Gender, Sexual Orientation, and Backlash in the Labor Market

Marina Mileo Gorsuch

March 2014

I. Introduction

The persistent differences in earnings based on sex may result from the role of identity in forming gendered behavioral expectations. Economic and psychological theories suggest that men react negatively to women acting in traditionally masculine ways because their own masculine identity is threatened (Akerlof and Kranton 2000; Tajifel and Turner 1979). That is, when women follow the behavioral prescriptions for men, it adversely affects men's identity. Men therefore react with a "backlash" against women (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). This backlash may have an economic impact, because many behaviors needed for success in the workplace, such as negotiating for a salary or being self-promoting, are viewed as traditionally masculine.

If the backlash effect varies by sexual orientation, especially if certain subgroups are exempt from the backlash, it suggests a large role for identity in the backlash effects. While there is evidence that gay men earn less than heterosexual men, gay women appear to earn substantially more than heterosexual women (Black, Makar, Sanders, and Taylor 2003; Black, Gates, Sanders, and Taylor 2000; Blandford 2003; Berg and Lien 2002; Badgett 1995). One explanation consistent with this pattern is that heterosexual women are expected to behave according to gender prescriptions, but gay women are not. This is consistent with the theories of social identity formation, where group identity is formed in contrast only to the "relevant comparison group" rather than all other groups (Tajifel and Turner 1979, p. 41).

This paper manipulates the sexual orientation and the use of more masculine adjectives of fictional job applicants to explore the role that identity plays in the backlash effect. The project first seeks to establish if resumes that use masculine language inspire backlash against female job

applicants in a laboratory setting and if this backlash varies by the sexual orientation of the applicant. The second stage of this project will examine if the laboratory findings extend to an audit study when these resumes are used to apply to real job ads. This project connects theories of social identity formation and evidence of the backlash effect to differences in labor market outcomes based on sex. It highlights a possible mechanism for earnings differences between heterosexual men and women: the role of identity in forming gendered expectations.

The results of the laboratory experiment indicate that heterosexual women who use masculine language on resumes are subject to backlash, but gay women who use masculine language are not. Men, both gay and heterosexual, do not experience any backlash when they use masculine language on their resumes. Application of a finite mixture model reveals two latent classes, one that has a strong backlash effect and one that has none. Respondent characteristics, especially being male and holding more conservative political views, predict a higher posterior probability of being in the latent class with the strong backlash effect.

This article is organized as follows. Section II motivates the study by describing earnings differences based on sex and sexual orientation and discussing two competing theoretical models that could explain these differences. Section III describes the laboratory experiment that examines if resumes that use masculine adjective inspire backlash. Section IV describes the econometric techniques used to analysis the data. Section V reports the results of the experiment, which found that heterosexual women are subject to backlash when their resumes use masculine adjectives, but gay women are not. Latent class analysis suggests that the overall effect is the average between one subgroup with both a strong backlash reaction and another subgroup without a backlash reaction. Section VI describes the audit study and reports the results. Section VII discusses the conclusions and implications of this study.

II. Background and Motivation

A. Earnings Differences and Backlash in the workplace

There are substantial differences in earnings based on sex and sexual orientation. Women who are full-time year-round workers earn 20% less than their male counterparts (Bureau of Labor Statistics 2012). Differences in observed human capital do not fully account for these differences (for example, see Blau and Kahn (2000) and Blau and Kahn (2007)).

Using population representative surveys that identity gay and lesbian people, numerous studies have found that gays and lesbians have different labor force outcomes than heterosexuals. Gay men earn less than heterosexual men and lesbians earn more than heterosexual women, even after adjusting for marital status, children, and occupational choice (Black, Makar, Sanders, and Taylor 2003; Black, Gates, Sanders, and Taylor 2000; Blandford 2003; Berg and Lien 2002; Badgett 1995).

In particular, Black, Makar, Sanders, and Taylor (2003) find that at the mean level of experience for lesbian and bisexual women earn 20% more than heterosexual women. Gay and bisexual men earn 16% less than heterosexual men. Notably, at low levels of experience gay women earn less than heterosexual women, but the return to experience is higher resulting in a higher average at the mean. The authors argue that these results are consistent with a human capital story, where gay men invest less in market orientated human capital than straight men and gay women invest more than straight women. The authors note one inconsistency in this theory - gay men have more education than straight men, despite their fathers having the same education as straight men, which the authors note is not entirely consistent with gay men acquiring less human capital than straight men.

Men negotiate salaries more often than women, contributing to the persistent earnings difference between men and women (Babcock, Laschever, Gelfand, and Small 2003; Babcock and Laschever 2003). This pattern is observed in both laboratory experiments and in observational studies. For example, a survey of recent MBA graduates revealed that 57% of men negotiated their initial salary, compared to only 7% of women (Babcock, Laschever, Gelfand, and Small, 2003; Babcock and Laschever 2003).These results hold up in the laboratory as well; Babcock, Gelfand, Small, and Stayn (2006) found gender differences in willingness to negotiate in a laboratory experiment. They informed their study participants they would be paid between $3 and $10 for playing Boggle, a word game. Afterwards, they offered each participant $3 in private. If the participants asked for more money, they were paid what they requested. Men were much more likely to request to be paid more; this difference persisted even when respondents were explicitly told the amount was negotiable (Small, Gefland, Babcock, and Gettman 2007).

The different rates of negotiation may in part be a logical response to the fact that women are more harshly punished by potential employers for attempting to negotiate (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). 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 (Amanatullah and Morris 2010).

Likewise, other experiments show that women are punished for 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).

This project explores the backlash effect as alternative mechanisms for the differences in earnings based on sex and sexual orientation. The backlash effect may decrease women's earnings, either directly (through the backlash itself) or by influencing how women act. That is, women may earn less than men if they act in traditionally masculine ways because they are viewed as less desirable colleagues; likewise, women may earn less if they try to avoid the backlash by not acting in traditionally masculine ways because they miss out on promotions or raises.

Moreover, the backlash effect could also explain the differences in earnings based on sexual orientation if we take into account the role of identity. I describe in the next section how theories of social identity suggest that not all women or all men are held to the same behavioral standard. Rather, only certain subgroups are used as a contrast group in forming an identity and therefore only these subgroups are subject to behavioral prescriptions. That is, behavior that may result in a backlash for heterosexual women may be neutral or beneficial to gay women. Likewise, gay men could be viewed as an out-group by heterosexual men and therefore subject to behavioral prescriptions similar to heterosexual women.

A model of statistical discrimination can also explain earning differences, but would have different predictions regarding the impact of acting in traditionally masculine way than theories

of social identity. A model of statistical discrimination suggests that the differences in employment outcomes result from the employer's belief that differences in sex and sexual orientation are associated with different levels of desired personality characteristics. As I show in subsection C, this model predicts that when gay and heterosexual women act in traditionally masculine ways, the employer's perception of them should move in the same direction (that is, benefit or harm both groups). Therefore, while the backlash effect predicts that traditionally masculine behavior will be harmful to heterosexual women but neutral or beneficial to gay women, the statistical discrimination model predicts that this behavior will either be beneficial or harmful to *both* gay and heterosexual women. Likewise, the backlash effect predicts that traditionally masculine behavior will be beneficial to heterosexual men but harmful to gay men, while the statistical discrimination model predicts that this behavior will either be beneficial or harmful to both gay and heterosexual men

B. Prescriptive stereotypes and social identity

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). In the above example, men defined themselves in contrast to women. But, perhaps heterosexual men define themselves in contrast to heterosexual women or gay men. That is, while gay women are clearly a subset of women and are a different group than heterosexual men, they may not be the group that heterosexual men define themselves against (or that heterosexual women define themselves with). If in fact gay women are not a relevant comparison group to heterosexual men, heterosexual men would suffer identity loss when gay women follow male behavioral prescriptions. Gay women would therefore not be punished for deviating from prescribed behavior for women as a whole. In contrast, gay men could be viewed as an out-group by heterosexual men and therefore subject to behavioral prescriptions similar to heterosexual women. Gay men may be punished for engaging in traditionally male behaviors and rewarded for traditionally female behaviors.

Indeed, there is initial evidence that lesbians may be exempt from the "motherhood penalty" that heterosexual women experience. In a laboratory experiment, Peplau and Fingerhut

(2004) asked respondents to read a brief description of a consultant at a major consulting firm. The description identified the consultant's sex, sexual orientation, and parenthood status. Heterosexual women with children were rated as less competent and less career oriented than heterosexual women without children, whereas men with children were viewed as equally or more career oriented and competent than men without children, which replicates previous findings of a motherhood penalty for women (Correll, Benard, and Paik 2007; Cuddy, Fiske, and Glick 2004). However, lesbians did not suffer the motherhood penalty and were viewed as competent and career oriented as heterosexual men both with and without children. This suggests that people have different expectations for gay female workers than heterosexual female workers and that these expectations align more closely with expectations for heterosexual men. These differences could derive from stereotypes about gay women being more masculine, different expectations about specialization within same-sex couples, or different expectations about gay women's preferences for career attachment.

Examining if backlash varies by sexual orientation demonstrates if backlash may play a role in the observed earnings differences based on sex and sexual orientation. If gay women are exempt from the backlash effect that decreases heterosexual women's earnings and gay men subject to it, this could explain part of the observed earnings differences.

C. Descriptive stereotypes or statistical discrimination

In contrast to prescriptive stereotypes, descriptive stereotypes are stereotypes about how women **are**. These stereotypes may cause firms to anticipate a "lack of fit" between a 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 women have a particular distribution of personality characteristics (Arrow 1973; Phelps 1972).

