Testing the Backlash Effect in the Labor Market: A Resume Audit Study

Marina Mileo Gorsuch

A broad literature in social psychology has established that respondents react negatively when women engage in traditionally masculine actions in the workplace (Heilman and Chen 2005; Heilman, Wallen, Fuchs, and Tamkins 2004; Rudman and Glick 1999; Rudman 1998; Rudman and Glick 2001; Bowles, Babcock, and Lai 2006; Amanatullah and Morris 2010). This negative reaction is described as a "backlash effect" (Bowles, Babcock, and Lai 2006). In this study, I take the question of backlash outside of a laboratory environment to see if real employers have the same response as respondents in a laboratory to traditionally masculine actions.

In previous work, laboratory respondents rated perceived-heterosexual women who used masculine adjectives on a resume worse than when they used feminine adjectives. The resumes of perceived-heterosexual men and perceived-gay women were both immune to this effect. To examine if the backlash effect impacts a real-world job search, I used experimentally manipulated resumes to apply for 1,300 jobs publicly advertised in multiple cities and towns. I find that employers call back women who use traditionally masculine adjectives more than when they use traditionally feminine adjectives. Perceived-heterosexual men are called back less when they use traditionally masculine adjectives compared to when they use traditionally feminine adjectives. That is, employers do not have a backlash effect when viewing the same manipulation that inspired a backlash effect among laboratory respondents. In fact, employers have the reverse of a backlash effect: they prefer women with traditionally masculine adjectives and perceived-heterosexual men with traditionally feminine adjectives.

A broad literature in social psychology has established that respondents react negatively when women engage in traditionally masculine actions in the workplace, such as withdrawing altruistic behavior, being successful in a male occupation, and self-promotion in an interview (Heilman and Chen 2005; Heilman, Wallen, Fuchs, and Tamkins 2004; Rudman and Glick 1999; Rudman 1998; Rudman and Glick 2001).

One example of this pattern is how respondents react to women who attempt to negotiate for a higher salary (Bowles, Babcock, and Lai 2006; Amanatullah and Morris 2010). In particular, both men and women in an experimental setting were less likely to want to work with women who negotiated and described them as less nice and more demanding, although equally competent. While men were also viewed as less nice and more demanding when they negotiated, there was no corresponding change in male respondents' willingness to work with them. This suggests that women are penalized for negotiating because they violated a prescription of femininity: niceness (Bowles, Babcock, and Lai 2006). In this literature, the negative effect of negotiating is termed a "backlash." Women appear to be aware of this backlash; women report a lower ceiling on the amount they can ask for in a negotiation before appearing "pushy" or "demanding" to a manager again in an experimental setting (Amanatullah and Morris 2010).

Because the backlash effect has the potential to contribute to earnings differences between men and women, it is important to understand how laboratory results translate into the real labor market. While these experimental studies use a wide variety of traditionally masculine actions, it has not been clearly established if real employers respond in the same way that experimental subjects respond. One potential difference between real employers and laboratory respondents is that employers have a stronger financial incentive to employ people who will be productive employees. Real employers may also know more than laboratory respondent about

whether or not acting in traditionally masculine ways is beneficial to their business' production. For example, both laboratory respondents and real employers may have a negative feeling upon observing norm-violating behavior. But, if real employers are more cognizant of or sensitive to a potential effect on productivity, they may be less likely to act on those feelings than a laboratory respondent. Thus, women engaging in masculine actions may have a different outcome in a real labor market setting than in a laboratory.

In previous work, laboratory respondents rated perceived-heterosexual women who used masculine adjectives on a resume worse than when they used feminine adjectives. The resumes of perceived-heterosexual men and perceived-gay women were both immune to this effect (Gorsuch unpublished). In this study, I take the question of backlash outside of a laboratory environment to see if real employers have the same response as laboratory respondents to traditionally masculine actions.

To test if real employers have the same response as laboratory respondents, I applied for publicly advertised jobs using resumes that vary on sex, sexual orientation, and the use of traditionally masculine or feminine adjectives. I find that employers call back heterosexual women who use traditionally masculine adjectives **more** than when they use traditionally feminine adjectives. Heterosexual men are called back **less** when they use traditionally masculine adjectives compared to when they use traditionally feminine adjectives. That is, employers do not have a backlash effect when viewing the same manipulation that inspired a backlash effect among laboratory respondents. In fact, they have the reverse of a backlash effect: they reward heterosexual men who use feminine adjectives and heterosexual women who use masculine adjectives.

Theoretical framework

Prescriptive stereotypes, stereotypes about how a woman **ought to be**, may cause firms to react negatively when women violate these stereotypes (Heilman 2001; Rudman and Phelan 2008). Because many of the traits that are viewed as necessary for a success in the workplace are not seen as desirable for women, prescriptive stereotypes adversely affect women who attempt to succeed in the workplace by acting in more traditionally masculine ways (Rudman and Phelan 2008; Gill 2004).

Akerlof and Kranton (2000) suggest that men react negatively to women acting in traditionally masculine ways because their own masculine identity is threatened. That is, when a member of a different group (women) acts like a member of their group (men), it adversely affects their own identity as men. Likewise women's identity may also be threatened when a member of their group acts contrary to their behavioral prescriptions because it threatens the group's identity as women. That is, "a person learns a set of values (prescriptions) such that her actions should conform with the behavior of some people and **contrast with that of others**." (Akerlof and Kranton 2000, p. 728 emphasis added). In this case, men's identity is based on their behavior differing from that of women; when women act in ways that are traditionally male, men's identity suffers and they may respond by imposing a punishment.

However, it is also important to consider which group (or groups) defines another group's identity. As Tajifel and Turner (1979) note that in their foundational theory of social identity formation "in-groups do not compare themselves with every cognitively available out-group: the out-group must be perceived as **a relevant comparison group**" (p. 41 emphasis added). Because gay women are perceived as having many traits in common with heterosexual men (Ahmed et al 2013), gay women are likely not the relevant contrast group for heterosexual men. This would

suggest that gay women would be exempt from the backlash effect. Indeed, previous work found that gay women are exempt from the backlash effect (Gorsuch unpublished).

In contrast to prescriptive stereotypes, descriptive stereotypes are stereotypes about how people of different sex and sexual orientations **are**. These stereotypes may cause firms to anticipate a "lack of fit" between a heterosexual female applicant and a job that is perceived to require masculine traits (Heilman 1995). Descriptive stereotypes could be thought of as a form of statistical discrimination where firms believe different subgroups have particular distributions of personality characteristics (Arrow 1973; Phelps 1972).

Under statistical discrimination, acting in a traditionally masculine manner may be a signal of having particular traits. In the case of an applicant using traditionally masculine adjectives to describe themselves, the employer will update their assessment of the applicant's personality characteristics based on their self-description. The posterior assessment of an applicant's personality after seeing the choice of adjective will depend on the employer's beliefs about the overall distribution of personality traits based on sex and sexual orientation as well as the cost of using masculine adjectives for each subgroup. For example, in a laboratory setting managers reported that women are less "hostile to others" than men in general, but women who are "successful managers" are viewed as more hostile to others than men who are similarly described (Heilman et al 1995). Importantly, a single descriptor will impact the assessment of numerous personality characteristics – being described as a "successful manager" also increased the perceived competence and independence of both men and women (Heilman et al 1995).

Employers may also want employees who have a bundle of traits – for example, firms may value an employee who can be collegial with co-workers but competitive towards

competitors. Indeed, numerous different personality characteristics are associated with success in the workplace (among many, Heckman et al 2006, Borghans et al 2008).

Respondents in a laboratory setting, even those who have been led to believe their responses will be used by hiring managers, could value elements of a an applicant differently than employers who believe they are evaluating a real applicant. For example, respondents are paid a flat rate for participating in the experiment regardless of their recommendation and will not actually be impacted by the hiring decision. In contrast, firms in a resume audit study must spend time contacting the applicant they select and are actually considering hiring this applicant. Firms may therefore place more value on the applicant's productivity than laboratory respondents. Likewise, laboratory respondents may evaluate the applicant as a coworker rather than an employee and therefore may place less value on the expected productivity of the applicant. Laboratory respondents may also be more influenced by feelings of altruism towards applicants, whereas firms are more concerned with productivity. Moreover, the firms in the audit study may be in a better position to estimate an applicant's potential productivity because they have more experience hiring people and are more knowledgeable about the production function at their firm. This means that laboratory respondents and real employers may have different updating processes after seeing a traditionally masculine adjective.

