

The University of Chicago

**Inspecting Job Hiring Discrimination Against  
Men Entering Female-Dominated Occupations**

Cole Weaver

August 2023

*A paper submitted in partial fulfillment of the requirements for the Master of Arts degree  
in the Master of Arts Program in the Social Sciences*

Faculty advisor and preceptor: Murilo Ramos

# Inspecting Job Hiring Discrimination Against Men Entering Female-Dominated Occupations

Cole Weaver

August 2023

## **Abstract**

This thesis uses an audit study field experiment to examine potential hiring discrimination against male job candidates entering traditionally female-dominated fields. The experiment revealed evidence of discrimination *against* male candidates in some female-dominated fields, but *for* male candidates in others. Notably, the descriptive findings suggest that occupations emphasizing manual labor may exhibit a preference for male candidates. However, it is important to note that these descriptive findings are purely observational in nature and require further analysis to establish causal relationships. Furthermore, the study contributes to the literature by shedding light on the nuanced heterogeneity within specific occupations regarding responses to male candidates. While the research does not delve into the underlying causes or mechanisms driving such discrimination, it emphasizes the presence of bias in hiring practices. These findings underscore the need for further investigation into the complex dynamics influencing gender-based employment disparities in traditionally female-dominated sectors.

## Introduction

Over the past few decades, the American male-female dynamic has undergone significant changes. Women are currently achieving greater success relative to men in various domains such as college attendance and wages ([Fry, 2022]). One important aspect of this new gender landscape is that many traditionally male jobs - such as factory work - are on the decline, while many traditionally female jobs - including nursing, teaching, and secretarial work - are on the rise. Despite these factors, many men have not yet transitioned from dying to growing industries ([Ngai and Petrongolo, 2017]; [Delfino, 2021]).

While it is hard to predict the job market, these trends indicate that in the near future men might more commonly enter fields where they are the minority. As a result, relatively new labor market frictions and talent misallocations may appear in the form of job hiring discrimination against men ([Hsieh et al., 2019]; [Schaede and Mankki, 2022]). Thus, in order to peek into this potential future, this thesis uses a field experiment to broadly inspect potential hiring discrimination against male job candidates entering fields that are traditionally female-dominated.

Perhaps surprisingly, this hiring bias against men entering “women’s jobs” has already been established through previous economics and social science research ([Kline et al., 2022]; [Arceo-Gomez and Campos-Vazquez, 2014]; [Yavorsky, 2019]; [Booth and Leigh, 2010]; [Zhou et al., 2013]; [Berson, 2012]; [Albert et al., 2011]; [Birkelund et al., 2019]; [Adamovic and Leibbrandt, 2023]; [Carlsson, 2011]; [Riach and Rich, 2006]). However, these studies usually have a narrow focus, only inspecting one or two specific jobs such as ‘secretary’ or ‘retail worker’ ([Lang and Spitzer, 2020]). Or, if the experiment studies a broad range of jobs, it does not focus on female-dominated jobs and thus is not well suited for a detailed examination of hiring biases against men ([Kline et al., 2022]).

As a natural extension of the existing literature, this paper examines a wide range of female-dominated occupations, at various gender ratio levels (ex: 95% women, 80% women, etc.). By studying these numerous fields and industries, this field experiment provides a broader picture than has been previously available. Optimistically, the information from this study might help men sort into more welcoming jobs in the future - and particularly discriminatory occupations might warrant further academic or governmental investigation. In addition, by studying fields at various levels of female participation rates, we are able to take a descriptive look at whether there is a drop off in discrimination after a certain threshold of gender ratio balance.

The results of this study show evidence of discrimination against male job candidates in some female-dominated fields, and for male candidates in others. In particular, the level of manual labor involved in the job seems to be an important factor, with more manual labor correlating with a higher preference for men. In contrast, an occupation’s gender ratio (i.e., percentage female) seems to not matter.