In the case of statistical discrimination, acting in a traditionally masculine manner may be a signal of having the desired masculine traits.

Consider a model of statistical discrimination with signaling where employers want to hire people who have a personality trait that is perceived as masculine, such as being competitive. Employers hire women from two different groups (either gay or heterosexual: $g = G, H$). The applicant's true type (competitive or demure: $\tau_i = C, D$) is not observed by the employer. Employers do observe a signal of the applicant's true type, for example if an applicant describes themselves as an "aggressive go-getter" or a "nurturing team leader" (signal of being aggressive or nurturing: $T_i = A, N$) is observable to the employer. The cost of sending an aggressive signal has some distribution, where the average cost is higher for demure applicants than competitive ones ($c_D > c_C$). The applicant will send the aggressive signal if the increase in probability of being hired multiplied by their expected present value of the job exceeds c_i. On average the competitive women have lower cost to sending the aggressive signal, so they will be more likely to send an aggressive signal ($\mu_{C,g} > \mu_{D,g}$ for $g = G, H$).

The employer has a prior belief about the proportion of competitive personality types among gay and heterosexual women (π_G, π_H). The employer also knows the probability of an aggressive signal from competitive and demure applicants ($\mu_{C,g}, \mu_{D,g}$). The employer forms a posterior belief about the probability of an applicant being demure after observing their group and their signal. For example, the following equations show the posterior probability of being demure for a gay woman given her signal:

$$P(D \mid A \cap G) = \frac{\mu_{D,G}(1-\pi_G)}{\mu_{D,G}(1-\pi_G) + \mu_{C,G}\pi_G} \qquad (1)$$

$$P(D \mid N \cap G) = \frac{(1-\mu_{D,G})(1-\pi_G)}{(1-\mu_{D,G})(1-\pi_G)+(1-\mu_{C,G})\pi_G} \qquad (2)$$

With algebraic manipulation, it is straightforward to show that if $\mu_{C,G} > \mu_{D,G}$, then $P(D \mid A \cap G) < P(D \mid N \cap G)$. That is, the firms believe that gay women who send an aggressive signal are less likely to be demure than gay women who send a nurturing signal. (For brevity, I omit the proof that an equilibrium exists for a range of prior beliefs).

Likewise, firms believe that heterosexual women who send an aggressive signal are less likely to be demure than heterosexual women who send a nurturing signal: (3)

$$P(D \mid A \cap H) = \frac{\mu_{D,H}(1-\pi_H)}{\mu_{D,H}(1-\pi_H)+\mu_{C,H}\pi_H} < \frac{(1-\mu_{D,H})(1-\pi_H)}{(1-\mu_{D,H})(1-\pi_H)+(1-\mu_{C,H})\pi_H} = P(D \mid N \cap H)$$

The change in the firm's perception of a woman's probability of being demure when she sends an aggressive signal compared to when she does not depends on the group she is in. For heterosexual women, the gain is $\frac{\mu_{D,H}(1-\pi_H)}{\mu_{D,H}(1-\pi_H)+\mu_{C,H}\pi_H} - \frac{(1-\mu_{D,H})(1-\pi_H)}{(1-\mu_{D,H})(1-\pi_H)+(1-\mu_{C,H})\pi_H}$ whereas for gay women it is $\frac{\mu_{D,G}(1-\pi_G)}{\mu_{D,G}(1-\pi_G)+\mu_{C,G}\pi_G} - \frac{(1-\mu_{D,G})(1-\pi_G)}{(1-\mu_{D,G})(1-\pi_G)+(1-\mu_{C,G})\pi_G}$. The difference in the gain to using masculine adjectives depends on $\pi_H, \pi_G, \mu_{D,G}, \mu_{C,G}, \mu_{D,H}$, and $\mu_{C,H}$.

If firms' preference for competitive personality traits is monotonically increasing, an applicant who is less likely to be demure is always more attractive to the firm. (The possibility that the preference may be non-monotonic is relevant in a future extension of this project when race in manipulated instead of sex.)

As set up here, the aggressive signal is a sign of being competitive, but it just as easily could signal other business traits, such as being committed to the labor force. While not shown here, using masculine adjective could also be a negative signal – it could indicate lack of

understanding of social norms. The results are the reverse of those shown here: the aggressive signal harms both gay and heterosexual women.

Equations 1, 2, and 3 were set up to apply to women. However, they can just as easily apply to men. In this case, employers have a prior belief about how competitive gay and heterosexual men are. If sending an aggressive signal is more costly on average to demure applicants within each group, they will be less likely to send an aggressive signal for a given payoff. Employers will form a posterior belief that an applicant is demure based on their group and their signal. As in Equations 1, 2, and 3, sending an aggressive signal decreases the employer's belief that both gay and heterosexual men are demure. The magnitude of the difference in the perceived probability of being demure between the nurturing and aggressive signal will depend on the employer's prior beliefs and the equilibrium probabilities of sending the aggressive signal for each type by group combination.

Therefore, the signaling model predicts that even if employers believe that the distribution of a desired trait differs based on sex and sexual orientation, signaling that trait should move the probability of being hired in the same direction for gay and heterosexual women and in the same direction for gay and heterosexual men. (There are unlikely exceptions, such as a corner solution; for example if $\pi_i = 1$ or 0, then the signal has no impact for group i and no one in group i will send it.) Therefore, while the backlash effect predicts that an aggressive signal would harm heterosexual women and benefit gay women, statistical discrimination predicts that it would either harm or benefit both gay and heterosexual women. Likewise, the backlash effect predicts that an aggressive signal would benefit heterosexual men and harm gay men, while statistical discrimination predicts that it would either harm or benefit both gay and heterosexual men.

D. Testable hypotheses of these two models

By manipulating sex, sexual orientation, and how masculine an applicant's resume is, this project tests the role of identity in the earnings differences between heterosexual men and women. If the reaction to violating stereotypes for men and women varies by sexual orientation, this strongly suggests that identity plays a key role in the backlash effect.

If employers hold descriptive stereotypes, they will anticipate different behaviors based on sex and sexual orientation (see Ahmed, Andersson, and Hammarstedt (2013) for a discussion). If these stereotypes are the driving mechanism for earnings differences based on sex and sexual orientation, signaling masculinity should have the same directional effect for both gay and heterosexual women, although of different magnitudes.

However, if *prescriptive* stereotypes, how a group of people ought to be, varies by sexual orientation, this suggests that only certain subgroups are the relevant contrast group. If there is a group that is not the relevant out-group for heterosexual men (gay women, for example) and this group is exempt from the backlash effect, this is strong evidence for the role of identity in the backlash effect and earnings differences based on sex.

This theoretical framework leads to three testable hypotheses:

Consistent with signaling	***Consistent with social identity***
When female applicants perform masculine actions, both gay women and heterosexual women will be viewed more positively (or less positively) than those who do not.	When heterosexual female applicants perform masculine actions, they will be viewed less positively than those who do not. When gay female applicants perform masculine actions, they will be viewed as positively or more positively than those who do not.
When male applicants perform masculine actions, both gay men and heterosexual men will be viewed more positively (or less positively) than those who do not.	When heterosexual male applicants perform masculine actions, they will be viewed more positively than those who do not. When gay male applicants perform masculine actions, they will be viewed less positively than those who do not.

III. Laboratory Experiment

To test the role of identity or prescriptive stereotypes versus statistical discrimination, I test if resumes with more or less masculine language are viewed differently based on the sex and sexual orientation of the applicant.

I created ten base resumes formed as a compilation of resumes from recent college graduates who publically listed their resume on Indeed.com, similar to the compilation resumes used in Bertrand and Mullainathan (2004). The ten base resumes were created from randomly selected resumes of people with a recently awarded bachelor's degree in biology from those listed on Indeed.com on a specific date (Oct 30th, 2013) in Durham, NC. Each compilation resume is created from randomly selected elements of each randomly selected resume. That is, a resume contains the university name from one resume, job title and description from another, another job from a third, etc. An annotated example resume is included in Appendix 1.

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 more or less masculine. In the pre-test, one group of respondents on Mechanical Turk (described in more detail below) 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 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.