The elements described above, prescriptive stereotypes, statistical discrimination, signaling, complementarity between traits, and different incentives between laboratory respondents and firms are incorporated into the model below where evaluators seek to hire employees that maximize Equation 1. In this model, the employer knows three things about the applicant: their sex, sexual orientation, and what type of adjective they used to describe themselves on their resume.

$$v = w * E[f(Collegial, Competetive)|\ sex, sexual\ orientation, adjective] + \quad (1)$$

$$(1-w) * \beta * I(female\ \&\ heterosexual\ \&\ masculine\ adjective)$$

Where $w \in [0,1], \frac{\partial f}{\partial Collegial} > 0, \frac{\partial f}{\partial Competitive} > 0$ and $\begin{cases} \frac{\partial^2 f}{\partial Competitive^2} < 0 \text{ and } \frac{\partial^2 f}{\partial Collegial^2} < 0 \\ \text{and/or} \\ \frac{\partial^2 f}{\partial Collegial \partial Competitive} > 0 \end{cases}$

The first half of Equation 1 consists of the expected productivity of an employee based on their personality characteristics. The evaluator does not observe the applicant's true value of collegiality and competitiveness, but they estimate these values from the applicant's sex, sexual orientation, and use of adjectives. There is a production function *f* where applicants who possess collegiality and competitiveness are more productive. The production function shows decreasing marginal productivity to both personality traits and/or complementarities between the two traits; for most typical production functions, this implies that employers prefer employees with both traits. The second component of Equation 1 consists of the prescriptive stereotypes or identity threat that occurs when heterosexual women use traditionally masculine language; β measures if people have a negative feeling when heterosexual women violate norms. Then, w is the weight that the evaluator assigns to productive component of Equation 1 while *(1-w)* is the weight assigned to the identity threat. For example, if respondents in a laboratory setting place less value on their expectation of the employees productivity than a real employer, then $w_{lab\ respondent} < w_{employer}$.

The social psychology literature suggests that the following condition holds for laboratory respondents:

$$v|_{female\ \&\ heterosexual\ \&\ masculine\ adjective} < v|_{female\ \&\ heterosexual\ \&\ feminine\ adjective}$$

However, if real employers place different value on production versus identity threat or if employers form expectations about productivity differently than respondents in a laboratory, these relationships may not hold for real employers. Because many actions needed for success in the workplace are traditionally masculine, the backlash effect has the potential to contribute to earnings differences between men and women. It is therefore important to understand how laboratory results translate into the real labor market.

Numerical example

The following example is a specific form of Equation 1 that shows how a difference in the weight laboratory respondents and real employers assigned to the first and second portions of Equation 1 would predict different backlash responses. In this example, $f(Collegial, Competetive) = Collegial^{1/2} + Competetive^{1/2}$. This production function shows decreasing marginal productivity of both personality characteristics. The two personality characteristics are not complements.

Equation 2 shows the evaluation of an employer while Equation 3 shows the evaluation of a laboratory respondent, where employers place a higher weight (.9) on the expected production than laboratory respondents (.1).

$$v_{Employer} = .9 * E\left[Collegial^{1/2} + Competetive^{1/2} | \, sex, sexual\ orientation, adjective\right] + .1 * \beta * I(female \,\&\, heterosexual \,\&\, masculine\ adjective) \quad (2)$$

$$v_{Lab\ respondent} = .1 * E\left[Collegial^{1/2} + Competetive^{1/2} | \, sex, sexual\ orientation, adjective\right] + .9 * \beta * I(female \,\&\, heterosexual \,\&\, masculine\ adjective) \quad (3)$$

The following table shows the distributions for *Collegial* and *Competitive* for an individual by sex, sexual orientation, and adjective. Within each cell, the two variables are independent ($Collegial \perp Competetive | \, sex, sexual\ orientation, adjective$)

Collegial *Competitive*	LGBT	Non-LGBT		LGBT	Non-LGBT
"I'm assertive"	*Collegial*~ U(5, 15) *Competitive*~ U(15, 25)	U(0, 10) U(20, 30)		U(5, 15) U(15, 25)	U(10, 20) U(10, 20)
"I'm nurturing"	U(15, 25) U(5, 15)	U(10, 20) U(10,20)		U(15, 25) U(5, 15)	U(20, 30) U(0, 10)
	Male			Female	

If both employers and respondents believe the distributions outlined above, they form expectations about the productivity of an application. For example, for an applicant who is male, not LGBT, and describes himself as "assertive," the expected $f(Collegial, Competetive)$ is

$$\int_{0}^{10}\int_{20}^{30} (Collegial^{1/2} + Competetive^{1/2}) \frac{1}{10}\frac{1}{10}\, dCompetetive\, dCollegial = 7.1$$

Using this result in Equations 2 and 3, employers and laboratory respondents form ν. The following table shows the employer and laboratory respondent ν for each cell:

$\nu_{Employer}$ $\nu_{Lab\ respondent}$	LGBT	Non-LGBT		LGBT	Non-LGBT
"I'm assertive"	$\nu_{Employer} = 6.8$ $\nu_{Lab\ respondent}$=.76	6.4 .71		6.8 .76	6.9-.1* β .77-.9* β
"I'm nurturing"	6.8 .76	6.9 .77		6.8 .76	6.4 .71
	Male			Female	

The "backlash effect" will occur in the lab if a laboratory respondent prefers a perceived-heterosexual woman who uses traditionally feminine language to one who uses masculine language. This will happen if .77-.9* β <.71, or equivalently, if $\beta > .07$. However, an employer

would only prefer a perceived-heterosexual women who uses traditionally feminine language to those who use masculine language if $6.9 - .1 * \beta < 6.4$, which happens if $\beta > 5.5$. That is, it takes a much larger identity threat (β) to make it worth it for the employer to engage in backlash. There are many values of β where the laboratory respondent would engage in backlash and the real employer would not, even if both the lab respondent and the employer have the same information about the production function of the firm and form the same expectations. Simply having a different weight on production and identity could result in lab respondents engaging in the backlash effect and not real employers.

It is important to note that under these conditions, the negative reaction to norm violation does negatively impact perceived-heterosexual women even though employers do not demonstrate the backlash effect. Employers prefer a perceived-heterosexual man who uses traditionally feminine language over a perceived-heterosexual woman who uses traditionally masculine language, even though they have the same expected productivity. That is, perceived-heterosexual women are still harmed by prescriptive stereotypes, even when they are not strong enough to result in a backlash.

Resume Audit Study: Experimental Manipulation

Previous research found that resumes using traditionally masculine language inspire backlash against perceived-heterosexual women, while men and perceived-gay women are exempt (Gorsuch unpublished). I use this same approach in an audit study, where I apply to publicly advertised jobs with resumes that are manipulated on sex, sexual orientation, and the use of traditionally masculine or feminine language. I examine if the resumes are contacted for interviews at different rates. Resume audit studies are a well-established method to examine discrimination in a real labor market setting and have been used to analyze discrimination based

on race, sexual orientation, parental status, and many other characteristics (Bertrand and Mullainathan 2004; Tilcsik 2011; Correll, Benard, and Paik 2007).

To create realistic resumes with the experimental manipulation, I first created a resume bank from recent college graduates who publicly listed their resume on Indeed.com.[1] I selected resumes that were listed in each city I would later apply for jobs; that is, the education and work histories of the compilation resumes were drawn from the area they were applying for a job. I used the computer program developed by Lahey and Beasley (2009) to create resumes with randomly selected entries for work history and education. In total, there are 117 work histories and 42 educations. Each resume contains two work histories, one extra-curricular activity, and one education entry. The education and work histories are balanced across the experimental manipulations, as shown in Appendix 2.

The objective statement of the resume, a common feature of resumes of recent college graduates, includes adjectives that are either masculine or feminine. The masculine adjectives are aggressive, enterprising, assertive, bold, confident, self-starter, achiever, and dynamic. The feminine adjectives are nurturing, caring, sympathetic, kind, supportive, encouraging, helpful, and cooperative. These adjectives were selected from a pre-test that determined which adjectives are perceived as masculine. In the pre-test, one group of respondents on Mechanical Turk viewed adjectives that are supposedly from a resume and answered the question "How likely is it that the applicant male?" Another group rated the same adjectives on how likely the applicant was female. As Figure 1 shows, adjectives that were viewed as relatively more likely to come from a male applicant by one group were viewed as less likely to come from a female applicant by the other group. This suggests that the manipulation will be effective; that is, using adjectives

[1] Indeed.com is a large online job board with job ads and resumes of job seekers. It is free to post and view resumes. Employers pay a "pay per click" fee for job ads that they post (Indeed.com 2014)

perceived as the most feminine and least masculine will signal traditionally feminine characteristics. Likewise, using adjectives perceived as the most masculine and least feminine signals traditionally masculine characteristics.

