our pooled analysis, we find that in more female-dominant occupations, White women received more callbacks than their male counterparts, whereas Black women did not experience a similar gender preference. Figure 6 graphs the coefficients for the percent female variable across three meta-regression models, each varying only by sample: all applicants, White applicants, and Black applicants.

As shown in Figure 6, a statistically significant outcome was observed for White applicants in contrast to the non-significant results for Black applicants. This result supports the assertion that the intersection of race and gender influences the magnitude and direction of gender bias. In other words, during the first stage of screening, as observed through audit studies conducted in the United States, the preference for female job candidates applies to White applicants but not to Black applicants. Thus, it is crucial to consider the race/ethnicity of job candidates when examining gender (non)discrimination.

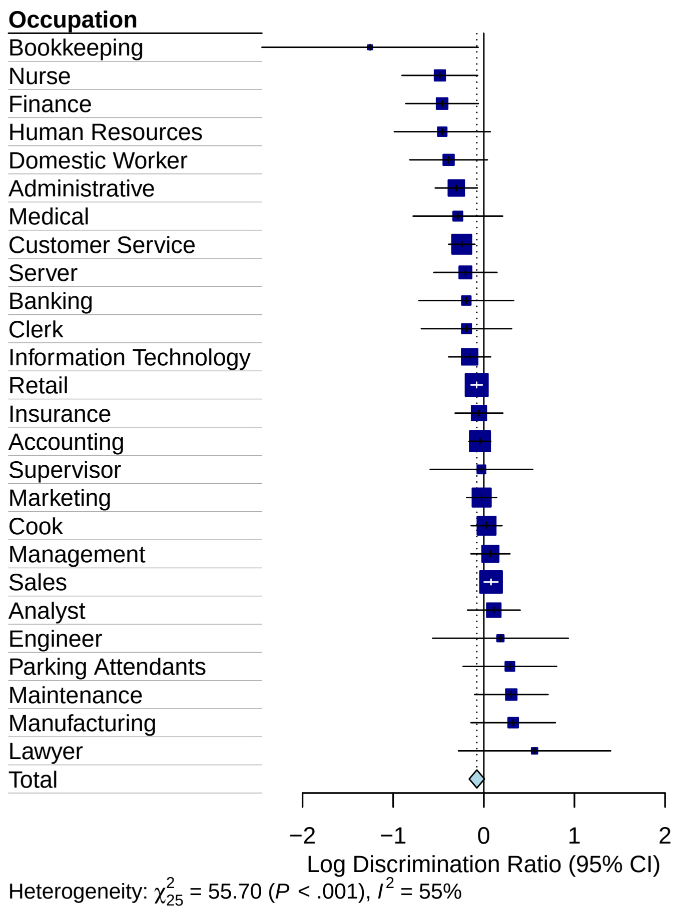



**Figure 7:** Gender discrimination among White applicants across occupations. *Note:* Negative ratios reflect a preference for female applicants and positive ratios reflect a preference for male applicants.

The forest plot in Figure 7 presents the gender discrimination ratios for White applicants by occupation. As in the pooled data, there is a negative-to-positive trend across occupations as they shift from relatively more female to male-dominated. Occupations at the top of the forest plot (e.g., nursing and administrative positions) tend to be female-dominated. The log discrimination ratios for these occupations are mostly negative, indicating that resumes from women received more callbacks than resumes from men. Occupations at the bottom of the plot (e.g., maintenance work and manufacturing) tend to be more male-dominated. The log discrimination ratios for these occupations are mostly positive, indicating that resumes from men received more callbacks than resumes from women. Despite this occupation-specific variation, the results from the occupational-level data for White applicants yield a statistically significant ($p < 0.1$) discrimination ratio of –0.08. Thus, a weighted average of the occupation-specific-effect sizes among White job candidates across the different studies signals a preference for female job applicants.

