

Three Essays on Gender, Population Studies, and Labor Economics

by

Marina Mileo Gorsuch

Public Policy Studies
Duke University

Date: _____

Approved:

Seth Sanders, Supervisor

Amar Hamoudi

Kenneth Dodge

Elizabeth Ananat

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in Public Policy Studies
in the Graduate School
of Duke University

2015

ABSTRACT

Three Essays on Gender, Population Studies, and Labor Economics

by

Marina Mileo Gorsuch

Public Policy Studies
Duke University

Date: _____

Approved:

Seth Sanders, Supervisor

Amar Hamoudi

Kenneth Dodge

Elizabeth Ananat

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in Public Policy Studies
in the Graduate School of
Duke University

2015

Copyright by
Marina Mileo Gorsuch
2015

Abstract

In this dissertation, I examine three questions on gender and labor economics. The first two questions are inspired by a broad literature in social psychology which 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).

I test two hypotheses related to this literature. First, I examine if resumes that use masculine adjectives inspire backlash against female job applicants in a laboratory setting and if this backlash varies by the sexual orientation of the applicant. Second, 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 a laboratory setting, I replicate the backlash effect and also show that it only affects perceived-heterosexual women. In a resume audit study, I find the reverse of a backlash effect: employers call back women who use traditionally masculine adjectives more than when they use traditionally feminine adjectives.

The third question examines the time men spent on childcare during the recession of 2007-2009. The recession provides a sudden change in the employment opportunities of men relative to women in the United States. Using the American Time

Use Survey and the linked Current Population Survey, I show that this lopsided shock to employment opportunities was accompanied by an increase in the average amount of time men spent on childcare. In particular, men's average time on physical care for children increased during the recession; this is an element of childcare that men perform less than women. I decompose the total change in average time on childcare into behavioral, compositional, and between group change. A behavioral change among employed men accounted for the majority of the total increase in the average time spent on childcare; among men who are out of the labor force, the increase is entirely due to compositional changes.

Contents

Abstract	iv
List of Tables.....	ix
List of Figures	xii
1. Gender, sexual orientation, and backlash.....	1
1.1 Theoretical framework.....	5
1.2 Numerical example	10
2. Testing the backlash effect in the laboratory.....	14
2.1 Experimental manipulation	14
2.2 Laboratory study: Empirical framework	21
2.2.1 Regression framework.....	21
2.2.2 Finite mixture model.....	23
2.3 Laboratory study: Results	25
2.3.1 Characteristics of respondents.....	25
2.3.2 Evaluating if the treatment was salient to respondents.....	27
2.3.3 Replicating the “backlash” effect for women.....	29
2.3.4 Does “backlash” vary by sexual orientation?.....	33
2.3.5 Finite mixture model.....	38
2.3.6 Differences for openly LGBT resumes.....	41
2.4 Conclusions from laboratory study	45
3. Testing the backlash effect in the labor market	49
3.1 Experimental manipulation	49

3.1.1 Weaknesses of resume audit studies	52
3.2 Resume audit study: Empirical framework.....	53
3.3 Resume audit study: Results.....	54
3.3.1 Description of jobs and employers	54
3.3.2 Is there a backlash effect?	57
3.3.3 The ad text: Is it a match?	60
3.3.4 Differences in callback rates by sex and sexual orientation	65
3.4 Conclusions from audit study	67
4. Decomposing the increase in men’s time on childcare during the recession.....	71
4.1 Introduction.....	71
4.2 Background literature	74
4.3 Data source	79
4.3.1 Measures of childcare	80
4.4 Stylized facts about the 2007-2009 recession, employment, and childcare	81
4.5 Decomposing the total change in childcare	87
4.5.1 Compositional changes within employment status	87
4.5.2 Behavioral, compositional, and between group change	92
4.5.3 Time spent on other activities.....	98
4.6 Conclusion from examining changes in time spent on childcare	101
Appendix 1: Resume example for laboratory experiment	105
Appendix 2: Difference in hireability measures for resumes with LGBT activity	106
Appendix 3: Resume example for audit study	108

Appendix 4: Balance of audit study	109
Appendix 5: Balance of laboratory experiment	111
Appendix 6: Regression results for time on childcare	112
Appendix 7: Decomposition method for change in time on childcare.....	117
Appendix 8: Analysis of panel data	119
Works Cited	122
Biography.....	129

List of Tables

Table 1 Results of an OLS regression of hireability variables on the type of adjective used on the resume, a female indicator, and their interactions. Outcome variables could take on values from 0 to 100.	32
Table 2: Results of an OLS regression of hireability variables on the type of adjective used on the resume, a female indicator, and their interactions Outcome variables could take on values from 0 to 100.	33
Table 3: Results of an OLS regression of hireability variables on the type of adjective used on the resume, an LGBT indicator, and their interactions. Outcome variables could take on values from 0 to 100.	35
Table 4: Results of an OLS regression of hireability variables on the type of adjective used on the resume, an LGBT indicator, and their interactions. Outcome variables could take on values from 0 to 100.	36
Table 5: Results of an OLS regression of hireability variables on the type of adjective used on the resume, an LGBT indicator, and their interactions. Outcome variables could take on values from 0 to 100.	38
Table 6: Results of a FMM regression of how strongly respondents would recommend the applicant on the type of adjective used on the resume, a female indicator, and their interactions.....	39
Table 7: Results of a FMM regression of how strongly respondents would recommend the applicant on the type of adjective used on the resume, an LGBT indicator, and their interactions.....	40
Table 8: Results of an OLS regression of how “pushy” the respondent rated the resume on an LGBT indicator. Outcome variables could take on values from 0 to 100.	42
Table 9: Results of an OLS regression of how much the respondent was “willing to work with” the applicant on an LGBT indicator. Outcome variables could take on values from 0 to 100.	42
Table 10: Results of an OLS regression of how “useful” the respondent rated the applicant’s work history on an LGBT indicator. Outcome variables could take on values from 0 to 10.	44
Table 11: Examples of common jobs applied for within each type of job.....	55

Table 12: 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.....	60
Table 13: Educational attainment for males who are 15 and over by employment status before and during the recession.....	89
Table 14: Decomposition of the total difference in the minutes of primary childcare during the recession between men with children in their households who were either unemployed or employed in the CPS survey and left the labor force by the time of the ATUS survey compared to those who were out of the labor force in both surveys.....	91
Table 15: Decomposition of the total change in time spent on primary childcare by men with children in their households.....	94
Table 16: Decomposition of the total change in time spent on physical care by men with children in their households.....	96
Table 17: Change in time spent on five major activities during and after the recession by men with children in their household. “Other activities” is the difference between the total time in a day and those five major activities.....	99
Table 18: Change in time spent on five major activities during and after the recession by employed men with children in their household.....	100
Table 19: Change in time spent on four subparts of work and work-related activities during and after the recession by employed men with children in their household.....	101
Table 20: Selected coefficients from regression of men’s time on childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables.....	113
Table 21: Selected coefficients from regression of women’s time on childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables	114
Table 22: Selected coefficients from regression of men’s time on childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables.....	115

Table 23: Selected coefficients from regression of men’s participation in childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables 116

Table 24: Selected coefficients from regression of men’s time on childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables..... 116

Table 25: Selected coefficients from regression of men’s self-reported time on housework per week on an indicator variable for 1990 or 1991 sample years, a time trend, and respondent fixed effects. 120

List of Figures

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.	16
Figure 2: Self-reported age and education among MTurk respondents N=878.....	25
Figure 3: Self-reported household income and race/ethnicity among MTurk respondents. N=878	26
Figure 4: MTurk respondents' answers to questions from the General Social Survey about gender roles and same-sex marriage. N=878	26
Figure 5: Assessment of the applicant's personality. N=878 Outcome variables could take on values from 0 to 100.	27
Figure 6: Assessment of the applicant's hireability (for resumes with the non-LGBT activity). N=416 Outcome variables could take on values from 0 to 100.	30
Figure 7: Assessment of applicant's hire-ability (for resumes of female applicants). N=416 Outcome variables could take on values from 0 to 100.....	34
Figure 8: Assessment of the applicant's hire-ability (for resumes of male applicants). N=416 Outcome variables could take on values from 0 to 100.....	37
Figure 9: Types of jobs that were applied for. N=1,334	54
Figure 10: Proportion of resumes called back by category of job. N=1,334	55
Figure 11: Distribution of the types of adjectives used in the text of the job advertisements applied to.....	57
Figure 12: Callback rate by sex and type of adjective used on resumes without the LGBT activity. N=678.....	58
Figure 13: Callback rate by sexual orientation and type of adjective used on female resumes. N=673	59
Figure 14: Callback rate by sex, sexual orientation for job ads with traditionally masculine adjectives. N=204.....	62

Figure 15: Callback rate by sex, sexual orientation, and adjective for job ads with traditionally masculine adjectives. N=204.....	63
Figure 16: 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.	64
Figure 17: Callback rate by sex, sexual orientation, and adjective used on resume for job ads with feminine language. N=395.	65
Figure 18: Callback rate by sex and perceived sexual orientation. N=1,334.....	66
Figure 19: Unemployment Rate for those 16 years and older by sex (L) and by sector (R)	82
Figure 20: Unemployment Rate (not seasonally adjusted) for men by marital status and presence of children.....	83
Figure 21: Minutes of primary childcare	84
Figure 22: Minutes of primary childcare for those with positive values	85
Figure 23: Time spent on types of childcare by men with children in household.....	86
Figure 24: Time spent on physical care for children	87

1. Gender, sexual orientation, and backlash

A broad literature has established that laboratory 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 finding 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).

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; Tajfel and Turner 1979). That is, when women follow the behavioral prescriptions for men, it adversely affects men's identity. However, it is important to note that group identity is formed in contrast only to the "relevant comparison group" rather than all other groups (Tajfel and Turner 1979, p. 41). If there are different subgroups within women, only those who are the relevant comparison group would trigger the backlash effect. In particular, gay women are perceived as having many traits in common with heterosexual men (Ahmed et al 2013) and therefore may not be the relevant comparison group. If gay women are not a relevant comparison group for heterosexual men, they may be exempt from the backlash effect. Consistent with this hypothesis is the evidence on earnings differences based on sexual orientation: 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).

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 this program of research, I test two sets of hypotheses. In the first study, I examine 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. In the second 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 language.

To test the first hypothesis, I ask respondents in a laboratory setting to evaluate resumes that vary on sex, sexual orientation, and the use of traditionally masculine or feminine adjectives. The results of the laboratory experiment indicate that women **without** an LGBT activity on their resume who use masculine language on resumes are

subject to backlash by male respondents, but women **with** an LGBT activity on their resume who use masculine language are not. Men, both with and without the LGBT activity, do not experience any backlash when they use masculine language on their resumes. These results are entirely consistent with the social identity approach: male respondents act more favorably when heterosexual women, their relevant comparison group, use feminine language rather than masculine language. Likewise, a group of women who are not their relevant comparison group, gay women, are immune to this effect. Application of a finite mixture model reveals two latent classes among male respondents: a large group that displays a strong backlash effect and a much smaller group that displays no backlash effect.

To test if real employers have the same response as respondents in a laboratory to traditionally masculine actions, I conducted a second study in which I submitted fictional applications 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.

1.1 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 pattern would suggest that gay women would be exempt from the backlash effect.

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 alter 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 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 feel more altruistic towards the fictional applicants. 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. The difference in experience may mean 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, employer knows three things about the applicant: their sex, sexual orientation, and what type of adjective they used to describe themselves on their resume.

$$v = w * E[f(\textit{Collegial}, \textit{Competitive}) | \textit{sex}, \textit{sexual orientation}, \textit{adjective}] + (1) \quad (1 - w) * \beta * I(\textit{female \& heterosexual \& masculine adjective})$$

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

The first half of Equation 1 consists of the expected productivity of an employee based on his or her personality characteristics. Evaluators do 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 form implies that employers will 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.

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

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

However, this condition may not hold when respondents evaluate gay women, either because gay women may not be men's relevant comparison group or gay women's use of a masculine adjective may contain different information about productivity. This hypothesis is entirely consistent with the evidence on earnings differences based on sexual orientation: 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). If gay women are not subject to the backlash effect, it could play a role in the observed earnings differences based on sexual orientation.

Additionally, 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.

1.2 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, $f(\text{Collegial}, \text{Competitive}) = \text{Collegial}^{1/2} + \text{Competitive}^{1/2}$. 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 in this example) on the expected production than laboratory respondents (.1).

$$v_{\text{Employer}} =$$

$$(2) \quad .9 * E \left[\text{Collegial}^{1/2} + \text{Competitive}^{1/2} \mid \text{sex}, \text{sexual orientation}, \text{adjective} \right]$$

$$+ .1 * \beta * I(\text{female} \ \& \ \text{heterosexual} \ \& \ \text{masculine adjective})$$

$$v_{\text{Lab respondent}} =$$

$$(3) \quad .1 * E \left[\text{Collegial}^{1/2} + \text{Competitive}^{1/2} \mid \text{sex}, \text{sexual orientation}, \text{adjective} \right]$$

$$+ .9 * \beta * I(\text{female} \ \& \ \text{heterosexual} \ \& \ \text{masculine adjective})$$

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 ($\text{Collegial} \perp \text{Competitive} \mid \text{sex}, \text{sexual orientation}, \text{adjective}$)

<i>Collegial</i> <i>Competitive</i>	LGBT	Non-LGBT		LGBT	Non-LGBT
"I'm assertive"	<i>Collegial</i> ~ U(5, 15) <i>Competitive</i> ~ 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 $f(\text{Collegial}, \text{Competitive})$ is

$$\int_0^{10} \int_{20}^{30} (\text{Collegial}^{1/2} + \text{Competitive}^{1/2}) \frac{1}{10} \frac{1}{10} d\text{Competitive} d\text{Collegial} = 7.1$$

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

v_{Employer} $v_{\text{Lab respondent}}$	LGBT	Non-LGBT		LGBT	Non-LGBT
“I’m assertive”	$v_{\text{Employer}} = 6.8$ $v_{\text{Lab respondent}} = .76$	6.4 .71		6.8 .76	$6.9 - .1 * \beta$ $.77 - .9 * \beta$
“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 $.77 - .9 * \beta < .71$, 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.

2. Testing the backlash effect in the laboratory

2.1 Experimental manipulation

I test the hypothesis that resumes that use masculine adjectives inspire backlash against female job applicants in a laboratory setting and whether the backlash effect varies by sexual orientation. To complete this test, I examine if resumes traditionally masculine or feminine 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 publicly 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 30, 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 masculine. In the pre-test, one group (N=85) of respondents on Mechanical Turk (MTurk is 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 (N=90) 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 finding 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.

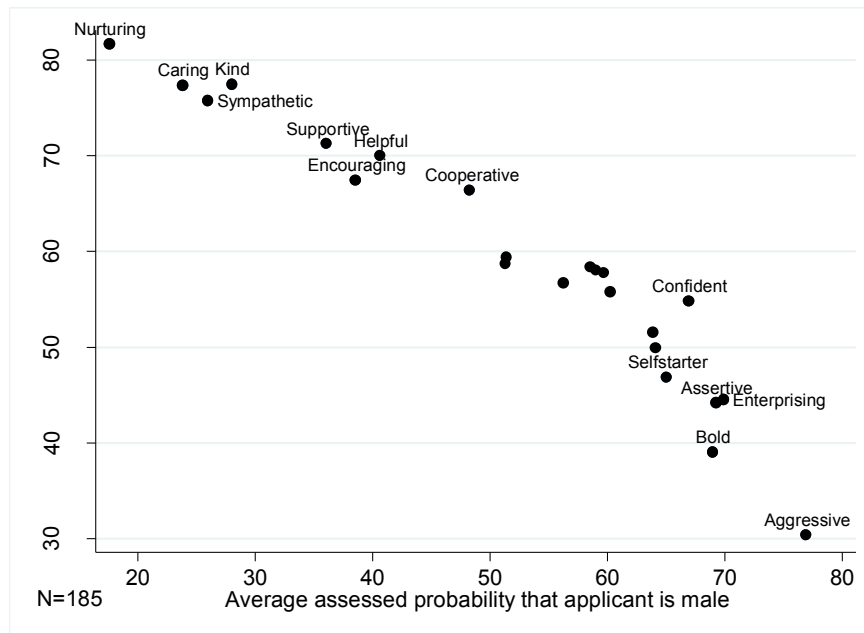


Figure 1: Results from a pre-test of adjectives. One group (N=85) reported how likely it was that the applicant was male (x-axis) while another group (N=90) reported how likely to applicant was female (y-axis). The adjectives with the strongest gender associations are labeled.

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 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, membership in an LGBT student group, and type of language used. The two 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 an LGBT 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, LGBT student group participation, and type of adjectives) once. For example, in one version of the questionnaire the first base resume was a man with masculine adjectives who was in an LGBT student group. In another version, the same base resume could be a woman with feminine adjectives

who was not in an LGBT student group. 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, he or she 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 5th 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.

This experimental design was approved by the Duke Institutional Review Board. All respondents were debriefed after they had completed the study.

2.2 Laboratory study: Empirical framework

2.2.1 Regression framework

To test if resumes with traditionally masculine or feminine language are viewed differently based on the sex and sexual orientation of the applicant, I will first examine if certain adjectives on a resume can inspire the “backlash” effect documented in the existing literature. To do this, I examine the difference between male and female resumes in the difference 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.

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

When Equation 4 is estimated on the resumes with the non-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, $\hat{\theta} + \hat{\delta}$ 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 3 employs the same diff-in-diff approach, but compares women's resumes with the LGBT activity to those with the non-LGBT activity. 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.

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

2.2.2 Finite mixture model

To examine if the effects found in Equations 4 and 5 are actually the average between two latent classes, I utilize finite mixture model analysis. That is, perhaps one group of people do not have a reaction when women use masculine adjectives and another group has a strong reaction. The analysis in Equation 4 would identify only the weighted average between the two groups. Finite mixture model analysis examines if there is heterogeneity in the sample based on unobserved characteristics.