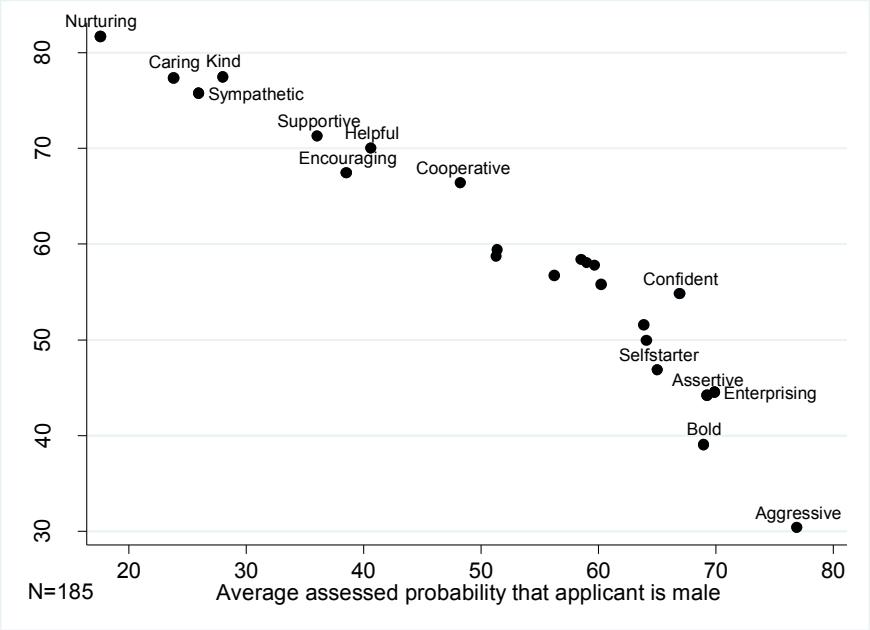


Figure 1*: Results from a pre-test of adjectives. One group (N=85) reported how likely it was that the applicant was male (x-axis) while another group (N=90) reported how likely to applicant was female (y-axis). The adjectives with the strongest gender associations are labeled.*

On the resume, the applicant's sex is indicated by the applicant's first name. The choice of a name is complicated by the fact that names also imply information about race, and there is evidence that how respondents view resumes of gay applicants varies by the race of the applicant (Pedulla 2012). For this project, I restrict to names that are more common among white people, and leave variation in the backlash effect by race to future work. The first names used in the manipulations are the five most common names for white babies born in California to high education parents. The female names are Katherine, Emma, Alexandra, Julia, and Rachel (Levitt and Dubner 2005). The male names are Benjamin, Samuel, Alexander, John, and William (Levitt and Dubner 2005). I selected the last names with the highest percent white in the 100 most

common last names from the 2000 Census. These last names are Wood, Sullivan, Myers, Peterson, Miller, Murphy, Fisher, Cox, Cook, and Long (Census 2012).

The resume manipulated perceived sexual orientation through a leadership position in a college group. Some resumes indicated the applicant held a leadership position in a lesbian, gay, bisexual, or transgender group, while others will indicate the applicant held a similar role in a non-LGBT organization. Tilcsik (2011) performed an audit study comparing callback rates for resumes of men that indicated they were the treasurer of a campus LGBT organization to those that indicated being the treasurer of a campus socialist organization. He found that 11.5% of the resumes with the socialist organization received a callback compared to 7.2% for the resume with LGBT organization. This suggests that using membership in a college LGBT organization is an effective way to signal sexual orientation.

I applied for jobs that were advertised on Craigslist, a major source of local job ads: Craigslist has over 1 million new job listings each month (Craigslist 2014). I applied for all jobs in the selected cities that were advertising a position of 20 or more hours per week where a recent college graduate with two years of part-time work experience would meet ad's listed requirements. If an ad required that the applicant have a particular degree, certificate, or experience, it was not included in the study. For example, it is common for restaurants to require an applicant have previous serving experience; any ad with this requirement was not included in the study. I applied for each eligible job that was advertised during the study period in the selected locations. I only applied to each company once, even if they listed multiple jobs during the study period. The industry and location of the jobs are balanced across the experimental manipulations, as shown in Appendix 2.

Each resume had a phone number with a local area code; this phone number functioned and appeared the same as a landline phone number, but was actually a VoIP (voice over IP) number. Each resume had an email address that was based on the applicant's full name; for example, the applicant named "Alexandra Myers" would have an email address similar to Alexandra.Myers24@gmail.com. Phone lines and email addresses were monitored daily. Finally, each resume contained a local address; each address was an apartment in a mid-range apartment complex located in the same city as the job the applicant was applying to.

The outcome of interest is any positive contact from an employer through email or phone. Positive contact includes an invitation to interview, a request for the applicant to contact the employer, or a request for more information. A response from the employer that simply indicated they received the application was not considered positive contact. To minimize the adverse effects on employers and on other job seekers, immediately after any positive contact from an employer, the fictional applicant informed the employer that they had received another offer and was no longer interested in the position.

This experimental design was approved by Duke Institutional Review Board.

Weaknesses of resume audit studies

While audit studies are a well-established and powerful method to examine discrimination in the labor market, they do have some weaknesses. First, they only assess discrimination or backlash at one specific point – the initial application during a job search. An audit study is less able to examine interactions that happen further along in employment, such as a salary negotiation during a yearly review. They also only examine publicly advertised jobs, although many jobs are obtained through social or family networks. Additionally, because of the use of the LGBT student organization as a signal of sexual orientation, this audit study only

assesses the backlash effect for recent college graduates. The effect may be different for women in prime child-bearing years.

Furthermore, audit studies will only reveal average discrimination, while it is the level of discrimination at the marginal firm that has welfare implications (Heckman 1998; Becker 1957). Finally, there may be unobserved characteristics that influence how employers react to an application. If these characteristics are related to the observed characteristics that are forced to be equal on the manipulated resumes, the employer may be acting on expectations about the unobserved characteristics rather than engaging in taste-based discrimination (Heckman 1998).

Despite these weaknesses, audit studies remain an important method to analyze discrimination. Unlike laboratory experiments, audit studies look at the behavior of real employers who believe they are making a real hiring decision. Unlike non-experimental data, audit studies allow the researcher to create equivalent resumes that vary on the manipulated characteristics. These strengths make audit studies a useful experimental method, despite the weaknesses discussed above.

Resume Audit Study: Empirical Framework

Regression Framework

To test if employers respond to resumes with traditionally masculine or feminine language differently based on the sex and sexual orientation of the applicant, I will examine if the difference in the callback rates with a linear probability model. The outcome variable (C_j) is an indicator for if resume *j* received a call from an employer.

$$C_j = \alpha + \beta * I(female\ applicant_j) + \theta * I(masculine\ adjective_j) + \delta * I(female\ applicant_j) * I(masculine\ adjective_j) + \pi * I(LGBT\ and\ male_j) + \varsigma * I(LGBT\ and\ male_j) * I(masculine\ adjective_j) + \sum_{c=1}^{C-1}\sum_{i=1}^{I-1} \tau_{c,i} * I(City_j = c) * I(Industry_j = i) + \sum_{w=1}^{W-1} \omega_k * I(Work_j = w) + \sum_{e=1}^{E-1} \omega_k * I(Educ_j = e) + \eta_j$$

In this regression, $\hat{\theta}$ estimates the change in proportion of perceived-heterosexual men receiving callbacks when they use traditionally masculine adjective. Likewise, $\hat{\theta} + \hat{\delta}$ estimates the change in proportion of women receiving callback when they use traditionally masculine adjectives (perceived heterosexual and perceived gay women are pooled together, for reasons that will become clear soon). If the same backlash effect that has been documented in laboratory setting extends to a job search, $\hat{\theta} + \hat{\delta}$ will be negative, indicating that women who use masculine adjective are called back less that those that use feminine adjectives. The difference between perceived-heterosexual men and women in the effect of using masculine adjectives is estimated by $\hat{\delta}$. The other difference in the difference, $\hat{\varsigma}$, estimates if men with the LGBT activity are impacted differently by using masculine adjectives than those without the LGBT activity.

Resume Audit Study: Results

Description of jobs and employers

In total, I applied to 1,334 jobs with manipulated resumes. Of these jobs, 269 applications (20.2%) resulted in positive contact. The jobs were categorized as falling into one of four types of jobs: office, labor or skilled trade, food service, and retail. The most common type of job was an office job.

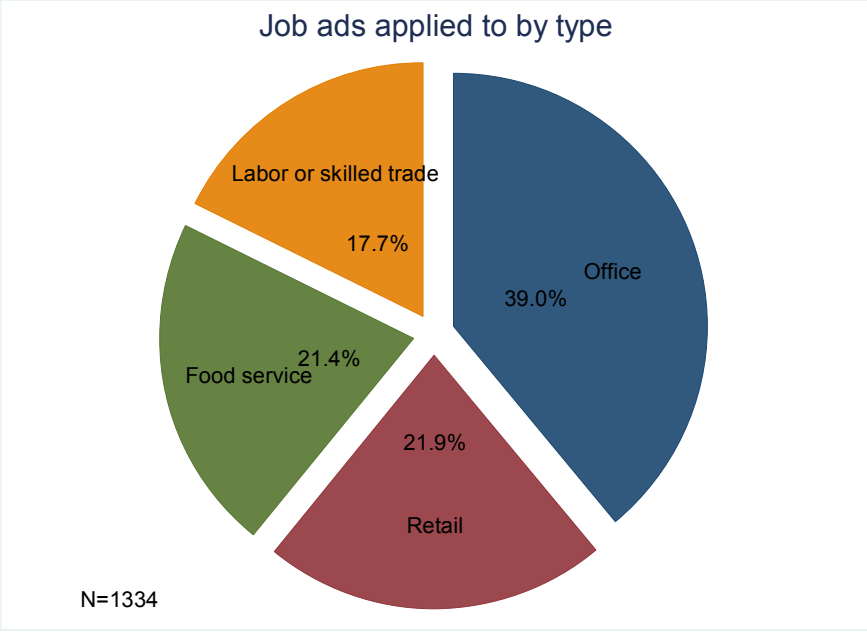


Figure 2*: Types of jobs that were applied for. N=1,334*

Within each category, there were a variety of jobs. The following table lists the most common jobs in each category:

Office	Food service	Retail	Labor or skilled trade
Receptionist/Front desk	Server	Customer service/Sales associate/Guest services	Manufacturing/ Production worker
Office manager/Office assistant/ Office coordinator	Host/Food runner	Assistant manager/ Team lead/Supervisor	Warehouse worker
Customer service representative	Barista	Kennel assistant/Pet resort specialist/Pet sitter	Lawn care technician
Project manager/Project coordinator/Project assistant	Assistant manager		Maintenance worker

***Table 1**: Examples of common jobs applied for within each type of job*

As shown in Figure 3, the callback rates varied by job category; food service had the highest callback rate (29%) and labor or skilled trade the lowest (12%).