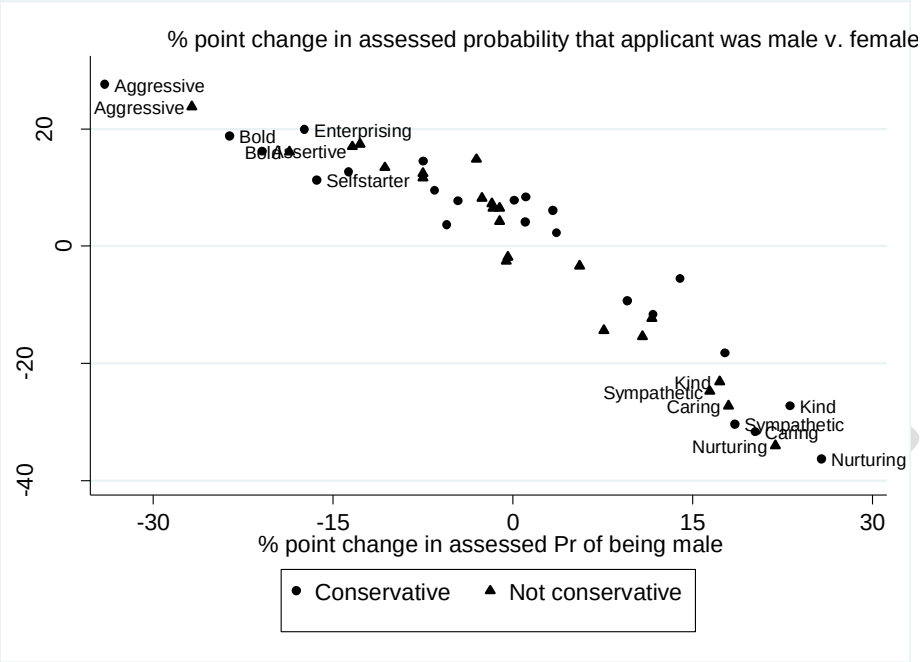


Figure 1*: Results from a pre-test of adjectives. One group reported how likely it was that the applicant was female (x-axis) while another group reported how likely to applicant was male (y-axis). Each observation was demeaned by respondent; each dot represents the average demeaned probability. The adjectives with the strongest gender associations are labeled. "Conservative" refers to those with conservative answers to questions about gender roles.*

In the pre-test, respondents with more conservative[1] answers to questions about gender roles based on the General Social Survey tended to have more extreme assessments of the adjectives. For example, an applicant who used "Aggressive" was viewed by conservative respondents as 28 percentage points more likely to be male than the average adjective and 34 percentage points less likely to be female. Non-conservative respondents thought that an applicant who used the word "Aggressive" was 24 percentage points more likely to be male and

[1] A respondent was considered "conservative" if 1. They agreed or strongly agreed with the statement that "A preschool child is likely to suffer if his or her mother works." Or 2. They disagreed or strongly disagreed with the statement "A working mother can establish just as warm and secure a relationship with her children as a mother who does not work."

27 percentage points less likely to be female. This again suggests that the manipulation is effective; those who hold more conservative views about gender roles have more extreme associations with the manipulations.

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 am planning to explore if backlash varies by race in the future. 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 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 recruited respondents on Amazon Mechanical Turk (MTurk) to assess the resumes on personality characteristics and level of perceived skill. MTurk is a marketplace that pays piece rate for small tasks completed online. Other studies have shown that respondents from these samples are not population representative (Mturk is skewed towards younger respondents), but that they are closer to a population representative distribution of race than typical college campus recruitment methods, and their responses are reliable (Buhrmester, Kwang, and Gosling 2011; Berinsky, Huber, and Lenz 2012; Horton, Rand, and Zeckhauser 2011). I restricted to respondents with an IP addresses in the United States and who had already successfully completed a specified number of tasks for other employers on MTurk.

The respondents were told they were helping a company sort resumes for an entry-level position for a college graduate who majored in biology. Unlike most fields, bachelor degrees in biology are neither over- nor under-represented among women; women earned 57.2% of all bachelor degrees in 2010 and 57.8% of all bachelor degrees in biological and agricultural sciences (National Center for Science and Engineering Statistics 2013). Concealing that the task is part of a research study reduced the chance that respondents alter their behavior to avoid appearing discriminatory or to "help" the researcher obtain the desired results. This concern is especially pertinent for workers on Mechanical Turk who appear to be more likely than other respondents to attempt to guess the desired interpretation behind experiments and alter their behavior correspondingly (Berinsky, Huber, and Lenz 2011).

Each respondent assessed ten resumes made up of two filler resumes and eight manipulated resumes. The eight manipulated resumes varied on sex, sexual orientation, and type of language used: gay, male, masculine adjective; gay, male, feminine adjective; heterosexual, male, masculine adjective; heterosexual, male, feminine adjective; gay, female, masculine

adjective; gay, female, feminine adjective; heterosexual, female, masculine adjective; and heterosexual, female, feminine adjective. The filler resumes helped disguise the manipulation by using neutral adjectives (flexible, adaptable, talented, and reliable) and reducing the proportion of resumes that are identifiable as a gay applicant. The two filler resumes were always presented first to the respondent. The following eight resumes were presented in a random order.

The respondents were asked to view each resume and then evaluate the job candidate on a number of characteristics. The survey was designed so that respondent had to stay on each resume page for a minimum of one minute. After viewing the whole resume for one minute, the respondents then rated the usefulness of the applicant's work and extracurricular activities on pages where they were shown only that section of the resume. The respondent then evaluated the applicant's personality, how strongly they would recommend the applicant, how willing the respondent would be to work with the applicant, the applicant's commitment to job, recommended salary, and likelihood of success (based on the outcome measures in Bowles, Babcock, and Lai (2006) and Correll, Benard, and Paik (2007)). After rating all ten resumes, respondents were asked for demographic information and questions about political ideology, including a question on the respondent's views towards lesbian, gay, bisexual, or transgender people and on gender roles based on questions from the General Social Survey.

There were ten versions of the questionnaire, so that each base resume was paired with each identity (the combination of sex, sexual orientation, and adjective) once. For example, in one version of the questionnaire the first base resume was a gay man with masculine adjectives. In another version, the same base resume was a gay man with feminine adjectives. While each respondent only sees each base resume once, each base resume is used with all of the

manipulations over the ten versions of the questionnaire. This experimental design allows for the inclusion of resume fixed effects and respondent fixed effects.

To increase the quality of the data analyzed, I use numerous methods to exclude respondents who could be a computer program answering questions randomly or a person who was not paying attention to the survey. First, I restricted respondents to only those with high accuracy on previously submitted tasks on MTurk. Second, I incorporated an "attention check" question in the survey. The directions above the question instructed the respondent to ignore the text of the question and instead type a specific word in the text box. If a respondent was clicking randomly or not reading the directions, they would not type the word into the text box. 79 respondents failed to type the correct word in the text box and are excluded from the analyses. Third, I asked respondents to indicate their sex in a text box; eight respondents put their age in the text box instead of their sex and one put a series of nonsensical letters – these respondents are excluded from the analyses. Fourth, I asked respondents to indicate if the applicant was male, female, or indeterminate; 24 respondents said the resume was of indeterminate sex or incorrectly identified the applicant's sex more than one time, so were excluded. Finally, if the respondent spent less than 26.2 minutes (the 5^{th} percentile) on the survey, they were excluded; this affected 30 respondents. Many of the excluded respondents were excluded for failing more than one of the quality checks. In total, 878 respondents passed all of the quality checks.

IV: Empirical Framework

To test if resumes with more or less masculine language are viewed differently based on the sex and sexual orientation of the applicant, I will first examine if resume can inspire the "backlash" effect documented in the existing literature. To do this, I examine the difference in the difference between men and women and between how resumes with masculine adjectives are

viewed compared to those with feminine adjectives. The outcome variable ($y_{i,r}$) is the respondent's (r) assessment of each applicant's (i) resume.

$$y_{i,r} = \alpha + \beta * \text{I(female applicant}_i) + \theta * \text{I(masculine adjective}_i) + \delta * \text{I(female applicant}_i) * \text{I(masculine adjective}_i) + \sum_{r=1}^{n-1} \tau_r * \text{I(respondent} = r) + \sum_{k=1}^{7} \omega_k * \text{I(base resume}_i = \text{k}) + \eta_{i,r} \quad (4)$$

When Equation 4 is estimated on the resumes without the LGBT activity, $\hat{\theta}$ estimates how using masculine adjectives impacts the perceived hire-ability of men and $\hat{\theta} + \hat{\delta}$ estimates how using masculine adjectives impacts the perceived hire-ability of women. $\hat{\delta}$ is the estimate of the difference in the difference between men and women in the perceived hire-ability between using masculine and feminine adjectives. To be consistent with the findings in the literature, to be consistent with the literature $\hat{\theta} + \hat{\delta}$ will be negative, indicating that women who use masculine adjective are perceived more poorly than those that use feminine adjectives. Additionally, to be consistent with the literature, $\hat{\delta}$ will be negative, indicating that the difference between women who use masculine adjectives relative to women who use feminine adjectives is more negative than the difference between men who use masculine adjective and those that use feminine adjectives.

To extend the analysis to include sexual orientation, Equation 5 employs the same diff-in-diff approach, but compares women's resumes with the LGBT activity to those without. When Equation 5 is estimated on women, $\hat{\lambda}$ estimates the difference between the masculine and feminine adjectives for the non-LGBT women's resumes. $\hat{\lambda} + \hat{\phi}$ is the estimate of the difference between the masculine adjective resume and feminine adjective resume for women with the LGBT activity. The difference in this difference $\hat{\phi}$ estimates if the women with the LGBT activity are impacted less than those without by using masculine adjectives relative to feminine adjectives.

$$y_{i,r} = \alpha + \gamma * \mathrm{I}(\text{LGBT activity}_i) + \lambda * \mathrm{I}(\text{masculine adjective}_i) + \phi * \mathrm{I}(\text{LGBT activity}_i) * \mathrm{I}(\text{masculine adjective}_i) + \sum_{r=1}^{n-1} \zeta_r * \mathrm{I}(\text{respondent} = r) + \sum_{k=1}^{7} \Psi_k * \mathrm{I}(\text{base resume}_i = \mathrm{k}) + \varepsilon_{i,r} \quad (5)$$

Finally, to examine if any backlash effect found in Equations 4 and 5 is actually the average between two latent classes, I utilize finite mixture model analysis. This maximum likelihood estimation estimates a set of coefficients for each latent class and also the proportion of the sample that falls in each latent class. I then examine if respondent characteristics predict the posterior probability of belonging to each latent class.