When focusing solely on Black applicants, a noteworthy disparity emerged. Figure 8 shows the parallel forest plot for Black applicants. Here, the discrimination

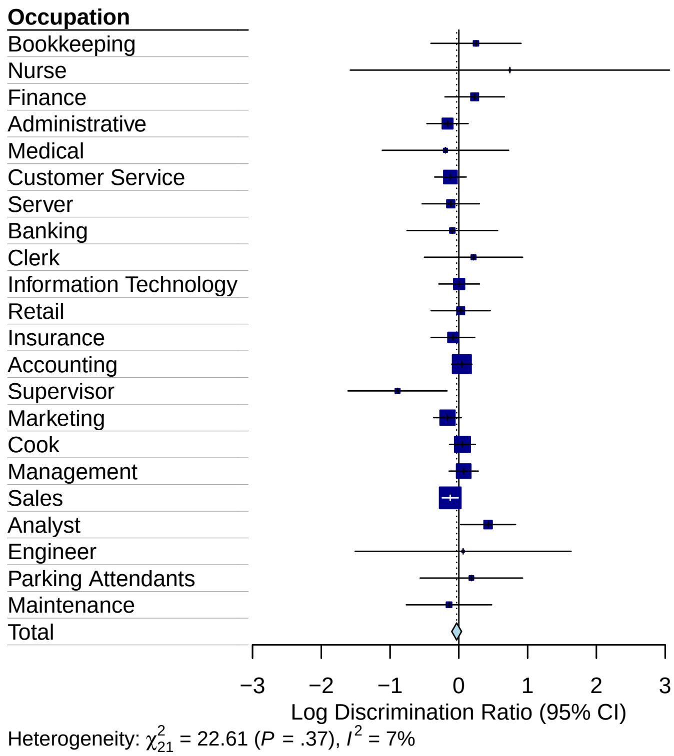



**Figure 8:** Gender discrimination among Black applicants across occupations. *Note:* Negative ratios reflect a preference for female applicants and positive ratios reflect a preference for male applicants.

ratio calculated from the occupational-level data for Black applicants (–0.03) is not statistically significant. Furthermore, the trend across occupations is less clear-cut. Some more female-dominated occupations have positive discrimination ratios (male candidates received more callback offers than female candidates), whereas some more male-dominated occupations have negative discrimination ratios (women received more callbacks than men). These nuanced dynamics highlight that the results of existing field experiments on gender bias generalize predominantly to White female applicants and Black female applicants follow somewhat different patterns.

It is true that the high callback rates for female applicants in female-dominant occupations are not found in experiments that have Black female applicants. However, caution is warranted when interpreting this pattern. The wide confidence intervals suggest that it is difficult to draw firm conclusions about the relationship between occupational characteristics and gender discrimination patterns for Black applicants. Overall, the results do not answer the question of whether the gender composition of an occupation predicting gender bias pattern only applies to the White female group. Based on our data and the available evidence, this pattern is likely to be only found among White women. However, clearly explaining the hiring patterns for Black applicants is not possible based on the limited, existing evidence. In the end, while differences in gender discrimination by race highlight the importance of considering the intersection of race and gender in studies of hiring discrimination, caution is warranted in interpreting these results.

In sum, the findings indicate that occupational characteristics, specifically the gender composition of an occupation, influence gender discrimination in hiring—the more female-dominated an occupation, the more female applicants are preferred. However, this influence varies by race, such that White female applicants in more female-dominated occupations benefit from a preference in hiring, whereas Black female applicants in female-dominated occupations do not experience a similar benefit. Further research is needed to unpack these complex dynamics.

### *Further Analysis*

We conducted several robustness checks. As shown in Figure 2, some of the sample studies on discrimination focus solely on gender, some focus on gender in conjunction with other characteristics, and some focus on other characteristics but control for gender. The full data set includes all studies—those that randomized or varied the gender of job applicants for each employer and those that blocked gender, meaning that they sent resumes from same-gendered job applicants (i.e., only female applicants or only male applicants) to each employer. We conducted a robustness check by rerunning the analyses with only those studies that randomized or varied the gender of job applicants. The overall results remained the same.

We also ran a model with an interaction term between race and gender to formally test for differences in the relationship between the gender composition of occupation and hiring outcomes across Black and White applicants. The interaction term was statistically significant at the $p < 0.05$ level. This finding reinforces our main analysis, indicating a meaningful difference in the relationship between the occupational gender composition and hiring outcomes across these racial groups (result tables for our robustness checks are included in the online supplement).