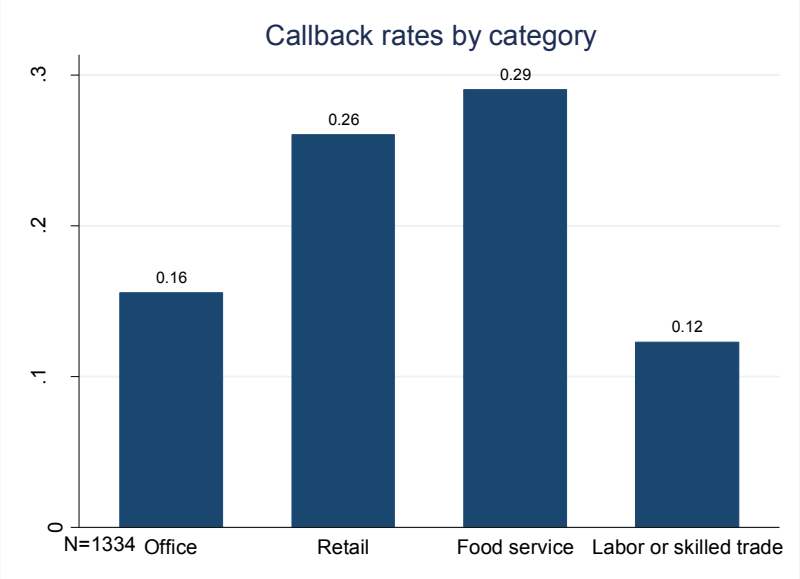


***Figure 3**: Average callback rates by job category. N=1,334*

The job advertisements themselves frequently used words that were traditionally feminine or masculine. Some advertisements said they were looking for a person who is "compassionate," "cooperative," or "warm." Others asked for an applicant who was "confident," "assertive," and "dynamic." Still others did not use adjectives or used neutral adjectives like "organized," "efficient," and "reliable." For example, one advertisement stated

> …we hope to find someone who can appreciate and help us to maintain a **feeling of welcome**, excellent customer service, unique and high-quality [redacted] cuisine, and our assertion that dining out is not just about "grabbing something to eat," but should delight and soothe all five senses. You should be professional but **warm**, **patient**, flexible, and able to interact **graciously** with all customers. Experience working in the hospitality or dining industry is preferred, but not required; we will gladly train the right candidate whose temperament and schedule are a good fit for our restaurant and small team.

While another advertisement stated they were looking for someone who is "**Driven**, **self motivated**, able to work **independently** with positive results." As displayed in Figure 4, the majority of job advertisements used neutral or no adjectives while about 1/3 used feminine adjectives:

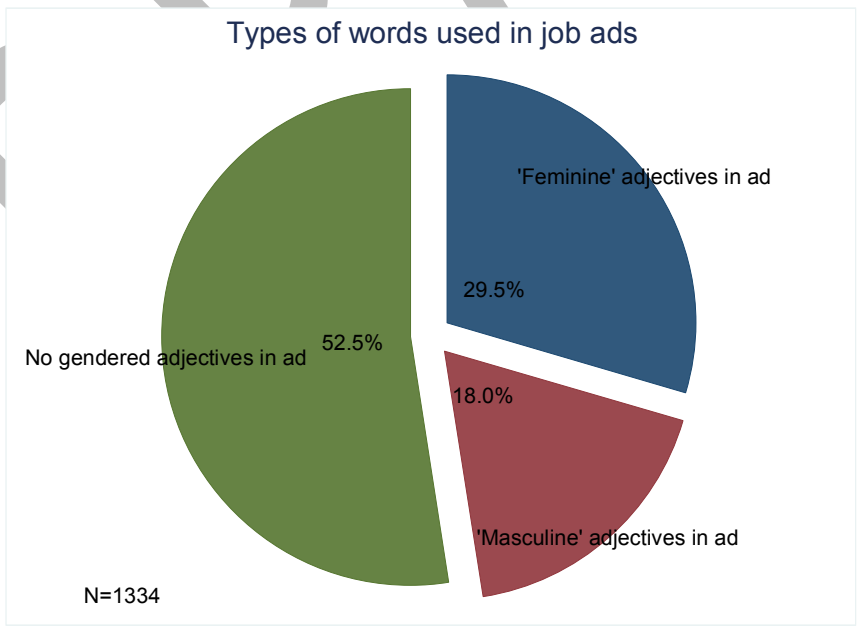


***Figure 4**: Distribution of the types of adjectives used in the text of the job advertisements applied to*

Is there a backlash effect?

Figure 5 shows the difference in callback rates between resumes with masculine and feminine adjectives. Perceived-heterosexual women have a **higher** callback when using masculine adjectives compared to feminine adjectives, while perceived-heterosexual men have a **lower** callback when using masculine adjectives compared to feminine adjectives. The differences have a p-value of .022 for male resumes and .073 for female resumes (robust errors clustered at the city by industry level). This is the exact opposite from the patterns found in the backlash literature, where women are rated more negatively when engaging in traditionally male behavior.

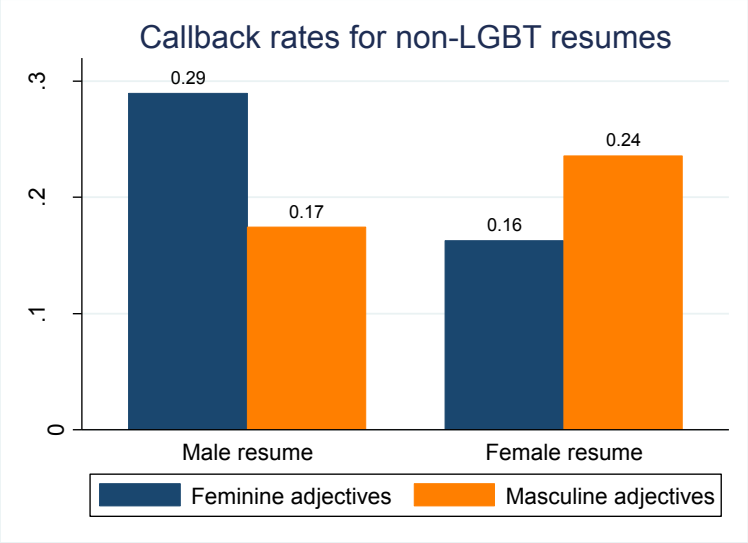


Figure 5*: Callback rate by sex and type of adjective used on resumes without the LGBT activity. N=678*

As shown in Figure 6, not only do perceived-heterosexual women benefit from using masculine adjectives while perceived-heterosexual men are harmed, there is also no difference in the pattern between perceived-gay and perceived-heterosexual women. Both of these patterns are opposite of the results in the laboratory experiments.

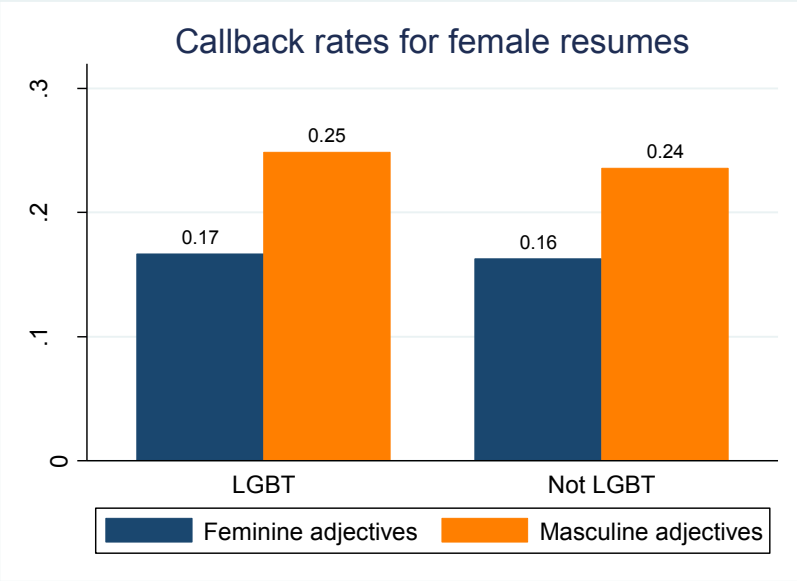


***Figure 6**: Callback rate by sexual orientation and type of adjective used on female resumes. N=673*

The following table shows the patterns from Figure 5 and 6 in a regression framework. The regression shown in Table 12 is a linear probability model with an indicator variable for receiving a callback as the dependent variable. Placing this question in a regression framework allows for numerous important controls: I include fixed effects for the work histories from the resume bank (total of 117 work histories), fixed effects for the education entry from the resume bank (42 total educations), and city by industry fixed effects. The independent variables include an indicator for female, gay male, assertive adjective, and their interactions.