V. Results

A. Characteristics of Respondents

The following graphs show the demographic characteristics of the respondents recruited through Mechanical Turk. The first two graphs show that the respondents tend to be young and well-educated: over 60% of the sample is under 35 and 50% has a Bachelor's degree or higher. The sample is represents both men and women well, with 52.5% of the sample being female.

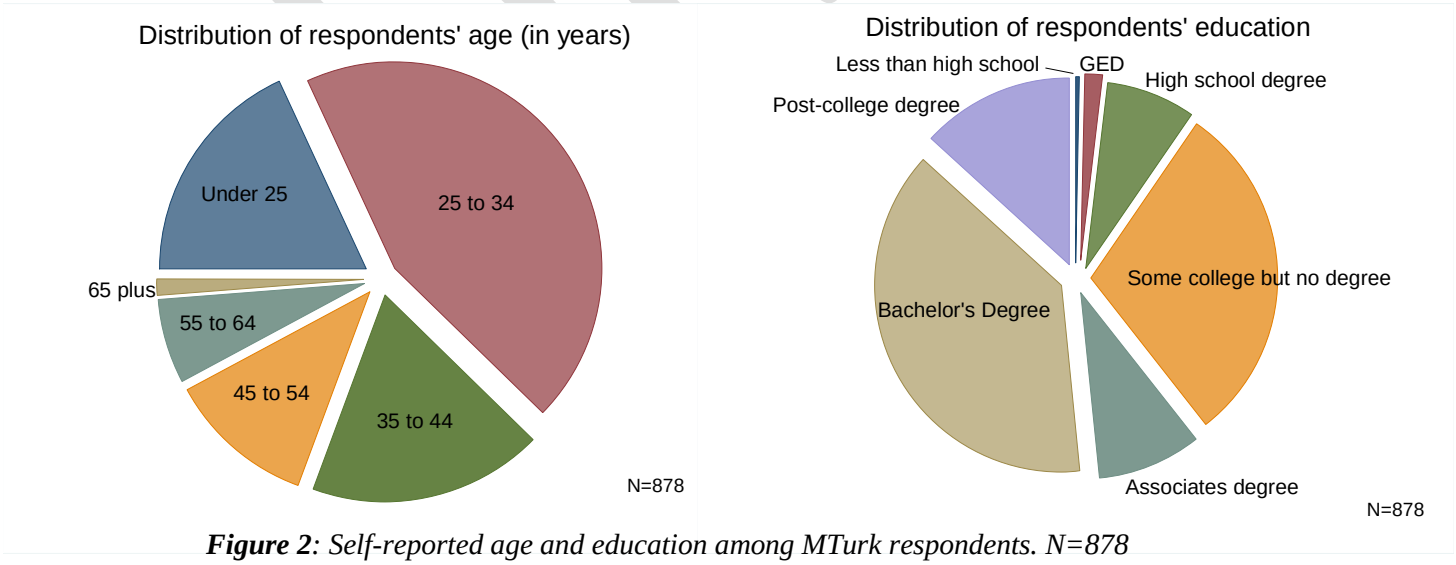


***Figure 2**: Self-reported age and education among MTurk respondents. N=878*

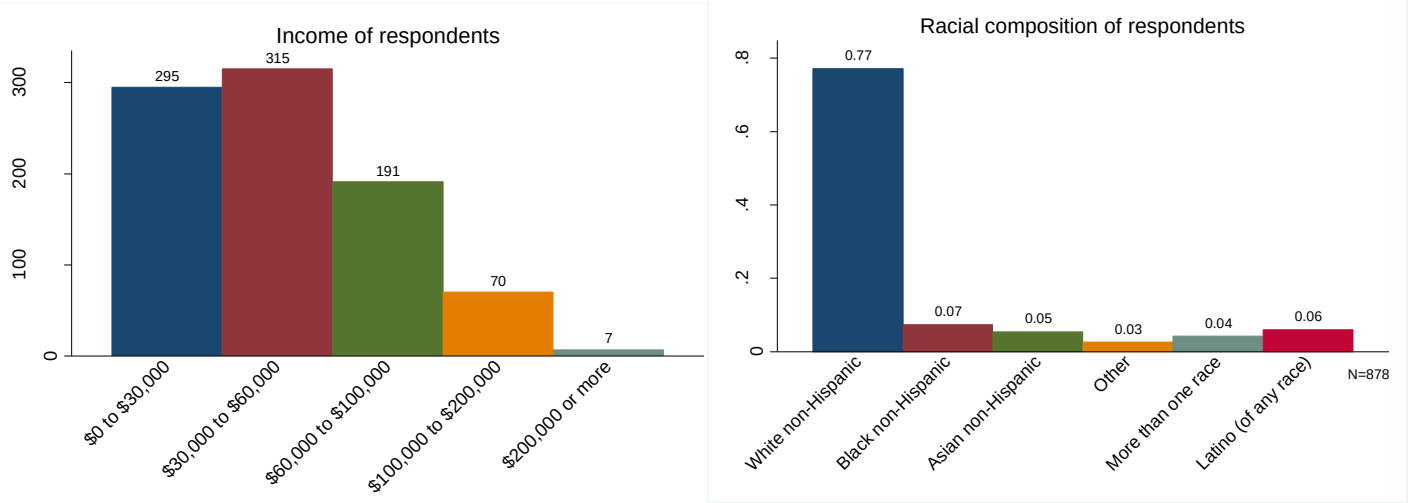


***Figure 3**: Self-reported household income and race/ethnicity among MTurk respondents. N=878*

The majority of respondents are non-Hispanic white, but with sizable portions that are non-Hispanic African American (7%), Asian (5%), and multi-racial (4%). The respondents hold predominantly liberal views; 74% of respondents agree or strongly agree that same-sex marriage should be legal. The vast majority of respondents live in households with an income of $60,000 or less.

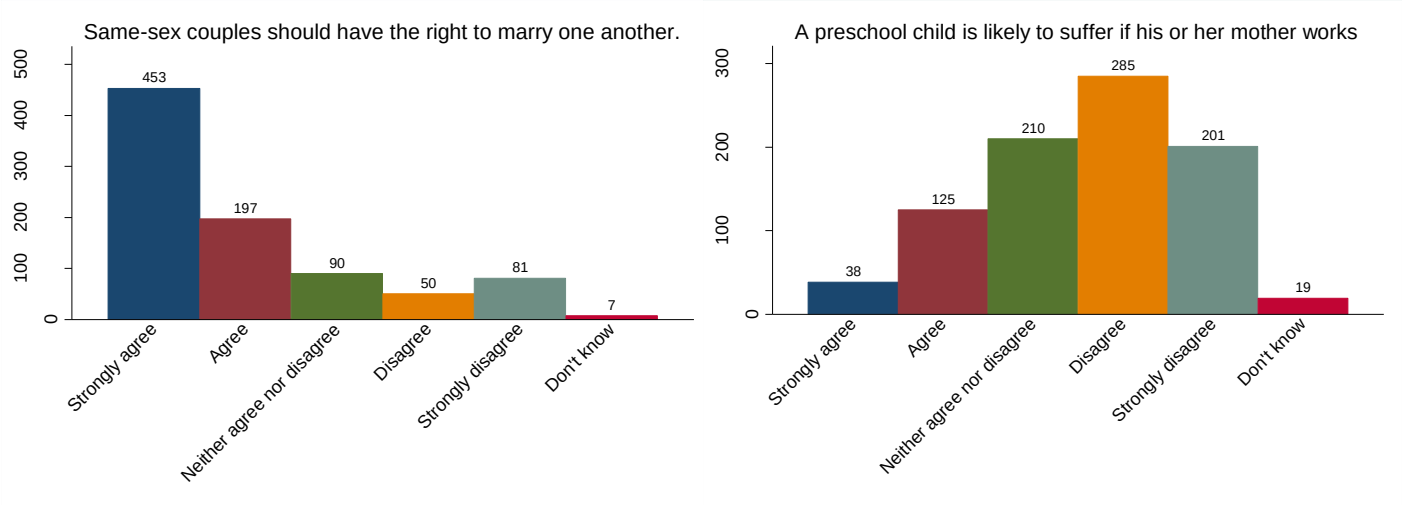


***Figure 4**:MTurk respondents' answers to questions from the General Social Survey about gender roles and same-sex marriage. N=878*

B. Evaluating if the treatment was salient to respondents

Prior to exploring the backlash effect, I first test if the use of adjective in the objective statement was salient to the respondent. I examine if the respondents evaluation of the applicant's personality is affected by the use of masculine adjectives (relative to feminine adjectives). Respondents evaluated eleven different personality characteristics on how well they described the applicant (from 0 to 100). The following graphs show that within each sex by sexual orientation subgroup, the use of the masculine adjectives makes an applicant appear less kind and passive and more pushy and confident.