## Discussion

Based on audit studies from the United States, some scholars have found no evidence of gender-based discrimination, whereas others have emphasized that gender-based discrimination persists. To unpack these inconclusive results, we conducted a meta-analysis of field experiments analyzing differences in callback rates for men and women. Reflecting the mixture of positive forecasts and concern about stalled progress toward gender equality (England et al. 2020) in the U.S. workforce, the initial meta-analysis of the full sample produced mixed results.

A series of meta-analyses examined in this study identified some of the factors responsible for these mixed results. Specifically, there are two main findings: the gender composition of an occupation and an applicant's race predict discrimination in callback rates. The results show that at the start of the hiring process, the gender composition of an occupation predicts bias. When a woman applies for a job in a female-dominant occupational field, she is more likely than her male counterpart to receive a callback. However, this pattern differs by race—in female-dominated occupations, White women applicants received more callbacks than White men, but Black women did not experience a parallel benefit. Both of these findings align with existing work on the persistence of gender segregation (Zhu and Grusky 2022) and racial inequality (Quillian et al. 2017) in the U.S. labor market.

This study has two main implications. First, our findings show almost no statistically significant effect of occupational-level characteristics on hiring discrimination. In contrast to our predictions that occupation-level status predicts levels of hiring discrimination based on gender, we find that occupational status—measured by wage, skill set, and symbolic prestige—does not correlate with gender bias in hiring.

However, we do find a gender compositional effect, where the ratio of female employees predicts gender bias in the initial hiring stage. Such findings have been theorized by models of tokenism and statistical discrimination. Our findings further support these theories and are aligned with empirical studies that argue that persisting sex segregation in the workforce shapes and is shaped by gender-based stereotypes (Breda et al. 2020; Charles and Bradley 2002). The finding that gender composition predicts hiring bias can relate to other outcomes relevant to gender inequality. One important example is what scholars call the power of having control over the work schedule (Clawson and Gerstel 2014). Job qualities, including scheduling control, are important characteristics of a job that shape employment pathways that diverge greatly based on gender (Damaske 2011; Damaske and Frech 2016). Thus, future studies should continue highlighting social problems related to stereotype reinforcement and explaining mechanisms that link hiring and gender inequality. Drawing upon gender, work, and family literature, new studies have room to contribute to the literature on hiring discrimination and gender inequality by specifying social problems related to gender-based stereotype reinforcement.

Second, our finding that there is an interplay between occupational gender composition and race parallels the results of ongoing studies that investigate both race and gender to explain hiring discrimination (Bertrand and Mullainathan 2004; Chavez, Weisshaar, and Cabello-Hutt 2022; Deming et al. 2016). This indicates that theories of employer biases based on a single category of gender do not sufficiently explain patterns of hiring demonstrated through field experiments. Only when the multidimensionality of gender and racial stereotypes is considered can we adequately assess the changes and continuities of hiring inequality at the initial screening stage. Thus, new experiments will benefit—theoretically and empirically—if they use intersectionality as the guiding framework.

This study has certain limitations. The data set is composed of field experiments, and because of this, the results of meta-analyses can only reflect the results of the field experiments that take place at the initial point of the screening stage. Most audit studies involve sending fictitious resumes for consideration for entry-level positions but do not follow up on the callbacks received (Gaddis 2018). Although elements of the typical design of field experiments (focusing on entry-level positions or mixing fictitious resumes with real applicants) reduce the risk that an employer will become aware of the experiment, the narrow focus on the screening stage limits the scope of the findings of such studies. Because most field experiments do not include later stages of the hiring process, meta-analyses using the experiments cannot produce information on the patterns of actual hiring outcomes. Limited studies that do examine the final point of the screening stage—the job offer—present more substantial discrimination than when only examining the callback stage (Quillian, Lee, and Oliver 2020). Therefore, the patterns observed at the initial screening stage may underrepresent the extent of hiring discrimination throughout the process. Nonetheless, the initial screening process has implications for hiring discrimination because of its impact on the later stages of the hiring process. Thus, more field experiments need to be conducted, even if the field experiments may have to take place at the initial stage of hiring.