As Table 2 shows, the negative effect of using masculine adjective for heterosexual men (the omitted group) is statistically significant at the .1 level. Female resumes are more likely to be called back when they use masculine language than when they use feminine language (-.101+.180=.079, p-value of F-test is .04). Gay male resumes receive a lower callback rate, but do not receive a penalty for using masculine language (-.101+.089=-.012, p-value on F-test is .81). This echoes the patterns shown in Figures 5 and 6: perceived heterosexual men are called

back less when they use masculine language while women are called back more when they use masculine language.

	Callback
Masculine adjective on resume	-0.101*
	(0.055)
Gay male resume	-0.107**
	(0.050)
Masculine adjective and gay male resume	0.089
	(0.087)
Female resume	-0.112*
	(0.054)
Masculine adjective and female resume	**0.180****
	(0.080)
Observations	1,334
R-squared	0.209
Work Experience FE	Yes
Education FE	Yes
City by Industry FE	Yes
Errors Clustered	City by industry
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1*	

Table 2: *Results of a LPM regression of an indicator for receiving a callback on the type of adjective used on the resume, a female indicator, a gay male indicator, and their interactions. Controls include work experience fixed effects, education fixed effects, and city by industry fixed effects. Errors are robust and clustered at the city by industry level.*

The ad text: Is it a match?

Half of the job ads used language that was either traditionally feminine (30%) or traditionally masculine (18%). This is a large difference from the laboratory experiment, where the job description was devoid of any gendered language and was the same for all applicants. It is possible that the impact of using traditionally masculine or feminine language on a resume may vary by the text used in the ad. For example, one ad stated that they are looking for someone "warm" and "patient" while another said they wanted someone "driven" and "self-motivated."

When an applicant says they are an "aggressive self-starter," this may seem like a poor fit for the first employer and a good fit for the second.

Indeed, the adjectives used in the text of the job advertisement do play an important role in determining who is likely to get a callback. However, the match between adjectives on the resume and in the adjectives used in ad is not the driving factor. Rather, the words in the job ads appear to reveal the employer's preference for the sex and sexual orientation of the applicant.

For ads that used traditionally masculine language, the perceived-heterosexual male resumes received callbacks at a rate over twice that of straight women (p=.053) and higher than both perceived-gay men and perceived-gay women (p=.08 and p=.05 respectively). Perceived gay men and perceived gay women had higher call back rates than perceived heterosexual women, but these differences were not statistically significant.

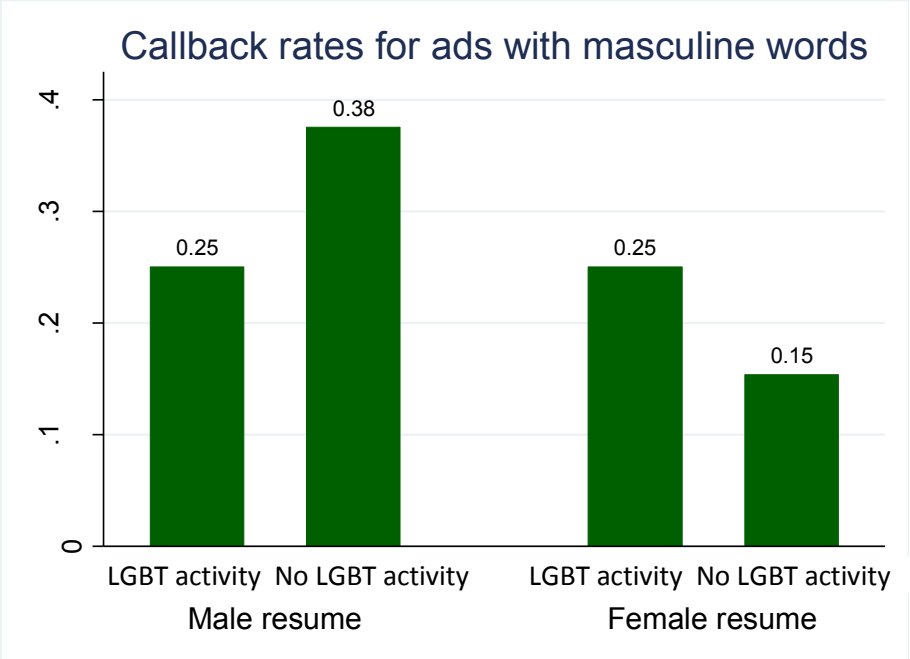


***Figure 7**: Callback rate by sex, sexual orientation for job ads with traditionally masculine adjectives. N=204.*

Moreover, perceived heterosexual women were unable to overcome their low callback rate for these jobs by using traditionally masculine adjectives. As shown in Figure 8, there is no

difference in the callback rate for perceived-heterosexual women between resumes with feminine adjectives and those with masculine adjectives. That is, when applying to a job that used traditionally masculine language, a perceived-heterosexual man who described himself with feminine language was still more likely to get a call back than a woman who described herself with masculine language.

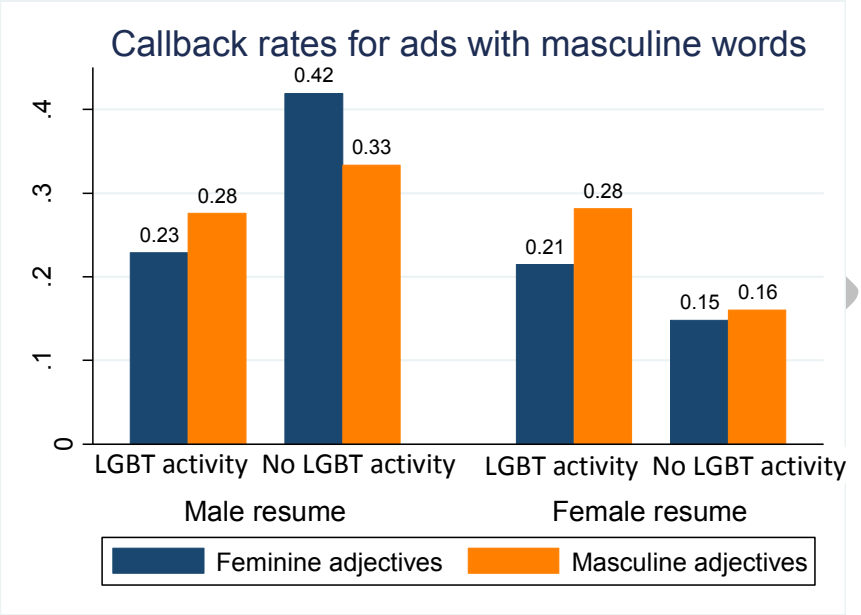


***Figure 8**: Callback rate by sex, sexual orientation, and adjective for job ads with traditionally masculine adjectives. N=204.*

Job ads without gendered language or with traditionally feminine language have less variation in callback rates, as displayed in Figure 9. However, perceived-heterosexual women do have an advantage for jobs described with traditionally feminine language (the difference between perceived-heterosexual women to all others has a p-value of .051 with robust errors clustered at industry by city level).

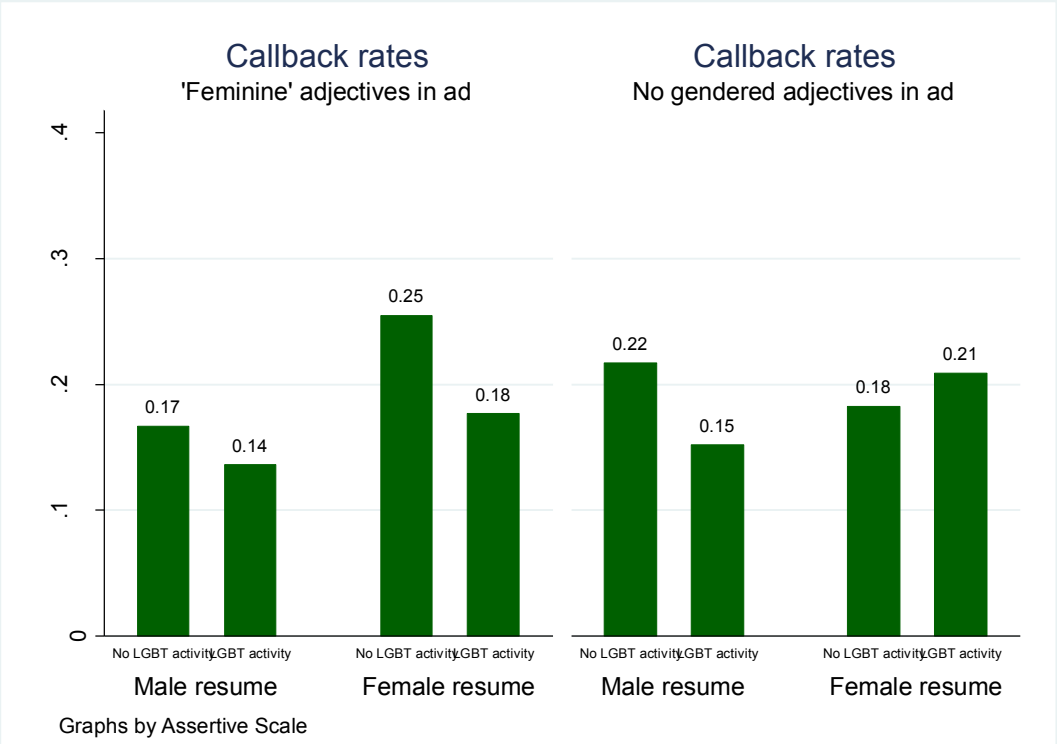


Figure 9*: Callback rate by sex and sexual orientation for job ads with feminine language or no gendered language. N=700 without gendered language. N=395 with feminine language.*