Suppose there are two classes of respondents, one with the relationship $y_{r,i} = \mathbf{x}_{r,i}\boldsymbol{\beta}_1 + \eta_{r,i}$ and the other $y_{r,i} = \mathbf{x}_{r,i}\boldsymbol{\beta}_2 + \omega_{r,i}$. Equation 4 shows the likelihood for respondent r, who could be in the first latent class with probability π_1 or in the second class with probability $\pi_2 = 1 - \pi_1$ (where $\pi_1 \in [0,1]$). Respondent r has I observations (I resumes reviewed by respondent r) distributed with pdf f_1 if respondent r is in group 1 and f_2 if they are in group 2. Assuming these I observations are independent¹, the product of their pdf represents the joint probability. The sum of the two products, weighted by the probabilities of being in group 1 and group 2 is the likelihood function for person r.

$$(4) \quad L_r(\pi_1, \pi_2, \boldsymbol{\beta}_1, \sigma_1^2, \boldsymbol{\beta}_2, \sigma_2^2 | \mathbf{y}_r) = \pi_1 * \prod_{i=1}^I f_1(y_{r,i} | \mathbf{x}_{r,i}, \boldsymbol{\beta}_1, \sigma_1^2) + \pi_2 * \prod_{i=1}^I f_2(y_{r,i} | \mathbf{x}_{r,i}, \boldsymbol{\beta}_2, \sigma_2^2)$$

¹ When applied, all data will be demeaned by respondent to account for dependence between observations within the same respondent.

If all N respondents are independent and identically distributed, the likelihood function is

$$(6) \quad L(\pi_1, \pi_2, \boldsymbol{\beta}_1, \sigma_1^2, \boldsymbol{\beta}_2, \sigma_2^2 | \mathbf{y}) = \prod_{r=1}^N \left[\pi_1 * \prod_{i=1}^I f_1(y_{r,i} | \mathbf{x}_{r,i}, \boldsymbol{\beta}_1, \sigma_1^2) + \pi_2 * \prod_{i=1}^I f_2(y_{r,i} | \mathbf{x}_{r,i}, \boldsymbol{\beta}_2, \sigma_2^2) \right]$$

Maximizing $\ln(L)$ by choosing $\pi_1, \pi_2, \boldsymbol{\beta}_1, \sigma_1^2, \boldsymbol{\beta}_2, \sigma_2^2$ will estimate a set of coefficients for each latent class and also the proportion of the sample that falls in each latent class. From this, I calculate the posterior probability that a respondent is in a specific latent class, shown here for being in group 1:

$$\begin{aligned} & \Pr(\text{Respondent } r \in \text{Group 1}) \\ &= \frac{\pi_1 * \prod_{i=1}^I f_1(y_{r,i} | \mathbf{x}_{r,i}, \boldsymbol{\beta}_1, \sigma_1^2)}{\pi_1 * \prod_{i=1}^I f_1(y_{r,i} | \mathbf{x}_{r,i}, \boldsymbol{\beta}_1, \sigma_1^2) + \pi_2 * \prod_{i=1}^I f_2(y_{r,i} | \mathbf{x}_{r,i}, \boldsymbol{\beta}_2, \sigma_2^2)} \end{aligned}$$

2.3 Laboratory study: Results

2.3.1 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.

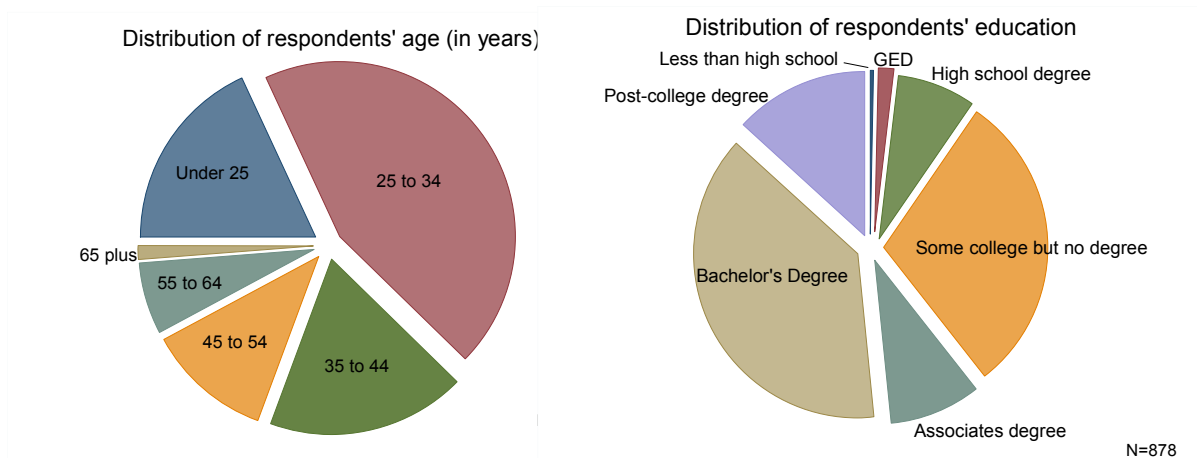


Figure 2: Self-reported age and education among MTurk respondents N=878

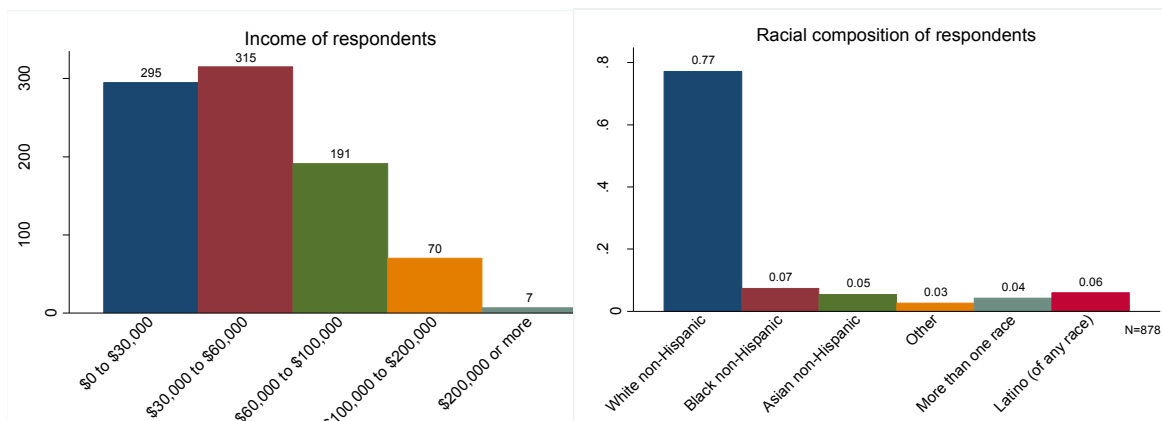


Figure 3: Self-reported household income and race/ethnicity among MTurk respondents. N=878

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 vast majority of respondents live in households with an income of \$60,000 or less. The respondents hold predominantly liberal views; 74% of respondents agree or strongly agree that same-sex marriage should be legal.

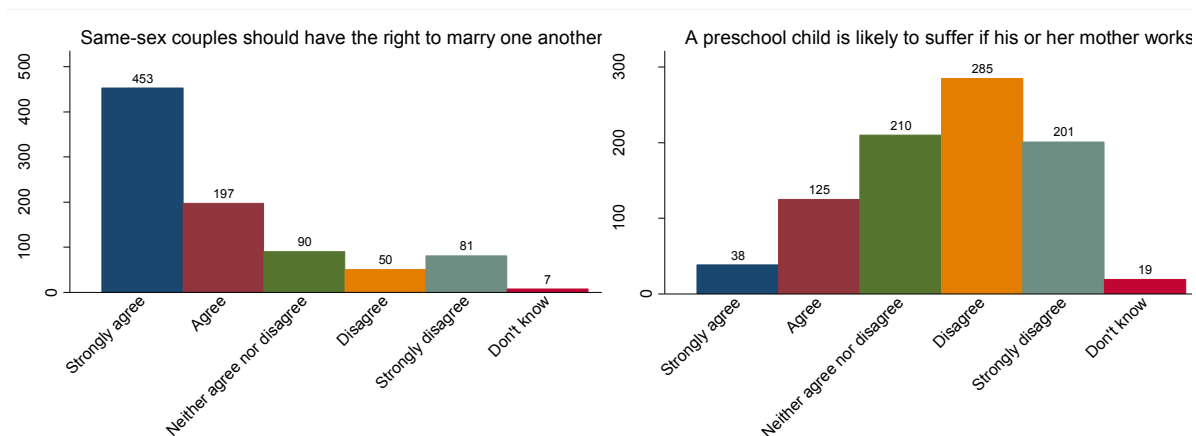


Figure 4: MTurk respondents' answers to questions from the General Social Survey about gender roles and same-sex marriage. N=878

2.3.2 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).

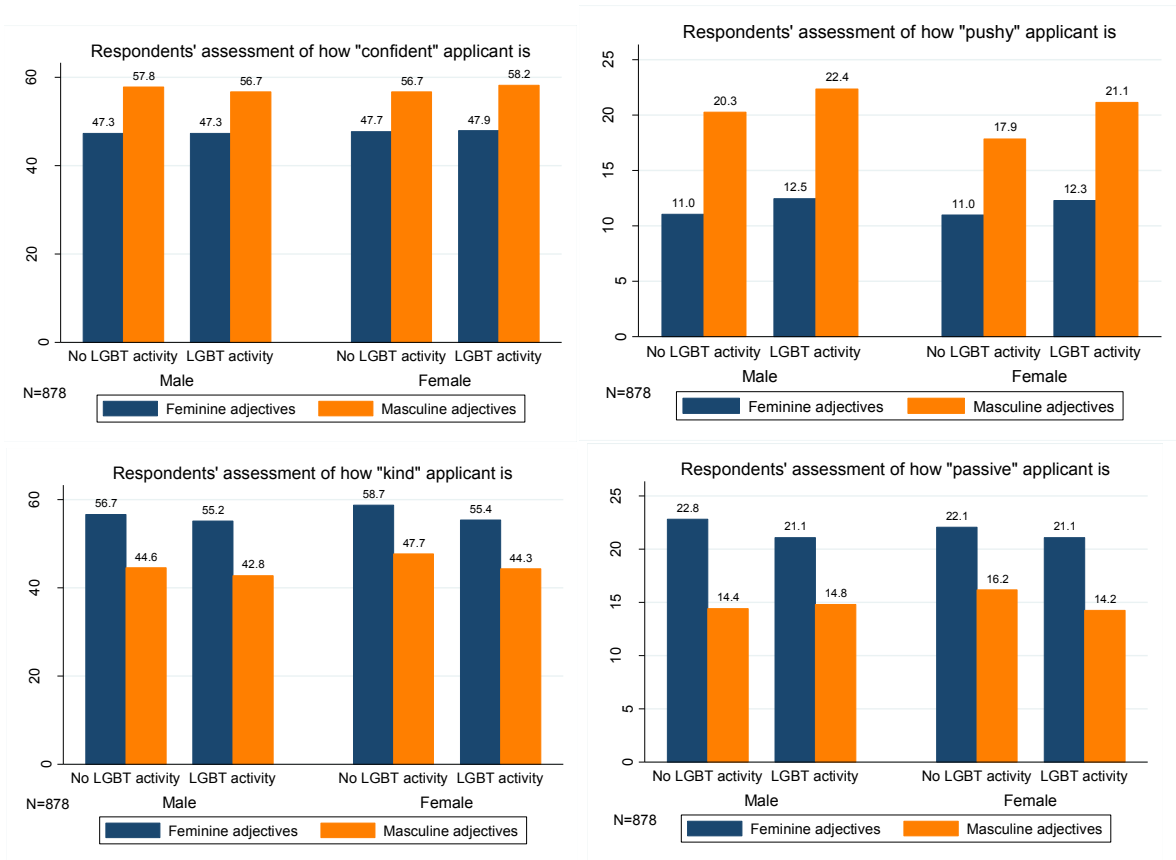


Figure 5: Assessment of the applicant's personality. N=878 Outcome variables could take on values from 0 to 100.

Respondents evaluated eleven different personality characteristics on how well they described the applicant (from 0 to 100). The Figure 5 shows 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.

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. 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 (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.

2.3.3 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 with the non-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 (robust standard errors clustered by respondent and respondent fixed effects).

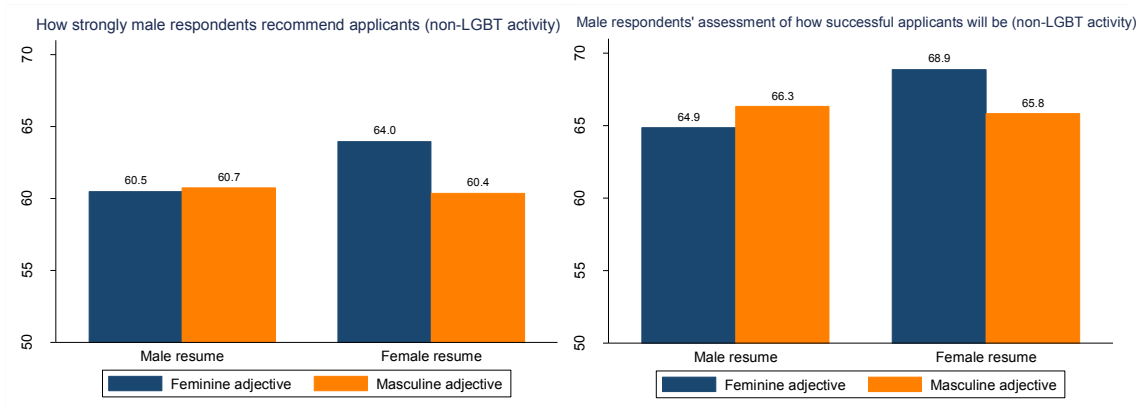


Figure 6: Assessment of the applicant's hireability (for resumes with the non-LGBT activity). N=416 Outcome variables could take on values from 0 to 100.

An important note is that women who use feminine language are rated more favorably 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” against women for using masculine adjectives could also be viewed as a premium for women using feminine adjectives. To examine this pattern 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 the non-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 pattern 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. This 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.

Table 1 Results of an OLS regression of hireability variables on the type of adjective used on the resume, a female indicator, and their interactions. Outcome variables could take on values from 0 to 100.

	Recommend	Successful	Salary	Committed	Willing to work
Masculine adjective	0.123 (1.423)	1.342 (1.259)	1.125 (1.070)	-1.387 (1.081)	-2.593** (1.286)
Female applicant	3.512** (1.388)	4.059*** (1.228)	2.113** (1.022)	1.590 (1.016)	3.923*** (1.132)
Masculine resume & female applicant	-3.461* (1.997)	-4.226** (1.685)	-2.488* (1.432)	-0.516 (1.514)	-2.128 (1.708)
Obs	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
<i>Robust standard errors in parentheses, clustered by respondent</i>					
<i>*** p<0.01, ** p<0.05, * p<0.1</i>					

Female respondents have a similar pattern among the coefficients' signs; however, they are not statistically significant. This pattern may reflect either a smaller effect among women or different proportions of latent classes resulting in a different average, as examined later.

Table 2: Results of an OLS regression of hireability variables on the type of adjective used on the resume, a female indicator, and their interactions Outcome variables could take on values from 0 to 100.

	Recommend	Successful	Salary	Committed	Willing to work
Masculine adjective	-0.344 (1.362)	0.538 (1.204)	0.402 (0.951)	-0.606 (1.094)	-6.290*** (1.256)
Female applicant	1.616 (1.350)	1.625 (1.156)	0.962 (0.904)	1.048 (1.051)	2.942*** (1.075)
Masculine adjective & female applicant	-1.054 (1.890)	-1.085 (1.541)	-0.621 (1.281)	-0.708 (1.425)	0.356 (1.575)
Obs	1,844	1,844	1,844	1,844	1,844
R-squared	0.628	0.607	0.556	0.614	0.589
Resumes	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
Cluster by Respondent	Yes	Yes	Yes	Yes	Yes
Respondents	Female	Female	Female	Female	Female

Robust standard errors in parentheses, clustered by respondent
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

2.3.4 Does “backlash” vary by sexual orientation?

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 with the non-LGBT activity who use feminine language are thought of as more successful and receive higher recommendations than those who use masculine language. This pattern 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.

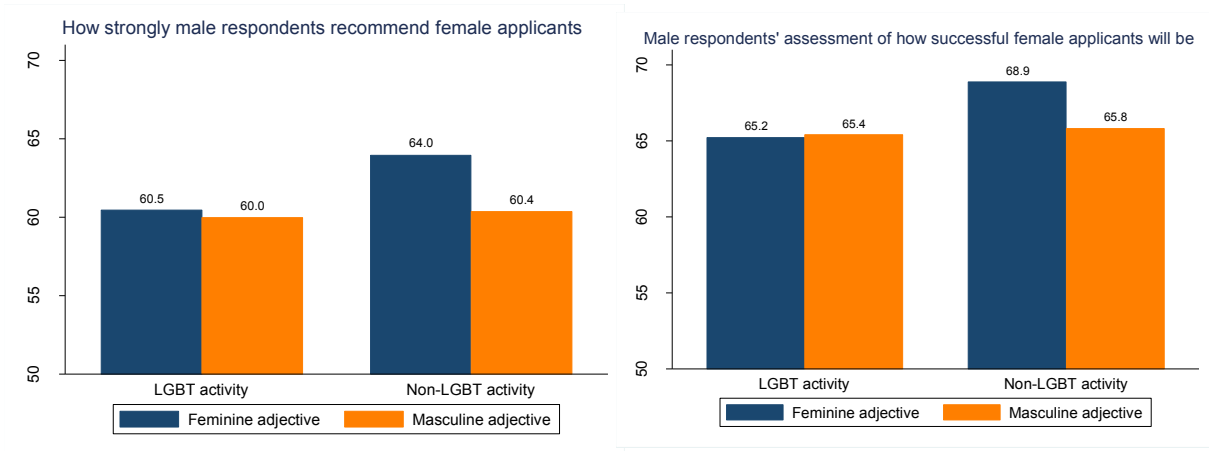


Figure 7: Assessment of applicant's hire-ability (for resumes of female applicants). N=416 Outcome variables could take on values from 0 to 100.

