Maternity Risk and the Lesbian Pay Gap: Evidence from the U.S. Decennial Census and American Community Survey

---

A thesis submitted in partial fulfilment of the requirements for the

Degree of Master of Commerce in Economics

in the University of Canterbury

by Hayden Skilling

Department of Economics and Finance

University of Canterbury

2014

Table of Contents

List of Tables

Acknowledgements

This thesis has posed some of the greatest challenges I have faced at the University of Canterbury. Overcoming these difficulties and completing this paper would not have been possible without the assistance of the university's talented academic staff. I would like to express my deepest gratitude to my supervisor, Dr. Eric Crampton, whose patience, guidance and support provided an invaluable contribution to this study. I would also like to thank my associate supervisor, Dr. Seamus Hogan for his valuable input and helpful suggestions. I am grateful for scholarships from the Kelliher Economics Foundation, Reserve Bank of New Zealand and the University of Canterbury, without which the completion of this thesis would not have been possible. Finally, I would like to thank all remaining staff for providing useful feedback, and family and friends for their unwavering support throughout this process.

Abstract

Prior research from the U.S. and abroad reveals a sizable lesbian earnings advantage over otherwise-similar heterosexual women. Using data from the 2000 U.S. Census and 2005-2010 American Community Surveys, we estimate traditional earnings equations and find robust evidence of a lesbian premium, corroborating the findings of previous studies. Using within-sample maternity incidence as an estimate of employers' forward-looking expectations, we then examine whether differences in the perceived likelihood of an employee requiring maternity leave, here-labelled 'maternity risk', contribute to the lesbian pay gap. Results from a direct assessment suggest that maternity risk adversely affects income, and that accounting for near-term differences in maternity risk reduces the lesbian premium by approximately ten to fifteen percent. Further analyses, using proxy variables for differential maternity risk, yield similar results. As such, the persistent finding of a lesbian earnings advantage in previous studies can be attributed, at least in part, to employers' aversion to maternity risk and its associated costs.

These findings are also of critical importance to the general labour-market discrimination literature. Given the adverse earnings effect of maternity risk, our analysis suggests that estimates of the well-established gender earnings disparity are likely to be considerably smaller when incorporating maternity risk into the analysis. Absent the ability to adequately control for maternity risk, strict attention should be paid to potential upward bias in estimated earnings differentials. Moreover, policymakers should consider the broader implications of maternity-leave policy on the labour-market outcomes of females. In this respect, maternity-leave policy may influence the hiring and promotion decisions of employers, thereby indirectly affecting sexual-orientation and gender equality in the labour market. However, further research in this area is still required, given the limitations inherent in the direct and indirect analyses.

List of Abbreviations

ACS	American Community Survey
ADA	Americans for Democratic Action
ALSWH	Australian Longitudinal Study on Women's Health
AUT	Association of University Teachers
CCHS	Canadian Community Health Survey
CDC	Centres for Disease Control and Prevention
CHIS	California Health Interview Survey
COPE	Committee on Political Education
CPS	Current Population Survey
ENDA	Employment Non-Discrimination Act
FIML	Full Information Maximum Likelihood
GSP	Gross State Product
GSS	General Social Survey
HLM	Hierarchical Linear Models
IMR	Inverse Mills Ratio
ISSP	International Social Survey Programme
LFS	Labour Force Survey
LISA	Longitudinal Integration Database for Health Insurance and Labour Market Studies
LOUISE	Longitudinal Database for Education, Income and Employment
MLE	Maximum Likelihood Estimation
NHSLS	National Health and Social Life Survey

OLS	Ordinary Least Squares
PUMS	Public-Use Microdata Sample
Q1	Quarter I
Q4	Quarter IV
UK	United Kingdom
U.S.	United States

1. Introduction

1.1. Introduction to the Study

Homosexual women earn more than heterosexual women. Theorists have proposed various potential reasons for the observed lesbian pay gap, such as differences in child-rearing status, labour supply, occupational choice and investments in human capital. Empirical studies testing the plausibility of a number of these theories generally find that an unexplained income differential in favour of lesbians exists, even after controlling for the proposed sources of the premium. If we take unexplained earnings differences as demonstrating labour-market discrimination, what are we to make of the persistent finding of a lesbian premium? Do employers discriminate in favour of lesbians, contrary to popular belief, or is the observed behaviour simply consistent with employer profit maximisation?

Petit (2007) proposes that, all else equal, employers may selectively hire workers to minimise expected costs of maternity leave. Specifically, employers are averse to bearing the interruption, replacement and maternity-benefit costs associated with employees' labour-force separation to give birth to or care for a child. The 'maternity-risk hypothesis' thus asserts that a reluctance to employ women exhibiting a greater likelihood of child-bearing results in a compensatory wage reduction to induce employers to bear the higher expected costs. Given that training and replacement costs are likely an increasing function of a position's salary, differences in the perceived likelihood of labour-market separation due to maternity, here-labelled 'maternity risk', may also result in discriminatory promotion practices. As lesbians represent a comparatively lower maternity risk relative to heterosexual women, a favourable lesbian pay gap may arise, even among women with equal productivity.[1]

Although previous studies have hinted at the importance of fertility and labour-force commitment in wage determination, no study has explicitly examined whether the forward-looking likelihood of temporary labour-market separation to bear children adversely affects earnings, thus giving rise to a lesbian premium. Using data from the 2000 United States (U.S.) Decennial Census, and 2005-2010 American Community Surveys (ACS), this study attempts to fill this void in the literature by estimating the effect of maternity risk on earnings and the lesbian pay gap. In addition to illuminating whether the observed premium can be partially

[1] The Pregnancy Discrimination Act of 1978 prohibits "sex discrimination on the basis of pregnancy" (U.S. Equal Employment Opportunity Commission, n.d.). Even if the law were strictly enforced, which is unlikely, the extent to which it would prevent discrimination on the basis of perceived maternity risk is minimal.

attributed to differences in maternity risk across demographic groups, the investigation reveals information of critical importance to the general labour-market discrimination literature. In particular, if maternity risk exerts a negative effect on earnings, holding all other factors constant, then existing studies examining the gender and sexual-orientation pay gaps may present biased estimates of discrimination. Moreover, policies affecting employers' maternity-cost incidence may influence hiring and promotion decisions, thereby affecting both labour-market prospects for females and earnings disparities across genders and sexual orientations. The implications of the analysis are thus of far-reaching consequence.

1.2. Overview

Using cohabitation to identify sexual orientation and marital status, we first assess whether a lesbian premium exists in three non-overlapping samples from the 2000 U.S. Census, pooled 2005-2007 ACS and pooled 2008-2010 ACS. Estimating traditional earnings equations, we uncover a statistically and economically significant lesbian pay advantage in all three samples. This finding is robust to the inclusion of controls for work and personal characteristics, region of residency, occupation and industry. Moreover, robustness checks show the presence of a lesbian pay gap is not conditional upon the dependent variable specification or the estimation method implemented. In the remainder of the thesis, we then examine whether this premium can be partially attributed to sexual-orientation differences in maternity risk.

Using intra-household relationships to infer fertility among householders, we allocate maternity rates by age group, sexual orientation, and marital status, applying within-sample maternity risk as an estimate of employers' forward-looking maternity-leave expectations. Including perceived maternity risk as an additional control in earnings regressions reveals mixed evidence as to the effect of maternity risk on earnings and the lesbian premium. We argue that prior promotion effects of maternity risk and sexual-orientation differences in accrued human capital, which we cannot disentangle, generate bias in the coefficient estimates. To minimise such bias, we allow observable human capital measures to differentially affect earnings by sexual orientation.

The results from this 'direct analysis' of the maternity-risk hypothesis strongly indicate a negative effect of maternity risk on earnings. This finding applies to both comparison groups, namely cohabiting and married heterosexual women, and is robust to various tests of model robustness. Importantly, the results also appear to be robust to further disaggregation in the maternity-risk allocation procedure. With respect to the magnitude of the pay gap, the majority of the results imply that accounting for near-term maternity risk reduces the lesbian premium by approximately ten to fifteen percent when evaluated at the mean level of potential experience.[2] Bias in the effect of potential experience and the inclusion of sexual-orientation interaction terms, however, pose problems in interpreting the results.

To supplement the direct analysis and circumvent its shortcomings, we assess whether contextual factors that likely affect the lesbian-heterosexual maternity gap alter the lesbian premium. Specifically, controlling for individual-level variables, as well as state-level ideology and other income-affecting factors, we examine whether state-level proxies exert the expected influence on the lesbian pay gap. The four proxies used include laws mandating insurance coverage of infertility treatment, domestic partnership laws, same-sex marriage bans, and lesbian prevalence rates.

Overall, the indirect tests of the maternity-risk hypothesis reveal considerable evidence in favour of the maternity-risk hypothesis. All four of the proxies generally affect the lesbian premium in a manner consistent with prior expectations. This finding withstands a host of robustness checks, including different dependent variable specifications, the addition of further individual and state-level controls, and alternative estimation techniques. Stratification by age and repeating the analysis for an analogous sample of males provide useful validity checks, demonstrating that mandated insurance coverage may not be an appropriate proxy for differential maternity risk. The estimated effect of lesbian prevalence also appears to be augmented by non-maternity-related factors. Although the results must be interpreted with caution, accrued maternity-risk effects appear to account for a significant portion of the lesbian premium over otherwise-similar heterosexual women.