As Figure 9 shows, perceived-heterosexual women have the highest callback rate for jobs advertised with traditionally feminine language. However, Figure 10 demonstrates that among these jobs it is actually women who used traditionally masculine adjective that had the highest callback. Among perceived-heterosexual men, it is those who use traditionally feminine language that are more likely to be called back. This reflects the same pattern as in the data as a whole: even among jobs ad that explicitly use traditionally feminine adjectives, perceived-heterosexual women are more likely to be called back when they use traditionally masculine language than when they use feminine language. Once again, real employers have a different response to the use of masculine adjectives than respondents in a laboratory setting.

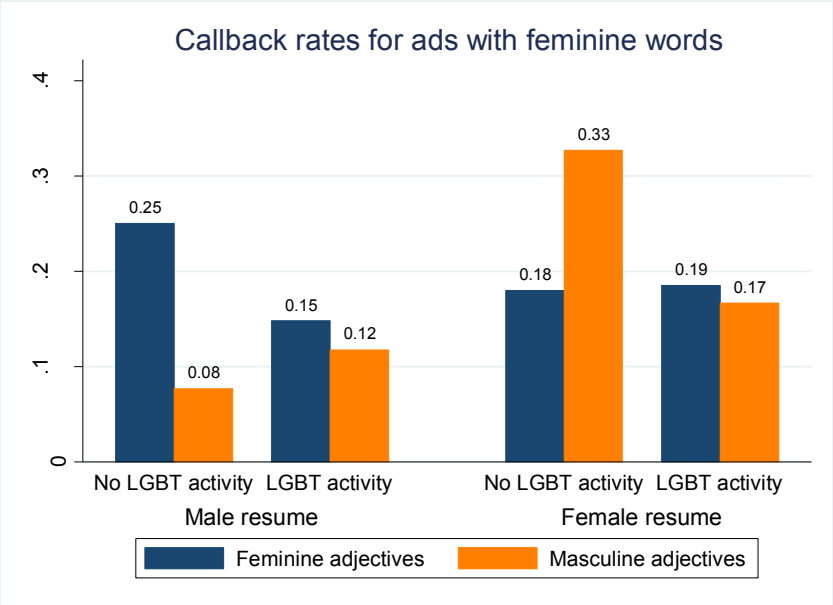


***Figure 10**: Callback rate by sex, sexual orientation, and adjective used on resume for job ads with feminine language. N=395.*

Differences in callback rates by sex and sexual orientation

While the main point of the audit study was to examine the backlash effect in a real labor market setting, it can also be used to examine difference in callback by sexual orientation. In an audit study in the United States, Tilcsik (2011) found a callback rate of 11.5% for male resumes with a non-LGBT extracurricular activity compared to 7.2% for those with an LGBT activity. Consistent with this, Figure 11 shows that among male resumes those with a non-LGBT activity had far more callbacks than those with the LGBT activity (*p* value of the difference is .035 with robust errors clustered at the city by industry level). The male resumes with the non-LGBT activity had a callback rate of 23% compared to only 17% for resumes with the LGBT activity. While the base response is much higher in this study, the difference in callback rates based on sexual orientation in Tilcsik (2011) is 37% while in this study it is 26%. The consistency between the two studies in this finding supports the credibility of the data overall. There was no difference between the LGBT and the non-LGBT activity among female resumes.

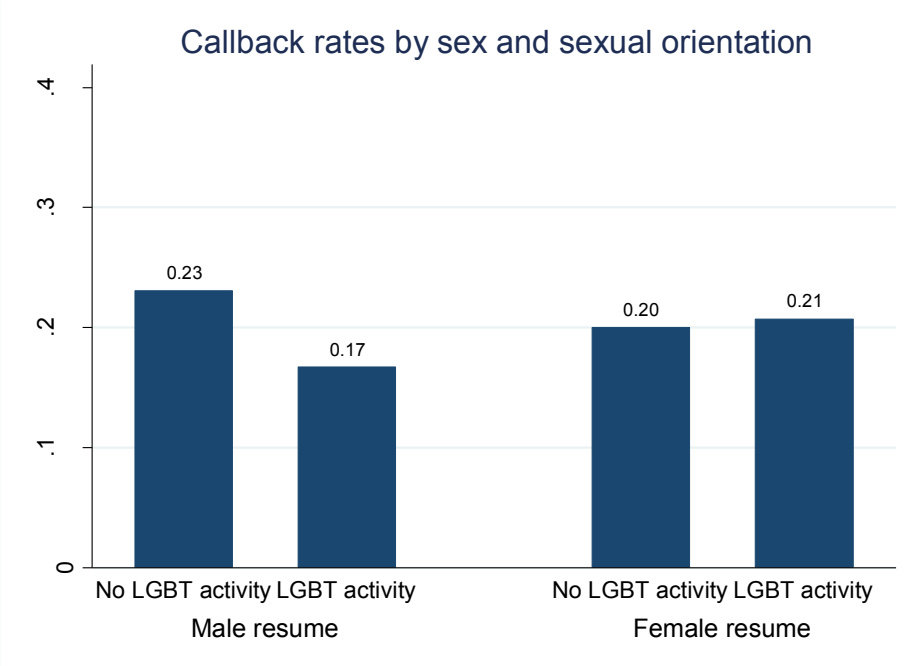


***Figure 11**: Callback rate by sex and perceived sexual orientation. N=1,334.*

Conclusions

Numerous laboratory studies in psychology have illustrated that when women act in traditionally masculine ways, men react with a "backlash" (Bowles, Babcock, and Lai 2006; Heilman and Chen 2005; Heilman, Wallen, Fuchs, and Tamkins 2004; Rudman and Glick 1999; Rudman 1998; Rudman and Glick 2001). However, it is left unclear how employers would respond to similar stimuli in a real labor market. Real employers may form different expectations about productivity from the use of traditionally masculine behavior or may value productivity and the threat to their identity differently than laboratory respondents.

In previous work, using masculine language on resumes provokes a backlash effect against perceived-heterosexual women (Gorsuch unpublished). However, when I took the question outside of the laboratory and applied for real jobs with the same manipulated resumes, I found the reverse pattern. Perceived-heterosexual women received **more** callbacks when they used masculine language than when they used feminine language. Even when the text of the job

ad used adjectives that are traditionally feminine, perceived-heterosexual women still benefited from using masculine language. Perceived-heterosexual men had the opposite pattern: they received fewer callbacks when they used traditionally masculine language.

These results are consistent with the idea that employers form expectations about productivity based on the use of masculine language that laboratory respondents either do not form or do not value as strongly as laboratory respondents. If employers want employees with bundles of attributes – an employee who is collegial **and** cooperative – it may be the applicant who sends the counter-stereotypical signal that will have the highest expected productivity. Even if both laboratory respondents and real employers respond negatively to norm-violating behavior, real employers may be overwhelmed by the change in expected productivity. This suggests that we must be cautious when extrapolating from results obtained in a laboratory setting to a real-world labor market.