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.

Table 3: Results of an OLS regression of hireability variables on the type of adjective used on the resume, an LGBT indicator, and their interactions. Outcome variables could take on values from 0 to 100.

	Recommend	Successful	Salary	Committed	Willing to work
Masculine adjectives	-3.489*** (1.336)	-2.961*** (1.115)	-1.425 (0.987)	-2.068** (1.025)	-4.849*** (1.177)
Resume with LGBT activity	-4.623*** (1.437)	-4.422*** (1.142)	-3.037*** (1.092)	-3.984*** (0.991)	-4.750*** (1.264)
Masculine adjective and LGBT activity	3.217* (1.835)	3.243** (1.580)	2.609* (1.415)	3.032** (1.388)	1.690 (1.658)
Obs	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, clustered by respondent
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Among female applicants who use feminine language, those without the LGBT activity are viewed more positively than women with the LGBT activity. 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 a more muted pattern.

Table 4: Results of an OLS regression of hireability variables on the type of adjective used on the resume, an LGBT indicator, and their interactions. Outcome variables could take on values from 0 to 100.

	Recommend	Successful	Salary	Committed	Willing to work
Masculine adjective	-1.369 (1.398)	-1.287 (1.030)	-0.536 (1.122)	-0.107 (0.926)	-5.921*** (1.168)
LGBT activity	-2.244 (1.403)	-1.565 (1.029)	-2.028* (1.160)	-1.353 (0.917)	-3.228*** (1.179)
Masculine adjective and LGBT activity	1.514 (1.913)	-0.151 (1.473)	0.713 (1.564)	0.596 (1.318)	0.0491 (1.637)
Obs	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, clustered by respondent
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The following graph examines resumes from male applicants, rather than female. For male resumes, both with the LGBT activity and with the non-LGBT activity, male respondents had slightly higher average ratings for those with the masculine adjectives, but the difference is not statistically significant.

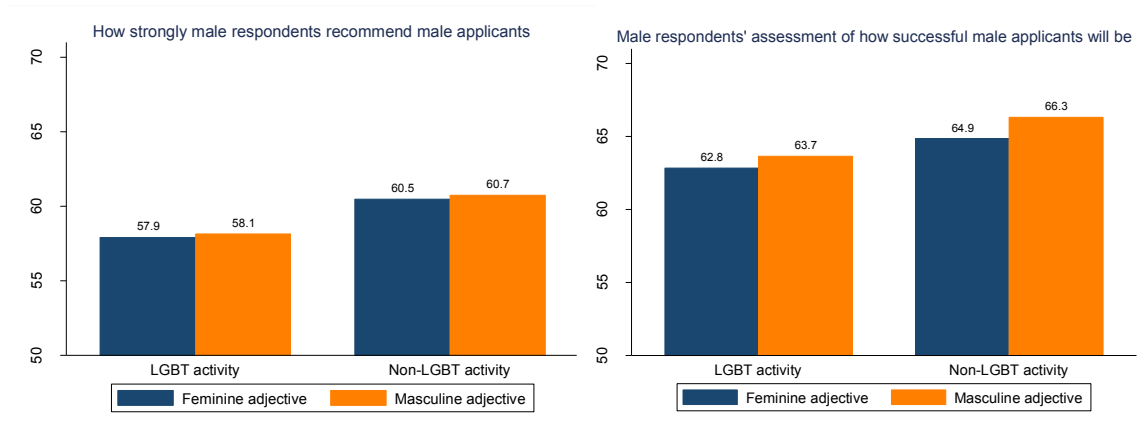


Figure 8: Assessment of the applicant's hire-ability (for resumes of male applicants). N=416 Outcome variables could take on values from 0 to 100.

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.

Table 5: Results of an OLS regression of hireability variables on the type of adjective used on the resume, an LGBT indicator, and their interactions. Outcome variables could take on values from 0 to 100.

	Recommend	Committed	Willing to work	Successful	Salary
Masculine adjective	0.240 (1.416)	-1.299 (1.088)	-2.478* (1.283)	1.439 (1.250)	1.252 (1.065)
LGBT activity	-4.003*** (1.437)	-4.591*** (1.103)	-4.280*** (1.303)	-2.884** (1.249)	-3.103*** (1.047)
Masculine adjective and LGBT activity	0.951 (1.894)	2.344 (1.442)	0.273 (1.718)	-0.0105 (1.692)	1.194 (1.402)
Obs	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
<i>Robust standard errors in parentheses</i>					
<i>*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$</i>					

The respondents' assessments of male applicants are not impacted by the use of masculine language. This pattern holds for both those with and without the LGBT activity.

2.3.5 Finite mixture model

To investigate if the premium for women without an LGBT activity who use feminine language (or the backlash to using masculine language) is the average between two latent classes, I apply the finite mixture model from the likelihood function in Equation 5. This analysis reveals two distinct classes among male respondents, but not among female respondents. In the following regressions, all the data is demeaned by respondent to account for the dependence between observations within the same respondent.

Among male respondents, there is one class that comprises 28% of the sample and shows no statistically significant difference between resumes of different sex and adjective combinations. The second latent class has a strong effect, with a large and statistically significant premium when women use feminine language relative to men who use feminine language and a large negative difference in the difference between men and women who use masculine language relative to feminine language.

Table 6: Results of a FMM regression of how strongly respondents would recommend the applicant on the type of adjective used on the resume, a female indicator, and their interactions.

	Recommend	Recommend	Recommend	Recommend
	<i>Latent class 1</i>	<i>Latent class 2</i>	<i>Latent class 1</i>	<i>Latent class 2</i>
Proportion in class	.28	.72	.21	.79
Masculine adjectives	-1.807 (1.658)	1.071 (1.909)	-1.504 (1.839)	0.0952 (1.753)
Female applicant	0.398 (1.672)	4.691** (1.900)	0.595 (1.812)	1.969 (1.753)
Female applicant & masculine adjectives	1.237 (2.197)	-5.859** (2.690)	2.391 (2.456)	-1.822 (2.479)
Constant	2.193*** (0.540)	0.714 (0.674)	2.278*** (0.622)	0.720 (0.625)
Observations	1,664	1,664	1,844	1,844
Resumes	No LGBT activity	No LGBT activity	No LGBT activity	No LGBT activity
Demeaned by respondent	Yes	Yes	Yes	Yes
Respondents	Male	Male	Female	Female

Standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In contrast, while there are two classes among female respondents, they are very similar to each other. Neither class has a statistically significant difference in the

difference between men and women who use masculine language relative to feminine language.

A similar pattern holds when examining the difference between resumes from female applicants with the non-LGBT activity and with the LGBT activity who either use masculine or feminine adjectives. Among male respondents, there is a smaller latent class that has no statistically significant difference between the LGBT activity and adjective combinations. However, the larger class has a strong effect: for women with the non-LGBT activity there is a large negative effect of using masculine adjectives (or a positive effect of using feminine adjectives). For those women with the LGBT activity, there is no difference between using the feminine and masculine adjectives.

Table 7: Results of a FMM regression of how strongly respondents would recommend the applicant on the type of adjective used on the resume, an LGBT indicator, and their interactions.

	Recommend	Recommend	Recommend	Recommend
	<i>Latent class 1</i>	<i>Latent class 2</i>	<i>Latent class 1</i>	<i>Latent class 2</i>
Proportion in class	.39	.61	.25	.75
Masculine adjectives	0.339 (1.571)	-6.064*** (2.214)	-0.245 (1.529)	-1.471 (1.812)
LGBT activity	0.182 (1.557)	-5.813*** (2.178)	-1.264 (1.450)	-3.946** (1.811)
Masculine adjective & LGBT activity	-0.544 (2.118)	5.426* (3.088)	-2.163 (2.248)	3.051 (2.580)
Constant	1.466*** (0.494)	0.605 (0.752)	0.252 (0.519)	0.528 (0.639)
Observations	1,664	1,664	1,844	1,844
Resumes	Female	Female	Female	Female
Respondent FE	Yes	Yes	Yes	Yes
Respondents	Male	Male	Female	Female

Standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

For female respondents, neither latent class has a statistically significant difference for resumes with masculine language versus feminine language among female applicants without the LGBT activity. Neither group has a statistically significant difference in the difference between masculine and feminine adjectives between those resumes with the non-LGBT activity and with the LGBT activity.

2.3.6 Differences for openly LGBT resumes

In the summary statistics and regression results presented above, the resumes with the LGBT activity were perceived differently than those without. 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 2). This result suggests that in addition to the finding that openly gay women are exempt from the backlash effect, there is also evidence of discrimination against resumes with LGBT activities on them.

Table 8: Results of an OLS regression of how “pushy” the respondent rated the resume on an LGBT indicator. Outcome variables could take on values from 0 to 100.

	Pushy	Pushy	Pushy	Pushy
LGBT activity	3.362*** (1.006)	1.966* (1.115)	1.290* (0.688)	1.415** (0.676)
Obs	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, clustered by respondent
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 9: Results of an OLS regression of how much the respondent was “willing to work with” the applicant on an LGBT indicator. Outcome variables could take on values from 0 to 100.

	Willing to work	Willing to work	Willing to work	Willing to work
LGBT activity	-3.088*** (1.044)	-1.949* (1.127)	-3.899*** (1.050)	-2.463** (1.044)
Obs	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, clustered by respondent
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

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 name². The first bullet point in the job description states “Planned and organized events that promoted diversity and raised awareness on various topics.” The entry includes other to demonstrate the magnitude of the role, such as “managed a committee of 10 to 12 members” (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

² Because the analysis was designed to examine differences *within* sexual orientation cells, the resumes were not designed to be identical *across* sexual orientation cells. Only one base resume had identical text for the non-LGBT and the LGBT versions of the resume.

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 regressions include base resume fixed effects; the base resumes have identical work experiences. 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.

Table 10: Results of an OLS regression of how “useful” the respondent rated the applicant’s work history on an LGBT indicator. Outcome variables could take on values from 0 to 10.

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

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, people did not want to work with them, and they were rated more negatively on the other hireability questions. Their extracurricular activity was viewed as less useful than an identical activity that is not LGBT related; this pattern 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 a male 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.

2.4 Conclusions from laboratory study

Numerous laboratory studies in have shown 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 not known if the backlash affects all women equally or if it varies by sexual orientation. If gay women are exempt from the backlash effect, this could play a role in the observed earnings differences based on sexual orientation, where 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).

Indeed, I found that using masculine language on resumes provokes a backlash effect against perceived-heterosexual women, but not perceived-gay women. Male respondents rate perceived-heterosexual female applicants who use masculine adjectives worse relative to when they use feminine adjectives. This backlash pattern does not appear for male applicants. The difference in the difference between male and female applicants who use masculine or feminine language 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, perceived-gay women are completely exempt from the backlash effect. Perceived-gay female applicants are **not** rated any more negatively when they use masculine adjectives relative to when they use feminine adjective. Latent class analysis suggests that this average effect is driven by a larger effect in a majority of male respondents tempered by a null reaction in a minority of male respondents.

Female respondents never engaged in the backlash effect against any group. Even when split into latent classes, both classes of female respondents rated perceived-heterosexual female applicants who used masculine language equally to those applicants who used feminine adjectives.

The finding that only male respondents engage in the backlash effect and only against heterosexual women is entirely consistent with theories of social identity - male respondents suffer identity loss when heterosexual women follow male behavioral prescriptions. Since gay women are not part of men's relevant comparison group, this would predict the finding that gay women do not experience the backlash effect. Likewise, female respondents never engaged in the backlash effect.

These results are inconsistent with the theory that a driving factor is information about productivity. If the impetus was a concern for productivity, male and female respondents should have similar reactions to the adjectives. For productivity to result in only male respondents engaging in the backlash effect and only against heterosexual women, male and female respondents would have to perceive dramatically different informational content in adjectives used specifically by heterosexual women.

Additionally, male resumes with an LGBT activity are penalized. These resumes are viewed more negatively on numerous personality characteristics and have lower rating on the hireability measures. Most importantly, their work history is viewed as less useful when compared to a resume with an identical work history, suggesting that the negative view about an LGBT activity on a man's resume spills over to the evaluation of other unrelated aspects of the resume. This pattern held even among laboratory respondents who reported progressive views on LGBT rights. This suggests that there is

a substantial and persistent penalty for men who include an LGBT activity on their resume.

Taken together, these results suggest that social identity makes heterosexual women in particular vulnerable to backlash when they engage in traditionally male behavior and gay men are particularly vulnerable to discrimination if they list an LGBT-affiliated activity on their resume. These findings are entirely consistent with observed earnings differences based on sexual orientation, where gay women make more money than similar heterosexual women while gay men make less money than similar heterosexual men. These results provide suggestive evidence that part of the observed earnings differences based on sex and sexual orientation are the result of these forms of backlash and discrimination.

3. Testing the backlash effect in the labor market

3.1 Experimental manipulation

To examine if the backlash effect exists when applying to real jobs, I applied for publicly advertised jobs using resumes manipulated on sex, sexual orientation, and the use traditionally masculine or feminine adjectives. 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). Laboratory experiments have many strengths, such as detailed outcome measures and information about the respondents, but they occur in an artificial setting and therefore their results may not generalize to a real labor market. In order to accurately assess discrimination, research must be done in a real labor market.

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.³ 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

³ 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)

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 4.

As in the laboratory study, the objective statement of the resume, a common feature of resumes of recent college graduates, includes adjectives that are either traditionally masculine or feminine. Likewise, the applicant's sex is indicated by the applicant's first name; the same names are used in the audit study as in the laboratory study. Just like the laboratory study, the resume manipulated perceived sexual orientation through a leadership position in a college group. The non-LGBT clubs were drawn from publicly listed resumes and selected randomly in the same manner as the work and education history.

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 4.

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 mitigate the negative impact on employers and other job seekers, immediately after any positive contact from an employer, the employer was informed that the applicant was no longer interested in the position.

This experimental design was approved by the Duke Institutional Review Board.

3.1.1 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.

3.2 Resume audit study: Empirical 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 callback from an employer.

$$\begin{aligned}
 C_j = & \alpha + \beta * I(\text{female applicant}_j) + \theta * I(\text{masculine adjective}_j) + \delta * I(\text{female applicant}_j) * I(\text{masculine adjective}_j) \\
 & + \pi * I(\text{LGBT and male}_j) + \varsigma * I(\text{LGBT and male}_j) * I(\text{masculine adjective}_j) \\
 (7) \quad & + \sum_{c=1}^{C-1} \sum_{i=1}^{I-1} \tau_{c,i} * I(\text{City}_j = c) * I(\text{Industry}_j = i) + \sum_{w=1}^{W-1} \omega_w * I(\text{Work}_j = w) + \sum_{e=1}^{E-1} \omega_e * I(\text{Educ}_j = e) + \eta_j
 \end{aligned}$$

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 δ . The other difference in the difference, $\hat{\zeta}$, estimates if men with the LGBT activity are impacted differently by using masculine adjectives than those without the LGBT activity.

3.3 Resume audit study: Results

3.3.1 Description of jobs and employers

In total, I sent fictional resumes to 1,334 publicly advertised jobs. 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.

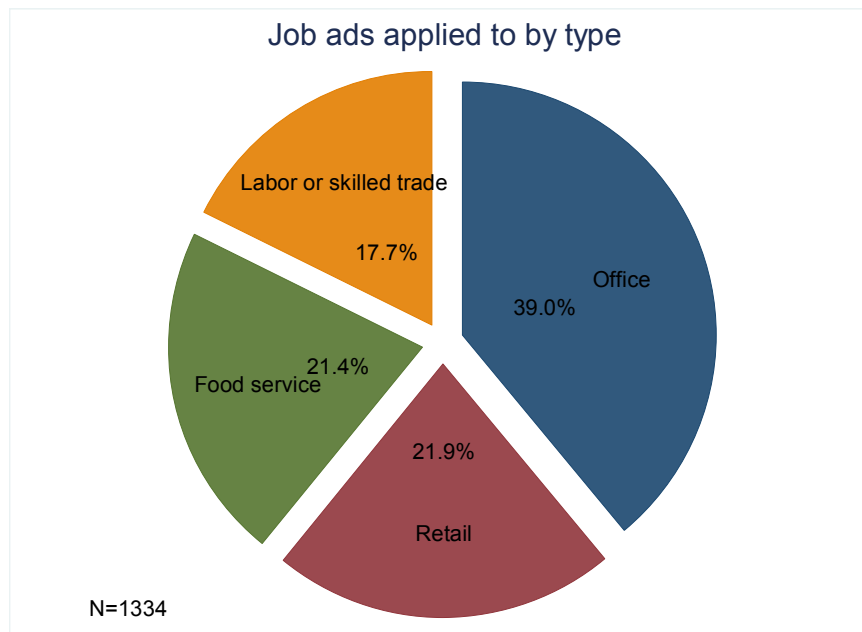


Figure 9: Types of jobs that were applied for. N=1,334

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

Table 11: Examples of common jobs applied for within each type of job

Office	Food service	Retail	Labor/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

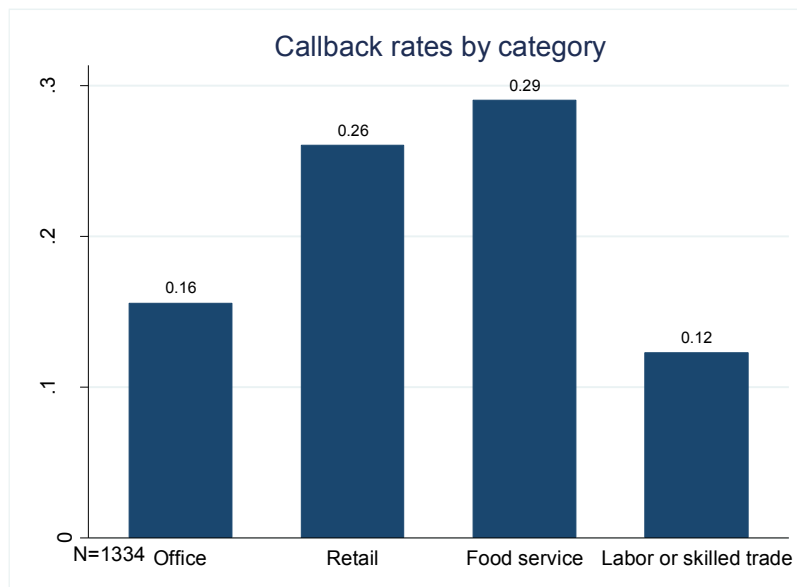


Figure 10: Proportion of resumes called back by category of job. N=1,334

As shown in Figure 10, the callback rates varied by job category; food service had the highest callback rate (29%) and labor or skilled trade the lowest (12%).