Extending the analysis to the gender pay gap also demonstrates an adverse effect of maternity risk, correcting for which is shown to reduce the male-female earnings disparity by approximately fifteen percent, on average. Therefore, a plausible interpretation of the results from this thesis is that apparent sexual-orientation and gender discrimination in the labour market can, at least in part, be explained by the rational response of employers to differential maternity risk among employees.

---

[2] In the context of this paper, "near-term maternity risk" refers to the likelihood of a female requiring maternity leave in the next one to five years.

The remainder of this thesis proceeds in the following manner. Section 2 provides a review of the empirical literature examining the lesbian pay gap, including a more detailed discussion of the purpose of this study. Section 3 examines whether a lesbian premium exists in each of the three samples, forming the base upon which the remainder of the paper builds. Sections 4 and 5, respectively, present the direct and indirect tests of the maternity-risk hypothesis. Section 6 provides a discussion of the findings, including implications for the gender pay gap, improvements and ideas for future research, and finally offers concluding remarks.

2. Literature Review

2.1. Current Explanations of the Lesbian Pay Gap

Although analysing the sexual-orientation pay gap appears prone to fewer problems than that of the gender pay gap, substantial heterogeneity still exists between lesbians and heterosexual women. Importantly, many income-determining factors, such as human capital, labour-force attachment and occupation, differ systematically by sexual orientation, leading to a sizable earnings differential. As previously noted, prior studies find that controlling for observable characteristics reduces, but does not eliminate, this differential, leaving an unexplained pay gap.[3] In this section, we discuss potential explanations of the lesbian premium as proposed in the literature and offer evidence regarding the plausibility of each theory.

Becker's (1971) "taste for discrimination" argument is often cited in studies examining the effect of sexual orientation on labour-market outcomes, despite discrimination against homosexuals being unable to explain the presence of a lesbian premium. In short, the theory implies that if homosexuality is stigmatised by customers, co-workers or employers, then we would expect to observe a lesbian earnings penalty.[4] Despite gay males repeatedly being shown to earn less than heterosexual males, as Becker's argument would suggest, few studies find evidence consistent with his proposition when examining the lesbian pay gap.[5] Considering this result in the context of the broader labour-market discrimination literature further suggests this is an anomalous finding. Specifically, field experiments by

[3] The magnitude of the lesbian pay gap estimated in previous studies is discussed in greater depth in Section 2.2.
[4] The extent to which employer discrimination translates into earnings differences depends on the level of market competition. For example, employer discrimination cannot persist in a perfectly competitive market.
[5] For an overview of the gay-male income penalty, see Badgett's (2006) summary. In the context of the lesbian pay gap, Ahmed and Hammarstedt (2010), Badgett (1995) and Carpenter (2008b) obtain results at least partially consistent with Becker's discrimination theory.

Weichselbaumer (2003, 2013), Drydakis (2011) and Ahmed et al. (2013a) show that lesbians are discriminated against in the hiring process, receiving fewer call-backs than otherwise-similar heterosexual women.[6] Badgett et al. (2009) summarise evidence from several recent surveys assessing the extent of discrimination against lesbians and gay males. They note that 15% to 43% of lesbian, gay and bisexual respondents experienced some form of workplace discrimination. Although the reliance on convenience samples and perceived discriminatory treatment reduces their applicability, the studies summarised by Badgett et al. (2009) provide further evidence that lesbians face adverse discrimination in the labour market. Explaining the observed lesbian premium thus represents an important area of research for academics and policymakers, particularly as the premium may be attenuated due to discriminatory treatment.

Traditional explanations of the lesbian pay gap attribute the observed premium to differences in child-rearing status and consequent differences in human capital accumulation. Using 1990 U.S. Census data, Black et al. (2000) estimate that 22% of partnered lesbians have children present in the household, compared with 38% of cohabiting heterosexual women and 59% of married heterosexual women. Similarly, using pooled 1989-1996 data from the U.S. General Social Survey (GSS), Blandford (2003) estimates that 25% of lesbians have borne children, compared with 62% of unmarried and 84% of married heterosexual women. On average, lesbians are thus less likely to have children than otherwise-similar heterosexual women, as we would expect. Children represent both an erosion of work-relevant human capital and time outside of the workforce. As a result, lesbians, with fewer children, enjoy higher wages than otherwise-similar heterosexual women. This explanation can be traced back to the motherhood earnings gap literature. Korenman and Neumark (1992) observe that women without children earn significantly more than women with children, even after accounting for observable characteristics. More recent studies by Budig and England (2001) and Zhang (2009) reach similar conclusions. On this basis, we would expect to observe a wage disparity in favour of lesbians relative to heterosexual women. Among several other studies, Black et al. (2003) test this theory by including the number of children present in the household as an additional regressor. They find that the number of children has a large negative effect on annual income but only modestly reduces the estimated lesbian premium. Jepsen (2007) divides her sample according to the presence of children and finds that the

[6] In contrast, a recent study by Baert (2013) finds weakly significant evidence of higher call-back rates among young lesbians relative to their heterosexual counterparts.

results do not differ substantially across groups. There is thus little evidence to support that the lesbian earnings premium can be attributed to differences in child-rearing status.

Black et al. (2003) argue that observed income differences may be due to lesbians' exhibiting a comparatively greater degree of specialisation in market production than heterosexual women. A plausible explanation is that lesbians, on average, will be in relationships that share household tasks more evenly than in heterosexual relationships, where household tasks are primarily undertaken by females (Kurdek, 1993). If a reduction in household responsibilities leads to greater labour supply, lesbians should accrue more work-related human capital. To the extent that the resulting unobservable differences in human capital are not captured by observable human capital measures, such as potential experience and education, earnings regressions are likely to find a lesbian premium. Moreover, studies that do not control for annual hours worked likely estimate larger sexual-orientation remuneration differences. Assuming that a partner's contribution to household tasks is inversely related to the number of hours devoted to market production, Jepsen (2007) includes partners' hours worked into her regressions to test the household specialisation argument.[7] She finds no economically significant effect of partners' hours worked on annual income, failing to corroborate the theory. Controlling for standard factors affecting labour supply, Tebaldi and Elmslie (2006) find that lesbian women display stronger labour-market commitment than both married and cohabiting heterosexual women, working longer hours and exhibiting greater willingness to partake in full-time employment. Klawitter (2011) reports similar results. Further, Klawitter (2011) and Elmslie and Tebaldi (2007) demonstrate that the lesbian premium is considerably smaller when controlling for annual hours worked. Their findings are consistent with the view that earnings differentials are, in part, driven by differences in the division of household labour between lesbian and heterosexual couples.

Another oft-cited explanation is that lesbians are more likely to invest in human capital for their careers. Becker (1991) makes the case that heterosexual women and lesbians will differ in terms of career choices and investments in human capital. Because heterosexual women anticipate forming a traditional household, and thus expect to participate more in non-market production than lesbians, they would invest less in work-relevant human capital, giving rise to a lesbian wage premium. Black et al. (2003) offer a similar explanation.

[7] Jepsen notes that "This approach relies on the debatable assumption that partner hours are exogenous to the woman's earnings." (Jepsen, 2007, p. 716). Arguably, the household specialisation theory is less relevant in the case of childless couples. Jepsen's results, however, hold for both women with and without children.

Daneshvary et al. (2009) provide an indirect test for this by examining whether previous marriage can explain the sexual-orientation wage disparity present in the 2000 Census data. A previously married lesbian may have anticipated forming a traditional household and thus early career decisions of previously married lesbians are more likely to mirror those of heterosexual women. They find that accounting for previous marriage significantly reduces the lesbian wage premium. Furthermore, allowing the effects of previous marriage to differ by sexual orientation reveals a larger negative wage effect for lesbians and a much smaller wage disparity, lending strength to Becker's argument.

Several other explanations build upon a similar premise. Berg and Lien (2002), for example, claim that lesbians suffer from a negative income effect due to having a partner with depressed earnings as a result of the gender pay gap. In this situation, lesbians respond by supplying more labour, giving rise to an earnings differential in a similar manner to the household production specialisation case. Peplau and Fingerhut (2004) stress that lesbians' greater labour-force attachment arises not only out of differences in household production and child-rearing responsibilities, but also due to their desire to be financially independent. That is, lesbians emphasise equal sharing of financial responsibilities and are less likely than heterosexual couples to pool financial resources (Peplau and Fingerhut, 2004). Involvement in a committed relationship therefore does not reduce a lesbian's need for full-time employment to the same extent as it does for a heterosexual woman. Ahmed et al. (2013b) and Badgett (2001) argue that greater career devotion and independence from family-related career interruptions are primarily responsible for stronger labour-force attachment among lesbians. As previously discussed, higher full-time employment rates and annual hours worked among lesbians provide evidence consistent with these explanations. Badgett (1995) offers an alternative method to assess whether the data are consistent with this theory by including an interaction between the lesbian indicator dummy and potential experience in her regression models. If lesbian absences from the labour force are fewer and shorter in duration, then potential experience should more accurately reflect actual experience for lesbians and systematically overstate that of heterosexual women, resulting in a positive coefficient on the interaction term. Badgett (1995) estimates a positive, but statistically insignificant, interaction-term coefficient, providing only weak support for the labour-force-attachment theory. Further studies by Jepsen (2007) and Daneshvary et al. (2008), however, obtain economically and statistically significant coefficients which substantiate Badgett's proposition.