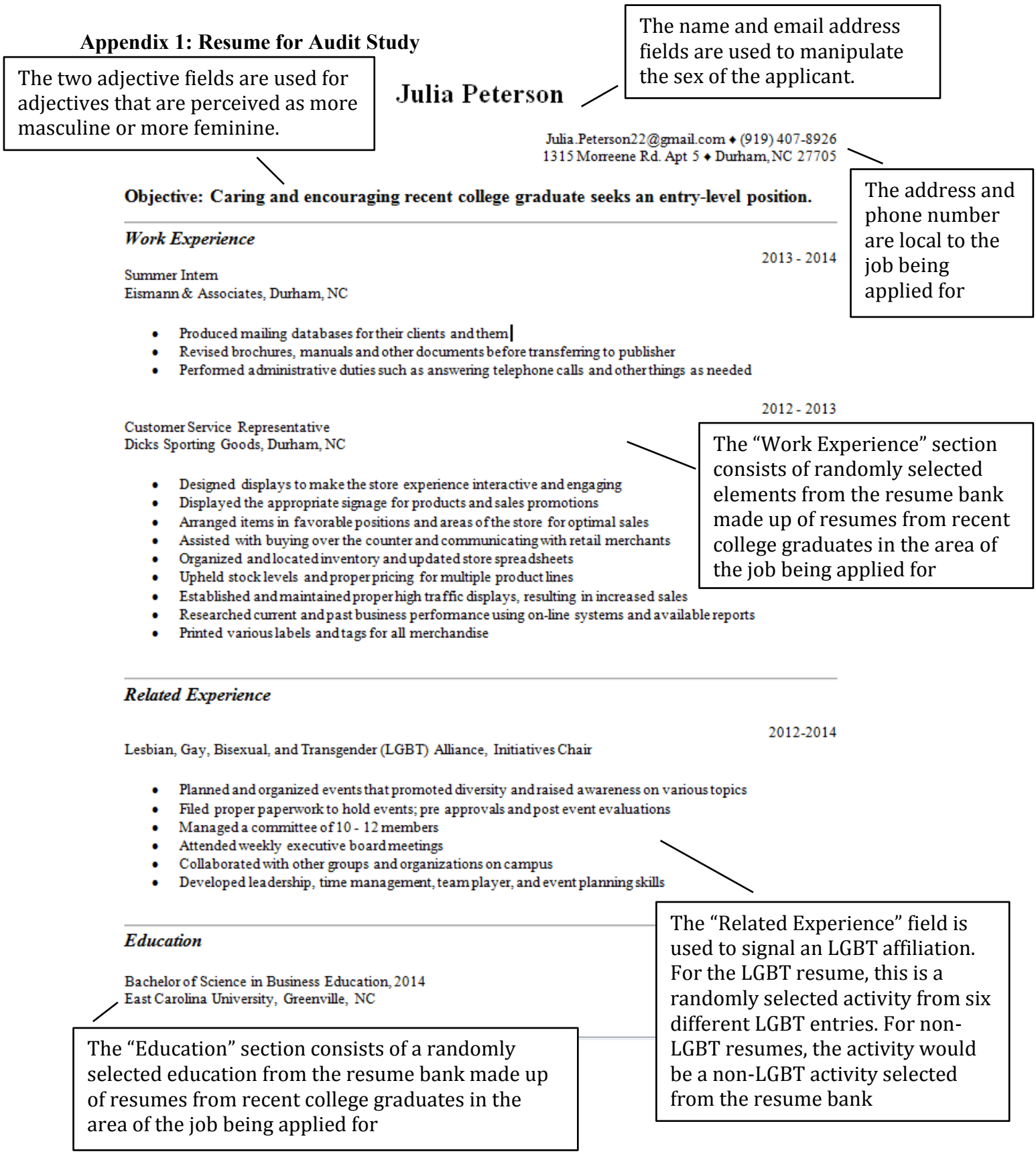
A finding consistent with the existing literature is that male resumes with an LGBT activity are penalized. The male resumes with an LGBT activity were called back at the lowest rate. This suggests that there is a substantial and persistent penalty for men who include an LGBT activity on their resume.

Finally, when job ads used traditionally masculine language, perceived-heterosexual women have very low callback rates and perceived-heterosexual men have very high callback rates. When perceived-heterosexual women use traditionally masculine language on their resume, there is no effect: they still have a callback rate less than half that of perceived-heterosexual men. This is consistent with that idea that employers are using traditionally masculine adjectives as a way of signaling for heterosexual male applicants rather than specific traits.

Appendix 1: Resume for Audit Study

The two adjective fields are used for adjectives that are perceived as more masculine or more feminine.

The name and email address fields are used to manipulate the sex of the applicant.



Julia Peterson

Julia.Peterson22@gmail.com ♦ (919) 407-8926
1315 Morreene Rd. Apt 5 ♦ Durham, NC 27705

The address and phone number are local to the job being applied for

Objective: Caring and encouraging recent college graduate seeks an entry-level position.

Work Experience

2013 - 2014

Summer Intern
Eismann & Associates, Durham, NC

- Produced mailing databases for their clients and them
- Revised brochures, manuals and other documents before transferring to publisher
- Performed administrative duties such as answering telephone calls and other things as needed

2012 - 2013

Customer Service Representative
Dicks Sporting Goods, Durham, NC

- Designed displays to make the store experience interactive and engaging
- Displayed the appropriate signage for products and sales promotions
- Arranged items in favorable positions and areas of the store for optimal sales
- Assisted with buying over the counter and communicating with retail merchants
- Organized and located inventory and updated store spreadsheets
- Upheld stock levels and proper pricing for multiple product lines
- Established and maintained proper high traffic displays, resulting in increased sales
- Researched current and past business performance using on-line systems and available reports
- Printed various labels and tags for all merchandise

The "Work Experience" section consists of randomly selected elements from the resume bank made up of resumes from recent college graduates in the area of the job being applied for

Related Experience

2012-2014

Lesbian, Gay, Bisexual, and Transgender (LGBT) Alliance, Initiatives Chair

- Planned and organized events that promoted diversity and raised awareness on various topics
- Filed proper paperwork to hold events; pre approvals and post event evaluations
- Managed a committee of 10 - 12 members
- Attended weekly executive board meetings
- Collaborated with other groups and organizations on campus
- Developed leadership, time management, team player, and event planning skills

The "Related Experience" field is used to signal an LGBT affiliation. For the LGBT resume, this is a randomly selected activity from six different LGBT entries. For non-LGBT resumes, the activity would be a non-LGBT activity selected from the resume bank

Education

Bachelor of Science in Business Education, 2014
East Carolina University, Greenville, NC

The "Education" section consists of a randomly selected education from the resume bank made up of resumes from recent college graduates in the area of the job being applied for

Appendix Figure 2: *Example of resume used in resume audit study. Each time a job was applied for, the work and education entries for the resume are randomly selected from a bank of publicly listed resumes from the area of the job being applied for. In total, there are 117 work histories and 42 educations. Three fields are used for the experimental manipulation (sex, masculine language, and LGBT affiliation); these fields are noted and described.*

Appendix 2: Balance of Audit Study

The following tables show the distribution of the eight manipulations across industry, city, and language used in the ad. The value in each cell shows the cell proportion. The chi-squared statistic of the test equal distribution is show for each table.

Manipulation	*Office*	*Retail*	*Food service*	*Labor or skilled trade*	*Total*
Female, no LGBT activity, and feminine adjective	5.02	2.40	2.85	2.17	12.44
Female, no LGBT activity, and masculine adjective	5.02	2.40	3.00	2.62	13.04
Male, no LGBT activity, and feminine adjective	4.57	3.07	3.07	1.72	12.44
Male, no LGBT activity, and masculine adjective	4.72	3.00	2.62	2.55	12.89
Male, LGBT activity, and feminine adjective	6.22	2.62	2.92	2.02	13.79
Male, LGBT activity, and masculine adjective	4.5	1.87	1.80	2.25	10.42
Female, LGBT activity, and feminine adjective	4.2	3.82	2.10	2.47	12.59
Female, LGBT activity, and masculine adjective	4.72	2.70	3.07	1.87	12.37
Total	38.98	21.89	21.44	17.69	100

Pearson chi2(21)= 24.3168 Pr = 0.278

Manipulation	*City 1**	*City 2*	*City 3*	*City 4*	*City 5*	*Total*
Female, no LGBT activity, and feminine adjective	6.37	1.8	1.65	0.6	2.02	12.44
Female, no LGBT activity, and masculine adjective	6.52	1.72	1.57	0.67	2.55	13.04
Male, no LGBT activity, and feminine adjective	6	1.65	1.87	0.67	2.25	12.44
Male, no LGBT activity, and masculine adjective	6.45	1.87	1.12	0.45	3	12.89
Male, LGBT activity, and feminine adjective	6.52	2.02	1.95	0.45	2.85	13.79
Male, LGBT activity, and masculine adjective	5.1	1.42	1.35	0.67	1.87	10.42
Female, LGBT activity, and feminine adjective	6.3	2.77	0.67	0.37	2.47	12.59
Female, LGBT activity, and masculine adjective	6.22	2.1	1.35	0.6	2.1	12.37
Total	49.48	15.37	11.54	4.5	19.12	100

Pearson chi2(28) = 24.1278 Pr = 0.675
**City 1 is the combination of three cities that share a Craigslist page*

Manipulation	*Feminine adjectives in ad*	*Masculine adjectives in ad*	*No gendered language in ad*	*Total*
Female, no LGBT activity, and feminine adjective	3.75	2.02	6.67	12.44
Female, no LGBT activity, and masculine adjective	3.9	1.87	7.27	13.04
Male, no LGBT activity, and feminine adjective	4.2	2.32	5.92	12.44
Male, no LGBT activity, and masculine adjective	3.9	2.47	6.52	12.89
Male, LGBT activity, and feminine adjective	4.05	2.62	7.12	13.79
Male, LGBT activity, and masculine adjective	2.55	2.17	5.7	10.42
Female, LGBT activity, and feminine adjective	4.05	2.1	6.45	12.59
Female, LGBT activity, and masculine adjective	3.15	2.4	6.82	12.37
Total	29.54	17.99	52.47	100

Pearson chi2(14) = 8.0917 Pr = 0.884

The following table show the p-value of the f-test that all coefficients are jointly zero after regressing an indicator for each manipulation on indicator variables for each education history (left column) and work history (right column). Only one regression out of 16 shows a p-value below .1.