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 11, the majority of job advertisements used neutral or no adjectives while about 1/3 used feminine adjectives.

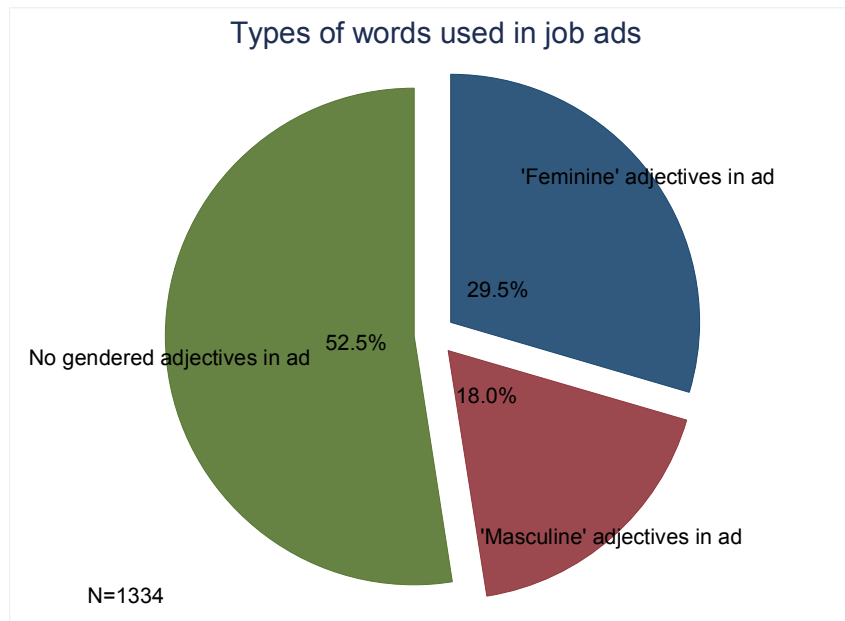


Figure 11: Distribution of the types of adjectives used in the text of the job advertisements applied to

3.3.2 Is there a backlash effect?

Figure 12 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.

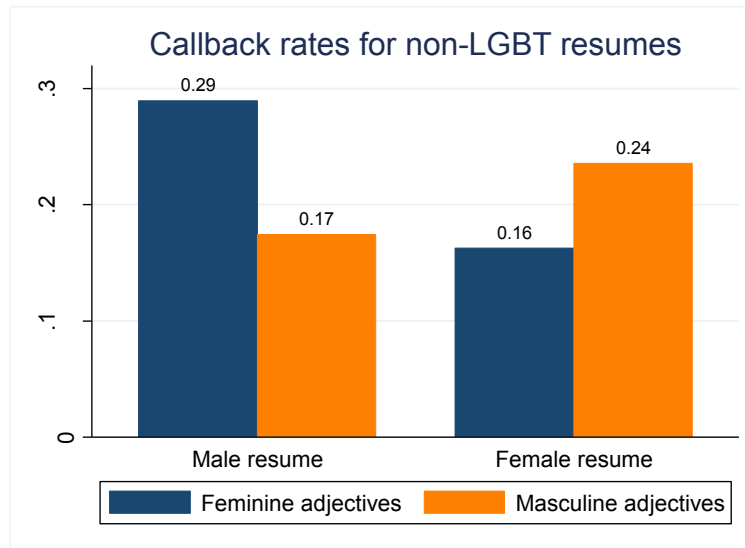


Figure 12: Callback rate by sex and type of adjective used on resumes without the LGBT activity. N=678

As shown in Figure 13, 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.

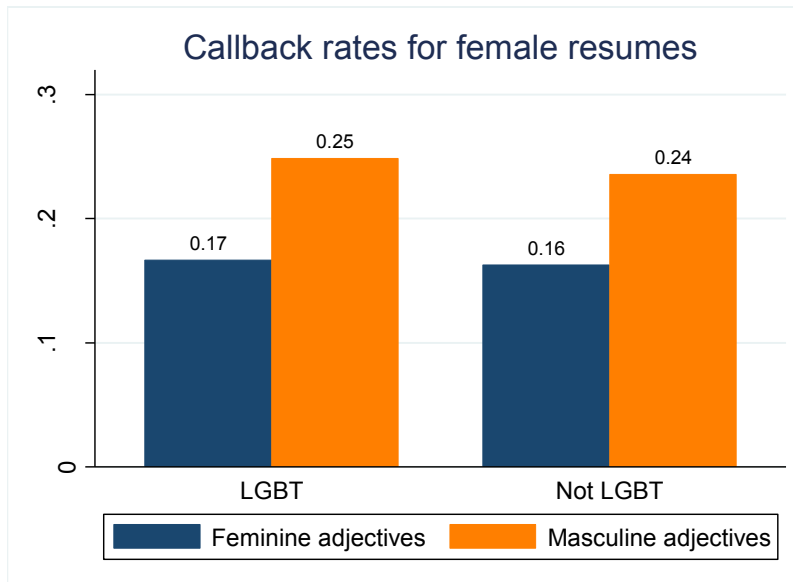


Figure 13: Callback rate by sexual orientation and type of adjective used on female resumes. N=673

The following table shows the patterns from Figure 12 and 13 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 12 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 12 and 13: perceived heterosexual men are called back less when they use masculine language while women are called back more when they use masculine language.

Table 12: 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.

	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
<i>Robust standard errors in parentheses</i>	
<i>*** p<0.01, ** p<0.05, * p<0.1</i>	

3.3.3 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.

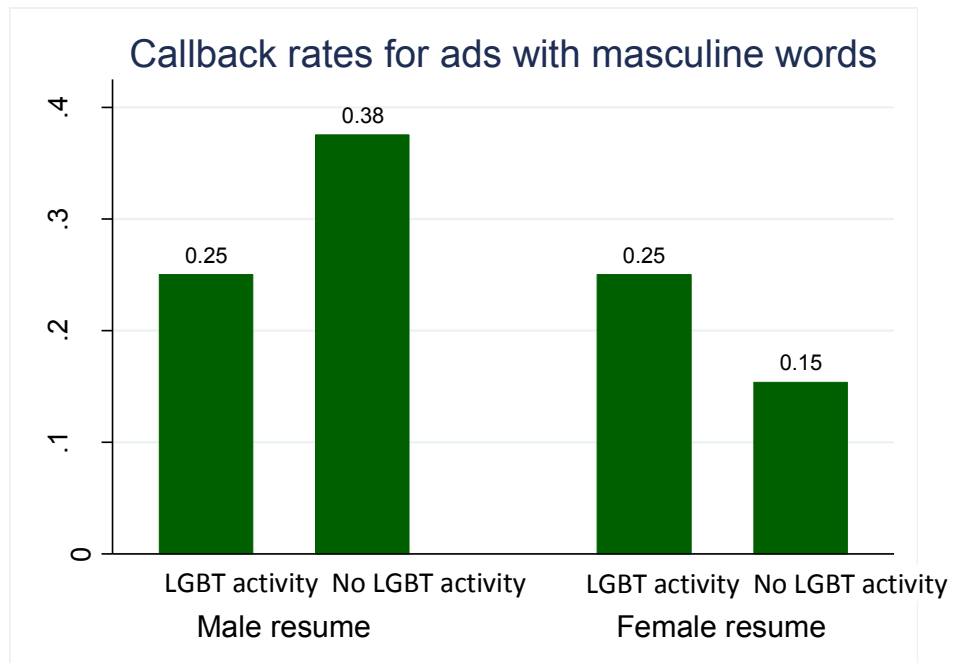


Figure 14: Callback rate by sex, sexual orientation for job ads with traditionally masculine adjectives. N=204.

Moreover, perceived heterosexual women were unable to overcome their low callback rate for these jobs by using traditionally masculine adjectives. As shown in Figure 15, 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.

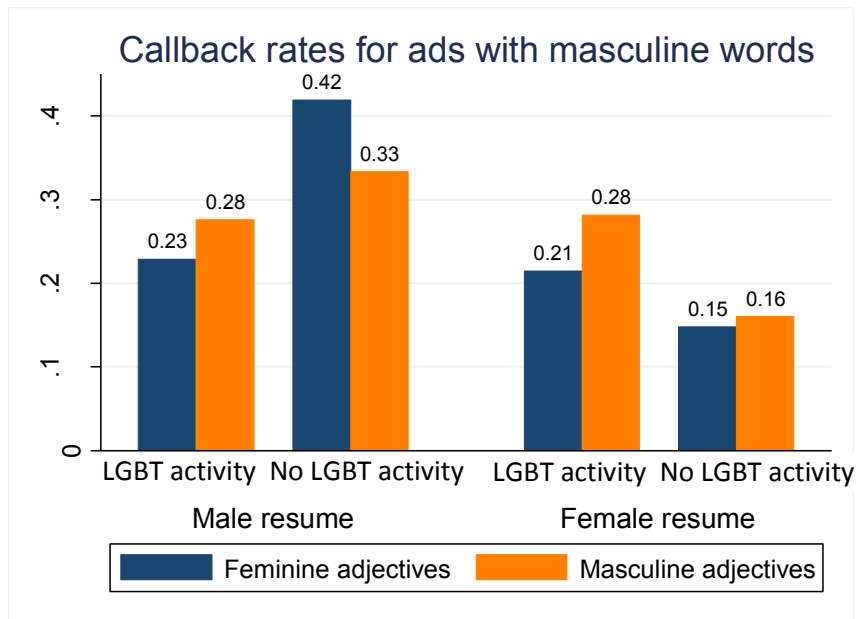


Figure 15: Callback rate by sex, sexual orientation, and adjective for job ads with traditionally masculine adjectives. N=204.

Job ads without gendered language or with traditionally feminine language have less variation in callback rates, as displayed in Figure 16. 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).

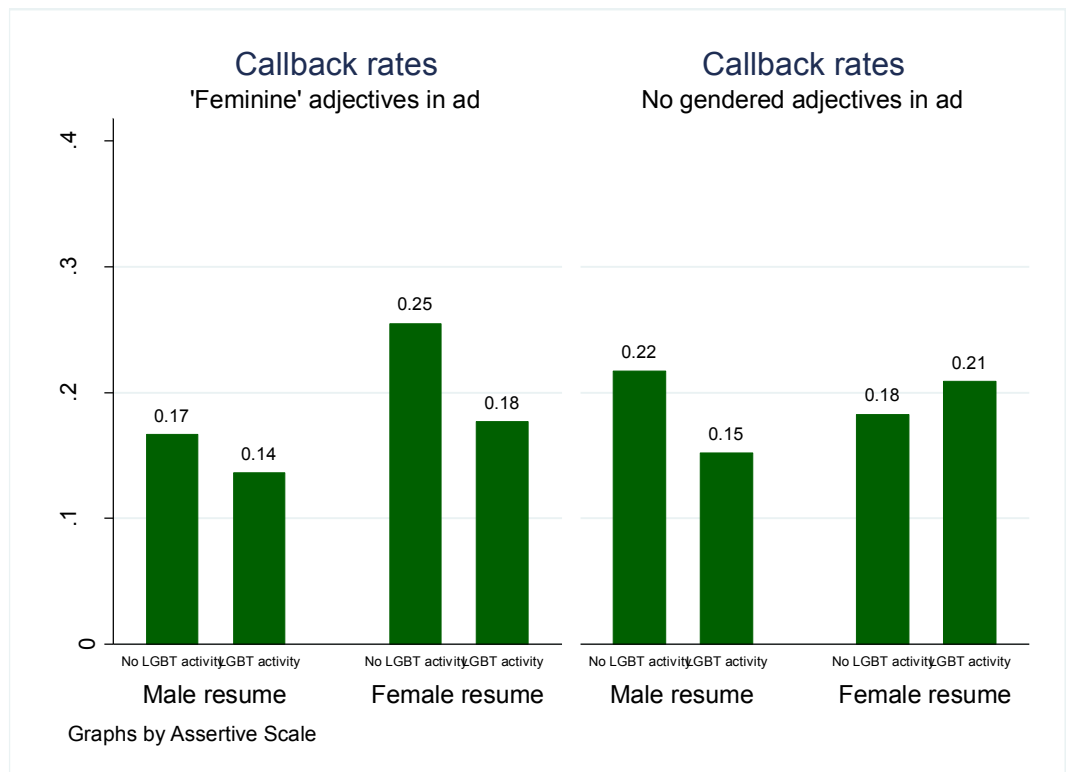


Figure 16: Callback rate by sex and sexual orientation for job ads with feminine language or no gendered language. N=700 without gendered language. N=395 with feminine language.

As Figure 16 shows, perceived-heterosexual women have the highest callback rate for jobs advertised with traditionally feminine language. However, Figure 17 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 results 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.

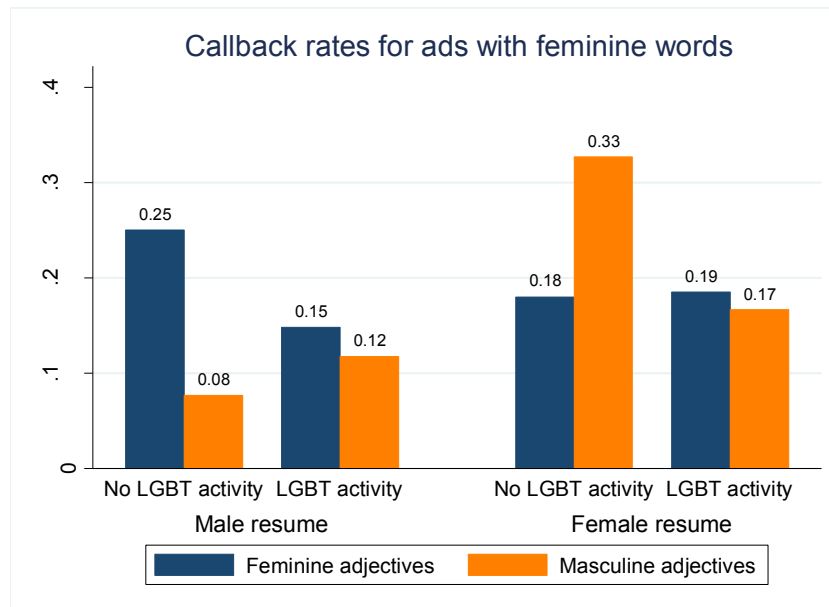


Figure 17: Callback rate by sex, sexual orientation, and adjective used on resume for job ads with feminine language. N=395.

3.3.4 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 18 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.

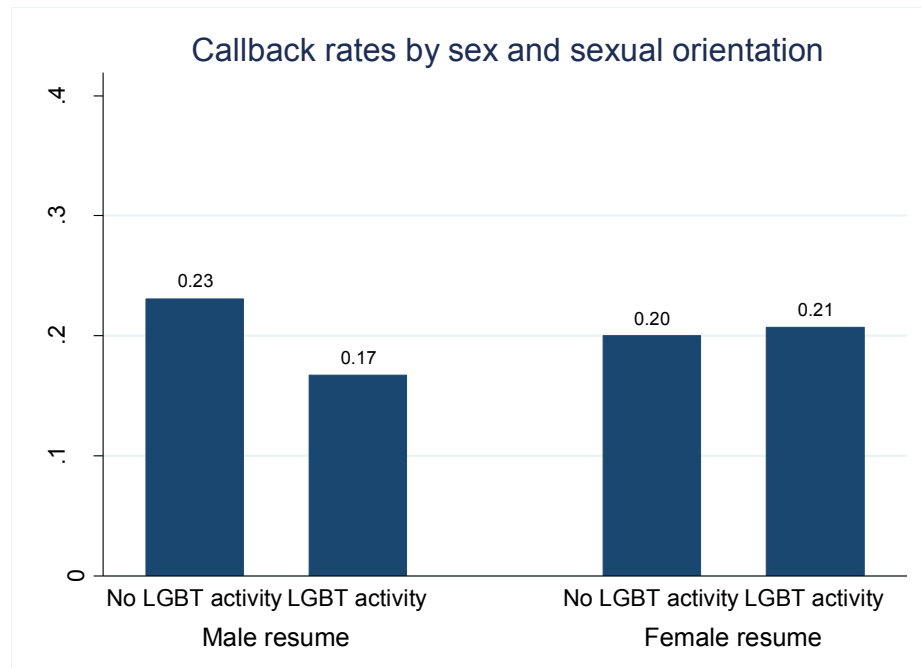


Figure 18: Callback rate by sex and perceived sexual orientation. N=1,334.

3.4 Conclusions from audit study

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 the first study, I found that in a laboratory setting using masculine language on resumes provokes a backlash effect for perceived-heterosexual women. Perceived-gay women are exempt from the backlash effect from male respondents. Equivalently, perceived-gay women miss out on the premium for using feminine adjectives.

However, when I took the question outside of the laboratory and applied for real jobs with the 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. Moreover,

while perceived-gay women were treated very differently than perceived-heterosexual women in the laboratory experiment, in the resume audit study both subgroups of women benefit from using masculine adjectives.

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. That is, both laboratory respondents and real employers may respond negatively to norm-violating behavior, but for real employers it appears to be overwhelmed by the change in expected productivity. This result suggests that we must be cautious when extrapolating from results obtained in a laboratory setting to a real-world labor market.

Further laboratory studies could help parse these differences; for example, informing the respondent that their payment will be tied to a measure of productivity could make respondents more motivated by the perceived productivity of the respondent. To test how the background of the respondent affects the results, I can examine if people with experience working in human resources have different responses in the laboratory setting than the general respondent population.