The within-occupation heterogeneity angle of this study unveils the diversity of responses towards male candidates within broad occupational categories. For example, while the term “nurse” may conjure up a single image in people’s minds, it encompasses a wide range of roles and responsibilities. The findings of this experiment demonstrate that employer preferences vary significantly even within specific fields. This serves as a valuable reminder to social scientists to not make unnecessary assumptions about homogeneity. Furthermore, by providing a detailed examination of sub-occupations, this study offers insights that can contribute to future research endeavors, such as identifying specific jobs to target when trying to identify underlying mechanisms of any discrimination that occurs.

The experiment for this paper was performed in the style of an audit / correspondence study, and consisted of creating fictitious resumes, randomly assigning gender to these resumes using gendered names, and then sending the resumes to employers. Other characteristics of the resume, such as work experience and GPA, were also randomly assigned. The important outcome variable was the response rate for resumes, i.e., whether an employer called or emailed in response to the application, which I tracked using CallRail (an online call tracking service) and various email accounts.

The statistical analysis for this experiment involves simply comparing mean response rates across male and female resumes. Since all other variables are controlled for on the resume, any resulting difference in response rates can be interpreted as causal. By definition, such a difference found using this method would be classified as discrimination in the eyes of American law. In comparison, Economists often go one step further and try to distinguish between statistical and taste-based sources of discrimination. However, in this experiment I do not aim to identify the core cause or underlying mechanisms of any discrimination observed.

The remainder of this thesis will be organized as follows. Section 2 includes a brief literature review to provide background information and context for the study. Section 3 describes the experimental design in detail, outlining the procedures and methods used to carry out the audit study. Section 4 presents the results of the experiment. Section 5 concludes.

## Literature review

*Defining discrimination* “Discrimination” has various definitions and interpretations. In American law, discrimination in the job hiring process can be defined as basing hiring decisions on presumptions about an applicant due to their sex, race, or age ([Kline et al., 2022]; [Kline and Walters, 2021]). In economics, discrimination is seen as differential treatment based on race or gender between two otherwise identical workers or job applicants ([Lang and Spitzer, 2020]; [Heckman, 1998]; [Guryan and Charles, 2013]). Economists also commonly differentiate taste-based and statistical discrimination - i.e., discriminatory behavior stemming from

prejudice, and discriminatory behavior stemming from valid statistical inference ([Lang and Spitzer, 2020]).

*Detecting discrimination* Prior economics research has used many methods to detect discrimination. The three most common techniques used for causally identifying discrimination in the job hiring process are regressions, audit studies, and correspondence studies ([Guryan and Charles, 2013])<sup>1</sup>. The earliest attempts in this strand of literature involved running regressions using a selection on observables approach, though these studies had an inherent omitted variable bias problem.

In response to this problem, audit studies became the next popular approach. This type of study involved sending trained applicants of different races and genders to interviews, and recording the different response rates. However, audit studies could not truly control for all relevant variables, such as the applicant actors' backgrounds. Thus, correspondence studies developed as a natural next step.

Correspondence studies involve creating fictitious application materials, randomly assigning the relevant characteristic(s), and then recording and comparing mean response rates. Resumes used to detect job hiring discrimination are the most common application of correspondence studies, though other applications exist as well ([Neyt et al., 2019]).

Note that correspondence studies are limited in nature. This specific experimental design can only detect an intermediate measurement of discrimination: contact rate disparities ([Heckman, 1998]; [Guryan and Charles, 2013]; [Bohren et al., 2022]). This does not take into account the various other links in the chain such as interview results, starting salary offers, and promotion differences - all of which are important for determining causal effects on wage differentials, a common object of interest in the discrimination literature. To address this, there have been one or two papers that attempt to project the results from correspondence studies to theoretical wage differentials ([Lanning, 2013]). However, as of now, correspondence studies are most commonly interpreted as simply inspecting "one link in the chain", and still provide valuable insight into questions regarding labor market discrimination.