Further explanations of the sexual-orientation earnings gap include Blandford's (2003) argument that lesbians are disproportionately represented in traditionally male-dominated occupations. Generally, these professions entail greater remuneration and hence, when failing to include occupational variables in wage decompositions, we would expect to see a positive coefficient on the lesbian indicator variable.[8] Blandford (2003) finds that moving from one-digit to two-digit occupational controls reduces the lesbian premium by approximately 25%, and argues that controlling for more subtle occupational clustering may further diminish the estimated earnings effect of sexual orientation. Conversely, Antecol et al. (2008), Daneshvary et al. (2009), and Carpenter (2005) find that although lesbians are over-represented in male-dominated professions, the apparent effect of occupational sorting on the lesbian premium is minimal.

Another candidate explanation relates to sexual-orientation differences in aggressiveness in wage negotiations. Babcock and Laschever (2003), among others, argue that gender differences in assertiveness may be partially responsible for the observed male-female pay gap. To the extent that hormonal or social differences result in lesbians exhibiting greater aggressiveness in seeking higher pay than their heterosexual counterparts, lesbians may be able to avoid some of the gender disadvantage, thereby resulting in the observed earnings premium. Frank (2006) presents a similar argument. Unfortunately, evidence supporting or refuting this theory is severely lacking.

Finally, Plug and Berkhout (2004) note that AIDS incidence among heterosexual women exceeds that of lesbians. To the extent that worker health translates into productivity differences, one should thus expect to observe a lesbian premium. Elmslie and Tebaldi (2012) analyse whether the reverse holds true for gay males. They find a weak positive correlation between negative HIV–related news and the gay-male penalty, but interpreting this finding in the context of their overall analysis reveals little support for statistical discrimination on the basis of HIV/AIDS incidence. Moreover, AIDS is sufficiently rare among both lesbians and heterosexual women that it is unlikely to significantly drive the sexual-orientation earnings differential. Given that mental and physical health likely differ by sexual orientation for reasons unrelated to HIV/AIDS, examining the effects of health status on the lesbian pay gap may represent a promising area for future research.

[8] In Section 2.3 we discuss the theoretical concerns relating to the inclusion of occupational controls in earnings regressions.

2.2. Estimates of the Lesbian Pay Gap

2.2.1. Evidence from the United States

Badgett (1995) was the first to employ an econometric analysis to examine the earnings effect of sexual orientation. Her study pooled 1989-1991 data from the U.S. GSS and classified lesbians as women with at least as many same-sex partners as opposite-sex partners.[9] Implementing Ordinary Least Squares (OLS) and controlling for several factors known to affect earnings, such as years of formal education, race, potential experience, geographic location, and occupation, she finds that lesbians are at an earnings disadvantage of approximately 21-30% as compared to heterosexual women.[10] Estimates obtained using the Heckman two-step procedure suggest a lesbian earnings disadvantage of 11% relative to unmarried heterosexual women and a premium of 15% over currently married women.[11] The lack of statistical significance among Badgett's estimates, however, renders her results inconclusive.

Several studies have followed Badgett (1995) in utilising data from the GSS.[12] The first of these was Badgett's (2001) update to her original study, which pooled data from the 1989-1994 GSS and 1992 National Health and Social Life Survey (NHSLS). Estimating a specification similar to her original study, absent the sexual-orientation interaction term, and relying on the same sexual orientation definition, Badgett finds an insignificant lesbian earnings premium of roughly 2-11%. Black et al. (2003), using 1989-1996 GSS data, replicate and substantially extend Badgett's pioneering work. Using Badgett's sexual orientation definition, Maximum Likelihood Estimation (MLE) reveals a lesbian earnings advantage of 6% over unmarried heterosexual women, and 9% compared to married women, although their estimates are also statistically insignificant. In contrast to Badgett (1995), Black et al. find that invoking alternative definitions of sexual orientation considerably affects their results. Specifically, considering sexual behaviour over the past year or five years, they find that

---

[9] Technically, this was Badgett's definition of lesbian/bisexual; however, she makes no distinction between the two, most likely due to concerns over sample size.

[10] Badgett includes the interaction between sexual orientation and potential experience as an additional regressor.

[11] Throughout this paper, we refer to unmarried heterosexual women as "cohabiting heterosexual women" when they are identified via the cohabitation procedure, and "unmarried heterosexual women" when sexual orientation is determined via other methods. We elaborate on the cohabitation procedure in Section 3.1. As the majority of this paper focuses on females, we often refer to cohabiting heterosexual women and married heterosexual women as "cohabiting heterosexuals" and "married heterosexuals", respectively.

[12] All cited studies examining the GSS restrict their samples to full-time employed individuals to minimise distortions resulting from differences in hours worked and lower hourly wages generally received by part-time employees. Further details pertaining to each study are provided in Appendix A, Table A.1.

annual earnings of lesbians exceed those of otherwise-similar heterosexual women by 22-40%, a result that is highly robust to equation specification. Blandford (2003), with the same dataset, estimates similar models using OLS and the Heckman two-step procedure. He defines "open" lesbians as those women who are behaviourally lesbian based on sexual experiences in the past year (or five years if an individual reports having had no partners in the past year), and are not currently married. This final restriction limits the analysis to lesbians who are most likely to be perceived as such by employers. These differences appear to have a minimal impact on the results, as Blandford's findings suggest that lesbians earn between 15 and 38% more than heterosexual women, mirroring Black et al. (2003).

Three additional published studies use GSS data to examine the lesbian pay gap, although the analyses presented diverge more materially from Badgett's original (1995) paper. Using data spanning 1991-1996 and defining sexual orientation based on behaviour over the previous five years, Berg and Lien (2002) estimate an ordered-probit type (MLE) model, controlling for broadly-defined occupational categories, human capital proxies, and an incomplete set of geographic dummies. Their regression model suggests lesbians earn 13-47% more than heterosexual women, reflecting the uncertainty surrounding the point estimate of 30%. Carpenter (2005) supplements his primary analysis of California Health Interview Survey (CHIS) data by considering the sensitivity of earnings differentials found in the GSS to the choice of time period analysed. Employing 1988-2000 GSS data, and classifying lesbians as women with exclusively same-sex relations in the past five years, he finds a lesbian premium of 21-31%, with temporal differences in both the size and significance of the estimates. Cushing-Daniels and Yeung (2009) analyse data from 1988-2006, offering separate analyses for different sub-periods and definitions of sexual orientation. Models estimated via OLS over the periods 1991-1996 and 1988-2006 produce results consistent with previous studies, with lesbians earning 9-15% more than heterosexual women.[13] An identical specification for the 1998-2004 period, however, reveals a small lesbian earnings disadvantage. Moreover, results obtained through Heckman Full Information Maximum Likelihood (FIML) estimation suggest that lesbians receive 13-14% lower annual compensation than married heterosexual women, and 5-7% higher earnings than unmarried heterosexual women, a finding inconsistent with even Badgett's original (1995) paper.

[13] Unlike many of the previous studies, Cushing-Daniels and Yeung fail to obtain statistically significant lesbian earnings differentials in any of their specifications.

Despite the initial popularity of the GSS, many recent studies examining the sexual-orientation earnings gap have used Public-Use Microdata Sample (PUMS) data from the 1990 and 2000 U.S. Decennial Census. The first such study, by Klawitter and Flatt (1998), used a 5% sample from the 1990 Census to analyse the effects of state and local anti-discrimination policies on income and the lesbian pay gap. Similarly to subsequent studies using census data, they define lesbians as females cohabiting with an "unmarried partner" of the same sex. In addition to the standard controls used in the aforementioned studies, Klawitter and Flatt account for state-level social factors that may affect earnings differentials in their OLS regressions. Among all employed women, they estimate a lesbian premium of 3-16% relative to cohabiting heterosexual women and 11-23% compared with married heterosexual women, depending to some extent on the presence and type of anti-discrimination policies. Importantly, however, they find that restricting the sample to full-time, full-year workers eliminates any statistically significant earnings difference. Clain and Leppel (2001), using 1-in-1000 data from the 1990 Census, implement the Heckman two-step procedure on all full-time employed individuals. Their models, which are obtained using stepwise specification search, notably include controls for partner's income and interactions between the lesbian indicator dummy and several variables, making interpretation of their results difficult. However, they find that lesbians are not at an earnings advantage over partnered heterosexual females unless they live in the Midwest or with dependents. Including women not cohabiting with a partner in their sample, they generally find a statistically significant lesbian premium, but this varies substantially across observable characteristics.