The difference between the laboratory and the audit study is also consistent with the idea that laboratory respondents are more altruistic than real employers – specifically, that male respondents may act altruistically towards perceived-heterosexual feminine-acting women. The difference between the two studies could also be driven by

other elements of the experimental designs. For example, the laboratory respondents evaluated job candidate for one type of job: an entry level position with a life sciences company. In the audit study, the job type varied, with the most common job being in an office setting. It may be that the difference in the results of the two studies is driven by the type of job. Likewise, laboratory respondents may be more likely to be aware of the actual motive of the experiment and alter their behavior, while an employer in an audit study would have a harder time figuring out that they are in an experiment. These differences all present future research questions.

In both studies, a consistent finding is that male resumes with an LGBT activity are penalized. These resumes are viewed more negatively on numerous personality characteristics, have lower rating on the outcome variables, and their work history is viewed as less useful. This pattern held even among laboratory respondents who reported progressive views on LGBT rights. In the audit study, 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.

4. Decomposing the increase in men's time on childcare during the recession

4.1 Introduction

In the past century, women in the United States dramatically increased the time they spent in the formal labor market. Men are slowly increasing the time they spend on childcare, although women still spend considerably more time on childcare than men. Recent debates continue to highlight women's struggle to maintain a successful career while raising a family (Slaughter 2012), which some argue is caused by the unequal amount of time women spend on childcare (Becker 1985; Hoschchild 1989). To understand the conflict facing women in the United States, it is crucial to examine the other half of the equation: the amount of time men spend on childcare and what factors are associated with changes in the amount of childcare men perform.

In this paper, I examine the change in amount of time men spent on childcare during the recession of 2007-2009. The recession of 2007-2009 provides a sudden and concentrated change in the employment opportunities of men relative to women in the United States. When the male unemployment rate for those over 16 peaked in October 2009 at 11.2%, the female unemployment rate was only 8.7%. I show that this lopsided shock to employment opportunities was accompanied by an increase in the average amount of time men spent on childcare. The majority of the total increase was due to a behavioral change by employed men. In particular, men spent more time on average on

physical care for children, an element of childcare that men have historically performed less often than women.

I use a lagged employment indicator to decompose the total change into the change that is due to behavioral change, change in the composition of each employment group, and change between employment groups. *Behavioral change* refers to an actual change in a man's behavior: for example, men spending more time on childcare during the recession than before indicates a behavioral change. *Compositional change* refers to changes in the composition of the group. The group of employed men before the recession is compositionally different from the group of employed men during the recession. Like traditional decomposition, my technique separates total change into changes that occur *within* a group (e.g. employed men spending more time on childcare) and *between* groups (e.g. employed men who become unemployed and then spend more time on childcare). But the traditional method, which uses repeated cross-sectional data, is unable to distinguish between behavioral and compositional changes within a group. Because the recession did not affect men's employment status randomly, the composition of the employed will be different during the recession than prior to the recession. The lagged employment indicator that I use takes this fact into account and allows for an estimation of the behavioral and compositional components of the "within" change along with the "between" change. In this way, the lagged employment

indicator can provide a more nuanced picture of overall change than the traditional method of decomposition.

This paper builds directly on Casper and O'Connell's (1998) examination of father's involvement in childcare during the 1991 recession. The authors found that more fathers in dual-earner couples were reported as care provider for preschool children during the 1991 recession than before in the Survey of Income and Program Participation. However, instead of a dichotomous outcome variable, I use the outcome measure used in studies of long run changes in men's provision of childcare: time spent on childcare recorded in time diaries (Bianchi et al. 2006; Sandberg and Hofferth 2001; Sayer et al. 2004). That is, this paper analyzes the short run impact of a sudden change in the macroeconomic environment as in Casper and O'Connell (1998), but uses the more sensitive outcome variables often used to study long run changes in the provision of childcare.

This paper also builds on Berik and Kongar (2013), who found that the gap in unpaid labor between married men and married women decreased during the 2007-2009 recession. In this paper, I link each American Time Use Survey (ATUS) respondent to their employment status in the Current Population Survey (CPS) to identify men who recently changed employment status. With this more detailed data, I decompose the total change found by Berik and Kongar (2013) into behavioral, compositional, and between group change. I find that a behavioral change among employed men accounted

for 73.9% of the total increase in time spent on childcare. However, among men who are out of the labor force, the increase in time spent on childcare is entirely due to compositional changes.

The remainder of this chapter is structured as follows. Section 2 describes men's increase in childcare provision over the past decades and how men's childcare provision changes during recessions. Section 3 describes the ATUS and the measures of childcare used in the analysis. Section 4 presents stylized facts about the 2007-2009 recession's impact on employment and childcare. Section 5 describes the changes in the composition of each employment status during the recession and decomposes the total increase in time on childcare into behavioral, composition, and between group change. Section 6 examines time spent on other activities. Section 7 offers discussion of the results.

4.2 Background literature

Historically, women have spent much more time on childcare than men. Men and women also perform different types of childcare; mothers spend a much higher proportion of their total childcare time on physical care for children, whereas fathers spend a higher proportion of their time engaging in activities with children (Berik and Kongar 2011; Craig 2006). Parents report that they enjoy engaging in activities with their children more than providing general childcare (Robinson 1993).

In the last fifty years, men in the United States have increased the time they spend on childcare. From 1965 to 2000, a period that experienced changing social norms

and a dramatic increase in women's time spent in the labor market, men and women both changed the amount of time they spent on childcare (Bianchi et al. 2006). Mothers, married mothers, and married fathers all increased the time they spent on primary childcare, but fathers provided an increasing share of the childcare during this time. That is, fathers responded to the long run changes in social norms and in the employment opportunities of their wives by increasing the time they spent on primary childcare. Fox et al (2013), Sayer, Bianchi, and Robinson (2004), Sandberg and Hofferth (2001), and Bianchi (2011) all find similar results when examining changes in the time children spent with parents over similar time periods.

Economic theory also highlights numerous ways in which a *short run* change in men's work opportunities, such as a recession, could impact the time they spend on childcare. First, when a man experiences unemployment, he has more time to devote to non-work tasks. Aguiar, Hurst, and Karabarbounis (2013) find that 50% of foregone hours of work during the 2007-2009 recession went to leisure (with the largest components being sleep and watching television), 30% to household tasks, and 5% to childcare for both men and women. A very small amount of forgone time is spent on searching for a new job. Likewise, becoming unemployed is associated with men spending more time on household production; this result is found both when examined with panel data (Gough and Killewald 2011) and cross sectional data (Burda and Hamermesh 2010). More broadly, as male work opportunities worsen, the opportunity

cost of men's time decreases. Men spend more time on other activities, because they are now less costly. This change is even more dramatic if the man gains the comparative advantage in household labor and specializes in household production (Becker 1981). Consistent with the hypothesis that the 2007-2009 recession altered how people spent time, Starr (2014) finds evidence of the "added worker effect," wherein women took on more breadwinner responsibilities when their husbands/partners lost their jobs. Specifically, she finds that women with non-employed partners had higher employment rates during the recession than before, while all other groups had lower employment rates during the recession.

Parents may also bargain about how to spend their time. Broadly, bargaining frameworks suggest that when men's job opportunities worsen relative to women's, men's bargaining power drops (for example, McElroy and Horney (1981), Lundberg and Pollack (1993), Lundberg and Pollack (1994), and a broader framework in Browning, Bourguignon, Chiappori, and Lechene (1994)). Empirical explorations have found evidence that control over income and other measures of bargaining power impacts consumption, division of labor, and children's education (for example, Browning et al 1994; Rangel 2006; Thomas 1990). This suggests that bargaining power may impact the division of household labor and investment in children — both elements that would impact the time men spend on childcare.

Moreover, childcare, unlike other types of household production, is less flexible (Kimmel and Connelly 2007; Coltrane 2000; Hook 2010). While yard work can be put off until the weekend, parents cannot leave a young child unsupervised. This fact could make childcare specifically sensitive to bargaining power because the only question is who will perform the task. When or if it will be performed is nonnegotiable. Bloemen and Stanca (2014) find a positive association between a wife's wage and her husband's time spent on childcare. However, total childcare remains constant — as the wife's wage increases, her husband's increase in childcare exactly offsets her own decrease in childcare. In contrast, the authors find that total housework decreases as the wife's wage increases, because the wife decreases her time on housework and her husband does not increase his time on housework enough to make up the difference. The finding that total childcare remained constant despite changes in women's wages is consistent with the idea that childcare is less discretionary than other tasks, and therefore potentially more subject to bargaining power.

Additionally, when male employment opportunities worsen, total family income may fall, even if the "added worker" increases her time in the labor market. As a result, if a family was purchasing childcare services, the family may shift away from these arrangements toward providing childcare themselves. Likewise, expectations about the future, how much a man values the future, as well as men's preferences for how to spend time generally will play a role in decisions about how to spend time on any given

day. If childcare is an investment that will pay off at a future date (i.e., by producing a happy productive adult child who will support their parents) childcare is more attractive than other activities that do not have a future payoff, such as watching television. However, if a man does not value the future as much, this payoff may not be enough to shift his behavior from more immediately enjoyable activities toward childcare.

Finally, sociological theory notes that social context may contribute to the persistent difference in the amount of time men and women spend on childcare (Coltrane 2000 contains a review). That is, there may be stigma for not conforming to social norms regarding the provision of childcare, childcare may be part of the “performance” of gender, or men and women may have acquired different levels skill with household labor, all of which may impose additional costs on a nontraditional division of labor. Since childcare is traditionally viewed as a feminine task, social context could mitigate other incentives for men to do more childcare.

Previous literature has found that men’s time on childcare does respond to short-run changes in economic opportunities. For example, men’s participation in childcare increased during the 1991 recession (Casper and O’Connell 1998) and the 2007-2009 recession (Berik and Kongar 2013). However, these studies leave it unclear whether the change in childcare came exclusively from the extra time that unemployed men would have spent working or if men whose employment status did not change responded to

the broader economic environment by spending more time on childcare. That is, a change at the average could come from a small proportion of the population (men who became unemployed) who spent dramatically more time on childcare or it could come from a large proportion of the population (men who remained employed) with a more modest increase in time spent on childcare. If the latter is true and men are responding to the broader economic environment, this suggests that economic incentives overwhelm patterned behavior regarding childcare. This paper addresses this question by decomposing the total change into between group change, behavioral change, and changes in the composition of each employment group.

4.3 Data source

In this paper, I use American Time Use Survey (ATUS) data from 2003 to 2011. The ATUS is an ongoing monthly survey of a subsample of the Current Population Survey (CPS) that occurs two to five months after the respondent's outgoing CPS survey. The ATUS contains detailed information on how the respondents spent a designated 24 hour period. The ATUS selects a stratified sample of households in the CPS in order to be able to compare households across demographic characteristics, such as race and the presence of children. One member of the household who is 15 or older is randomly selected to participate in the ATUS. The ATUS interviewer asks the respondent to describe in their own words their survey day starting at 4 am. The interviewer then classifies the activities into a coding structure with 17 major activities.

These major activities are disaggregated into more than 100 intermediate activities and more than 400 detailed activities (Abraham et al 2011). The ATUS includes numerous important demographic characteristics of the respondent and members of their household, a link to the respondent's employment status in their outgoing CPS interview, and their current employment status.

I restrict the analysis to weekdays only, because men have different patterns of time on childcare on weekdays and weekends. For example, employed men spend more time on primary childcare on weekends, while unemployed men spend less. Results for weekends are similar to weekdays but muted (available upon request). The ATUS oversamples weekend days, so this reduces the number of respondents significantly (Abraham et al 2011).

4.3.1 Measures of childcare

The ATUS includes numerous measures of time spent on childcare, and the broadest categories are time spent on primary care and time spent on secondary care. It is important to differentiate between types of childcare when considering a possible change in time use patterns (Wang and Bianchi 2009). Bianchi (2000) describes the distinction between direct care for children, which she describes as potentially "valuable" and an "investment" in children, and secondary childcare, which is more passive, with the parent being "on call" (p. 405).

Primary care occurs when the respondent's primary activity involves directly caring for a child or activities that promote a child's welfare. Examples of this type of care include physical care, reading to a child, playing with a child, or helping a child with homework. Secondary childcare includes caring for a child while also doing something else; for example, secondary childcare occurs if a respondent is working and also has a child in his or her care. Secondary care is measured in the ATUS for children under 13, while primary care is measured for those under 18.

The analysis in this paper focuses on time spent on primary childcare. This measure of primary childcare totals the number of minutes per day the respondent spent on primary childcare, including children who live in the same household as the respondent and those who live in a different household. For example, if a man co-resides with one child and not another, this measure of primary childcare counts equally the time he spent with either child. Analyses presented in the text address men who live with at least one child; results for all men are in Appendix 6.

4.4 Stylized facts about the 2007-2009 recession, employment, and childcare

The 2007-2009 recession (December 2007 to June 2009) was marked by a sudden and dramatic increase in the unemployment rate across the United States. The national unemployment rate doubled from 5.0% of the labor force in December 2007 to its peak at 10.0% in October 2009. The rate of unemployed, under-employed, and discouraged workers grew from 8.8% in December 2007 to 17.2% at its peak in October 2009.

The unemployment rate for men was much higher than women during the 2007-2009 recession: the male unemployment rate peaked in October 2009 at 11.2% while the female unemployment rate was 8.7%. The unemployment rate in predominantly male sectors increased more than in predominantly female sectors. The difference was so striking that economists and commentators referred to the recession as a “man-cession” (Wall 2009). Figure 1 shows the unemployment rates of men and women alongside the unemployment rates in manufacturing (72.7% male) and education (76.7% female) sectors. This pattern is similar to other measures of employment; for example the percent of men who were employed dipped during the 2007-2009 recession much more strongly than that of women.

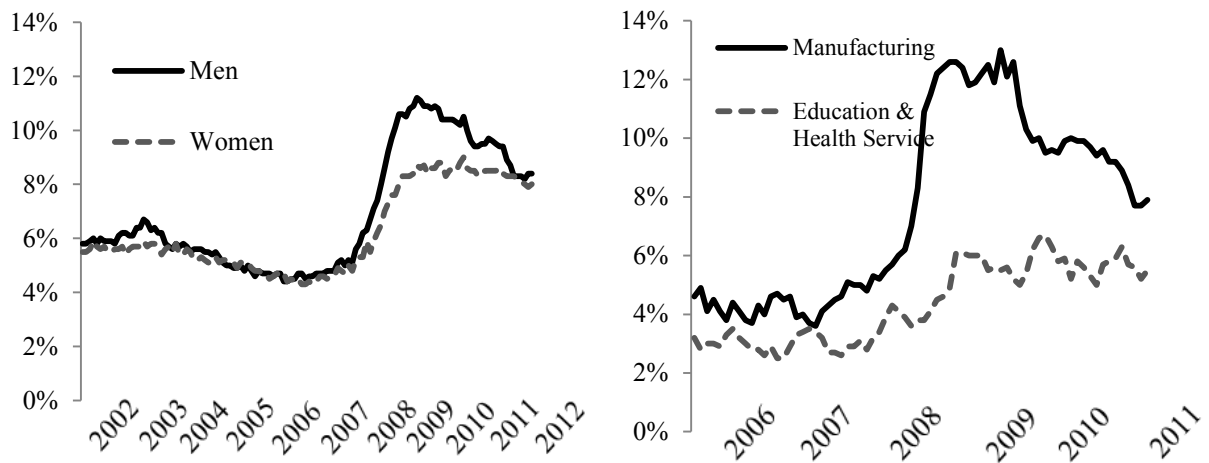
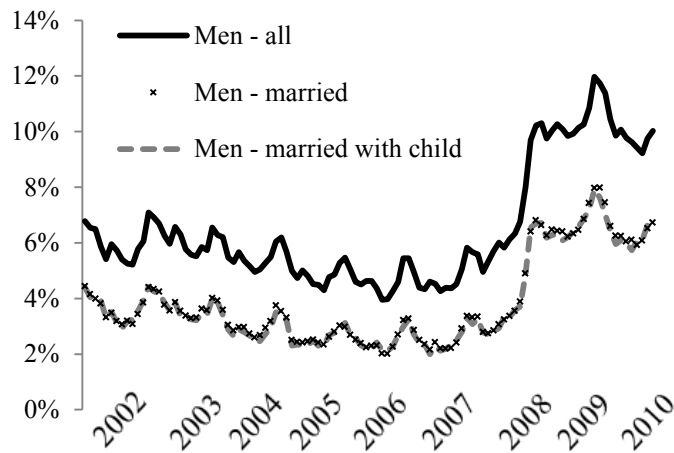


Figure 19: Unemployment Rate for those 16 years and older by sex (L) and by sector (R)

The pattern of male unemployment was similar for all men, married men, and married men with children. While married men and married men with children have

lower unemployment rates than men as a whole, their unemployment rate also increased dramatically during the recession.

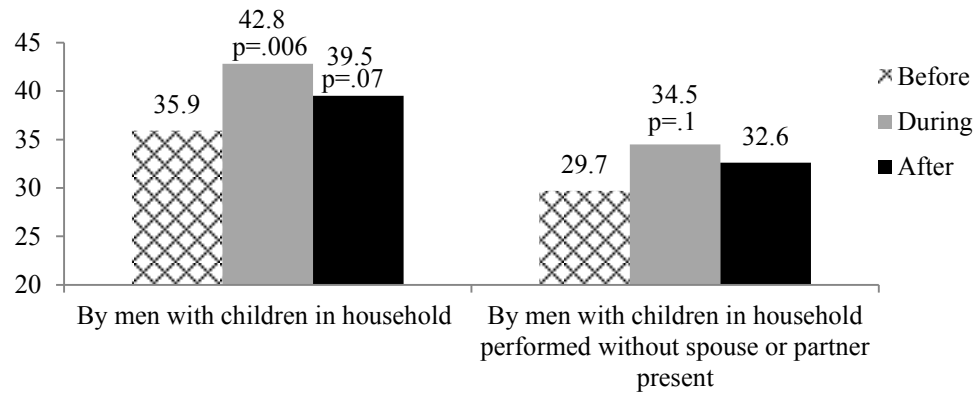


Source: Current Population Survey
 Figures use complex sampling weights provided by CPS

Figure 20: Unemployment Rate (not seasonally adjusted) for men by marital status and presence of children

The lopsided increase in unemployment during the 2007-2009 recession was accompanied by an increase in the average amount of time men spent on childcare. Figure 21 shows that the average time men who lived in a household with children spent on primary childcare increased by 6.9 minutes per day during the recession. This additional time overlapped somewhat with time spent by their spouse or partner; however, the right panel shows that the average time men spent caring for their children without their spouse or partner also increased by 4.8 minutes per day. As shown in Appendix 6, these increases persist after controlling for a linear year trend and

demographic characteristics. Women did not have a statistically significant change in the average amount of time they spent on childcare (Appendix 6). The proportion of men engaging in any childcare also did not change (Appendix 6).