## Experimental design

This study used an audit study design to investigate job hiring discrimination based on gender. Fictitious resumes were randomly generated en masse, then sent to employers hiring for a given job. Following common practice, gender was signaled using names, and work history and education were randomly assigned since these elements affect response rates ([Lahey and Beasley, 2018]).

Resumes were sent in matched pairs to increase statistical power, with at least a day's worth of difference between each application to avoid suspicion from the hiring company. However, approximately 2% of

---

<sup>1</sup>Note that "audit study" and "correspondence study" are often used interchangeably in practice. Aside from the literature review section, this paper will tend towards using the term "audit study", to follow common practice.

job posts expired before the second application could be sent. Due to the small rate, this should not be problematic for our analyses.

Four major occupational categories were chosen to receive resumes. These categories were chosen to include a diverse spread of industries, manual labor requirements, and gender ratios (ex: 70% female, 80% female, 90% female). Specific jobs and sub-occupations within each category were then further investigated. These decisions allowed this study to inspect 1) the potential heterogeneity of response rates within occupations, and 2) compare discrimination across occupations. The broad occupational categories included: nurse, secretary, social worker, and teachers. The specific jobs and sub-occupations are included in the results section.

In an attempt to obtain a broadly representative sample, five major cities were chosen to pull job listings from. These cities are spread across the country, with roughly comparable populations and job listing sizes. The cities included New York, Los Angeles, Chicago, Dallas, and Houston. However, this selection did not include rural areas and smaller cities, which would be good geographies for future research to inspect.

The process of generating resumes entailed three primary steps. First, names used to signal gender were drawn from the Kline et al 2022 paper<sup>2</sup>. Second, work experience and education histories were extracted from authentic resumes that corresponded to a particular occupation and location. Third, names, work experience, and education histories were then randomly assigned to the fabricated resumes.

In previous studies, this random assignment step has been performed using a resume randomizer tool developed by economists ([Lahey and Beasley, 2009]). This tool generates resumes with a diverse range of formatting based on user-defined inputs. However, this study did not require emphasis on formatting, as the application platform Indeed automatically normalizes resumes regardless of their original formatting.

To maximize efficiency and reduce technological complexities, the study employed the Indeed platform's Easy Apply feature. This feature streamlines the application process for multiple job postings of the same occupation by providing a common set of application questions for some postings, and for many others, allowing the user to submit an application with a single click after selecting "apply now". This process was automated in the study using a custom-made python script, which minimized human involvement and ensured consistency in the application procedure. The choice to only use "Easy Apply" job postings limits the external validity of this experiments results, but is a measured and worthwhile trade-off given the constraints of the experimenter.

Generating the fictitious resumes involved creating numerous online accounts of various types. To start unique combinations of first and last names were used to create avatars for the fictitious resumes. To

---

<sup>2</sup>Due to statistical power constraints, only white names were used. Future research would do well to incorporate names from other ethnicities.

track individual responses, distinct email addresses were generated using the gmail dot method (for more on this, click here). CallRail was then used to create tracking phone numbers for each unique name-job pairing. To ensure consistency in the application process, Indeed accounts were created for each name using a corresponding phone number and email address for the proper job category.

After creating the fictitious resumes and corresponding avatars, the resumes were sent to the desired job postings over the course of several weeks. Responses were then tracked through various means such as incoming calls and texts on CallRail, as well as emails sent to the designated gmail dot combinations. These methods of tracking allowed for a comprehensive evaluation of response rates.

Limitations of this study include limited tech support and computing power, which resulted in the study being more manual labor-intensive than necessary. Additionally, the study was not able to inspect response rates to the level of detail desired, such as accounting for different ethnicities and a wider spread of jobs. Also, this study did not tease out any underlying mechanisms that could help explain the source of discrimination in the hiring process. However, despite these limitations, this study still provides valuable insight into potential discrimination based on gender in the hiring process across several major job categories.

## Results

Table 1 presents a concise overview of the summary statistics by city. The distribution of resumes sent across the five selected cities was relatively even. This consistency across cities suggests that the findings are not heavily influenced by regional variations.