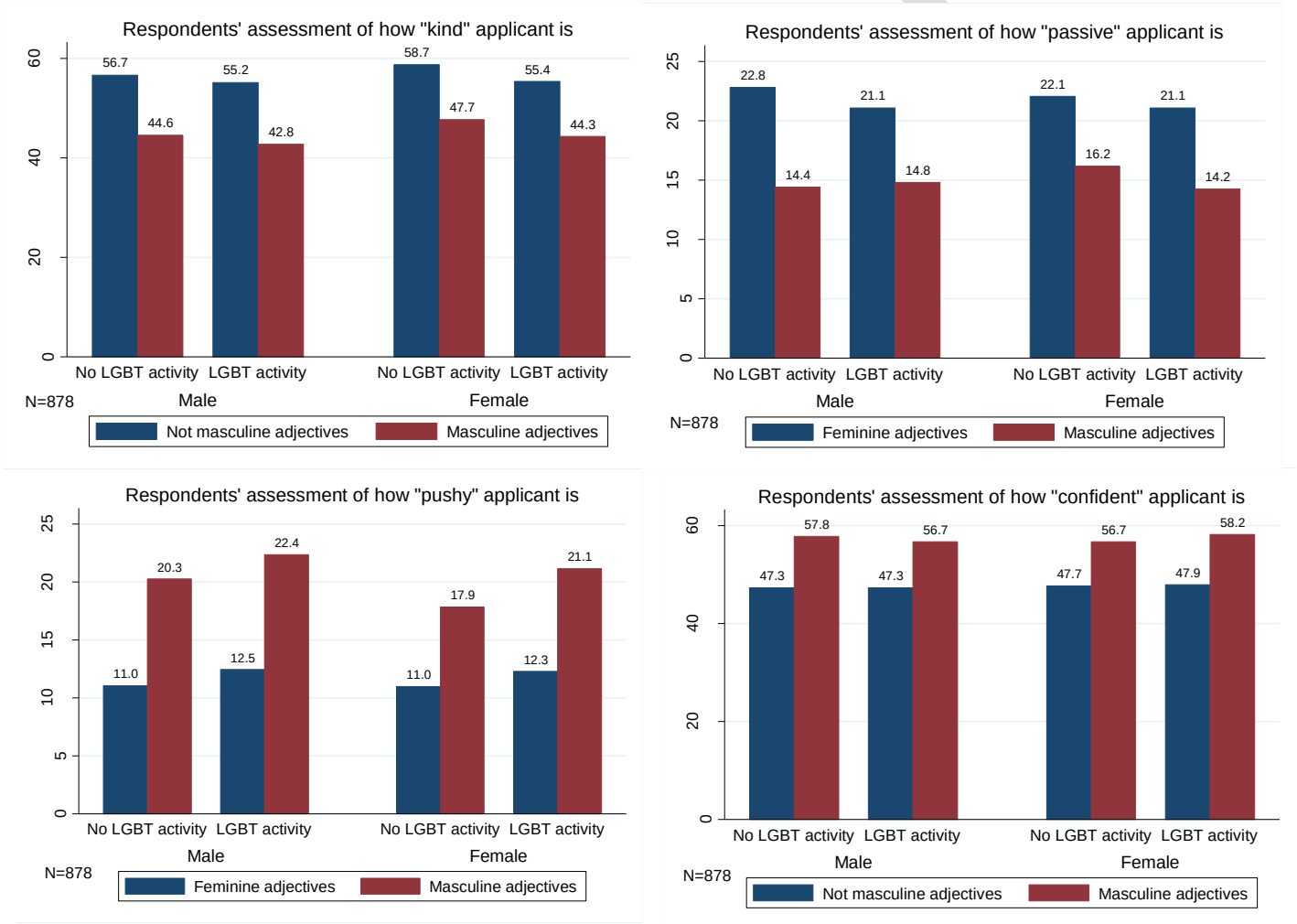


***Figure 5**: Assessment of the applicant's personality. Respondents reported that applicants who used masculine adjectives were more pushy and confident and less kind and passive within each sex by sexual orientation subgroup. N=878*

The results shown above suggest that the use of adjectives in the objective statement is effective – the respondents noticed and responded to the treatment. In each sex by sexual orientation subgroup, the difference in the perceived personality characteristics between the masculine and feminine adjectives is statistically significant at the .001 level (robust standard errors, clustered at the respondent level, with respondent fixed effects). The use of masculine adjectives strongly and consistently impacts how the respondent views an applicant's personality.

The evaluations of "pushy" and "passive" had large masses on zero; to address this, I also evaluated dichotomous (positive or zero) versions of the "pushy" and "passive" measures in a logit model. The evaluations of "kind" and "confident" had masses on multiples of ten; I created ten bins (e.g., from 0 to 9, 10 to 19, and so on) for the "kind" and "confident" measures that I evaluated in an ordered logit model. The results of the logit and ordered logit models mirror those described above and were all significant at the .001 level (robust standard errors, clustered at the respondent level). The results for all eleven personality characteristics are statistically significant in each of the four subgroups and follow the same pattern; these results are available upon request.

C. Replicating the "backlash" effect for women

To replicate the results established in the literature that women experience a "backlash" when they engage in traditionally male behavior, I examine if using masculine adjectives has a different impact on measures of hireability for female applicants than male applicants (for now, restricted to applicants without the LGBT activity).

As the graphs below demonstrate, on average male respondents rate male applicants who use masculine adjectives equally with those who use feminine adjectives. Male respondents view

female applicants who use masculine language as less successful (p=.066) and recommend them less (p=.089) than women who use feminine language (clustered robust standard errors and respondent fixed effects).

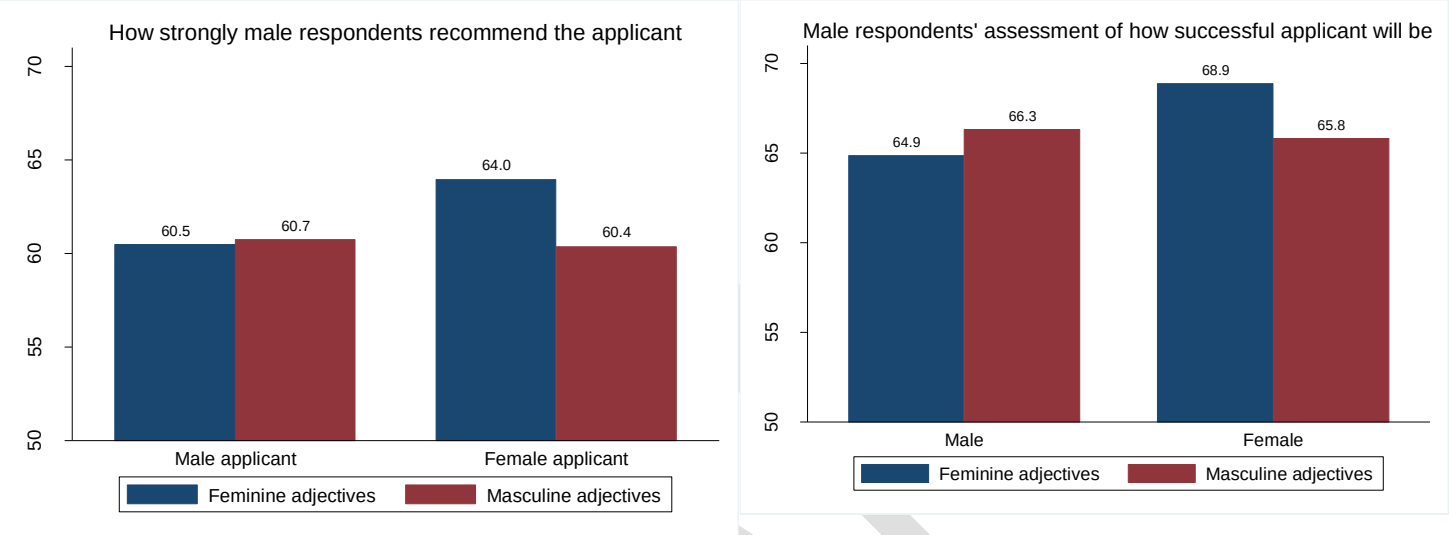


***Figure 6**: Assessment of the applicant's hireability (for resumes without the LGBT activity). Male respondents reported that female applicants who used masculine adjectives would be less successful and recommended them less than female applicants who used feminine language. N=416*

An important note is that women who use feminine language are rated better on these measures than any other group, including men of both adjective types. This is consistent with Bowles, Babcock, and Lai (2006), who find that women who did not negotiate were rated more highly than both men who negotiated and men who did not negotiate. This suggests that the "backlash" for against women for using masculine adjectives could also be viewed as a premium for women using feminine adjectives. To examine this more closely, I compared male and female applicants on the filler resumes, which used gender neutral adjectives (the first two resumes in any survey were had neutral adjectives with no LGBT activity to help the respondent get used to the survey). Female filler resumes were also rated more highly than male filler resumes on salary and how successful the applicant will be (difference in means are significant at the .05 level, with respondent fixed effects and robust clustered errors at the respondent level).

This suggests some preference for female applicants except when they use masculine adjectives, rather than a preference for women who use feminine adjectives over all other groups. However, this analysis must be interpreted cautiously, because the filler resumes were intended to help the respondent adjust to the experiment rather than be used in the analysis.

The following table shows that the difference in male respondents' rating of an applicant between women who use masculine adjectives and those who use feminine adjectives is more negative than this difference for men on three outcome measures: how much the respondent recommended the applicant, how likely the respondent thought the applicant would be successful, and the recommended starting salary. Once again, women who use feminine adjectives are rated higher than men who use feminine adjectives for four of the five outcome measures. The again emphasizes the point that the larger difference between masculine and feminine language for women could be viewed as a premium for women who use feminine language or a backlash to using masculine adjectives. Female respondents showed no evidence of the backlash effect, as shown in Appendix 3.