Figures use complex sampling weights provided by ATUS-X. The pvalue shows the result for testing for a difference from the before value using robust standard errors and observations clustered by state. N=12,224

Figure 21: Minutes of primary childcare

Examining the time spent on primary childcare conditional on having a positive value demonstrates a similar pattern as the unconditional measures. Among men who engaged in any primary childcare, the average amount of primary childcare increased by 11.4 minutes during the recession. Likewise, the average amount of primary childcare without their spouse or partner present increased by 10.6 minutes.

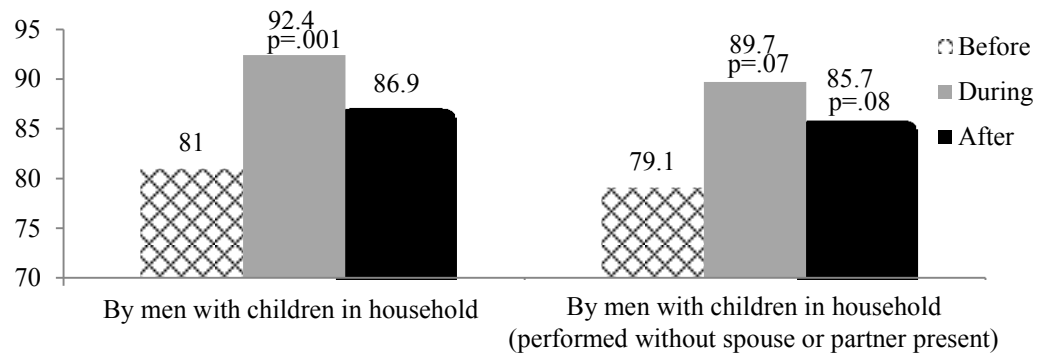


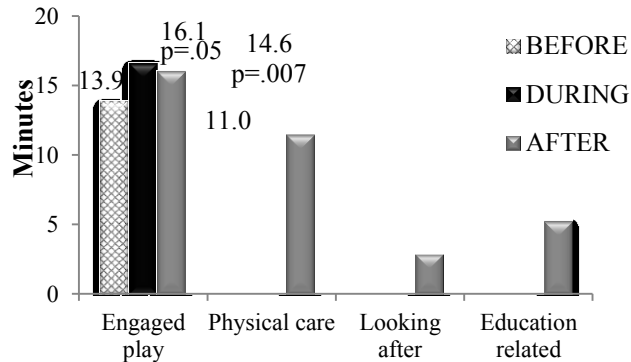
Figure 22: Minutes of primary childcare for those with positive values

Men and women have historically engaged in different activities when they care for children. Mothers spend a much higher proportion of their total childcare time on physical care for children, while fathers spend a higher proportion on engaging activities with children (Berik and Kongar 2011; Craig 2006).

To analyze these more specific types of childcare, I created more detailed time use variables within primary childcare. The first variable looks at engaged interaction with children, including playing with children, arts and crafts, playing sports with children, talking to or listening to children, and attending children's events. The second variable is comprised of physical care for children, such as putting a child to bed, changing a diaper, bathing, feeding, or rocking a child. The third variable focuses on more passive care, i.e., looking after, watching, or supervising a child without interacting with the child. The final variable sums the time spent on activities related to education, reading, or teaching the child. These distinctions will show whether the

increase in male time spent on childcare included these forms of engaged interaction or whether the increase focused on basic physical care.

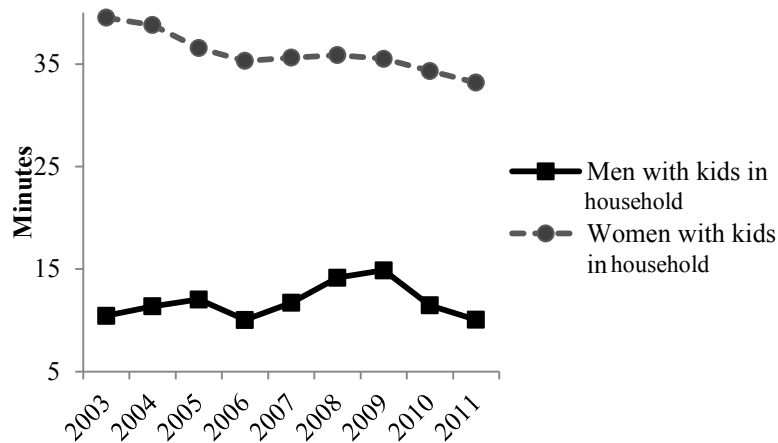
Figure 23 shows that the average time men spent on physical care increased by 3.6 minutes during the recession. As shown in Appendix 6, the increase on physical care remains statistically significant after controlling for a linear year trend and demographic characteristics.



Figures use complex weights provided by ATUS-X. Statistically significant differences noted with pvalue for testing the difference from the before value using robust standard errors and observations clustered by state. N=12,224

Figure 23: Time spent on types of childcare by men with children in household

The figure below shows the average time spent on physical care for children by year. While women’s average time on physical childcare steadily decreased, men’s average increased suddenly in 2007 and continuing through 2009. In 2010 and 2011, men’s average returned to the pre- recession levels of physical childcare.



Figures use complex weights provided by ATUS-X. N=12,224 (men) N=16,916 (women)

Figure 24: Time spent on physical care for children

An increase in the average amount of time men spend performing physical care for children is striking because it demonstrates that men are not just increasing the time they spend playing with their children, but also they are engaging in the less pleasurable aspects (Robinson 1993) of childcare that they have historically performed less often (Berik and Kongar 2011; Craig 2006).

4.5 Decomposing the total change in childcare

4.5.1 Compositional changes within employment status

A common approach to determining the source of a broad change is to decompose the total change into a within group change and a between group change. With this approach, the increase in time spent on childcare during the recession could be a “between group” change, because more men are unemployed, and unemployed men spend more time on childcare than employed men. The increase could also be a “within

group” change, wherein men who remained employed increased the time spent on childcare.

However, the within/between approach cannot be easily applied in to the 2007-2009 period because the composition of each employment category changed during the recession. The unemployment rate for men 16 years and older more than doubled from 5.1% of the labor force in December 2007 to its peak at 11.2% in October 2009. That is, many men who were employed (or out of the labor force) prior to the recession became unemployed during the recession; these new entrants to unemployment likely differ from those who were unemployed prior to the recession. Similarly, discouraged workers who left the labor force may have also changed the composition of those out of the labor force may have changed during the recession.

The following table shows the educational distribution for men who are unemployed⁴ or out of the labor force in the CPS survey before the recession (1/2003 to 11/2007) and during the recession (12/2007 to 6/2009). Figures use complex weights provided by CPS. The educational attainment within each employment group changed during the recession. For example, the percent of unemployed men with a high school

⁴ The CPS and ATUS include a separate category for respondents who are employed, but have not worked in the past week. This includes, for example, respondents who cannot work because of a labor dispute. For the analyses of time use, it would not be appropriate to include these men among the employed, because they spend their time in a manner similar to the unemployed. Employed men who worked in the past week and who have a child in their household spent 435 minutes on the average weekday at work and related activities. Those who were employed, but did not work in the past week, spent only 44 minutes on work and related activities. The unemployed spent 33 minutes at work and related activities. Thus, for this analysis, I include respondents who are employed but not working with the unemployed.

degree increased by 2.2 percentage points during the recession. Likewise, the percent of men out of the labor force with a high school degree or more increased by 2.1 percentage points during the recession. That is, men who were unemployed or out of the labor force were more highly educated during the recession than before.

Table 13: Educational attainment for males who are 15 and over by employment status before and during the recession.

		Before	During	Significance
Unemployed men	Less than high school	23.2%	22.0%	***
	High school	33.0%	35.2%	***
	Associates degree	23.7%	24.5%	***
	College plus	20.1%	18.3%	***
Men not in labor force	Less than high school	37.3%	35.2%	***
	High school	27.6%	27.9%	*
	Associates degree	19.8%	20.9%	***
	College plus	15.3%	16.0%	***

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

Source: Current Population Survey

The significance column denotes the statistical significance of the difference between “Before” and “During” with robust standard errors. Figures use complex weights provided by CPS
 N=260,359 (unemployed) N=1,171,744 (out of the labor force)

Because the composition of each employment group changed during the recession with regard to observed characteristics, it likely also changed with regard to unobserved characteristics that influence time spent on childcare. To address the composition of each employment group directly, I link each ATUS respondent to a lagged measure of employment status from his last CPS survey. This link reveals the role played by the changing composition of each employment group and its effect on the “within” increase in childcare. For example, men who are not in the labor force in the

ATUS survey performed 27.2 minutes of primary childcare during the recession compared to only 24.4 prior to the recession: a “within” increase. During the recession, those who were either unemployed or employed in the CPS survey and left the labor force by the time of the ATUS survey spent 67.7 minutes on childcare, compared to only 22.0 minutes by those who were out of the labor market during both the CPS and the ATUS. Thus the “within” increase from 24.4 to 27.2 was driven by new entrants.

Table 13 shows that men who were out of the labor force during the recession were more educated than those who were out of the labor force prior to the recession. This fact suggests that different education levels and other demographic characteristics may explain a portion of the 45.7 minute difference in childcare between the men recently out of the labor and those who remained out of the labor force.

Using a decomposition technique developed by Gelbach (2009), Table 14 illustrates that demographic characteristics (education, race, marital status, age, number of children in the household, and age of youngest child) account for 29.7 minutes of the 45.7-minute difference between men who recently left the labor force and those men who remained out of the labor force. Of those characteristics, education, marital status, and age each explain roughly 20% of the total difference. The recent entrants were more highly educated, more likely to be married, and slightly older, all attributes associated with spending more time on primary childcare.

Table 14: Decomposition of the total difference in the minutes of primary childcare during the recession between men with children in their households who were either unemployed or employed in the CPS survey and left the labor force by the time of the ATUS survey compared to those who were out of the labor force in both surveys.

	Minutes	Percent of total difference
Total difference	45.7	100%
Total unexplained difference	16.0	35%
Total explained difference	29.7	65%
<i>Education</i>	8.9	19%
<i>Race and ethnicity</i>	-.5	-1%
<i>Marital status</i>	9.9	22%
<i>Number of children</i>	.02	0%
<i>Age</i>	8.0	18%
<i>Age of youngest child in household</i>	3.3	7%

Note: Figures use complex weights provided by ATUS-X

The pattern shown in Table 14 highlights the importance of the composition of each employment group. The increase in the amount of time men who are out of the labor force spent on childcare is driven by the addition of educated, married, older men who recently left the labor force. This increase is not a “between” change, wherein employed men who drop out of the labor force begin performing childcare at the same level as other men who are out of the labor force. Nor is this a “within” change in behavior, wherein men who have always been out of the labor force change the amount of time they spend on childcare. Instead, the increase in the average time spent on childcare results from a change in the composition of men who are out of the labor force.

4.5.2 Behavioral, compositional, and between group change

By using the lagged employment status, the total change in average minutes of childcare can be separated into the following components:

$$\underbrace{\sum_{i=N,E,U} [P_i^D [P_{N,i}^D (M_{N,i}^D - M_i^B) + P_{E,i}^D (M_{E,i}^D - M_i^B) + P_{U,i}^D (M_{U,i}^D - M_i^B)]]}_{\text{Within group change}} + \underbrace{(M_i^B - M^B)(P_i^D - P_i^B)}_{\text{Between group change}}$$

Where P_i^B is the proportion of men who are in employment status i before the recession and $P_{j,i}^D$ is the proportion of men during the recession who were in employment status j in the lagged CPS survey conditional on being in employment status i in the current ATUS survey (that is, $P_{N,i}^D + P_{E,i}^D + P_{U,i}^D = 1$). M_i^B is the average time spent on childcare by men before the recession in employment status i and $M_{j,i}^D$ is the average time spent on childcare by men during the recession who were in employment status j in the lagged CPS survey and are in employment status i in the current ATUS survey. The derivation of the decomposition equation is shown in detail in Appendix 7.

The first three elements of the decomposition equation together form the “within” change commonly examined. For example, when $i=N$, the sum of these three elements compares the time men out of the labor force spent on childcare prior to the recession and during the recession. However, this sum can be further decomposed. The first element compares men who were out of the labor force prior to the recession to those who were out of the labor force during the recession in the ATUS and also two to five months prior to the ATUS. The second and third elements compare men who were

out of the labor force prior to the recession to those who were out of the labor force during the recession in the ATUS but were employed or unemployed two to five months prior to the ATUS. That is, the first element is the best estimate of a behavioral change while the second and third elements are compositional. The final element is the “between” change due to the proportion of men out of the labor force; that is, because men who are out of the labor force spend a different amount of time than the average man on childcare, having more men out of the labor force will change the average time spent on childcare.

Table 15 shows the results of applying the decomposition equation. The vast majority (73.9%) of the total change in average time on childcare was a behavioral change by employed men. The large behavioral change by employed men suggests that men are responding to the broader economic environment and not just their own employment status.

Table 15: Decomposition of the total change in time spent on primary childcare by men with children in their households.

		Change	% change
Employed	Behavioral	5.13 minutes	73.9%
	Compositional	-0.51	-7.3%
	<i>Not in labor force → Employed</i>	-.99	-14.3%
	<i>Unemployed → Employed</i>	0.48	7.0%
	Between	-0.05	-0.7%
Unemployed	Behavioral	1.54	22.2%
	Compositional	0.34	4.9%
	<i>Employed → Unemployed</i>	1.59	22.9%
	<i>Not in labor force → Unemployed</i>	-1.25	-18.0%
	Between	0.11	1.6%
Not in labor force	Behavioral	-0.30	-4.3%
	Compositional	0.69	9.9%
	<i>Employed → Not in labor force</i>	0.15	2.2%
	<i>Unemployed → Not in labor force</i>	0.54	7.7%
	Between	-0.02	-0.3%
	Total	6.9	100.0%

Note: Figures use complex weights provided by ATUS-X

A behavioral change among the unemployed was the second largest component of the total increase in average time on childcare. That is, not only are there more unemployed men during the recession (the “between” change, which was small and positive), but unemployed men also began spending more time on childcare. Compositional effects were also large among the unemployed. The influx of the formerly employed men increased the time unemployed men spent on childcare by the same amount as the behavioral change among the unemployed; however, this increase was offset by the entrance of men who were previously out of the labor force.

Among those men not in the labor force, the change was entirely compositional and not behavioral or between employment group. Without the more detailed

decomposition, it would appear that this was a behavioral change among those out of the labor force, but the increase was in fact due to new entrants to the group.

These conclusions highlight two further results. First, the result that the largest component of the total change in average time on childcare was a behavioral change among employed men suggests that men responded to the broader economic environment and not just their own employment statuses when it comes to providing childcare. Indeed, between group change accounted for very little of the total increase in childcare. This is consistent with Aguiar et al (2013) who found that only 5% of forgone work hours during the 2007-2009 recession were spent on childcare. In contrast, the same decomposition performed on time at work (results available upon request) revealed that 68.6% of the total decrease in time at work was due to men becoming unemployed. Second, consideration of the composition of each employment group is essential when performing a decomposition. The composition of small groups will be strongly impacted by a shock – like the 2007-2009 recession. In this case, the changing compositions of unemployed men and men who are out of the labor force play an important role in the change of time spent on childcare within these groups.

Table 16 shows the same decomposition as Table 15 for the increase in physical care for children. The pattern is very similar: the largest proportion of the increase in average time on physical care comes from a behavioral change among employed men. The second-largest proportion of the increase comes from a behavioral change among

the unemployed. Again, both the unemployed and those out of the labor force are influenced strongly by changes in their composition. Among those out of the labor force, the increase in time on physical care is due to the changing composition. Likewise, among the unemployed, the entrance of formerly employed men drives the time on physical care up, but this is countered by the influx of men who were out of the labor force and became unemployed.

Table 16: Decomposition of the total change in time spent on physical care by men with children in their households.

		Change	% change
Employed	Behavioral	1.48 minutes	41.6%
	Compositional	.13	3.6%
	<i>Not in labor force → Employed</i>	-.08	-2.2%
	<i>Unemployed → Employed</i>	0.21	5.8%
	Between	-0.04	-1.0%
Unemployed	Behavioral	1.32	37.3%
	Compositional	0.28	7.8%
	<i>Employed → Unemployed</i>	.75	21.2%
	<i>Not in labor force → Unemployed</i>	-.47	-13.3%
	Between	-0.02	-0.5%
Not in labor force	Behavioral	-0.12	-3.3%
	Compositional	0.53	14.8%
	<i>Employed → Not in labor force</i>	0.10	2.8%
	<i>Unemployed → Not in labor force</i>	0.43	12.0%
	Between	-0.01	-0.4%
	Total	3.6 minutes	100.0%

Note: Figures use complex weights provided by ATUS-X

Due to data limitations, this decomposition approach has a weakness: while I can assess the compositional effects of new entrants to each employment group, I cannot determine the effects of departures. That is, the ATUS always occurs after the final CPS interview, so I do not know the employment status of the ATUS respondents after their

interview. Among the employed, the largest compositional change is not new entrants to the employed, but rather employed men who become unemployed or leave the labor force during the recession. Unfortunately, I cannot identify those who will become unemployed or leave the labor force after their ATUS survey. Thus, in Tables 15 and 16, I estimate the behavioral change among the employed by comparing the time employed men (including those who will not be employed during the recession) spent on childcare prior to the recession to employed men during the recession who were also employed two to five months prior to their ATUS interview.