	Education history	*Work history*
Female, no LGBT activity, and feminine adjective	0.7192	0.4682
Female, no LGBT activity, and masculine adjective	0.8724	0.15
Male, no LGBT activity, and feminine adjective	0.547	0.5694
Male, no LGBT activity, and masculine adjective	0.2311	0.5637
Male, LGBT activity, and feminine adjective	0.8053	0.5729
Male, LGBT activity, and masculine adjective	0.4528	**0.0311**
Female, LGBT activity, and feminine adjective	0.2823	0.1279
Female, LGBT activity, and masculine adjective	0.2429	0.5217

Works Cited

Akerlof, G. A., & Kranton R.E. (2000) "Economics and identity" *Quarterly Journal of Economics* , 115(3), 715-753.

Ahmed, Ali M., Lina Andersson, and Mats Hammarstedt (2013) "Are gay men and lesbians discriminated against in the hiring process?" *Southern Economic Journal,* 79(3), 565-585.

Amanatullah, E. T.; Morris, M.W. (2010) "Negotiating gender roles: Gender differences in assertive negotiating are mediated by women's fear of backlash and attenuated when negotiating on behalf of others." *Journal of Personality and Social Psychology*, 98(2), 256-267

Arrow, K. J. (1973), "The theory of discrimination", in O. Ashenfelter and A. Rees (eds.), *Discrimination in Labor Markets*, Princeton, NJ: Princeton University Press.

Babcock, L., Gelfand, M., Small, D., & Stayn H. (2006) "Gender differences in the propensity to initiate negotiations" D.D. Crèmer, M. Zeelenberg, J.K. Murnighan (eds.), *Social Psychology and Economics*, Lawrence Erlbaum, Mahwah, NJ, 239–259.

Babcock, L., & Laschever S. (2003) *Women Don't Ask* Princeton University Press, Princeton, NJ

Babcock, L., Laschever, S., Gelfand, M., & Small D. (2003) "Nice girls don't ask" *Harvard Business Review*, 81, 14–15.

Badgett, L. (1995) "The wage effects of sexual orientation" *Industrial and Labor Relations Review,* 48(4), 726-739.

Becker, G. (1957) *The Economics of Discrimination*. Chicago: University of Chicago Press,

Berg, N. & Lien, D. (2002) "Measuring the effect of sexual orientation on income: Evidence of discrimination?" *Contemporary Economic Policy*, 20(4), 394-414.

Berinsky, A., Huber G., & Lenz, G. (2012) "Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk" *Political Analysis*, 20(3), 351-368.

Bertrand, M. & Mullainathan, S. (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.

Black, D.A., Maker, H.R., Sanders, S., & Taylor, L.J. (2003) "The earnings effect of sexual orientation" *Industrial and Labor Relations Review*, 56(3), 449-469.

Black, D., Gates, G., Sanders, S., & Taylor, L. (2000) "Demographics of the gay and lesbian population in the United States: Evidence from available systematic data sources" *Demography* 37(2), 139-154.

Blandford, J.M. (2003) "The nexus of sexual orientation and gender in the determination of earnings" *Industrial and Labor Relations Review*, 56(4), 622-642.

Blau, F. D. & Kahn, L.M. (2007) "The Gender Pay Gap: Have Women Gone as Far as They Can?" *Academy of Management Perspectives* , 21(1), 7-23.

Blau, F.D. & Kahn, L.M. (2000) "Gender Differences in Pay" *The Journal of Economic Perspectives*, 14(4), 75-99.

Borghans, L., Duckworth, A.L., Heckman, J.J., & Weel B. (2008) "The Economics and Psychology of Personality Traits" *The Journal of Human Resources* 43(4) 972-1059

Bowles, H.R., Babcock, L., and Lai, L. (2007) "Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask" *Organizational Behavior and Human Decision Processes,* 103(1) May, 84-103.

Buhrmest, M., Kwang, T., & Gosling, S.D. (2011) "Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data?" *Perspectives on Psychological Science*, 6(1), 3-5.

Bureau of Labor Statistics (2012) "Women in the Labor Force: A Databook" http://www.bls.gov/cps/wlf-databook-2011.pdf

Census Bureau (2012) "File A" *Genealogy Data*: *Frequently Occurring Surnames from Census 2000* http://www.census.gov/genealogy/www/data/2000surnames/

Correll, S.J., Benard, S., & Paik, I. (2007). "Getting a job? Is there a motherhood penalty?" *American Journal of Sociology*, 112, 1297-1339.

Craigslist (2014) "Craigslist Factsheet" http://www.craigslist.org/about/factsheet

Cuddy, A., Fiske, S., and Glick, P. (2004) "When professionals become mothers, warmth doesn't cut the ice" *Journal of Social Issues*, 60(4), 701-718.

Gill, M.J. (2004) "When information does not deter stereotyping: Prescriptive stereotyping can foster bias under conditions that deter descriptive stereotyping" *Journal of Experimental Social Psychology,* 40, 619-632.

Gorsuch, Marina Mileo "Gender, Sexual Orientation, and Backlash in the Labor Market: A Laboratory Experiment," Duke University

Hall, E.V., & Livingston, R.W. (2012) "The hubris penalty: Biased responses to "Celebration" displays of black football players" *Journal of Experimental Social Psychology* 48(4), 899–904

Heckman, J.J. (1998) "Detecting Discrimination" *The Journal of Economics Perspectives* 12(2) 101-116.

Heilman, M.E., Block, C.J., & Martell, R.F. (1995) "Sex Stereotypes: Do They Influence Perceptions of Managers?" *Journal of Social Behavior and Personality* 10(4) 237-252

Heilman, M.E. & Chen, J.J. (2005) "Same behavior, different consequences: Reactions to men's and women's altruistic citizenship behavior" *Journal of Applied Psychology*, 90(3), 431-441.

Heilman, M.E., Wallen, A.S., Fuchs, D., & Tamkins, M.M. (2004) "Penalties for success: Reactions to women who succeed at male gender-typed tasks" *Journal of Applied Psychology*, 89(3), 416-427.

Heilman, M.E. (2001) "Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder" *Journal of Social Issues*, 57(4), 657-674.

Heilman, M.E. (1995) "Sex stereotypes and their effects in the workplace: What we know and what we don't know." *Journal of Social Behavior & Personality*, 10(6), 3-26.

Horton, J.J., Rand, D., & Zeckhauser, R.J. (2011) "The online laboratory: conducting experiments in a real labor market" *Experimental Economics* 14, 399-425

Indeed.com (2014) "About Indeed" http://www.indeed.com/about#our-company

Lahey, J.N. & Beasley R.A. (2009) "Computerizing Audit Studies" *Journal of Economic and Behavioral Organization* 70(3), 508-514.

Levitt, S.D. & Dubner, S.J. (2005) *Freakonomics* Harper Collins: New York City, NY.

Livingston, R.W. & Pearce, N.A. (2009) "The teddy-bear effect: Does having a baby face benefit black chief executive officers?" *Psychological Science*, 20(10), 1229-1236

Livingston, R.W., Rosette, A.S., & Washington, E.F. (2012) "Can an agentic black woman get ahead? The impact of race and interpersonal dominance on perceptions of female leaders" *Psychological Science*, 23(4), 354-358

National Center for Science and Engineering Statistics. (2013) "Table 11. Women as a percentage of all bachelor's recipients, by major field group: 1966–2010" *Science and Engineering Degrees: 1966–2010*

Pedulla, D. S. (2012). "The positive consequences of negative stereotypes: Race, sexual orientation, and the job application process. Working Paper 7" *Center for the Study of Social Organization*, http://www.tessexperiments.org/data/PedullaPaper.pdf

Peplau, L. & Fingerhut, A. (2004) "The paradox of the lesbian worker" *Journal of Social Issues*, 60(4), 719-735.

Phelps, E.S. (1972). "The statistical theory of racism and sexism". *American Economic Review* 62, 659–661.

Rudman, L. (1998) "Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management" *Journal of Personality and Social Psychology*, 74(3), 629-645.

Rudman, L. & Glick P. (1999) "Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers" *Journal of Personality and Social Psychology*, 77(5), 1004-1010.

Rudman, L. A., & Glick, P. (2001). "Prescriptive gender stereotypes and backlash toward agentic women" *Journal of Social Issues*, 57, 732–762.

Rudman, L. & Phelan, J.E. (2008) "Backlash effects for disconfirming gender stereotypes in organizations" *Research in Organizational Behavior*, 28, 61-79.

Savani, R. & von Stengel B. (2014) "Game theory Explorer: Software for the Applied Game Theorist" *Computational Management Science*

Small, D.A., Gelfand, M., Babcock, L., & Gettman, H. (2007) "Who goes to the bargaining table? The influence of gender and framing on the initiation of negotiation." *Journal of Personality and Social Psychology*, 93(4), 600-613.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. Austin, W.G., & Worchel, S. (eds.) *The social psychology of intergroup relations*, 33-47.

Tilcsik, A. (2011) "Pride and prejudice: Employment discrimination against openly gay men in the United States" *American Journal of Sociology*, 117(2), 586-626.