Many studies have used more recent data from the 2000 Census (5% sample). Jepsen (2007), for instance, estimates a range of OLS models to assess various theories of earnings differentials. Her base specification, which is similar to previous studies, suggests lesbian premiums of 9-14% over married heterosexual women and 10-17% relative to cohabiting heterosexual women, for those in full-time employment.[14] In models with several interaction terms, lesbians appear to earn significantly more than other women, although the published results preclude evaluation of the premium at the means of interacted variables. Arabsheibani et al. (2007) stratify their analysis of the lesbian premium by education, job sector, region, age, and employment status. Their OLS analysis of the U.S. Census data implies a lesbian premium of 3-12% over heterosexual couples, with the premium being greater for older

[14] In robustness tests, Jepsen relaxes this constraint to include part-time employed women. She finds the results are robust to their inclusion.

women, those employed full-time or in the private sector, and among women with lower educational attainment. Antecol et al. (2008) decompose the sexual-orientation pay gap among full and part-time employed white women via two alternative techniques. Their Blinder-Oaxaca decompositions suggest a lesbian premium of 4% over cohabiting heterosexual women, but no difference relative to married women.[15] Adopting the decomposition method of DiNardo, Fortin, and Lemieux (1996), they find lesbians receive 8% higher hourly pay than heterosexual women at the lower end of the wage distribution, but there is minimal divergence among higher-earning women. Daneshvary et al. (2008) analyse the effects of education on the sexual-orientation wage gap. Using Blinder-Oaxaca decompositions, they find evidence of a lesbian premium of approximately 8-10% among women without a bachelor's degree, but a wage advantage of only 2% at higher levels of education. Their OLS models further suggest that the premium is greater among women with higher levels of potential experience. Daneshvary et al. (2009), using the same dataset, investigate the impact of a previous marriage on the lesbian wage premium. They find a premium amounting to 4-8% for never-married lesbians. Having had a previous marriage reduces the premium to 2-4% over cohabiting and single women, and the wage advantage over married women disappears altogether.[16]

Gates (2009) combines the 1% and 5% PUMS from the 2000 U.S. Census to examine the effects of sexual-orientation anti-discrimination policies on the wages of lesbians. Restricting his sample to full-time employed women and interacting sexual orientation with the policy variable in OLS regressions, Gates finds that lesbians earn 4-8% higher wages than other women in states with sexual-orientation anti-discrimination laws, while absent such policies the premium drops to 3-6%. Two final studies using the 5% PUMS estimate Hierarchical Linear Models (HLM), or 'multilevel models' to analyse the effects of state-level factors on the lesbian premium. Controlling for individual factors and state-level gay tolerance, Baumle and Poston (2011) estimate lesbian premiums of 4% and 8-9% over married and cohabiting heterosexual women, respectively, among all employed women. Klawitter (2011) updates her co-authored 1998 paper using more recent data from the 2000 Census, superior methodology (multilevel models and quantile regression), and a number of additional robustness tests. She finds a significant lesbian advantage in hourly wages of 3%

---

[15] See Blinder (1973) and Oaxaca (1973) for more information on this approach.
[16] Both studies by Daneshvary et al. restrict their analysis to full-time employed women, although in their latter study the authors report robustness checks including part-time workers.

and 7% relative to married and cohabiting heterosexual women, respectively, while comparisons of annual income yield respective premium estimates in excess of 16% and 9%.

Other studies employ more unique datasets. Carpenter's (2005) primary analysis uses data from the 2001 CHIS, allowing him to classify lesbians based on self-reported sexual orientation. Across several specifications, he finds that hourly wages of lesbians are between 6% lower and 4% higher than otherwise-similar heterosexual women, although none of the estimates are statistically significant. This study may not be representative of the U.S. as a whole, due to the highly liberal nature of California.[17] Thus, the absence of a sexual-orientation earnings differential in California is not necessarily indicative of the same on a national scale. Elmslie and Tebaldi (2007), using data from the 2004 Current Population Survey (CPS), adopt the cohabitation procedure in defining sexual orientation. Implementing the Heckman two-step procedure, they find no significant hourly wage differences between lesbians and married women and a lesbian premium of roughly 2-8% compared with cohabiting heterosexual women. Importantly, Elmslie and Tebaldi find that failing to control for hours and weeks worked produces qualitatively similar results to previous studies, with the estimated lesbian premium jumping to approximately 19% in one specification.[18]

2.2.2. International Evidence

The abundant evidence of a lesbian earnings premium in the U.S. has spawned a growing literature establishing whether this phenomenon extends to other countries. Plug and Berkhout (2004) provide one of the first international examinations of the sexual-orientation earnings gap using 1998-2000 survey data of recent Dutch university graduates. Their study departs from the U.S. analyses along a number of important dimensions. Namely, they define lesbians based on self-identified sexual preferences,[19] their sample includes only young women with high educational attainment, and their OLS analysis combines both males and females. Among all employed females, they find no significant sexual-orientation differences in

[17] Although greater gay-friendliness should reduce sexual-orientation discrimination and increase the premium, it may also enhance lesbian maternity incidence, reducing the premium. The direction of the bias induced by the use of a California-specific sample is thus indeterminate.

[18] Carpenter (2004) compares household incomes of different-sex and same-sex couples using independent 1996-2000 data from the Centres for Disease Control and Prevention (CDC). He finds that same-sex female households obtain lower incomes than different-sex couples. The empirical findings from this study are not discussed in detail, nor are they included in Table A.1, as household income differences are significantly affected by the earnings of male partners in heterosexual households.

[19] The relevant question stated "Concerning your sexual preference, what do you prefer?". Respondents could choose between three alternatives: 1) only men; 2) only women; and 3) both men and women.

monthly or hourly earnings across several specifications. Restricting the analysis to full-time employed individuals, however, reveals a lesbian premium of 3-4%. Carpenter (2008b) similarly studies both the pecuniary and non-pecuniary effects of sexual orientation among young women using confidential data obtained from the Australian Longitudinal Study on Women's Health (ALSWH). Using self-reported sexual orientation and interval regression, Carpenter estimates a number of models by progressively adding controls for human capital, health status, occupation, and other factors, along with several robustness checks. He finds that the lesbians in his sample are significantly disadvantaged in the labour market, receiving 24-31% lower compensation than otherwise-similar heterosexual women. Frank (2006) also analyses a non-representative sample obtained from the United Kingdom (UK) Association of University Teachers (AUT) 2000-2001 survey of employees at six British universities. Pooling data for both genders and using self-reported sexual orientation, Frank's OLS results suggest an insignificant lesbian premium of approximately 8% among all staff. Limiting the sample to academics reveals a statistically significant premium of 17%, but including further controls for occupational rank renders the estimated premium (5-14%) insignificant.

Arabsheibani et al. (2004) provide the first such analysis of nationally representative UK data by pooling the 1996 Quarter I (Q1) to 2001 Quarter IV (Q4) waves of the Labour Force Survey (LFS). They classify lesbians via the cohabitation procedure and stratify their OLS analysis by region and age. Their results show that lesbians receive 10-13% higher hourly wages than single women, and 3-21% greater hourly remuneration than partnered heterosexual women, with the estimates being larger among older women and non-Londoners. Arabsheibani et al. (2005) extend their 2004 study by including four additional quarters of data (1996Q1-2002Q4) and opting for Blinder-Oaxaca decompositions over OLS. Their results are consistent with their previous study, finding a lesbian wage advantage of 8% and 9% over coupled and all heterosexual women, respectively. As noted in the previous section, Arabsheibani et al. (2007) examine the lesbian premium in the U.S., disaggregated by several factors known to influence wages. In their paper, they concurrently examine 1996Q1-2004Q4 LFS data with a comparable model specification and report qualitatively similar findings to those in the U.S. and their earlier studies. Specifically, their results reveal that lesbians receive 2% lower to 12% higher wages than heterosexual women, with larger premiums observed among older and less educated women, part-time workers, and individuals employed in the public sector.

Carpenter (2008a) pools data from the 2003 and 2005 Canadian Community Health Surveys (CCHS) in order to assess whether the results obtained from U.S. studies carry over to Canada. Using self-reported sexual orientation and OLS, he estimates a lesbian premium of 16-17%, a result highly consistent with those found in U.S. studies. Re-estimating his models by relationship status, Carpenter observes this premium is predominantly driven by a substantial lesbian advantage among partnered women (43%), while that for single lesbians is virtually non-existent (1%). Using 1994 data from the International Social Survey Programme (ISSP), Heineck (2009) offers a truly international examination of the earnings effects of sexual orientation.[20] Applying the Heckman two-step procedure and defining homosexuality based on exclusivity of same-sex relations in the past five years, Heineck estimates the lesbian premium among all employed females to be approximately 10% in both his gender-pooled and female-specific regressions. His estimates, however, are not statistically significant. Laurent and Mihoubi (2012) evaluate the wage effects of sexual orientation in the French labour market by aggregating 1996-2007 data from the French Employment Survey. Although data collection issues preclude use of the "spouse" response to identify same-sex couples, the authors classify homosexual households as those comprising two adults of the same gender reporting a friendship and imposing several filters to minimise misclassification.[21] Implementing OLS and the Heckman two-step procedure, they estimate that lesbians employed in the private sector earn 2% and 4% more than cohabiting and married heterosexual women, respectively. This earnings advantage disappears almost entirely when considering females in public sector employment, with lesbians boasting a mere 1% earnings advantage over their married heterosexual counterparts and receiving no such premium relative to cohabiting heterosexual women.