	Recommend	Successful	Salary	Committed	Willing to work
Masculine adjective	0.123	1.342	1.125	-1.387	-2.593**
	(1.423)	(1.259)	(1.070)	(1.081)	(1.286)
Female applicant	3.512**	4.059***	2.113**	1.590	3.923***
	(1.388)	(1.228)	(1.022)	(1.016)	(1.132)
Masculine resume & female applicant	**-3.461***	**-4.226****	**-2.488***	**-0.516**	**-2.128**
	(1.997)	(1.685)	(1.432)	(1.514)	(1.708)
Obs (Resume x respondent)	1,664	1,664	1,664	1,664	1,664
R-squared	0.624	0.613	0.579	0.642	0.597
Applicants	No LGBT activity	No LGBT activity	No LGBT activity	No LGBT activity	No LGBT activity
Respondent FE	Yes	Yes	Yes	Yes	Yes
Base resume FE	Yes	Yes	Yes	Yes	Yes
Clustered by respondent	Yes	Yes	Yes	Yes	Yes
Respondents	Male	Male	Male	Male	Male
Robust standard errors in parentheses					
**** p<0.01, ** p<0.05, * p<0.1*					

D. Sexual orientation and backlash

The following graphs show that male respondents think that female applicants with the LGBT activity and masculine adjectives will be as successful and recommend them equally to those with feminine adjectives. As before, women without the LGBT activity who use feminine language are thought of as more successful and receive higher recommendations than those who use masculine language. This suggests that gay women are treated more like men – either they miss out on the premium that women without the LGBT activity receive when using feminine adjectives or they are exempt from the backlash.

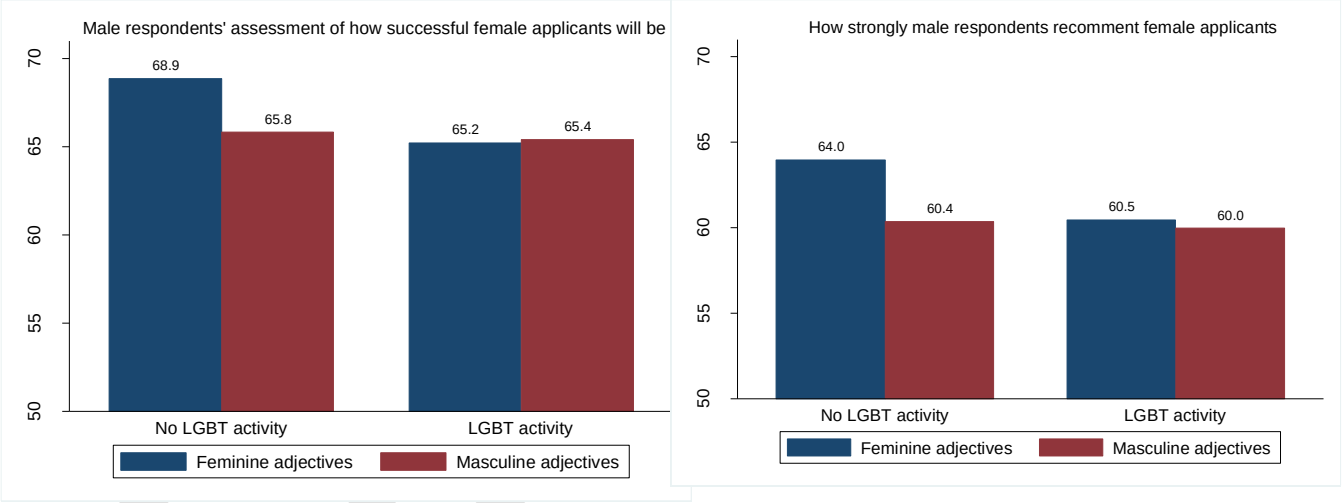


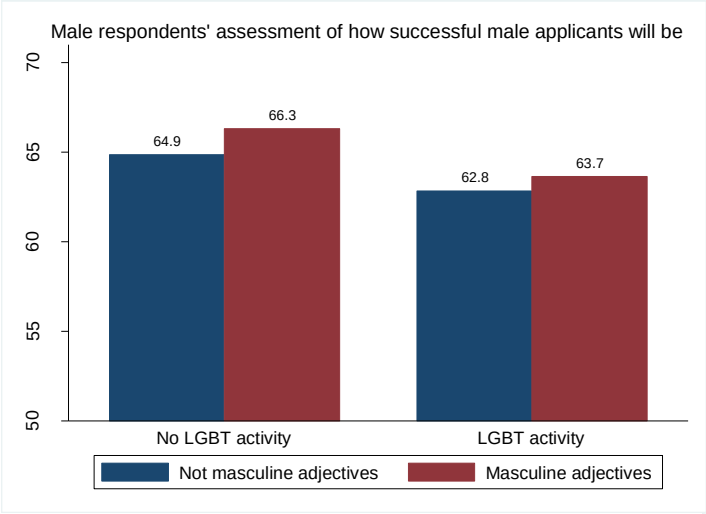
***Figure 7**: Assessment of the applicant's hire-ability (for resumes of female applicants). Male respondents reported that female applicants without the LGBT activity who used masculine adjectives would be less successful and recommended them less than female applicants who used feminine adjectives. No such pattern holds for the resumes with the LGBT activity. N=416*

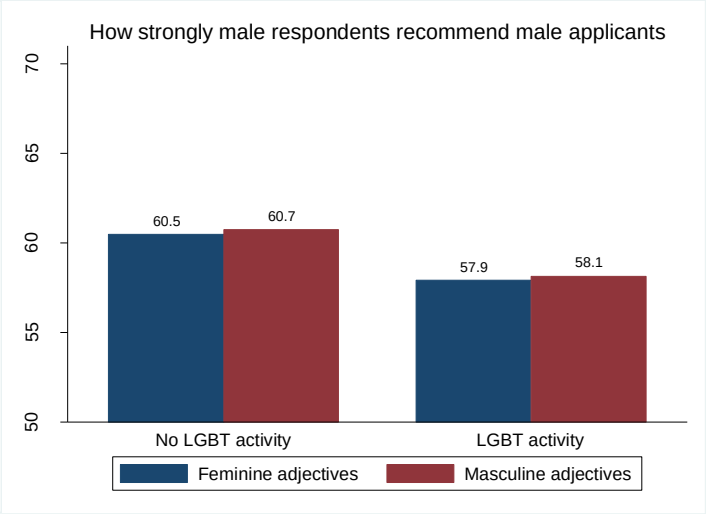
The following table shows that the difference between women without the LGBT activity who use masculine adjectives and those who use feminine adjectives is more negative than this difference for women with the LGBT activity on four outcome measures: how much the respondent recommended the applicant, how likely the respondent thought the applicant would be successful, the recommended starting salary, and how committed the applicant is.

	Recommend	Successful	Salary	Committed	Willing to work
Masculine adjectives	-3.489***	-2.961***	-1.425	-2.068**	-4.849***
	(1.336)	(1.115)	(0.987)	(1.025)	(1.177)
Resume with LGBT activity	-4.623***	-4.422***	-3.037***	-3.984***	-4.750***
	(1.437)	(1.142)	(1.092)	(0.991)	(1.264)
Masculine adjective and LGBT activity	3.217*	3.243**	2.609*	3.032**	1.690
	(1.835)	(1.580)	(1.415)	(1.388)	(1.658)
Obs (Resume x Respondent)	1,664	1,664	1,664	1,664	1,664
R-squared	0.596	0.629	0.542	0.642	0.588
Resumes	Female	Female	Female	Female	Female
Respondent FE	Yes	Yes	Yes	Yes	Yes
Base resume FE	Yes	Yes	Yes	Yes	Yes
Clustered by Respondent	Yes	Yes	Yes	Yes	Yes
Respondents	Male	Male	Male	Male	Male
Robust standard errors in parentheses					
**** p<0.01, ** p<0.05, * p<0.1*					

Female applicants with the LGBT activity who use feminine language are viewed more poorly than women without the LGBT who use feminine language. As with men, this emphasizes the point that the larger difference between masculine and feminine language for women without the LGBT activity could be viewed as a premium for women without the LGBT activity who use feminine language rather than as a backlash to using masculine adjectives. Once again, female respondents showed no evidence of the backlash effect, as shown in Appendix 3.

The following graph examines resumes from male applicants, rather than female. For male resumes, both with and without the LGBT activity, male respondents had slightly higher average ratings for those with the masculine adjectives, but the difference is not statistically significant.





The regression results mirror the results in the summary statistics above. Generally, there is no effect of using masculine adjectives for men without the LGBT activity (except for willing to work with). Additionally, the difference between the resume with the LGBT activity and without in the difference between the masculine adjective and feminine adjectives tended to be positive, but not significant.

	Recommend	Committed	Willing to work	Successful	Salary
Masculine adjective	0.240	-1.299	-2.478*	1.439	1.252
	(1.416)	(1.088)	(1.283)	(1.250)	(1.065)
Resume with LGBT activity	-4.003***	-4.591***	-4.280***	-2.884**	-3.103***
	(1.437)	(1.103)	(1.303)	(1.249)	(1.047)
Masculine adjective and LGBT activity	0.951	2.344	0.273	-0.0105	1.194
	(1.894)	(1.442)	(1.718)	(1.692)	(1.402)
Obs (Resume x Respondent)	1,664	1,664	1,664	1,664	1,664
R-squared	0.616	0.646	0.577	0.624	0.582
Resumes	Male	Male	Male	Male	Male
Respondent FE	Yes	Yes	Yes	Yes	Yes
Resume FE	Yes	Yes	Yes	Yes	Yes
Cluster by Respondent	Yes	Yes	Yes	Yes	Yes
Respondents	Male	Male	Male	Male	Male
Robust standard errors in parentheses					
**** p<0.01, ** p<0.05, * p<0.1*					

This pattern, where neither men with nor without the LGBT activity are impacted by the use of masculine language, suggests that expectations about masculine behavior are more powerful for women than for men.