To assess the potential impact of changes in composition on the estimate of behavioral change among employed men, I control for demographic characteristics when comparing the time spent on childcare pre-recession by employed men and during the recession by employed men who were employed two to five months prior to their ATUS interview. After including race, education, marital status, age, age squared, age of youngest child, number of household children, and a year trend employed men still increased time spent on childcare during the recession by 6.4 minutes (compared to a raw increase of 7.8 minutes). That is, the demographic variables explain 1.4 minutes of the increase. Applying a Gelbach decomposition reveals that age of the youngest child is the largest component of the 1.4 explained minutes followed by education and marital status. Employed men during the recession had younger children, were more highly educated, and were more likely to be married – all components associated with more

time on childcare. Thus, the estimate of behavioral change among employed men in Table 15 contains some compositional change. However, this is the best estimate of behavioral change without longitudinal data or information about the ATUS respondents' employment status after their ATUS interview.

The cross-sectional nature of the ATUS dataset is a weakness, so I examined if a panel dataset with a coarser measure of time use, but with longitudinal data would yield similar results. Appendix 8 contains an analysis of panel data which finds that employed men reported more hours of housework during the 1990-1991 recession.

4.5.3 Time spent on other activities

As men's employment opportunities decrease, they may have spent more time on other activities. Table 17 shows that of the largest uses of time (childcare, sleep, socializing/leisure, work and related activities, and housework), only average time on childcare has a statistically significant increase and work a statistically significant decrease. When the different time use categories are estimated with seemingly unrelated regression to allow for testing of coefficients across equations, the coefficient on the during indicator for time on childcare is statistically significantly different from that for time on household activities ($p < .001$). This suggests, similar to the findings of Kimmel and Connely (2007), that childcare behaves differently from housework.

Table 17: Change in time spent on five major activities during and after the recession by men with children in their household. “Other activities” is the difference between the total time in a day and those five major activities.

	Primary childcare	Work & related activities	Sleep	Socializing, relaxing, & leisure	Household activities	Other activities
During	6.94** (2.41)	-12.4+ (7.01)	3.97 (3.86)	2.76 (4.85)	-4.02 (2.52)	2.75 (6.60)
After	3.64+ (1.98)	-32.5*** (6.60)	4.16 (3.69)	6.83 (4.76)	3.13 (2.37)	14.72** (5.64)
Observations	12,224					
Population	Men with children in their household					
<i>Robust standard errors in parentheses, clustered by state</i>						
<i>*** p<0.001, ** p<0.01, * p<0.05, + p<0.1, regressions use complex weights provided by ATUS-X</i>						

To check whether the increase in average time on childcare among the employed was a result of a coarse measurement of employment, I examine the average time employed men spent on these five activities. If, for example, men employed prior to the recession spent more time at work than those employed during, that finding would suggest that the measure of employment used was too coarse to pick up a more subtle change that occurred. However, Table 18 shows that average time at work did not have a statistically significant change during the recession among the employed. This result suggests that the increase in time on childcare among employed men is not attributable to poor measurement of employment status.

Table 18: Change in time spent on five major activities during and after the recession by employed men with children in their household.

	Primary childcare	Work & related activities	Sleep	Socializing, relaxing, & leisure	Household activities	Other activities
During	6.22** (2.33)	-2.42 (6.17)	7.37** (2.83)	-4.50 (4.45)	-1.65 (2.28)	-5.03 (4.74)
After	2.70 (2.05)	-5.54 (6.82)	3.77 (4.46)	-5.96 (5.67)	-0.13 (2.65)	5.15 (5.02)
Observations	9,603					
Population	Employed men with children in their household					
<i>Robust standard errors in parentheses, clustered by state</i>						
<i>*** p<0.001, ** p<0.01, * p<0.05, + p<0.1, regressions use complex weights provided by ATUS-X</i>						
<i>"Other activities" is the difference between the total time in a day and those five major activities</i>						

Additionally, employed men's type of employment may have changed during the recession, even though employed men spent the same average amount of time on work and work-related activities. For example, an employed man may face reduced hours at his main job and therefore supplement with a second job, with informal income generating activities, or by searching for an additional job. This could mean that the increase in average time on childcare among employed men was due to a change in the type of employment. To examine this possibility, I split the total minutes spent on work and work-related into four subparts: main job, secondary job, job search activities, and other income generating activities (for example, shoveling a neighbor's driveway for pay).

Table 19: Change in time spent on four subparts of work and work-related activities during and after the recession by employed men with children in their household.

	Minutes at main job	Minutes at second job	Minutes on job search activities	Minutes of other income generating activities
During	-1.339 (6.138)	-1.251 (1.666)	0.675 (0.551)	-0.503 (0.535)
After	-3.944 (6.415)	-0.862 (1.426)	0.347 (0.316)	-1.081* (0.457)
Observations	9,603			
Population	Employed men with children in their household			
<i>Robust standard errors in parentheses, clustered by state</i>				
<i>*** p<0.001, ** p<0.01, * p<0.05, + p<0.1, regressions use complex weights provided by ATUS-X</i>				

These four activities sum to the total time spent on work and work-related activities in Table 18. As Table 19 shows, there was no statistically significant change in the average time spent on the four types of work during the recession. This suggests that employed men did not have a dramatic change in these four types of work during the recession.

4.6 Conclusion from examining changes in time spent on childcare

This chapter examined the circumstances under which men spend more time on childcare. Men's employment opportunities worsened relative to women's during the 2007-2009 recession, and the average time men spent on childcare increased. The increase was relatively large: over 10% of the difference in the time men and women in households with children spent on childcare prior to the recession.

A large contribution of this study is the detailed decomposition of the total increase in the average time spent on childcare into behavioral, compositional, and between group change. The vast majority of the increase in average time spent on childcare was a behavioral change among employed men. This finding suggests that employed men may have responded to the broader economic environment and not their own individual employment status. That is, these employed men's anticipated family income or opportunity cost may have decreased even if their current employment was stable. Likewise, employed men may have lost bargaining power within the family because their employment became more tenuous.

Indeed, the increase in average time spent on primary childcare by men principally consisted of time spent on physical care, an element of childcare men historically have performed to a far less extent than women (Berik and Kongar 2011; Craig 2006). Moreover, physical care for children is the type of childcare that parents report enjoying the least (Robinson 1993). The fact that men spent more time on average on physical care for children suggests that even the childcare activity with the most entrenched gendered division responds to economic pressure.

In contrast, the change in average time spent on childcare among men who were out of the labor force was entirely driven by the composition of the group. That is, what appeared to be a behavioral change among men out of the labor force was in fact due to the changing composition of the group. New entrants to this group spent more time on

childcare, raising the group's average; men who had been out of the labor force for longer did not increase the time they spent on childcare during the recession. The new entrants were more highly educated, more likely to be married, and slightly older: all characteristics associated with spending more time on primary childcare. The decomposition technique used here underscores the importance of accounting for new entrants to small groups who may change the composition of the group.

The increase in the average time men spent on childcare suggests that even in light of historical patterns of behavior, a change in employment opportunities is associated with a change in how much time men spend on childcare. Economic theory highlights numerous reasons why men would have spent more time on childcare during the recession: On average, the recession decreased men's opportunity cost for spending time on childcare, their family income, and their bargaining power relative to women.

Taken together, these results highlight an important pattern. We know that in the long run fathers have slowly increased the time they spend on childcare (Bianchi, Robinson, and Milkie 2006; Fox et al 2012; Sayer, Bianchi, and Robinson 2004; Sandberg and Hofferth 2001; Bianchi 2011). However, the results in this paper suggest that men also respond to economic incentives in the short run. Not only did men's average time on childcare increase during the recession, but the type of childcare men performed was physical care, the aspect of childcare men historically have done the least. Moreover, the increase in average time on childcare was driven by a behavioral change among

employed men, not by men changing employment statuses who became unemployed or left the labor force. Thus, we saw an increase in the average time men spent on the type of childcare with the most uneven gender division and this increase was associated with a behavioral change among employed men rather than men becoming unemployed. This finding provides strong evidence that short term economic pressure overwhelms patterned behavior regarding childcare.

Appendix 1: Resume example for laboratory experiment

John Long
[Redacted]
Greenville, NC
[Redacted]
JohnLong@[Redacted]

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

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

Objective: Confident and enterprising recent college graduate pursuing a career as a biologist

Experience

Research Technician - Miller Lab, UNC School of Medicine, Chapel Hill, NC
May 2012 to September 2012

- Conducting research under a post doctorate fellow on Klebsiella Pneumoniae

Customer Specialist - Best Buy, Raleigh, NC
April 2011- April 2012

- Provided excellent customer service to people of all backgrounds
- Managed transactions accurately and ethically
- Met sales goals in a fast -paced environment

Hollister & Aeropostale Sales Associates/Customer Service
2006 - 2010 (College breaks)

- Demonstrated a high level of selling and customer service skills
- Achieved sales goals and used company tools to develop strong selling skills and reinforce the brand vision

Related Activities

LGBTQ Alliance, Initiatives Chair, East Carolina University - Greenville, NC
April 2012 to December 2012

- 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

Education

B.S. in Biology, 2012
East Carolina University - Greenville, NC

The "Related Experience" field is used to signal an LGBT affiliation. If this were a non-LGBT resume, the student group name would be a similar non-LGBT group.

Appendix 2: Difference in hireability measures for resumes with LGBT activity

The difference in hireability measures between resumes with and without an LGBT activity within each sex by adjective combination:

	Recommend	Recommend	Recommend	Recommend
LGBT activity	-1.172 (1.142)	-2.252* (1.219)	-3.444*** (1.215)	-2.446** (1.202)
Obs	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 (0.922)	-1.747* (0.924)	-2.749*** (0.875)	-3.181*** (0.928)
Obs	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 (0.812)	-0.960 (0.814)	-2.209** (0.861)	-1.808** (0.836)
Obs	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 (0.996)	-1.578 (1.034)	-3.227*** (0.987)	-2.013** (1.011)
Obs	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

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix 3: Resume example for audit study

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

Julia Peterson

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

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

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

Work Experience

Summer Intern
Eismann & Associates, Durham, NC

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

- 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

Customer Service Representative
Dicks Sporting Goods, Durham, NC

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

- 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 sale
- Assisted with buying over the counter and communicating with retail merchandise
- 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

Related Experience

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

2012-2014

- 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

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

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

Appendix 4: 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	<i>Office</i>	<i>Retail</i>	<i>Food service</i>	<i>Labor or skilled trade</i>	<i>Total</i>
<i>Female, no LGBT activity, and feminine adjective</i>	5.02	2.40	2.85	2.17	12.44
<i>Female, no LGBT activity, and masculine adjective</i>	5.02	2.40	3.00	2.62	13.04
<i>Male, no LGBT activity, and feminine adjective</i>	4.57	3.07	3.07	1.72	12.44
<i>Male, no LGBT activity, and masculine adjective</i>	4.72	3.00	2.62	2.55	12.89
<i>Male, LGBT activity, and feminine adjective</i>	6.22	2.62	2.92	2.02	13.79
<i>Male, LGBT activity, and masculine adjective</i>	4.5	1.87	1.80	2.25	10.42
<i>Female, LGBT activity, and feminine adjective</i>	4.2	3.82	2.10	2.47	12.59
<i>Female, LGBT activity, and masculine adjective</i>	4.72	2.70	3.07	1.87	12.37
<i>Total</i>	38.98	21.89	21.44	17.69	100

Pearson chi2(21)= 24.3168 Pr = 0.278

Manipulation	<i>City 1*</i>	<i>City 2</i>	<i>City 3</i>	<i>City 4</i>	<i>City 5</i>	<i>Total</i>
<i>Female, no LGBT activity, and feminine adjective</i>	6.37	1.8	1.65	0.6	2.02	12.44
<i>Female, no LGBT activity, and masculine adjective</i>	6.52	1.72	1.57	0.67	2.55	13.04
<i>Male, no LGBT activity, and feminine adjective</i>	6	1.65	1.87	0.67	2.25	12.44
<i>Male, no LGBT activity, and masculine adjective</i>	6.45	1.87	1.12	0.45	3	12.89
<i>Male, LGBT activity, and feminine adjective</i>	6.52	2.02	1.95	0.45	2.85	13.79
<i>Male, LGBT activity, and masculine adjective</i>	5.1	1.42	1.35	0.67	1.87	10.42
<i>Female, LGBT activity, and feminine adjective</i>	6.3	2.77	0.67	0.37	2.47	12.59
<i>Female, LGBT activity, and masculine adjective</i>	6.22	2.1	1.35	0.6	2.1	12.37
<i>Total</i>	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	<i>Feminine adjectives in ad</i>	<i>Masculine adjectives in ad</i>	<i>No gendered language in ad</i>	<i>Total</i>
<i>Female, no LGBT activity, and feminine adjective</i>	3.75	2.02	6.67	12.44
<i>Female, no LGBT activity, and masculine adjective</i>	3.9	1.87	7.27	13.04
<i>Male, no LGBT activity, and feminine adjective</i>	4.2	2.32	5.92	12.44
<i>Male, no LGBT activity, and masculine adjective</i>	3.9	2.47	6.52	12.89
<i>Male, LGBT activity, and feminine adjective</i>	4.05	2.62	7.12	13.79
<i>Male, LGBT activity, and masculine adjective</i>	2.55	2.17	5.7	10.42
<i>Female, LGBT activity, and feminine adjective</i>	4.05	2.1	6.45	12.59
<i>Female, LGBT activity, and masculine adjective</i>	3.15	2.4	6.82	12.37
<i>Total</i>	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.

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

Appendix 5: Balance of laboratory experiment

The following table shows the distribution of base resumes (identical work history, education, font, and style) by the experimental manipulations. The value in each cell shows the cell proportion.

Manipulation	<i>Base</i>				
	<i>Resume 1</i>	2	3	4	5
<i>Female, no LGBT activity, and feminine adjective</i>	1.3	1.22	1.27	1.24	1.21
<i>Female, no LGBT activity, and masculine adjective</i>	1.25	1.25	1.3	1.22	1.27
<i>Male, no LGBT activity, and feminine adjective</i>	1.22	1.27	1.24	1.21	1.21
<i>Male, no LGBT activity, and masculine adjective</i>	1.25	1.3	1.22	1.27	1.24
<i>Male, LGBT activity, and feminine adjective</i>	1.21	1.28	1.27	1.25	1.25
<i>Male, LGBT activity, and masculine adjective</i>	1.24	1.21	1.21	1.28	1.27
<i>Female, LGBT activity, and feminine adjective</i>	1.21	1.21	1.28	1.27	1.25
<i>Female, LGBT activity, and masculine adjective</i>	1.27	1.24	1.21	1.21	1.28
<i>Total</i>	9.95	9.98	9.99	9.95	9.98

	6	7	8	9	10	<i>Total</i>
<i>Female, no LGBT activity, and feminine adjective</i>	1.21	1.28	1.27	1.25	1.25	12.5
<i>Female, no LGBT activity, and masculine adjective</i>	1.24	1.21	1.21	1.28	1.27	12.5
<i>Male, no LGBT activity, and feminine adjective</i>	1.28	1.27	1.25	1.25	1.3	12.5
<i>Male, no LGBT activity, and masculine adjective</i>	1.21	1.21	1.28	1.27	1.25	12.5
<i>Male, LGBT activity, and feminine adjective</i>	1.3	1.22	1.27	1.24	1.21	12.5
<i>Male, LGBT activity, and masculine adjective</i>	1.25	1.25	1.3	1.22	1.27	12.5
<i>Female, LGBT activity, and feminine adjective</i>	1.25	1.3	1.22	1.27	1.24	12.5
<i>Female, LGBT activity, and masculine adjective</i>	1.27	1.25	1.25	1.3	1.22	12.5
<i>Total</i>	10.01	9.99	10.05	10.08	10.01	100

Pearson chi2(63) = 3.3240 Pr = 1.000

Appendix 6: Regression results for time on childcare

Figure 21 showed that men who lived in a household with children spent 6.9 more minutes on average on childcare during the recession. To see whether men's raw increase persists after accounting for other characteristics, I employ the following regression:

$$y_i = \mathbf{X}_i\boldsymbol{\beta} + year_i\xi + I(DURING_i)\delta + I(AFTER_i)\gamma + \varepsilon_i$$

In this framework, the amount of time a person spends on childcare (y_i) is a function of their demographic characteristics (\mathbf{X}), a linear year trend, and an indicator variable for if the time diary occurred during the recession. The coefficient on the indicator variable for being in the recession, $\hat{\delta}$, will show if there was a change in the amount of childcare provided by men as a whole during the recession. The demographic variables include education, race, ethnicity, marital status, number of children in the household, age of the youngest child, age, and age squared. All regressions use the complex sampling weights provided by the ATUS-X. All regressions use robust standard errors and the observations are clustered by state.

Table 20: Selected coefficients from regression of men’s time on childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables

	Total minutes of primary childcare	Minutes of primary childcare when spouse or partner is not present	Total minutes of primary childcare	Minutes of primary childcare when spouse or partner is not present
During recession	8.204** (2.670)	7.037* (2.837)	16.18*** (3.981)	16.84** (6.078)
After recession	6.159 (3.847)	6.663* (3.068)	14.49* (6.864)	16.75** (5.770)
Observations	12,224	12,224	6,234	5,251
R-squared	0.136	0.076	0.085	0.046
Population	Men with children in household		Men with children in household with positive values on outcome variable	

Robust standard errors in parentheses, observations clustered by state
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$
Control variables include education, race, ethnicity, marital status, number of children in household, age of youngest child, linear year trend, age, and age squared. Regressions use complex weights from ATUS-X.

The above table demonstrates that the average time men in households with children spent on childcare increased during the recession. This finding is consistent across specifications; the increase remains regardless of whether childcare is measured in total or restricted to only when a man’s spouse or partner is not present. Likewise, the increase is consistent whether measured for all men in households with children or just those who spent a positive amount of time on childcare.