Ahmed and Hammarstedt (2010) use Swedish register data from the Longitudinal Database for Education, Income and Employment (LOUISE) to compare 2003 annual income of all lesbians living in a civil union with those of married heterosexual women. Their OLS results, which are stratified by residence in a metropolitan area, suggest that lesbians earn between 20% less and 8% more than married heterosexuals, with the largest income penalties being observed in non-metropolitan areas. Inspecting the quantile regression estimates also

---

[20] Included in the analysis were U.S., Australia, Ireland, Poland and Bulgaria. In addition to standard controls seen in prior studies, Heineck includes country fixed-effects to control for international earnings differences.

[21] These filters include removing all couples where either member is a student, apprentice, farmer, or retiree; applying minimum and maximum age constraints; retaining couples only where both members are French; and including minimum and maximum income restrictions.

reveals a large and significant earnings disadvantage for lesbians, particularly among those at the lower end of the income distribution. A subsequent study by Ahmed et al. (2013b) uses 2007 register data from the extended LOUISE database, renamed LISA (Longitudinal Integration Database for Health Insurance and Labour Market Studies), permitting comparisons of both annual income and full-time monthly income. Closely following the method of Ahmed and Hammarstedt (2010), they find lesbians in a civil union earn between 4% less and 3% more than their married heterosexual counterparts, based on full-time monthly earnings. Moreover, their complete results reveal larger lesbian premiums when comparing annual income, as well as in the public sector and at the upper end of the earning distribution.

2.3. Theoretical Considerations

Evidently, the majority of studies emanating from both the U.S. and abroad uncover a positive and significant lesbian pay gap, although this varies substantially depending on the country and dataset analysed, as well as the model specification and estimation technique implemented. In this section, we highlight several shortcomings pertaining to a number of previous studies which reduce the applicability of their findings. We then discuss the purpose of this thesis, with specific reference to how it builds upon the previous literature.

2.3.1. Deficiencies in Prior Research

Although prior research provides evidence of a clear and convincing lesbian earnings premium, several contributing studies suffer from deficiencies which limit their dependability in informing debate and policy decisions. As noted by Elmslie and Tebaldi (2007), the major shortcoming of many studies is that they incur bias in coefficient estimates on lesbian indicator variables by failing to account for differences in hours worked. Studies seeking to attribute the residual pay gap to discrimination thus overstate true sexual-orientation remuneration differences. As shown in Table A.1, estimates of the lesbian earnings premium are markedly higher among studies lacking controls for annual hours worked. A second potential problem in the current literature is that inadequate consideration is given to marital status and its effect on labour-market outcomes. Elmslie and Tebaldi (2007) argue that marriage significantly affects labour-market outcomes and this effect must be separated out before addressing concerns relating to sexual-orientation discrimination. Several studies, however, include a marital status variable and fail to articulate its importance in interpreting the pay gap or neglect its inclusion altogether.

Some datasets, such as the GSS, suffer from the additional problem of containing a small sample, even when data is pooled across a number of years. As a result, estimates obtained using these datasets are highly sensitive to the time period analysed, as discussed in Black et al. (2003) and Cushing-Daniels & Yeung (2009). Another major drawback of some studies is that their narrow research population hinders inferences relating to the overall population of interest.[22] For example, region-specific results do little to illuminate nationwide earnings differentials, while an earnings disadvantage for young lesbians does not imply an analogous penalty among older women.[23] A further problem present in the literature is the inconsistent treatment of occupational choice. Blandford (2003) argues that due to the unequal occupational distribution of heterosexual women and lesbians, studies failing to incorporate individuals' occupation likely bias earnings differentials upwards as a result of occupational remuneration differences. Conversely, Ransom and Oaxaca (2005) suggest that occupational segregation may itself be evidence of labour-market discrimination and, in the absence of selectivity corrections, including occupational controls may bias earnings differentials downwards.[24] These divergent views must be considered when deciding on an appropriate model specification. Finally, present empirical studies have not considered maternity risk and its consequent wage effect, resulting in potentially misleading inference.

2.3.2. Purpose of the Study

In this study, we attempt to build on the prior literature by explicitly accounting for the various shortcomings highlighted above. To remove potential biases associated with differing labour supply, we control for differences in annual hours worked in annual income regressions. While data limitations prohibit extensive use across all samples, where available the log of hourly wages is used as the dependent variable in robustness checks. Following Daneshvary et al. (2009), we stratify the analysis by marital status to allow for the presence of

---

[22] Carpenter (2005), Carpenter (2008b) and Frank (2006) are good examples of such studies. Although these studies do not allow generalisations pertaining to the entire population, which we cite here as a shortcoming, they provide a valuable contribution to the literature in examining a subset of the population of interest.

[23] Absent unobservable differences in human capital, the finding of an earnings disadvantage for young lesbians may be somewhat inconsistent with the maternity-risk hypothesis. However, as discussed in Chapter 4, sexual-orientation differences in human capital are likely directly related to potential experience and thus age, prohibiting such inference on the basis of an age-stratified analysis.

[24] Ransom and Oaxaca (2005) argue this in the context of the gender pay gap, although the same argument applies to the sexual-orientation pay gap.

a marriage premium or penalty.[25] Consideration is also given to the potential implications of prior marital history in determining maternity incidence in Section 4.3.[26] We avoid the robustness problems associated with a small dataset by obtaining large samples from the U.S. Census and American Community Surveys. Implicitly, the choice of datasets also removes the difficulties inherent in extrapolating results from a non-representative sample to the population level. Finally, the issue of endogeneity is not exclusive to occupational choice and as such should not be treated in isolation. Carpenter (2005) for instance, expresses concern that choices relating to occupation, geographic location and educational attainment may reflect responses to real or perceived discrimination and including these variables in regressions may thus lead to an “over-controlling” problem. Following Carpenter (2005), we account for this possibility in the initial analysis by estimating a number of models, incrementally adding controls which become progressively less exogenous.

This thesis attempts to provide an insight into the source of the lesbian pay gap, using more recent data relative to much of the current literature. In order to do so, a key purpose of this study is to investigate whether maternity risk adversely affects labour-market outcomes. Petit (2007) argues that employers may selectively hire employees in order to lower expected maternity-leave costs. Wage discrimination thus arises out of reluctance to employ women of greater maternity risk in highly paid jobs. Alternatively, differential treatment may be a result of compensatory pay reductions to induce the employer to bear higher maternity risk. In each case, the resulting earnings discrepancy represents a rational response on behalf of the employer.[27]

The second objective of this paper is to ascertain whether the lesbian earnings premium can, at least in part, be attributed to sexual-orientation differences in maternity risk. The 'maternity-risk hypothesis' asserts that maternity risk adversely affects earnings and thus sexual-orientation differences in maternity risk contribute to the observed lesbian premium. To our knowledge, no current empirical study examines the effect of maternity risk on income

[25] The results of unreported Chow tests support this partitioning of the sample, as the test statistic is statistically significant at conventional levels in each case. Combining all partnered heterosexual women into one group may thus impose unnecessary restrictions on the coefficient estimates.

[26] We do not include prior marital history as a control in regression analyses as selection into divorce may pose problems in the analysis, as noted by Daneshvary et al. (2009).

[27] In addition to reduced earnings, discrimination on the basis of maternity risk may manifest itself in employment discrimination. Namely, if employers exhibit an aversion to bearing maternity risk, one would expect to see lower employment rates among women with greater perceived levels of maternity risk. This study focuses on the earnings effect of maternity risk, while an examination of the employment effect is left for future research.

or the lesbian premium. This suggests that current studies, when controlling for number of children present in a household, are only incorporating one aspect of the sexual-orientation maternity differential. A significant void in the literature is thus evident, which may have implications for current analyses of the gender and sexual-orientation pay gaps. Importantly, if maternity risk is shown to adversely affect wages, present studies which do not control for maternity risk may produce upward-biased estimates of gender and sexual-orientation wage inequality. In Appendix B, we discuss the available evidence regarding the maternity-risk hypothesis in previous studies.

Finally, with this thesis, we hope to contribute to a strand of the literature which is currently in its infancy. Although policies affording rights to homosexuals and attitudes towards homosexuality vary considerably across states, few studies exploit this cross-state variation to examine sexual-orientation earnings differences. To our knowledge, only four such published studies exist, among which only two implement multilevel analyses similar to those estimated herein.[28] Furthermore, the timing of this thesis affords us the benefit of using more recent ACS data in addition to the census data used in previous studies.

3. The Lesbian Earnings Premium

3.1. Data Sources and Sample

As previously noted, this study relies on Public-Use Microdata Sample (PUMS) data from the 2000 U.S. Decennial Census (5% sample) and 2005 to 2010 American Community Surveys. The 2000 Census is the largest available dataset that allows an assessment of sexual orientation, while providing detailed information on labour-market outcomes (such as employment status, weeks worked and income) and demographic characteristics (such as age, education and race). Following the 2000 Census, the annual American Community Survey replaced the long form of the decennial census, providing more timely and often more complete information regarding socioeconomic status and demographics. The ACS PUMS provides information on approximately 1% of the U.S. population, requiring the data to be pooled across years to obtain a sample comparable to the 2000 Census. Due to differences in the censoring, surveying and reporting procedures implemented by the U.S. Census Bureau

[28] The indirect analysis contained in Section 5 was originally estimated using multilevel models. However, the complex model specifications led to difficulties in estimation when using the full dataset. As such, the use of multilevel models in this thesis is confined to assessing the robustness of findings to the choice of estimation technique.

across years, which may affect the compatibility of the data, we split the ACS data into two samples so as to combine only those years in which similar procedures were adopted.[29] This also facilitates a test of whether the results are invariant to the choice of sample period, which has often not been the case in previous studies, such as those using GSS data. The three resulting samples are the 2000 Census, pooled 2005-2007 ACS, and pooled 2008-2010 ACS.