Reflecting technological development, new studies conducting audit studies must reflect newer screening processes. With the dramatic development of artificial intelligence in the era of Big Data, there has been a rapid increase in the utilization of new technologies for applicant selection. Research conducted by Jobscan (Myers 2023), a recruitment solutions company, showed that 98.8 percent of Fortune 500 companies use applicant tracking systems—software that makes hiring decisions by using algorithms to automatically filter submitted resumes. Recent studies have started to investigate the fairness of such algorithm-based hiring (Raghavan and Barocas 2019), but thus far, there are few experimental studies on hiring discrimination—based on race, gender, sexual orientation, and parental status—resulting from AI-based screening processes (for more discussion on this gap in the literature, see Pedulla 2018: 1498–1499). It will be critical for future research on hiring discrimination to investigate the details of gender bias in the context of technological innovation.

Finally, this study has policy implications. Our findings about intersectionality and the multidimensionality of gender and race in the hiring process send an important policy message, particularly related to equal employment laws. Crenshaw (1989: 139) first coined the term intersectionality based on her evidence from legal cases and her argument that Black women are legally, policywise, and theoretically erased when scholars and policymakers depend on the "single-axis analysis that distorts" the experiences of Black women. In particular, policies that aim to solve issues of gender- and race-based discrimination (separately) often result in reinforcing existing marginalizations (i.e., Black women) who will fall into the crack that is neither noticed nor protected by the policies. Thus, equal employment law should consider the fact that hiring processes must consider diversity from the viewpoint of intersectionality and ground the law from the starting point that discriminations are experienced in a compound way rather than resulting from a single axis of inequality.

## Notes

1 The remaining 24 percent of the studies did not manipulate gender but did control for gender by sending resumes from either only female applicants or only male applicants to each job opening. We later conducted robustness checks related to this issue.

2 Within 37 studies, 12 studies did not contain data on occupations, thus we emailed the study authors and requested breakdowns of the specific jobs targeted in the field experiments. The final occupational-level data set included data points from 23 studies.

3 There were two studies that described the occupation category to which they sent resumes as a whole. For example, business/research/data analysts can be categorized

into three separate occupation groups. However, based on the study description, we used "management analyst" information as crosswalk. Two studies were coded using this method. In addition, data for hourly median wage were missing for one occupation (elementary and middle school teachers), so we used an average of the median hourly wage of occupations that have a similar median annual wage to said occupation.

4 The occupational skill-level variable is an index of occupational education, which is a measure of the percentage of occupational incumbents' education level (one or more years of post-secondary education completed) created using the 1989 General Social Survey and the 1990 U.S. Census. Following Yu and Kuo (2017), we use this as a measure of the overall skill levels of workers within occupations.

## References


Arrow, Kenneth J. 1973. "The Theory of Discrimination." Pp. 3–33 in *Discrimination in Labor Markets*. Princeton University Press.

Bertrand, Marianne and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94(4):991–1013. https://doi.org/10.1257/0002828042002561.

Bielby, William T. and James N. Baron. 1986. "Men and Women at Work: Sex Segregation and Statistical Discrimination." *American Journal of Sociology* 91(4):759–99. https://doi.org/10.1086/228350.

Borenstein, Michael, Larry V. Hedges, Julian P. T. Higgins, and Hannah R. Rothstein. 2010. "A Basic Introduction to Fixed-Effect and Random-Effects Models for Meta-Analysis." *Research Synthesis Methods* 1(2):97–111. https://doi.org/10.1002/jrsm.12.

Botelho, Tristan L., and Melody Chang. 2022. "The Evaluation of Founder Failure and Success by Hiring Firms: A Field Experiment." *Organization Science* 34(1):484–508. https://doi.org/10.1287/orsc.2022.1592.

Breda, Thomas, Elyès Jouini, Clotilde Napp, and Georgia Thebault. 2020. "Gender Stereotypes Can Explain the Gender-Equality Paradox." *Proceedings of the National Academy of Sciences* 117(49):31063–69. https://doi.org/10.1073/pnas.2008704117.

Budig, Michelle J. 2002. "Male Advantage and the Gender Composition of Jobs: Who Rides the Glass Escalator?" *Social Problems* 49(2):258–77. https://doi.org/10.1525/sp.2002.49.2.258.

Budig, Michelle J. and Melissa J. Hodges. 2010. "Differences in Disadvantage: Variation in the Motherhood Penalty across White Women's Earnings Distribution." *American Sociological Review* 75(5):705–28. https://doi.org/10.1177/0003122410381593.

Campero, Santiago and Roberto M. Fernandez. 2019. "Gender Composition of Labor Queues and Gender Disparities in Hiring." *Social Forces* 97(4):1487–516. https://doi.org/10.1093/sf/soy097.