Table 1: Summary Statistics By City

	A. Resumes sent			B. Responses in 30 days		
	Female	Male	Total	Call or text	Email	Total
New York	1,715	1,772	3,487	255	623	878
Chicago	1,788	1,771	3,559	271	578	849
Los Angeles	1,790	1,811	3,601	318	608	926
Dallas	1,886	1,717	3,603	303	604	907
Houston	1,697	1,760	3,457	298	517	815
<i>N</i>	8,876	8,831	17,707	1,445	2,930	4,375

Table 2: Summary Statistics By Sub-Occupation

	A. Resumes sent			B. Responses in 30 days		
	Female	Male	Total	Call or text	Email	Total
Nurses						
Acute Care Nurses	473	447	920	81	147	228
Advanced Practice Psychiatric Nurses	495	498	993	68	160	228
Clinical Nurse Specialists	525	529	1,054	78	193	271
Critical Care Nurses	552	518	1,070	86	197	283
Nurse Anesthetists	520	489	1,009	101	172	273
Nurse Midwives	473	459	932	68	178	246
Nurse Practitioners	504	529	1,033	82	166	248
Registered Nurses	466	451	90	183	273	
Secretaries and Administrative Assistants						
Executive	481	500	981	64	172	236
Legal	446	469	915	66	120	186
Medical	503	484	987	94	148	242
Social Workers						
Child, Family, and School	520	473	993	95	162	257
Healthcare	473	492	965	85	133	218
Mental Health and Substance Abuse	493	1,004	85	151	236	
Early Education Teachers						
Preschool	460	478	938	81	158	239
Kindergarten	486	498	984	64	158	222
Middle School	478	444	922	59	146	205
Preschool (Special Education)	528	562	1,090	98	186	284
<i>N</i>	8,876	8,831	17,707	1,445	2,930	4,375

Table 2 presents the summary statistics for the sub-occupations included in the study. It provides a breakdown of the number of resumes sent by gender, as well as the number of responses received, categorized by type. The overall response rate achieved was approximately 25%. It is worth noting that the number of male and female applications sent was almost perfectly balanced, indicating a fair representation of both genders in the study. Furthermore, the data reveals that employers were more likely to respond via email compared to phone calls or text messages.

Table 3 is the central focus of this paper, providing insights into mean response rates by gender for each sub-occupation, along with the corresponding differences in response rates. Overall, the response rates were relatively similar across the sub-occupations. However, the broad occupations exhibit different gender preferences within specific sub-occupations. This suggests that hiring practices vary within occupational categories, highlighting the nuanced and complex nature of gender preferences in the job market.



Table 3: Mean Response Rates

	Female	Male	Difference
Occupation and sub-occupation			
Nurses			
Acute Care Nurses	0.203	0.295	-0.092**
Advanced Practice Psychiatric Nurses	0.275	0.185	0.090**
Clinical Nurse Specialists	0.261	0.253	0.008
Critical Care Nurses	0.236	0.295	-0.060*
Nurse Anesthetists	0.265	0.276	-0.011
Nurse Midwives	0.249	0.279	-0.029
Nurse Practitioners	0.270	0.212	0.058*
Registered Nurses	0.253	0.344	-0.090*
Secretaries and Administrative Assistants			
Executive	0.241	0.240	0.001
Legal	0.215	0.192	0.023
Medical	0.247	0.244	0.003
Social Workers			
Child, Family, and School	0.240	0.279	-0.039
Healthcare	0.235	0.217	0.017
Mental Health and Substance Abuse	0.243	0.227	0.016
Early Education Teachers			
Preschool	0.259	0.251	0.008
Kindergarten	0.253	0.199	0.054*
Middle School	0.238	0.205	0.034
Preschool (Special Education)	0.252	0.269	-0.017

Significance levels were produced by regressing “male name” on “callback”.