Similarly to previous U.S. Census analyses, we categorise the relationship status of respondents based on the cohabitation procedure.[30] Specifically, female respondents are classified as lesbians when they reside with a female householder who identifies them as their "unmarried partner", bypassing other classification options for non-relatives, including "roomer/boarder" and "housemate/roommate". Cohabiting heterosexuals are identified when the householder classifies a different-sex individual within the household as their "unmarried partner", while married heterosexuals are identified by selection of the "husband/wife" tick-box. Although this thesis primarily compares the labour-market outcomes of the three aforementioned groups (individuals cohabiting with a partner or spouse), a collectively exhaustive set is formed for completeness, by creating a fourth group, here-labelled "no co-residential relationship". This group comprises all individuals not represented in the former categories, including children, single persons, coupled individuals living in separate households and cohabiting couples where neither partner is the householder. Implementing this identification procedure for each dataset yields the sample sizes contained in the first column of Table A.2. This shows, for example, that prior to any deletions, the 5% PUMS from the 2000 Census contains 7,211,710 females, of which 32,756 are classified as lesbians and 220,847 as cohabiting heterosexuals.[31] Due to computation constraints, fifteen percent random samples are taken of married women and women not partaking in a co-residential

[29] For example, significant formatting changes in the questionnaire in 2008 likely affected the reporting of same-sex spouses and same-sex unmarried partners, reducing the compatibility of the datasets (U.S. Census Bureau, 2009). The effects of the global financial crisis further motivate the split in the ACS sample to pre-2008 and post-2008.

[30] Appendix C outlines alternative methods to identify lesbians within social science data and discusses the advantages and disadvantages of using the cohabitation procedure. As noted in the appendix, the results in this thesis strictly apply only to coupled females.

[31] Although lesbians comprise only 0.99 to 1.20 percent of the coupled female population in the PUMS data, this is consistent with lesbian prevalence in other datasets, such as the GSS and NHSLS (Daneshvary et al., 2009). Moreover, as discussed in Appendix C, the lesbians in the sample are those most likely to be considered partnered lesbians by their employers, so potential undercounting of the lesbian population should not pose significant problems.

relationship.[32] Table A.2. details the composition of the remaining sample following this and subsequent restrictions.

The research population is then restricted to females between the ages of eighteen and sixty-five years inclusive. Following Daneshvary et al. (2008), both partners' observations are dropped if either partner is less than eighteen years of age. To further reduce unobservable heterogeneity, all residents of group quarters are excluded from the sample. Moreover, as discussed in Appendix C, several studies express concern regarding measurement error in enumerating same-sex couples from the PUMS files. Specifically, the edit and allocation procedure followed by the U.S. Census Bureau in the 2000 Census and subsequent American Community Surveys may result in a contamination of the same-sex couples group. To provide assurance that the categories are not contaminated by miscoding of sex or an incorrect relationship allocation, both partners' observations are dropped when either partner's sex, age, relationship to the householder, or marital status is allocated by the U.S. Census Bureau. Table A.2 emphasises the potential importance of this final restriction; excluding couples with allocated marital status significantly decreases the size of the lesbian sample while only marginally affecting the other subsamples. Finally, we exclude all individuals who were at no stage employed in the civilian labour force or reported zero wage/salary income in the previous year.[33] This exclusion entails removing approximately fifteen percent of lesbians, twenty percent of cohabiting heterosexuals, and thirty-two percent of married women from the sample. Our final sample from the 2000 Census consists of 12,204 lesbians, 163,834 cohabiting heterosexual women, and 243,299 married women. Table A.2, Column (6) contains analogous information for all three samples.

3.2. Summary Statistics

Table 3.1 provides summary statistics for lesbians, cohabiting heterosexuals and married women within each sample. In the interest of brevity, when discussing the summary statistics we focus predominantly on 1999 values, although we also highlight similarities and differences across samples.[34] Unless otherwise stated, the explanations in this section are 'on

[32] These samples are taken for each year prior to pooling datasets to ensure adequate representation across years.
[33] Those serving in the military and institutionalised individuals are therefore not part of the research population. Similarly to Klawitter (2011), we impose no further restrictions on class of worker as discrimination on the basis of sexual orientation or maternity risk may affect employment choices.
[34] The 2000 Census PUMS contains information for the 1999 calendar year, while the ACS PUMS contain information for the twelve months preceding survey completion. All real income figures account for this

Table 3.1. Descriptive statistics by sample and sexual orientation

	2000 Census			2005-2007 ACS			2008-2010 ACS		
Variables	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals
Work characteristics									
Annual real income	47,175	29,540	33,831	50,982	30,288	37,243	50,497	30,291	38,472
	(46,705)	(29,510)	(34,173)	(49,318)	(29,094)	(38,310)	(49,731)	(29,245)	(38,789)
Weeks worked	46.74	44.31	44.76	46.70	44.23	45.00	46.88	45.26	45.85
	(10.78)	(13.07)	(12.76)	(11.08)	(13.46)	(12.66)	(10.31)	(12.11)	(11.48)
Hours worked per week	41.47	38.77	36.93	41.29	38.31	36.83	40.52	37.64	36.69
	(10.17)	(9.82)	(11.21)	(10.71)	(10.08)	(11.49)	(10.43)	(10.08)	(11.41)
Personal characteristics									
Age	38.13	33.46	41.79	40.92	35.02	44.14	41.82	35.28	44.93
	(9.78)	(10.36)	(10.46)	(10.56)	(11.47)	(10.65)	(11.14)	(11.64)	(10.82)
Potential experience	18.54	15.42	23.22	21.02	16.67	25.13	21.86	16.77	25.77
	(9.60)	(10.59)	(10.84)	(10.38)	(11.77)	(11.09)	(10.94)	(12.03)	(11.35)
No high school	0.07	0.14	0.09	0.04	0.10	0.06	0.03	0.09	0.05
	(0.26)	(0.35)	(0.29)	(0.19)	(0.30)	(0.24)	(0.17)	(0.28)	(0.22)
High school	0.16	0.31	0.29	0.16	0.30	0.26	0.15	0.27	0.23
	(0.36)	(0.46)	(0.45)	(0.36)	(0.46)	(0.44)	(0.36)	(0.44)	(0.42)
Some college	0.24	0.28	0.24	0.22	0.26	0.22	0.23	0.28	0.22
	(0.43)	(0.45)	(0.43)	(0.41)	(0.44)	(0.41)	(0.42)	(0.45)	(0.42)
Associate	0.08	0.08	0.09	0.09	0.09	0.11	0.09	0.10	0.11
	(0.27)	(0.27)	(0.29)	(0.28)	(0.29)	(0.31)	(0.29)	(0.30)	(0.31)
Bachelor's	0.25	0.14	0.19	0.27	0.17	0.22	0.26	0.19	0.24
	(0.43)	(0.34)	(0.39)	(0.44)	(0.38)	(0.42)	(0.44)	(0.39)	(0.42)
Postgraduate	0.21	0.05	0.10	0.23	0.07	0.13	0.24	0.07	0.15
	(0.40)	(0.22)	(0.30)	(0.42)	(0.25)	(0.34)	(0.43)	(0.26)	(0.35)
Hispanic	0.08	0.11	0.08	0.09	0.12	0.09	0.10	0.13	0.09
	(0.27)	(0.31)	(0.28)	(0.28)	(0.33)	(0.28)	(0.30)	(0.34)	(0.29)
White	0.84	0.78	0.84	0.87	0.80	0.84	0.87	0.80	0.84
	(0.36)	(0.42)	(0.37)	(0.34)	(0.40)	(0.37)	(0.33)	(0.40)	(0.37)
Black	0.08	0.11	0.07	0.05	0.08	0.06	0.06	0.09	0.06
	(0.27)	(0.31)	(0.25)	(0.22)	(0.28)	(0.24)	(0.23)	(0.28)	(0.24)
Asian	0.01	0.02	0.04	0.02	0.03	0.05	0.02	0.03	0.05
	(0.11)	(0.14)	(0.19)	(0.13)	(0.16)	(0.21)	(0.12)	(0.16)	(0.22)

Table 3.1 (continued)