In contrast, the following table shows that the average time women in households with children spent on childcare did not change during the recession.

Table 21: Selected coefficients from regression of women’s time on childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables

	Total minutes of primary childcare	Minutes of primary childcare when spouse or partner is not present	Total minutes of primary childcare	Minutes of primary childcare when spouse or partner is not present
During recession	0.365 (2.880)	1.093 (3.121)	-3.097 (4.387)	-2.669 (4.520)
After recession	0.494 (3.806)	2.152 (3.836)	-0.400 (4.912)	0.134 (4.670)
Observations	16,916	16,916	12,715	12,371
R-squared	0.260	0.217	0.175	0.127
Population	Women with children in household		Women with children in household with positive values on outcome variable	

Robust standard errors in parentheses, observations clustered by state

**** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$*

Control variables include education, race, ethnicity, marital status, number of children in household, age of youngest child, linear year trend, age, and age squared. Regressions use complex weights from ATUS-X.

To check whether the result shown in Table 20 is solely due to a change in the composition of men living in households with children, I ran the same regressions on all men. The results are consistent with those found above.

Table 22: Selected coefficients from regression of men’s time on childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables

	Total minutes of primary childcare	Minutes of primary childcare when spouse or partner is not present	Total minutes of primary childcare	Minutes of primary childcare when spouse or partner is not present
During recession	3.676** (1.329)	2.890+ (1.535)	15.42*** (4.073)	15.16* (6.779)
After recession	3.261+ (1.735)	2.588 (1.644)	16.06** (5.804)	14.70** (5.441)
Observations	27,040	27,040	6,829	5,745
R-squared	0.130	0.107	0.021	0.021
Population	Men	Men	Men with positive values on outcome variables	

Robust standard errors in parentheses, observations clustered by state

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Control variables include education, race, ethnicity, marital status, number of children in household, indicator for no children, linear year trend, age, and age squared. Regressions use complex weights from ATUS-X.

As shown in Table 23, the proportion of men engaging in any childcare did not change during the recession.

Table 23: Selected coefficients from regression of men’s participation in childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables

	Any primary care	Any secondary care	Any primary care	Any secondary care
	<i>Probit</i>		<i>Logit</i>	
During	0.0409 (0.0526)	-0.0335 (0.0813)	0.0622 (0.0900)	-0.0397 (0.143)
After	0.0109 (0.0516)	-0.0954 (0.0335)	0.00903 (0.0878)	-0.132 (0.147)
Observations	12,220	9,076	12,220	9,076
Population	Men with children in household	Men with children under 13 in household	Men with children in household	Men with children under 13 in household

Robust standard errors in parentheses, observations clustered by state

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Controls include education, race, ethnicity, marital status, number of children in the household, age of the youngest child, linear year trend, age, and age squared. Regressions use complex weights provided by ATUS-X.

Secondary childcare is only measured for children under 13.

Table 24 shows that the average time spent on physical care by men with children in their household increased during the recession.

Table 24: Selected coefficients from regression of men’s time on childcare on an indicator variable for during the recession, and indicator variable for after the recession, and control variables

	Engaged play	Physical care	Looking after	Education Related
During	3.197 (2.363)	3.822** (1.353)	0.0468 (0.712)	0.137 (0.878)
After	3.132 (2.929)	1.277 (1.864)	1.009 (1.064)	-0.0348 (0.899)
Observations	12,224	12,224	12,224	12,224
R-squared	0.055	0.107	0.008	0.028
Population	Men with children in household			

Robust standard errors in parentheses, observations clustered by state

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$

Control variables include education, race, ethnicity, marital status, number of children in household, age of youngest child, linear year trend, age, and age squared. Regressions use complex weights from ATUS-X.

Appendix 7: Decomposition method for change in time on childcare

To decompose the total change in average time spent on childcare, I start with the basic difference between the average minutes of childcare during the recession (M^D) and before the recession (M^B).

$$\Delta = M^D - M^B$$

Each average can be re-written as the weighted average of the childcare performed by men in each employment status. P_i^t is the proportion of men who are in employment status i in period t ; M_i^t refers to the average minutes of childcare performed by men in employment status i in period t .

$$\begin{aligned} \Delta &= P_E^D M_E^D + P_U^D M_U^D + P_N^D M_N^D - [P_E^B M_E^B + P_U^B M_U^B + P_N^B M_N^B] \\ \Delta &= [P_E^D M_E^D - P_E^B M_E^B] + [P_U^D M_U^D - P_U^B M_U^B] + [P_N^D M_N^D - P_N^B M_N^B] \end{aligned}$$

Now, add and subtract $P_i^D M_i^B$ for each employment status.

$$\begin{aligned} \Delta &= P_E^D [M_E^D - M_E^B] + M_E^B [P_E^D - P_E^B] + P_U^D [M_U^D - M_U^B] + M_U^B [P_U^D - P_U^B] + P_N^D [M_N^D \\ &\quad - M_N^B] + M_N^B [P_N^D - P_N^B] \end{aligned}$$

The minutes spent on childcare during the recession by each employment status (M_i^D) can be separated out by a lagged employment measure. That is, $M_i^D = P_{E,i}^D M_{E,i}^D + P_{U,i}^D M_{U,i}^D + P_{N,i}^D M_{N,i}^D$. Where $P_{j,i}^D$ is the proportion of men during the recession who were in employment status j in the lagged CPS survey conditional on being in

employment status i in the current ATUS survey (that is, $P_{N,i}^D + P_{E,i}^D + P_{U,i}^D = 1$), and $M_{j,i}^D$ is the average time spent on childcare by men during the recession who were in employment status j in the lagged CPS survey and are in employment status i in the current ATUS survey.

$$\text{So, } M_i^D - M_i^B = P_{E,i}^D(M_{E,i}^D - M_i^B) + P_{U,i}^D(M_{U,i}^D - M_i^B) + P_{N,i}^D(M_{N,i}^D - M_i^B)$$

Substituting this in for each $M_i^D - M_i^B$ gives

$$\Delta = \sum_{i=N,E,U} [P_i^D [P_{N,i}^D(M_{N,i}^D - M_i^B) + P_{E,i}^D(M_{E,i}^D - M_i^B) + P_{U,i}^D(M_{U,i}^D - M_i^B)] + (M_i^B)(P_i^D - P_i^B)]$$

Finally, because $(P_E^D - P_E^B) + (P_U^D - P_U^B) + (P_N^D - P_N^B) = 0$, add $-M^B[(P_E^D - P_E^B) + (P_U^D - P_U^B) + (P_N^D - P_N^B)]$. This gives Eq. 1:

$$\Delta = \sum_{i=N,E,U} [P_i^D [P_{N,i}^D(M_{N,i}^D - M_i^B) + P_{E,i}^D(M_{E,i}^D - M_i^B) + P_{U,i}^D(M_{U,i}^D - M_i^B)] + (M_i^B) - M^B)(P_i^D - P_i^B)]$$

Appendix 8: Analysis of panel data

The ATUS contains rich data on time use, but it does not track the same person over time. The Panel Study of Income Dynamics (PSID) contains only very coarse measures of time use, but has multiple observations on the same respondent over time. I use the PSID to show that the time male heads of household report spending on housework increased during the 1990-1991 recession⁵. This echoes Casper and O'Connell (1998), who analyzed cross-sectional data and found an increase fathers who were care providers for preschool children during the 1990-1991 recession. The consistency between cross-sectional and panel data during the 1990-1991 recession suggests that the analyses of the ATUS for the 2007-2009 recession in the main text are credible.

The PSID began in 1968 with a nationally representative sample of 5,000 families; these individuals and their descendants have been followed over time. The PSID asks respondents how much time they spend on housework in the average week. Stylized questions about time use tends to be biased upwards (Hofferth 2000) and inconsistent between a self-report and a proxy-report: husbands report spending more time on housework than their wives say they do (Achen and Stafford 2005). A respondent fixed effect accounts for time-invariant reporting bias.

⁵ Beginning in 1999, the PSID became biennial. This makes it difficult to analyze the 2007-2009 recession, because the bulk of the recession occurred in 2008, a year with no PSID survey. However, both 1990 and 1991 are covered by the PSID.

To examine if male household heads increased their time on housework during the 1990-1991 recession, I employ the following regression: $y_{i,t} = \alpha + year_{i,t}\xi + I(DURING_{i,t})\delta + \sum_{j=1}^{n-1} I(i = j)\beta_j + \varepsilon_{i,t}$ In this framework, $y_{i,t}$ is the amount of time a person i spends on housework per week in year t . The time the respondent spent on housework is a function of a linear year trend, a respondent fixed effect, and an indicator variable for if the survey occurred during the recession (measured as 1990 or 1991, to conform to Casper and O’Connell). The coefficient on the indicator variable for being in the recession, $\hat{\delta}$, will show if there was a change in the amount of housework reported by men during the recession relative to their own average. In the following table a “male household head” is a man who was the head of household for all ten survey years (1986 to 1995).

Table 25: Selected coefficients from regression of men’s self-reported time on housework per week on an indicator variable for 1990 or 1991 sample years, a time trend, and respondent fixed effects.

Self-reported hours of housework in an average week (PSID from 1986 to 1995)			
During 1990 or 1991	0.321** (0.107)	0.324** (0.112)	0.306** (0.113)
Year trend	0.0320 (0.0201)	-0.00548 (0.0196)	-0.00498 (0.0198)
Mean	7.77	6.85	6.87
Obs (person x year)	33,436	18,087	17,809
Population	Male household heads	Male household heads with continuous employment	Male household heads with continuous employment and no missing housework values
Respondent FE	Yes	Yes	Yes
<i>Robust standard errors in parentheses, clustered by respondent</i>			
<i>*** p<0.001, ** p<0.01, * p<0.05, + p<0.1</i>			
Respondents are in the core family sample; results use yearly complex weights for core sample provided by PSID			

As shown in Table 25, male household heads reported spending more time on housework during the 1990-1991 recession relative to their own average. This pattern is consistent when restricted to men who were employed during all ten surveys. Thus, both cross-sectional data (Casper and O'Connell 1998) and panel data show an increase in male unpaid labor during the 1990-1991 recession. This is supportive, although not conclusive, evidence that the analysis in the main text is credible.

Works Cited

- Abraham, K. G., Flood, S. M., Sobek, M., & Thorn, B. (2011). American Time Use Survey Data Extract System: Version 2.4 [Machine-readable database]. Maryland Population Research Center, University of Maryland, College Park, Maryland, and Minnesota Population Center, University of Minnesota, Minneapolis, Minnesota. <http://www.atusdata.org>
- Achen, A. & Stafford, F. (2005) Data Quality of Housework Hours in the Panel Study of Income Dynamics:
Who Really Does the Dishes? PSID Technical Paper Series
- Aguiar, M., Hurst, E., & Karabarbounis, L. (2013). Time use during the Great Recession. *The American Economic Review*, 103(5), 1664-1696.
- 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,
- Becker, G. (1981). *A Treatise on the Family* Cambridge: Harvard University Press.
- Becker, G. (1985). Human Capital, Effort, and the Sexual Division of Labor *Journal of Labor Economics* 3(1), S33-S58.
- Berg, N. & Lien, D. (2002) “Measuring the effect of sexual orientation on income: Evidence of discrimination?” *Contemporary Economic Policy*, 20(4), 394-414.
- Berik, G., & Kongar, E., (2011) Time Use of Mothers and Fathers in Hard Times and Better Times. *Levy Economics Institute Working Paper*.
- Berik, G., & Kongar, E., (2013) Time Allocation of Married Mothers and Fathers in Hard Times: The 2007–09 US Recession, *Feminist Economics*, 19(3), 208-237
- 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.
- Bianchi, S. (2000). Maternal employment and time with children: Dramatic change or surprising continuity? *Demography*, 37(4), 401-414.
- Bianchi, S. (2011) Family Change and Time Allocation in American Families. *The Annals of the American Academy of Political and Social Science* 638, 21-44.
- Bianchi, S., Robinson J., & Milkie, M. (2006). *Changing Rhythms of American Family Life*. New York: Russell Sage Foundation.
- 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.
- Bloemen, H.G. & Stancanelli E. (2014). Market hours, household work, child care, and wage rates of partners: an empirical analysis. *Review of Economics of the Household*, 12(1), 51-81
- 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.
- Browning, Martin, Francois Bourguignon, Pierre-Andre Chiappori, Valerie Lechene. (1994) Income and Outcomes: A Structural Model of Intrahousehold Allocation. *The Journal of Political Economy*, 102(6), 1067-1096.
- 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.
- Burda, M. & Hamermesh D. (2010) Unemployment, market work and household production. *Economics Letters*, 107, 131-133.
- Bureau of Labor Statistics (2012) "Women in the Labor Force: A Databook" <http://www.bls.gov/cps/wlf-databook-2011.pdf>
- Casper, L., & O'Connell, M. (1998). Work, income, the economy, and married fathers as child-care providers. *Demography*, 35(2), 243-250.
- Census Bureau (2012) "File A" *Genealogy Data: Frequently Occurring Surnames from Census 2000* <http://www.census.gov/genealogy/www/data/2000surnames/>
- Coltrane, S. (2000). Research on household labor: Modeling and measuring the social embeddedness of routine family work. *Journal of Marriage and Family*, 62(4), 1208-1233.
- Correll, S.J., Benard, S., & Paik, I. (2007). "Getting a job? Is there a motherhood penalty?" *American Journal of Sociology*, 112, 1297-1339.

- Craig, L. (2006) Does Father Care Mean Fathers Share? A Comparison of How Mothers and Fathers in Intact Families Spend Time with Children. *Gender and Society*, 20(2), 256-281.
- 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.
- Fox, L., Han, W., Ruhm, C., & Waldfogel, J. (2013) Time for Children: Trends in Employment Patterns of Parents, 1967-2009. *Demography* 50, 25-49
- Gelbach, J. (2009) When Do Covariates Matter? And Which Ones, and How Much? *Social Science Research Network*
- 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.
- Gough, M. & Killewald, A. (2011) Unemployment in families: The case of housework. *Journal of Marriage and Family* 73(5), 1085-1100.
- 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.
- Hochschild, A. with Machung A. (1989) *The Second Shift* New York: Viking
- Hofferth, S. (2000) Family Reading to Young Children: Social Desirability and Cultural Biases in Reporting *Time-Use Measurement and Research* National Research Council, National Academy Press, 79-80.
- Hook, J. L. (2010). Gender Inequality in the Welfare State: Sex Segregation in Housework, 1965–2003. *American Journal of Sociology*, 115(5), 1480-1523.
- 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>
- Kimmel, J., & Connelly, R. (2007). Mothers' Time Choices Caregiving, Leisure, Home Production, and Paid Work. *Journal of Human Resources*, 42(3), 643-681.
- 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
- Lundberg, S. & Pollak, R. (1993). Separate Spheres Bargaining and the Marriage Market. *Journal of Political Economy*, 101(6), 988-1010.
- Lundberg, S. & Pollak, R., (1994) "Noncooperative Bargaining Models of Marriage," *American Economic Review*, 84(2), 132-137.
- McElroy, M. B. & Horney, M. J. (1981). Nash-bargained Household Decisions: Toward a Generalization of the Theory of Demand. *International Economic Review*, 22, 333-349.
- 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*

- Panel Study of Income Dynamics, public use dataset (2014). Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.
- 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.
- Rangel, M. (2006). "Alimony rights and intrahousehold allocation of resources", *Economic Journal*, 116(513), 627-58.
- Robinson, J. P. (1993). As We Like It. *American Demographics*, Feb. 44-48.
- 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.
- Sandberg, J.F. & Hofferth, S.L. (2001) "Changes in Children's Time with Parents: United States, 1981-1997" *Demography*, 38(3) 423-436.
- Sayer, L.C., Bianchi, S. M., & Robinson, J.P. (2004) Are Parents Investing Less in Children? Trends in Mothers' and Fathers' Time with Children *American Journal of Sociology*, 110 1-43
- Slaughter, A. (2012). Why Women Still Can't Have It All *The Atlantic*, July/August.

- 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.
- Starr, M. A. (2014). Gender, added-worker effects, and the 2007–2009 recession: Looking within the household. *Review of Economics of the Household*, 12(2), 1-27.
- 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.
- Thomas, D. (1990). Intra-Household Resource Allocation An Inferential Approach. *Journal of Human Resources*, 25(4), 635-664.
- Tilcsik, A. (2011) "Pride and prejudice: Employment discrimination against openly gay men in the United States" *American Journal of Sociology*, 117(2), 586-626.
- Wall, H. J. (2009) The 'Man-Cession' of 2008-2009: It's Big, but It's Not Great. *Federal Reserve Bank of St. Louis*
<http://www.stlouisfed.org/publications/re/articles/?id=1712>
- Wang, R. & Bianchi S. ATUS Fathers' Involvement in Childcare *Social Indicators Research* 93(1) 2009 141-145

Biography

Marina Mileo Gorsuch was born in Saint Paul, Minnesota on June 8, 1986 to her parents Janice Mileo and Thomas Gorsuch. She attended Earlham College in Richmond, Indiana and graduated with a Bachelor of Arts in Economics in 2008. She then attended Tufts University in Medford, Massachusetts where she received a Master of Science in Economics in 2010. She attended Duke University for her PhD in Public Policy and graduated in 2015.

Marina has published two articles. Her first is titled “Decomposing the Increase in Men’s Time on Childcare during the Great Recession” which was published in *Review of Economics of the Household*. She also published an article co-authored with Seth Sanders and Bei Wu titled “Tooth Loss in Appalachia and the Mississippi Delta Relative to Other Regions in the United States, 1999-2010” in the *American Journal of Public Health*.

Marina has received numerous honors and awards, including the Aleane Webb Dissertation Research Fellowship, a Poster Award from the Population Association of America, the multiple Summer Research Fellowships from Duke University, numerous Conference Travel Grants from the Sanford School of Public Policy at Duke University, and the Behavioral and Social Sciences Travel Award from the Gerontological Society of America. She is also a member of Phi Beta Kappa.