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Figure 1 presents a graph illustrating these differences in response rates between genders. The graph demonstrates that, on average, the mean response rate differences within broad occupations tend to hover around zero. However, what becomes evident is the substantial variation in gender preferences within each broad occupational category. Notably, the occupation of nurses exhibits the highest degree of variation, indicating a wide range of employer preferences when it comes to hiring male or female candidates. Conversely, the occupation of secretaries and administrative assistants displays the smallest variation, suggesting a more consistent gender preference (or lack thereof) among employers in these roles.

Figure 1: Mean response rate differences by general occupation

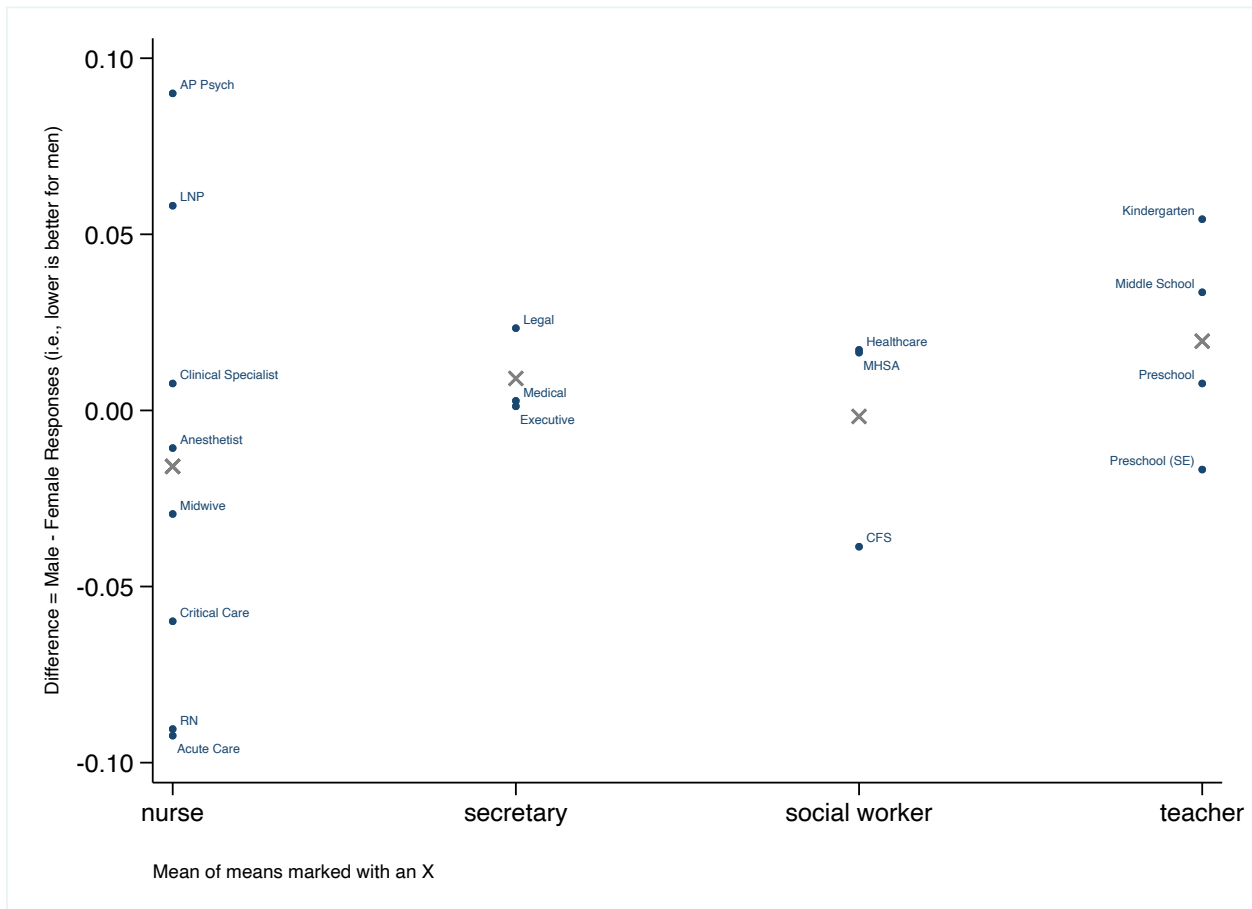


Table 4 explores the results just a bit further, offering descriptive comparisons among the sub-occupations. The table is sorted by the difference in response rates between genders. Notably, it becomes apparent that sub-occupations that exhibit a preference for male candidates tend to have higher manual labor scores, with the exception of Child, Family, and School Social Workers. Conversely, sub-occupations that display a neutral stance or a clear preference for female candidates tend to have lower manual labor scores. Interestingly, the column indicating the percentage of females in each sub-occupation does not demonstrate a noticeable pattern. This suggests that there may not be a discernible “drop-off point” of discrimination based on gender ratio alone.

Table 4: Descriptive Comparisons

Sub-occupation	Response Rate Difference	Manual Labor Score	Percent Female
Acute Care Nurses	-0.092**	87.9	91.2
Registered Nurses	-0.090*	98.1	86.7
Critical Care Nurses	-0.060*	79.4	81.1
CFS Social Workers	-0.039	17.8	86.8
Nurse Midwives	-0.029	62.4	94
Preschool Teachers (SE)	-0.017	54.9	94.4
Nurse Anesthetists	-0.011	80.2	59.6
Executive Secretaries and AAs	0.001	36.9	96.4
Medical Secretaries and AAs	0.003	30.1	95