Charles, Maria and Karen Bradley. 2002. "Equal but Separate? A Cross-National Study of Sex Segregation in Higher Education." *American Sociological Review* 67(4):573. https://doi.org/10.2307/3088946.

Charles, Maria and David B. Grusky. 2004. *Occupational Ghettos: The Worldwide Segregation of Women and Men*. Stanford, CA: Stanford University Press.

Chavez, Koji, Katherine Weisshaar, and Tania Cabello-Hutt. 2022. "Gender and Racial Discrimination in Hiring Before and During the COVID-19 Pandemic: Evidence from a Field Experiment of Accountants, 2018–2020." *Work and Occupations* 49(3):275–315. https://doi.org/10.1177/07308884221094539.

Choo, Hae Yeon and Myra Marx Ferree. 2010. "Practicing Intersectionality in Sociological Research: A Critical Analysis of Inclusions, Interactions, and Institutions in the Study of Inequalities." *Sociological Theory* 28(2):129–49. https://doi.org/10.1111/j.1467-9558.2010.01370.x.

Clawson, Dan, and Naomi Gerstel. 2014. *Unequal Time: Gender, Class, and Family in Employment Schedules*. New York: Russell Sage Foundation.

Correll, Shelley J., Stephen Benard, and In Paik. 2007. "Getting a Job: Is There a Motherhood Penalty?" *American Journal of Sociology* 112(5):1297–339. https://doi.org/10.1086/511799.

Cotter, David, Joan Hermsen, and Paula England. 2008. "Moms and Jobs: Trends in Mothers' Employment and Which Mothers Stay Home." Pp. 379–86 in *American Families: A Multicultural Reader*. New York: Routledge.

Cotter, David A., Joan M. Hermsen, and Reeve Vanneman. 2004. "Gender Inequality at Work." Pp. 107–37 in *The American people: Census 2000*. New York: Russell Sage Foundation.

Crenshaw, Kimberlé. 1989. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Anti-Discrimination Doctrine, Feminist Theory and Anti-Racist Politics." Pp. 139–167 in *University of Chicago Legal Forum*.

Damaske, Sarah. 2011. *For the Family? How Class and Gender Shape Women's Work*. New York: Oxford University Press.

Damaske, Sarah and Adrianne Frech. 2016. "Women's Work Pathways across the Life Course." *Demography* 53(2):365–91. https://doi.org/10.1007/s13524-016-0464-z.

Deming, David J., Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F. Katz. 2016. "The Value of Postsecondary Credentials in the Labor Market: An Experimental Study." *American Economic Review* 106(3):778–806. https://doi.org/10.1257/aer.20141757.

Deterding, Nicole M. and David S. Pedulla. 2016. "Educational Authority in the "'Open Door"' Marketplace: Labor Market Consequences of For-Profit, Nonprofit, and Fictional Educational Credentials." *Sociology of Education* 89(3):155–70. https://doi.org/10.1177/0038040716652455.

Eaton, Asia A., Jessica F. Saunders, Ryan K. Jacobson, and Keon West. 2020. "How Gender and Race Stereotypes Impact the Advancement of Scholars in STEM: Professors' Biased Evaluations of Physics and Biology Post-Doctoral Candidates." *Sex Roles* 82(3–4):127–41. https://doi.org/10.1007/s11199-019-01052-w.

England, Paula, Andrew Levine, and Emma Mishel. 2020. "Progress toward Gender Equality in the United States Has Slowed or Stalled." *Proceedings of the National Academy of Sciences* 117(13):6990–97. https://doi.org/10.1073/pnas.1918891117.

Flage, Alexandre. 2019. "Discrimination against Gays and Lesbians in Hiring Decisions: A Meta-Analysis." *International Journal of Manpower* 41(6):671–91. https://doi.org/10.1108/IJM-08-2018-0239.

Frederick, Carl. 2010. *A Crosswalk for Using Pre-2000 Occupational Status and Prestige Codes with Post-2000 Occupational Codes. 2010–3*. Madison, WI: Center for Demography and Ecology, University of Wisconsin-Madison.

Gaddis, S. Michael, ed. 2018. *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Cham: Springer International Publishing.