	2000 Census			2005-2007 ACS			2008-2010 ACS		
Variables	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals
Other race	0.05	0.07	0.04	0.04	0.07	0.04	0.03	0.06	0.03
	(0.21)	(0.25)	(0.21)	(0.20)	(0.25)	(0.20)	(0.18)	(0.23)	(0.18)
Mixed race	0.02	0.03	0.02	0.02	0.02	0.01	0.02	0.02	0.01
	(0.15)	(0.16)	(0.13)	(0.14)	(0.14)	(0.10)	(0.15)	(0.15)	(0.11)
Disabled	0.13	0.15	0.12	0.08	0.07	0.06	0.06	0.05	0.04
	(0.34)	(0.36)	(0.33)	(0.27)	(0.25)	(0.24)	(0.25)	(0.22)	(0.21)
English proficient	0.97	0.95	0.95	0.98	0.94	0.95	0.98	0.94	0.94
	(0.17)	(0.22)	(0.22)	(0.15)	(0.23)	(0.23)	(0.14)	(0.23)	(0.23)
U.S. citizen	0.99	0.98	0.97	0.99	0.97	0.97	0.99	0.97	0.97
	(0.12)	(0.15)	(0.17)	(0.10)	(0.17)	(0.17)	(0.09)	(0.17)	(0.17)
No. own children under 18	0.33	0.64	0.98	0.29	0.56	0.87	0.31	0.59	0.85
	(0.76)	(1.00)	(1.12)	(0.69)	(0.94)	(1.09)	(0.74)	(0.98)	(1.09)
U.S. residence									
Northeast	0.22	0.21	0.19	0.20	0.20	0.18	0.20	0.20	0.18
	(0.41)	(0.40)	(0.39)	(0.40)	(0.40)	(0.39)	(0.40)	(0.40)	(0.39)
Midwest	0.18	0.25	0.25	0.19	0.24	0.25	0.19	0.24	0.25
	(0.38)	(0.43)	(0.44)	(0.39)	(0.43)	(0.43)	(0.40)	(0.43)	(0.43)
South	0.30	0.31	0.35	0.32	0.32	0.36	0.34	0.32	0.36
	(0.46)	(0.46)	(0.48)	(0.47)	(0.47)	(0.48)	(0.47)	(0.47)	(0.48)
West	0.31	0.24	0.20	0.29	0.24	0.21	0.27	0.24	0.21
	(0.46)	(0.43)	(0.40)	(0.45)	(0.43)	(0.41)	(0.44)	(0.43)	(0.41)
Metropolitan residence	0.84	0.72	0.69	0.83	0.74	0.71	0.83	0.74	0.72
	(0.37)	(0.45)	(0.46)	(0.37)	(0.44)	(0.45)	(0.37)	(0.44)	(0.45)
Occupation									
Mgt. and professionals	0.49	0.28	0.40	0.52	0.31	0.45	0.52	0.33	0.47
	(0.50)	(0.45)	(0.49)	(0.50)	(0.46)	(0.50)	(0.50)	(0.47)	(0.50)
Service occupations	0.13	0.21	0.14	0.13	0.23	0.14	0.14	0.24	0.14
	(0.34)	(0.41)	(0.35)	(0.34)	(0.42)	(0.34)	(0.35)	(0.43)	(0.35)
Sales	0.24	0.38	0.36	0.25	0.36	0.34	0.25	0.35	0.32
	(0.43)	(0.48)	(0.48)	(0.43)	(0.48)	(0.47)	(0.43)	(0.48)	(0.47)
Construction	0.04	0.01	0.01	0.03	0.01	0.01	0.02	0.01	0.00
	(0.19)	(0.10)	(0.08)	(0.17)	(0.10)	(0.08)	(0.15)	(0.09)	(0.07)

Table 3.1 (continued)

	2000 Census			2005-2007 ACS			2008-2010 ACS		
Variables	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals
Production and transport.	0.10	0.12	0.09	0.07	0.09	0.06	0.07	0.07	0.05
	(0.29)	(0.32)	(0.28)	(0.25)	(0.28)	(0.24)	(0.25)	(0.26)	(0.23)
Industry									
Agriculture	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(0.07)	(0.08)	(0.09)	(0.07)	(0.09)	(0.09)	(0.07)	(0.09)	(0.09)
Construction	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	(0.16)	(0.13)	(0.13)	(0.15)	(0.13)	(0.13)	(0.13)	(0.12)	(0.12)
Manufacturing	0.10	0.13	0.11	0.08	0.09	0.08	0.08	0.08	0.08
	(0.30)	(0.33)	(0.31)	(0.27)	(0.29)	(0.27)	(0.27)	(0.27)	(0.26)
Wholesale trade	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	(0.17)	(0.16)	(0.15)	(0.16)	(0.15)	(0.15)	(0.14)	(0.14)	(0.14)
Retail trade	0.09	0.14	0.11	0.10	0.14	0.10	0.10	0.14	0.10
	(0.29)	(0.35)	(0.31)	(0.30)	(0.35)	(0.30)	(0.30)	(0.35)	(0.30)
Transport. and utilities	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.03	0.03
	(0.20)	(0.17)	(0.17)	(0.19)	(0.16)	(0.16)	(0.18)	(0.16)	(0.16)
Information	0.04	0.03	0.03	0.04	0.02	0.02	0.03	0.02	0.02
	(0.21)	(0.18)	(0.16)	(0.19)	(0.15)	(0.15)	(0.17)	(0.15)	(0.14)
Finance and insurance	0.06	0.08	0.09	0.07	0.08	0.09	0.07	0.08	0.09
	(0.24)	(0.27)	(0.28)	(0.26)	(0.28)	(0.29)	(0.25)	(0.27)	(0.28)
Professional and mgt.	0.11	0.09	0.08	0.11	0.10	0.09	0.11	0.10	0.09
	(0.31)	(0.29)	(0.27)	(0.31)	(0.30)	(0.28)	(0.32)	(0.30)	(0.29)
Education and health	0.30	0.23	0.36	0.34	0.27	0.39	0.34	0.29	0.41
	(0.46)	(0.42)	(0.48)	(0.47)	(0.44)	(0.49)	(0.47)	(0.45)	(0.49)
Arts and entertainment	0.08	0.13	0.06	0.07	0.13	0.06	0.08	0.14	0.06
	(0.27)	(0.34)	(0.24)	(0.26)	(0.34)	(0.23)	(0.27)	(0.34)	(0.23)
Other services	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	(0.20)	(0.19)	(0.20)	(0.19)	(0.19)	(0.20)	(0.19)	(0.19)	(0.20)
Public administration	0.06	0.04	0.05	0.07	0.04	0.05	0.07	0.04	0.05
	(0.24)	(0.19)	(0.22)	(0.25)	(0.19)	(0.22)	(0.26)	(0.20)	(0.22)
Number of observations	12,204	163,834	243,299	10,302	107,599	161,515	10,653	116,650	158,587

Notes: Standard errors in parantheses. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

average'. In 1999, the average real income of lesbians was $47,175, significantly more than any other category. Married women earned $33,831 per annum, while cohabiting heterosexuals received the lowest compensation, with average real income of $29,540. As real incomes have risen, this relative income gap has persisted, with lesbians earning the most, and cohabiting heterosexuals the least, within each sample. The statistics are also suggestive of greater labour-market commitment among lesbians compared to heterosexual women, with lesbians working longer hours each week and more weeks in the previous year.[35] Somewhat unsurprisingly, married women work fewer hours per week than cohabiting heterosexual women in each sample.

Personal characteristics vary noticeably between categories. In the 2000 Census sample, married women are the oldest at 41.79 years of age, while lesbians and cohabiting heterosexual women are significantly younger (38.13 and 33.46 years, respectively). Turning to the ACS samples reveals the same age ranking, although the average woman is considerably older than their Census sample counterpart. Educational attainment also differs across household types. Lesbians are the most highly educated, with twenty-five (twenty-one) percent of lesbians in the 1999 sample having a bachelor's (postgraduate) degree. The contrast with other women is marked, with only nineteen percent of cohabiting heterosexual women and twenty-nine percent of married heterosexual women obtaining at least a bachelor's degree. The situation at the opposite end of the education spectrum also highlights the significant disparity in educational attainment across categories. For instance, forty-five percent of cohabiting heterosexual women studied no further than high school, compared to only twenty-three percent of lesbians. Interestingly, in addition to an increase in overall education levels, the ACS data suggests a narrowing of the sexual-orientation education gap compared to both cohabiting and married heterosexual women. A comparison of potential experience levels largely mirrors the sexual-orientation age-patterning, with married women having the highest level of potential experience in each sample, followed by lesbians and cohabiting heterosexuals. Differences in potential experience thus do not appear to contribute to the unadjusted lesbian premium over married women.

---

difference. In Appendix D, Table D.1 we note how real income is calculated and provide the definition of each variable used in the analysis that follows.

[35] It is important to note that the values for weeks worked in the 2008-2010 sample are inferred using the mid-point of a categorical weeks variable. The descriptive statistics are therefore provided only for comparative purposes and the associated continuous variable is not used in subsequent regression analyses.

Regarding differences in ethnicity, cohabiting heterosexuals consist more heavily of Hispanic women than lesbians and married heterosexuals. In 1999, for example, eleven percent of cohabiting heterosexuals identified as being of Hispanic origin, relative to eight percent of lesbians and married women. Racial dissimilarities are also apparent, with an appreciably larger portion of lesbians and married women being white relative to cohabiting heterosexuals. Furthermore, in all three samples, married heterosexuals contain relatively more Asian women (four to five percent) than do lesbians (one to two percent) and cohabiting heterosexuals (two to three percent).