Clinical Nurse Specialists	0.008	37.1	90.1
Preschool Teachers	0.008	48.0	93.6
MHSA Social Workers	0.016	17.0	75.7
Healthcare Social Workers	0.017	18.3	73.4
Legal Secretaries and AAs	0.023	16.0	82.6
Middle School Teachers	0.034	31.7	64.6
Kindergarten Teachers	0.054*	38.4	91.8
Nurse Practitioners	0.058*	42.7	87.4
AP Psychiatric Nurses	0.090**	20.3	84.4

Significance levels were produced by regressing “male name” on “callback”.

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## Conclusion

It is important to acknowledge and address the constraints and limitations encountered in this study. Firstly, the limited availability of technical support and computing power impacted the extent to which automated job application could be implemented. As a result, the study required more manual labor than ideally desired. Additionally, due to resource limitations, the analysis of response rates was not able to delve into the level of detail desired, such as examining variations in response rates across different ethnicities or exploring a wider range of job types. It is crucial to recognize that the technique employed in this study only captures a single link in the broader chain of hiring processes, and therefore, the findings should be interpreted within this context. Furthermore, this study did not delve into the underlying mechanisms driving the observed discrimination, which would be a logical next step for further research in this area.

Lastly, it is worth noting that the approach used in this study may not be applicable in the future due to the constantly evolving online job board policies on scraping and automated bot systems. The effectiveness of the methodology relied on specific conditions and partial automation at the time of the study. As both policies and scraping and automation technologies evolve, it is important for future researchers to consider the potential challenges and changes in implementing similar experiments. Adapting to these evolving circumstances will be crucial in conducting robust and reliable research on hiring discrimination.