Gaddis, S. Michael, Edvard Larsen, Charles Crabtree, and John Holbein. 2021. "Discrimination Against Black and Hispanic Americans Is Highest in Hiring and Housing Contexts: A Meta-Analysis of Correspondence Audits." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3975770.

Galos, Diana Roxana and Alexander Coppock. 2023. "Gender Composition Predicts Gender Bias: A Meta-Reanalysis of Hiring Discrimination Audit Experiments." *Science Advances* 9(18):eade7979. https://doi.org/10.1126/sciadv.ade7979.

Ganzeboom, Harry B. G. and Donald J. Treiman. 2019. "International Stratification and Mobility File: Conversion Tools."

Hauser, Robert M. and John Robert Warren. 1997. "4. Socioeconomic Indexes for Occupations: A Review, Update, and Critique." *Sociological Methodology* 27(1):177–298. https://doi.org/10.1111/1467-9531.271028.

Hirsh, C. Elizabeth. 2009. "The Strength of Weak Enforcement: The Impact of Discrimination Charges, Legal Environments, and Organizational Conditions on Workplace Segregation." *American Sociological Review* 74(2):245–71. https://doi.org/10.1177/000312240907400205.

Huffman, Matt L. 2004. "Gender Inequality across Local Wage Hierarchies." *Work and Occupations* 31(3):323–44. https://doi.org/10.1177/0730888404266384.

Kanter, Rosabeth Moss. 1977. "Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women." *American Journal of Sociology* 82(5):965–90. https://doi.org/10.1086/226425.

Killewald, Alexandra and Jonathan Bearak. 2014. "Is the Motherhood Penalty Larger for Low-Wage Women? A Comment on Quantile Regression." *American Sociological Review* 79(2):350–57. https://doi.org/10.1177/0003122414524574.

Kreisberg, Nicole and Nathan Wilmers. 2022. "Blacklist or Short List: Do Employers Discriminate against Union Supporter Job Applicants?" *ILR Review* 75(4):943–73. https://doi.org/10.1177/00197939211036444.

Lamont, Michèle, Graziella M. Silva, Jessica Welburn, Joshua Guetzkow, Nissim Mizrachi, Hanna Herzog, and Elisa Reis. 2016. *Getting Respect: Responding to Stigma and Discrimination in the United States, Brazil, and Israel*. Princeton, New Jersey: Princeton University Press. https://doi.org/10.2307/j.ctv346qr9.

Lippens, Louis, Siel Vermeiren, and Stijn Baert. 2023. "The State of Hiring Discrimination: A Meta-Analysis of (Almost) All Recent Correspondence Experiments." *European Economic Review* 151:104315. https://doi.org/10.1016/j.euroecorev.2022.104315.

Mai, Quan D. 2023. "The Demographic Context of Hiring Discrimination: Evidence from a Field Experiment in 50 Metropolitan Statistical Areas." *Work and Occupations* 50(4):463–98. https://doi.org/10.1177/07308884221134470.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27(1):415–44. https://doi.org/10.1146/annurev.soc.27.1.415.

Mihut, Georgiana. 2022. "Does University Prestige Lead to Discrimination in the Labor Market? Evidence from a Labor Market Field Experiment in Three Countries." *Studies in Higher Education* 47(6):1227–42. https://doi.org/10.1080/03075079.2020.1870949.

Myers, Sydney. 2023. "2023 Applicant Tracking System (ATS) Usage Report: Key Shifts and Strategies for Job Seekers." Retrieved (https://www.jobscan.co/blog/fortune-500-use-applicant-tracking-systems/).

Neumark, D., R. J. Bank, and K. D. Van Nort. 1996. "Sex Discrimination in Restaurant Hiring: An Audit Study." *The Quarterly Journal of Economics* 111(3):915–41. https://doi.org/10.2307/2946676.

Nunley, John M., Adam Pugh, Nicholas Romero, and R. Alan Seals. 2015. "Racial Discrimination in the Labor Market for Recent College Graduates: Evidence from a Field Experiment." *The B.E. Journal of Economic Analysis & Policy* 15(3):1093–125. https://doi.org/10.1515/bejeap-2014-0082.

Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108(5):937–75. https://doi.org/10.1086/374403.