E. Discrimination against openly LGBT resumes

In the summary statistics and regression results presented above, the resumes with the LGBT activity were perceived differently. For example, as shown in the tables below, the applicants with the LGBT activity were perceived as more "pushy" in all four sex by adjective subgroups. Moreover, people were less willing to work with them and rated them more poorly on other hireability outcomes (shown in Appendix 3). This suggests that in addition to the main result of this paper, that openly gay women are exempt from the backlash effect, there is also evidence of discrimination against resumes with LGBT activities on them.

	Pushy	**Pushy**	**Pushy**	**Pushy**
LGBT activity on resume	3.362***	1.966*	1.290*	1.415**
	(1.006)	(1.115)	(0.688)	(0.676)
Obs (Resume by Respondent)	1,756	1,756	1,756	1,756
R-squared	0.760	0.736	0.794	0.796
Resumes	Female with masculine adjective	Male with masculine adjective	Female with feminine adjective	Male with feminine adjective
Respondent FE	Yes	Yes	Yes	Yes
Resume FE	Yes	Yes	Yes	Yes
Cluster by Respondent	Yes	Yes	Yes	Yes
Respondents	All	All	All	All
Robust standard errors in parentheses				
**** p<0.01, ** p<0.05, * p<0.1*				

	Willing to work	Willing to work	Willing to work	Willing to work
LGBT activity on resume	-3.088***	-1.949*	-3.899***	-2.463**
	(1.044)	(1.127)	(1.050)	(1.044)
Obs (Resume by Respondent)	1,756	1,756	1,756	1,756
R-squared	0.751	0.720	0.714	0.742
Resumes	Female with masculine adjective	Male with masculine adjective	Female with feminine adjective	Male with feminine adjective
Respondent FE	Yes	Yes	Yes	Yes
Resume FE	Yes	Yes	Yes	Yes
Cluster by Respondent	Yes	Yes	Yes	Yes
Respondents	All	All	All	All
Robust standard errors in parentheses				
**** p<0.01, ** p<0.05, * p<0.1*				

To examine if it is just the final outcomes that are impacted by the inclusion of an LGBT activity, I also examine if the LGBT extracurricular activities are perceived as less useful than equivalent activities. In order to control for the value of different extracurricular activities, I analyze a particular base resume where the extracurricular resume entry for the non-LGBT group is identical to the LGBT group, except for the name. Both entries describe the role as having "Planned and organized events that promoted diversity and raised awareness on various topics" and other details to demonstrate the magnitude of the role (see Appendix 1 for full text). The LGBT club was named the "LGBT Alliance" while the non-LGBT club was named "Student Activities Board."

After viewing the whole resume for one minute, respondents were asked to evaluate the usefulness of the applicant's extracurricular activities on a scale of one to ten on a page where only the respondents name, objective statement, and extracurricular activity were visible. Despite having the exact same detailed description of the role, the resumes with the non-LGBT club were rated as 2.84 on the useful scale while the LGBT version was only 2.51 (p=.06, robust

standard errors). Because the usefulness ratings only took on whole numbers, I also use an ordered logit model to examine this question. The odds ratio on the LGBT activity indicator is .75 (p=.03, robust standard errors; N=702), indicating again that the LGBT extracurricular activity is viewed as less useful than the identical entry for a non-LGBT group. Restricting to only those respondents who agree or strongly agree that same-sex marriage should be legal also results in a statistically significant difference: 2.96 for the non-LGBT group and 2.46 for the LGBT group (p=.02) and an odds ratio of .68 (p=.01).

This negative association also leaks over to the assessment of the applicant's work history for male applicants being evaluated by male respondents. After viewing the whole resume for one minute, the respondent is shown a page with only the applicant's name and work history (no extracurricular activity) and asked to rate the usefulness of the applicant's work history on a scale of one to ten. The following table shows that for male applicants evaluated by male respondents, having an LGBT activity on the resume results in a lower assessment of the usefulness of the applicant's work history. The following regressions include base resume fixed effects; the base resumes have identical work experiences.

	Usefulness of work history	**Usefulness of work history**	**Usefulness of work history**	**Usefulness of work history**
LGBT activity on resumes	-0.214**	0.0425	-0.0796	0.0107
	(0.0984)	(0.102)	(0.102)	(0.0954)
Obs (Resume x Respondent)	1,664	1,664	1,844	1,844
R-squared	0.679	0.654	0.646	0.651
Resumes	Male	Female	Male	Female
Respondent FE	Yes	Yes	Yes	Yes
Resume FE	Yes	Yes	Yes	Yes
Clustered by Respondent	Yes	Yes	Yes	Yes
Respondents	Male	Male	Female	Female
Robust standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1*				

Taken together, the above results suggest that having an LGBT related extracurricular activity is viewed negatively. The applicants with the LGBT activity were viewed as more pushy and people did not want to work with them. Their extracurricular activity was viewed as less useful than an identical activity that is not LGBT related; this remains true even among those who support same-sex marriage. Moreover, when men evaluate the usefulness of a male applicant's work history on a separate page from the extracurricular activity, they rate the work experience of an applicant who has the LGBT activity as less useful relative to the identical work history of another applicant without the LGBT activity. The spillover did not occur for female applicants or for female respondents.

F. Finite Mixture Model

VI. Future Analysis: Audit study

The results of the laboratory experiment suggest that resumes with masculine adjectives are perceived differently for different sex and sexual orientation combinations. To extend this analysis to a population-based study, similarly manipulated resumes will be used in an audit study.

Similar compilation resumes as described in the laboratory experiment will be used in the audit study, expect the resumes will be adjusted to be geographically and topically appropriate to the job advertisement (similar to Bertrand and Mullainathan 2004). To do this, I will purchase a resume bank from CareerBuilder.com. This bank contains all the entries on resumes currently uploaded to CareerBuilder.com. I will randomly select entries from the resume bank to create multiple compilation resumes appropriate for different geographical areas and fields. The same

experimental manipulation as in the laboratory experiment will be used to create resumes that vary by the sex and sexual orientation of the applicant.

Then, I will use the resumes to apply for advertised entry level jobs in multiple metropolitan areas. Comparing the callback rate in each sex by sexual orientation cell will demonstrate if the backlash effect occurs in a real world job search.

Moreover, following Tilcsik (2011), the analysis will examine the callback rates for heterogeneous effects when the job advertisement uses masculine adjectives or are in male dominated industries. Tilcsik (2011) found the largest difference in callback rates between heterosexual and gay male resumes for jobs advertised with masculine adjectives. A similar audit study in Sweden found that gay men faced larger difference in callback rates in male-dominated industries and gay women in female-dominated industries (Ahmed, Andersson, and Hammarstedt 2013). Likewise, Rudman and Glick (1999) found that the backlash effect was strongest in their laboratory experiment when the advertised job was described with feminine characteristics. This suggests that the combination of sex, sexual orientation, and masculine styled resumes may have different effects when applying to jobs advertisements in different industries or that contain masculine or feminine adjectives.

VII. Conclusions and Extensions

The persistent differences in earnings based on sex and sexual orientation may result from the role of identity in forming gendered behavioral expectations. Economic and psychological theories suggest that men react negatively to women acting in traditionally masculine ways because their own masculine identity is threatened (Akerlof and Kranton 2000; Tajifel and Turner 1979). Numerous 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).

In this paper, I argue that it is important to consider **which** group defines another group's identity. By manipulating the sex, sexual orientation, and masculinity of a resume, I first demonstrate that male respondents rate female applicants who use masculine adjectives worse relative to female applicants who use feminine adjectives (equivalently, they rate female applicants who use feminine adjectives better than those who use masculine adjectives). The pattern does not appear for male applicants. Moreover, the difference in the difference between male and female applicants is significant for three important hireability measures: how much the respondent would recommend the company hire the applicant, recommend starting salary, and their likelihood of success.

In contrast, gay women are exempt from the backlash effect from male respondents (or miss out on the premium for using feminine adjectives). This suggests that gay women are not a relevant comparison group to men, and that men do not suffer identity loss when gay women follow male behavioral prescriptions. More broadly, it suggests that identity plays a key role in the persistent earnings differences based on sex and sexual orientation.