## References

- [Adamovic and Leibbrandt, 2023] Adamovic, M. and Leibbrandt, A. (2023). A large-scale field experiment on occupational gender segregation and hiring discrimination. *Industrial Relations: A Journal of Economy and Society*, 62(1):34–59.
- [Albert et al., 2011] Albert, R., Escot, L., and Fernández-Cornejo, J. A. (2011). A field experiment to study sex and age discrimination in the madrid labour market. *The International Journal of Human Resource Management*, 22(02):351–375.
- [Arceo-Gomez and Campos-Vazquez, 2014] Arceo-Gomez, E. O. and Campos-Vazquez, R. M. (2014). Race and marriage in the labor market: A discrimination correspondence study in a developing country. *American Economic Review*, 104(5):376–380.
- [Berson, 2012] Berson, C. (2012). Does competition induce hiring equity?
- [Birkelund et al., 2019] Birkelund, G. E., Janz, A., and Larsen, E. N. (2019). Do males experience hiring discrimination in female-dominated occupations? an overview of field experiments since 1996.
- [Bohren et al., 2022] Bohren, J. A., Hull, P., and Imas, A. (2022). Systemic discrimination: Theory and measurement. Technical report, National Bureau of Economic Research.
- [Booth and Leigh, 2010] Booth, A. and Leigh, A. (2010). Do employers discriminate by gender? a field experiment in female-dominated occupations. *Economics Letters*, 107(2):236–238.
- [Carlsson, 2011] Carlsson, M. (2011). Does hiring discrimination cause gender segregation in the swedish labor market? *Feminist Economics*, 17(3):71–102.
- [Delfino, 2021] Delfino, A. (2021). Breaking gender barriers: Experimental evidence on men in pink-collar jobs.
- [Fry, 2022] Fry, R. (2022). Young women are out-earning young men in several us cities. *Young*.
- [Guryan and Charles, 2013] Guryan, J. and Charles, K. K. (2013). Taste-based or statistical discrimination: the economics of discrimination returns to its roots. *The Economic Journal*, 123(572):F417–F432.
- [Heckman, 1998] Heckman, J. J. (1998). Detecting discrimination. *Journal of economic perspectives*, 12(2):101–116.
- [Hsieh et al., 2019] Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and us economic growth. *Econometrica*, 87(5):1439–1474.

- [Kline et al., 2022] Kline, P., Rose, E. K., and Walters, C. R. (2022). Systemic discrimination among large us employers. *The Quarterly Journal of Economics*, 137(4):1963–2036.
- [Kline and Walters, 2021] Kline, P. and Walters, C. (2021). Reasonable doubt: Experimental detection of job-level employment discrimination. *Econometrica*, 89(2):765–792.
- [Lahey and Beasley, 2018] Lahey, J. and Beasley, R. (2018). Technical aspects of correspondence studies. In *Audit studies: Behind the scenes with theory, method, and nuance*, pages 81–101. Springer.
- [Lahey and Beasley, 2009] Lahey, J. N. and Beasley, R. A. (2009). Computerizing audit studies. *Journal of Economic Behavior & Organization*, 70(3):508–514.
- [Lang and Spitzer, 2020] Lang, K. and Spitzer, A. K.-L. (2020). Race discrimination: An economic perspective. *Journal of Economic Perspectives*, 34(2):68–89.
- [Lanning, 2013] Lanning, J. A. (2013). Opportunities denied, wages diminished: Using search theory to translate audit-pair study findings into wage differentials. *The BE Journal of Economic Analysis & Policy*, 13(2):921–958.
- [Neyt et al., 2019] Neyt, B., Vandenbulcke, S., and Baert, S. (2019). Are men intimidated by highly educated women? undercover on tinder. *Economics of Education Review*, 73:101914.
- [Ngai and Petrongolo, 2017] Ngai, L. R. and Petrongolo, B. (2017). Gender gaps and the rise of the service economy. *American Economic Journal: Macroeconomics*, 9(4):1–44.
- [Riach and Rich, 2006] Riach, P. A. and Rich, J. (2006). An experimental investigation of sexual discrimination in hiring in the english labor market. *The BE Journal of Economic Analysis & Policy*, 6(2).
- [Schaede and Mankki, 2022] Schaede, U. and Mankki, V. (2022). Quota vs quality? long-term gains from an unusual gender quota.
- [Yavorsky, 2019] Yavorsky, J. E. (2019). Uneven patterns of inequality: An audit analysis of hiring-related practices by gendered and classed contexts. *Social Forces*, 98(2):461–492.
- [Zhou et al., 2013] Zhou, X., Zhang, J., and Song, X. (2013). Gender discrimination in hiring: Evidence from 19,130 resumes in china. *Available at SSRN 2195840*.