Pedulla, David S. 2018. "How Race and Unemployment Shape Labor Market Opportunities: Additive, Amplified, or Muted Effects?" *Social Forces* 96(4):1477–506. https://doi.org/10.1093/sf/soy002.

Quadlin, Natasha. 2018. "The Mark of a Woman's Record: Gender and Academic Performance in Hiring." *American Sociological Review* 83(2):331–60. https://doi.org/10.1177/0003122418762291.

Quillian, Lincoln, John J. Lee, and Mariana Oliver. 2020. "Evidence from Field Experiments in Hiring Shows Substantial Additional Racial Discrimination after the Callback." *Social Forces* 99(2):732–59. https://doi.org/10.1093/sf/soaa026.

Quillian, Lincoln, Devah Pager, Ole Hexel, and Arnfinn H. Midtbøen. 2017. "Meta-Analysis of Field Experiments Shows No Change in Racial Discrimination in Hiring over Time." *Proceedings of the National Academy of Sciences* 114(41):10870–75. https://doi.org/10.1073/pnas.1706255114.

Raghavan, Manish, and Solon Barocas. 2019. "Challenges for Mitigating Bias in Algorithmic Hiring." Retrieved (https://brookings.edu/articles/challenges-for-mitigating-bias-in-algorithmic-hiring/).

Ridgeway, Cecilia L. 2014. "Why Status Matters for Inequality." *American Sociological Review* 79(1):1–16. https://doi.org/10.1177/0003122413515997.

Ridgeway, Cecilia L. 2019. *Status: Why Is It Everywhere? Why Does It Matter?* New York: Russell Sage Foundation.

Rivera, Lauren A., and András Tilcsik. 2016. "Class Advantage, Commitment Penalty: The Gendered Effect of Social Class Signals in an Elite Labor Market." *American Sociological Review* 81(6):1097–131. https://doi.org/10.1177/0003122416668154.

Schaerer, Michael, Christilene Du Plessis, My Hoang Bao Nguyen, Robbie C. M. Van Aert, Leo Tiokhin, Daniël Lakens, Elena Giulia Clemente, Thomas Pfeiffer, Anna Dreber, Magnus Johannesson, Cory J. Clark, and Eric Luis Uhlmann. 2023. "On the Trajectory of Discrimination: A Meta-Analysis and Forecasting Survey Capturing 44 Years of Field Experiments on Gender and Hiring Decisions." *Organizational Behavior and Human Decision Processes* 179:104280. https://doi.org/10.1016/j.obhdp.2023.104280.

Thomas, Kyla. 2018. "The Labor Market Value of Taste: An Experimental Study of Class Bias in U.S. Employment." *Sociological Science* 5:562–95. https://doi.org/10.15195/v5.a24.

Valentino, Lauren. 2020. "The Segregation Premium: How Gender Shapes the Symbolic Valuation Process of Occupational Prestige Judgments." *Social Forces* 99(1):31–58. https://doi.org/10.1093/sf/soz145.

Weisshaar, Katherine. 2018. "From Opt Out to Blocked Out: The Challenges for Labor Market Re-Entry after Family-Related Employment Lapses." *American Sociological Review* 83(1):34–60. https://doi.org/10.1177/0003122417752355.

Yu, Wei-hsin and Janet Chen-Lan Kuo. 2017b. “The Motherhood Wage Penalty by Work Conditions: How Do Occupational Characteristics Hinder or Empower Mothers?” *American Sociological Review* 82(4):744–69. https://doi.org/10.1177/0003122417712729.

Zhu, Ling and David B. Grusky. 2022. “The Intergenerational Sources of the U-Turn in Gender Segregation.” *Proceedings of the National Academy of Sciences* 119(32):e2121439119. https://doi.org/10.1073/pnas.2121439119.

**Acknowledgments:** We wish to thank Kangwook Lee, Ziqian Lin, Yuchen Zeng, and the members of the University of Wisconsin-Madison Gender Seminar (FemSem) for their helpful feedback. We also appreciate comments provided by the editors and reviewers of the Sociological Science. Support for this research was provided by the University of Wisconsin-Madison Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation (grant #AAI9746).

**So Yun Park:** Department of Sociology, University of Wisconsin–Madison. E-mail: spark553@wisc.edu
**Eunsil Oh:** Department of Sociology, University of Wisconsin–Madison.
E-mail: eunsil.oh@wisc.edu