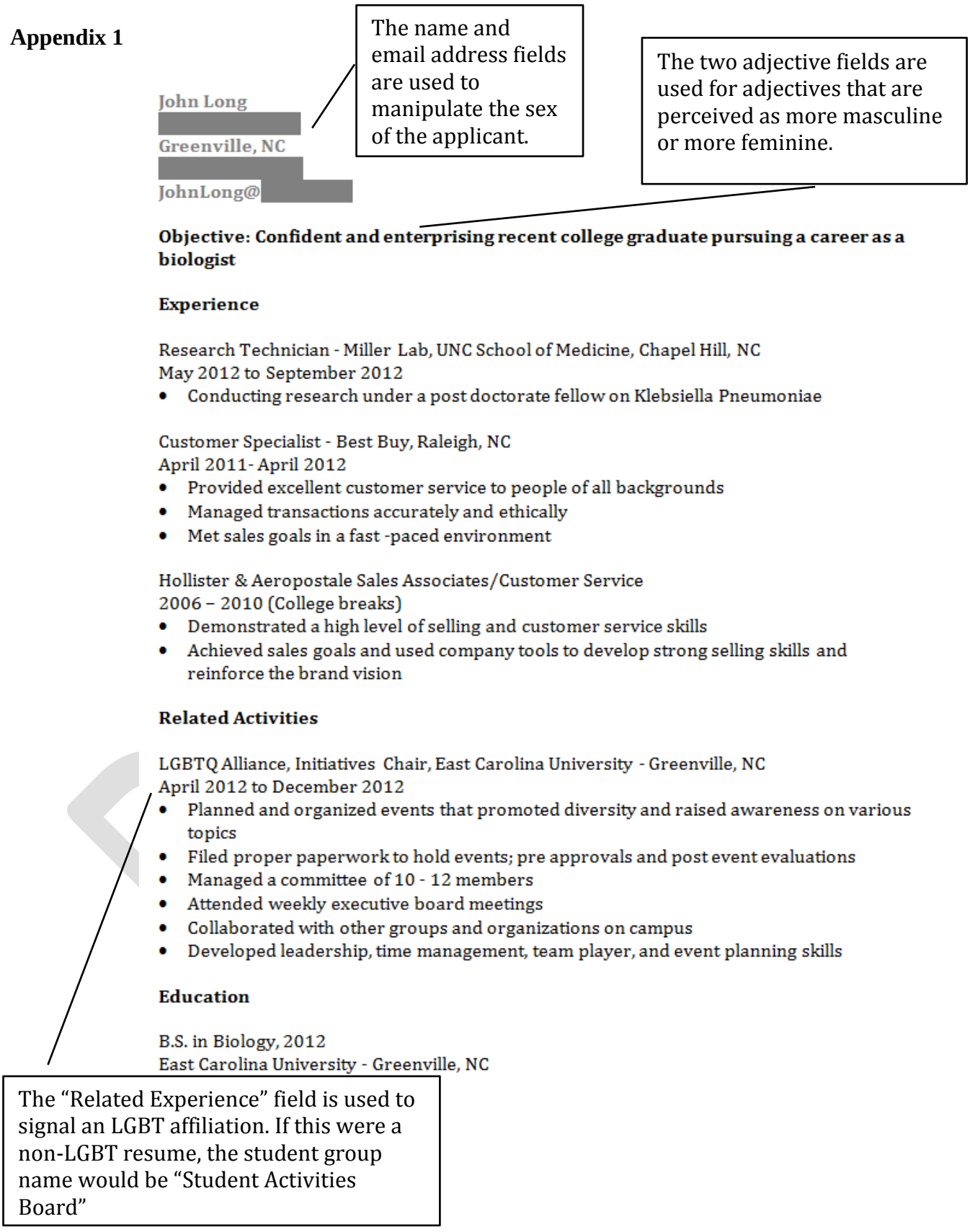
The results in this paper do not support a theory of earnings differences based on sex being driven by statistical discrimination. If earnings differences are driven by employers' perceptions of the distribution of personality characteristics for different sex and sexual orientation subgroups, a signal of those characteristics should move the employer's assessment of the applicant's hire-ability in the same direction (albeit with different magnitudes). In this paper, I find that masculine adjectives harm women without the LGBT activity (or feminine adjectives help them), but not when women list an LGBT activity on their resume.

Additionally, while not the central point of the paper, the analysis revealed striking differences in how a resume with an LGBT activity was viewed relative to one with a similar non-LGBT activity. Even among respondents who agreed with the statement that same-sex marriage should be legal, an LGBT extracurricular activity was viewed as less useful than an identical non-LGBT activity. Overall, the LGBT resumes within each sex by adjective subgroup were viewed as more pushy, respondents were less willing to work with those applicants, and they rated the applicants more poorly on most of the hireability measures. This suggests that even though the United States is experiencing a dramatic increase in support for same-sex marriage, this may not translate directly into reduced implicit discrimination in the labor market.

While results of the audit study are not yet complete, the audit study modeled closely on the laboratory study will provide a population-based estimate of the impact of the backlash effect.

Finally, recent work (Pedulla 2012) has found that a job applicant's sexual orientation is perceived very differently based on the race of the applicant. Likewise, traditionally masculine actions are perceived differently by sex and race combinations (Livingston and Pearce 2009; Livingston, Rosette, and Washington 2012). In future work, I will explore this dynamic further by examining the backlash effect for different race and sexual orientation combinations.

Appendix 1



Appendix Figure 1: *Example of a compilation resume. The entries in the resume are compiled from randomly selected publically listed resumes. Three fields are used for the experimental manipulation (sex, masculine language, and LGBT affiliation); these fields are noted and described.*

Appendix 3

	Recommend	Successful	Salary	Committed	Willing to work
Masculine adjective	-0.344	0.538	0.402	-0.606	-6.290***
	(1.362)	(1.204)	(0.951)	(1.094)	(1.256)
Female applicant	1.616	1.625	0.962	1.048	2.942***
	(1.350)	(1.156)	(0.904)	(1.051)	(1.075)
Masculine adjective and female applicant	-1.054	-1.085	-0.621	-0.708	0.356
	(1.890)	(1.541)	(1.281)	(1.425)	(1.575)
Obs (Resume by Respondent)	1,844	1,844	1,844	1,844	1,844
R-squared	0.628	0.607	0.556	0.614	0.589
Resumes	Not LGBT	Not LGBT	Not LGBT	Not LGBT	Not LGBT
Respondent FE	Yes	Yes	Yes	Yes	Yes
Base resume FE	Yes	Yes	Yes	Yes	Yes
Cluster by Respondent	Yes	Yes	Yes	Yes	Yes
Respondents	Female	Female	Female	Female	Female
Robust standard errors in parentheses					
**** p<0.01, ** p<0.05, * p<0.1*					

	Recommend	Successful	Salary	Committed	Willing to work
Masculine adjective	-1.369	-1.287	-0.536	-0.107	-5.921***
	(1.398)	(1.030)	(1.122)	(0.926)	(1.168)
LGBT activity	-2.244	-1.565	-2.028*	-1.353	-3.228***
	(1.403)	(1.029)	(1.160)	(0.917)	(1.179)
Masculine adjective and LGBT activity	1.514	-0.151	0.713	0.596	0.0491
	(1.913)	(1.473)	(1.564)	(1.318)	(1.637)
Obs (Resume x Respondent)	1,844	1,844	1,844	1,844	1,844
R-squared	0.624	0.629	0.615	0.560	0.617
Resumes	Female	Female	Female	Female	Female
Respondent FE	Yes	Yes	Yes	Yes	Yes
Base resume FE	Yes	Yes	Yes	Yes	Yes
Cluster by Respondent	Yes	Yes	Yes	Yes	Yes
Respondents	Female	Female	Female	Female	Female
Robust standard errors in parentheses					
**** p<0.01, ** p<0.05, * p<0.1*					

	Recommend	Recommend	Recommend	Recommend
LGBT activity	-1.172	-2.252*	-3.444***	-2.446**
	(1.142)	(1.219)	(1.215)	(1.202)
Obs (Resume x Respondent)	1,756	1,756	1,756	1,756
R-squared	0.771	0.728	0.738	0.765
Resumes	Female with masculine adjective	Male with masculine adjective	Female with feminine adjective	Male with feminine adjective
Respondent FE	Yes	Yes	Yes	Yes
Resume FE	Yes	Yes	Yes	Yes
Clustered by Respondent	Yes	Yes	Yes	Yes
Respondents	All	All	All	All
Robust standard errors in parentheses				
**** p<0.01, ** p<0.05, * p<0.1*				
	Committed	**Committed**	**Committed**	**Committed**
LGBT activity	-1.454	-1.747*	-2.749***	-3.181***
	(0.922)	(0.924)	(0.875)	(0.928)
Obs (Resume x Respondent)	1,756	1,756	1,756	1,756
R-squared	0.758	0.755	0.768	0.768
Resumes	Female with masculine adjective	Male with masculine adjective	Female with feminine adjective	Male with feminine adjective
Respondent FE	Yes	Yes	Yes	Yes
Resume FE	Yes	Yes	Yes	Yes
Clustered by Respondent	Yes	Yes	Yes	Yes
Respondents	All	All	All	All
Robust standard errors in parentheses				
**** p<0.01, ** p<0.05, * p<0.1*				
	Salary	**Salary**	**Salary**	**Salary**
LGBT activity	-0.635	-0.960	-2.209**	-1.808**
	(0.812)	(0.814)	(0.861)	(0.836)
Obs (Resume x Respondent)	1,756	1,756	1,756	1,756
R-squared	0.727	0.711	0.680	0.732
Resumes	Female with masculine adjective	Male with masculine adjective	Female with feminine adjective	Male with feminine adjective
Respondent FE	Yes	Yes	Yes	Yes
Resume FE	Yes	Yes	Yes	Yes
Clustered by Respondent	Yes	Yes	Yes	Yes
Respondents	All	All	All	All
Robust standard errors in parentheses				
**** p<0.01, ** p<0.05, * p<0.1*				

	Successful	Successful	Successful	Successful
LGBT activity	-1.282	-1.578	-3.227***	-2.013**
	(0.996)	(1.034)	(0.987)	(1.011)
Obs (Resume x Respondent)	1,756	1,756	1,756	1,756
R-squared	0.762	0.731	0.749	0.766
Resumes	Female with masculine adjective	Male with masculine adjective	Female with feminine adjective	Male with feminine adjective
Respondent FE	Yes	Yes	Yes	Yes
Resume FE	Yes	Yes	Yes	Yes
Clustered by Respondent	Yes	Yes	Yes	Yes
Respondents	All	All	All	All
Robust standard errors in parentheses				
**** p<0.01, ** p<0.05, * p<0.1*				

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.

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.

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.

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.

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

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

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.

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.