Another notable attribute highlighted by Table 3.1 is that disability prevalence among partnered women appears to be falling over time. Minor differences in questions regarding disability status across census/ACS years may, however, be responsible for this apparent change.[36] Nevertheless, a common feature among the samples is that married women are less likely to be disabled than other women, but such differences are relatively small. Lesbians also exhibit greater proficiency in English and are more likely to be U.S. citizens compared to cohabiting and married heterosexual women. In addition, the average number of children residing in the household differs across subsamples. Households occupied by married couples have the highest average number of children present (0.98 in 1999), followed by cohabiting heterosexuals (0.64). Lesbian households contain the fewest children (0.33), a finding which persists across samples.

In terms of regional grouping, married women exhibit a strong preference towards residing in the South. Lesbians, on the other hand, are relatively more likely to locate in the West of the U.S. and in metropolitan areas than other women, and display a relative aversion to living in the Midwest. In addition, the summary statistics are suggestive of occupational segregation between lesbians and other women. Lesbians are over-represented in construction and under-represented in sales positions. Four percent of the 1999 lesbian sample work in construction, compared with one percent of other women.[37] Moreover, only a quarter of the lesbians in each sample are employed in sales positions, compared with more than one-third of cohabiting and married heterosexual women. Cohabiting heterosexual women are also under-represented in management and professional occupations. Finally, the industry classifications display a strong tendency for cohabiting heterosexuals to undertake

---

[36] For example, additional disability questions were asked in the 2000 Census relative to subsequent American Community Surveys.

[37] Industry classifications and ACS data suggest a smaller sexual-orientation gap in this profession.

employment in retail trade and arts/entertainment, while married women are significantly more likely than other women to work in education and health. As the occupational categorisation suggests, lesbians are disproportionately employed in professional and management positions, which likely contributes to the lesbian earnings advantage.

3.3. Method

Evidently, lesbians earn significantly more income than heterosexual women. As the previous discussion highlights, however, differences in work and personal characteristics may be at least partially responsible for the observed lesbian premium. In this section, we present an analysis of the lesbian earnings premium in a similar framework to previous studies. Specifically, we test for the presence of such a premium in each of the three samples when controlling for observable individual characteristics, but ignoring forward-looking expectations of labour-market separation due to child-bearing and contextual factors which may affect income.

3.3.1. Base Analysis

In order to examine the lesbian pay gap, we perform comparative earnings analyses between lesbians and heterosexual women. Jepsen (2007) and Daneshvary et al. (2009) argue that confining the analysis to partnered lesbians and cohabiting heterosexuals leads to greater comparability due to the similarities between the two groups. That is, the women are of similar legal status, as well as having a domestic partner that can contribute to household tasks and share child-rearing responsibilities. However, the descriptive statistics in the previous section cast doubt on the assertion that cohabiting heterosexual women represent a better comparison group than married heterosexuals. Moreover, developments in the rights of same-sex couples, including legalisation of same-sex marriage in some states, have led to enhanced comparability between lesbians and married heterosexuals. As a result, we concurrently analyse the lesbian premium relative to cohabiting as well as married heterosexual women. To examine the effect of sexual orientation on earnings, we estimate earnings regressions of the following form:

$$\ln(R_i) = \boldsymbol{X}_i'\beta + \boldsymbol{Z}_i'\phi + Lesbian_i\gamma + \varepsilon_i \qquad (A)$$

where i indexes the observation, $\ln(R)$ is the natural log of real annual wage and salary income, and the vector $\boldsymbol{X}$ consists of work characteristics, including usual hours worked per week and five indicator variables for weeks worked in the previous year. $\boldsymbol{Z}$ is a vector of variables controlling for personal characteristics, region of residency, occupation and industry, the inclusion of which depends on the model to be estimated. More detail is provided below regarding the specification of individual regression equations. *Lesbian* is an indicator variable equal to one if the individual is a female cohabiting with a same-sex partner, and zero otherwise. We assume the error term, ε, is heteroskedastic but otherwise well-behaved.

It is important to note that the regressors contained in $\boldsymbol{Z}$, such as education, residential location and the presence of children are, to varying degrees, endogenous. As previously discussed, incrementally adding progressively more endogenous covariates to the regression models and presenting a range of premium estimates allows us to reduce concerns associated with the 'over-controlling' problem. Model (A1) represents the base regression model which includes the vector $\boldsymbol{X}$ of work characteristics and the lesbian indicator dummy.[38] Specification (A2) further controls for potential experience (defined as age - years of education - 5) and its square, an indicator for being of Hispanic origin, four dummies for race, as well as indicators for being proficient in English, a U.S. citizen, or disabled.[39] Model (A3) includes five additional indicators for highest educational attainment, while (A4) further adds three indicators for region of residency and a dummy variable for residing in a metropolitan area. Specification (A5) adds a control for the number of children under eighteen residing in the household and (A6), which represents the full-control model, includes four indicators for occupation and twelve for industry. As in subsequent analyses, we estimate each model separately with cohabiting and married heterosexual women as the comparison group, for each sample, and account for heteroskedasticity by computing robust standard errors in all regression models.

3.3.2. Robustness Checks

An important issue that must be addressed is that of comparability with previous studies. As is evident from Table A.1 and the discussion in Section 2, previous studies estimating the sexual-orientation pay gap have generally used a different dependent variable to that used

---

[38] For both pooled ACS samples, we also include two year dummies in all regression models throughout this thesis to account for year fixed-effects. For ease of exposition, these are excluded from the written equation specifications.

[39] As previously noted, the definitions of all discussed variables are provided in Appendix D, Table D.1.

herein. For example, past analyses of U.S. Census data have commonly analysed sexual-orientation differences in hourly wages, while GSS studies focus on the annual income gap among full-time employed women, but do not control for hours or weeks of work. To ensure that any estimated lesbian earnings advantage is not a by-product of the chosen specification, we re-estimate the above models with minor changes for two separate cases. In the first robustness check, we drop the vector of work characteristics and use the log of real hourly wages as the dependent variable.[40] This entails estimating regressions of the following form:

$$\ln(W_i) = \boldsymbol{Z}_i^{'}\phi + Lesbian_i\gamma + \varepsilon_i \qquad (B)$$

where $\ln(W)$ is the natural log of the real hourly wage and $\boldsymbol{Z}$ is as previously specified in each model. As a second robustness check, we drop the vector of work characteristics and restrict the analysis to women working full-time.[41] This can be written as:

$$\ln(R_i) = \boldsymbol{Z}_i^{'}\phi + Lesbian_i\gamma + \varepsilon_i \qquad (C)$$

where all variables are as previously defined and only full-time workers are included.

Consideration must also be given to the possibility of sample-selection bias due to wage and salary income being unobservable for unemployed individuals. Specifically, if unobservable characteristics affect both the likelihood of employment and subsequent compensation rates, then the zero-conditional-mean assumption will not hold and OLS will produce inconsistent estimates. As we are interested in assessing the partial effect of sexual orientation on earnings at the population level, the use of a non-random sample, in which we only observe the earnings of employed individuals, may result in inaccurate inference. Table 3.2 provides suggestive evidence that selectivity may be an important factor in driving apparent earnings differentials, with employment rates among lesbians exceeding those of cohabiting and married heterosexual women by four to six percentage points, and sixteen to seventeen percentage points, respectively.[42] To account for this possibility, we implement the Heckman two-step procedure (Heckman, 1979). In the first step, the probability of engaging in employment in the previous year is modelled using probit estimation, and the Inverse Mills

---

[40] This, however, precludes use of the 2008-2010 sample due to unavailability of a continuous measure of weeks worked.

[41] To minimise unobservable heterogeneity, we define full-time workers as individuals working at least thirty-four hours per week for a minimum of forty weeks in the previous year.

[42] Rates are calculated from Table A.2. Sample sizes in column (5) represent women eligible to participate in the labour force, while column (6) represents women employed at some stage in the previous year. The ratio of the two provides a measure of the employment rate for each subsample.

Table 3.2. Proportion of women employed by sample and household type

	2000 Census	2005-2007 ACS	2008-2010 ACS
Lesbians	0.8535	0.8421	0.8405
Cohabiting heterosexuals	0.8135	0.7984	0.7796
Married heterosexuals	0.6840	0.6820	0.6808

Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Ratio (IMR) is generated.[43] Generally, the factors included in the selection equation include those from the earnings equation, plus additional identifying variables, or exclusion restrictions, which do not appear in the earnings equation. In the second step, the IMR is included as an additional explanatory variable in the earnings equation.

Although the Heckman procedure controls for sample-selectivity, it can result in severe multicollinearity if sufficient exclusion restrictions are not defined. To mitigate this problem, the selection equation contains appropriate variables from the earnings regressions, as well as two identifying variables:

- The real income of the female's partner, to account for the reduction in labour supply associated with the income effect.
- An indicator for the presence of children under six in the household. Although number of children should capture the erosion of human capital due to prior labour-market absences, the presence of children under six may matter more in determining ability to seek employment.

The probability of employment is thus modelled as a function of the characteristics in $\boldsymbol{Z}$ included in the earnings equation, the lesbian indicator dummy, partner income and the presence of young children.[44] Each model from the base regressions defined in Equation A is therefore re-estimated in a two-step fashion as discussed above, which we specify as:

$$\ln(R_i) = \boldsymbol{X}_i^{'}\beta + \boldsymbol{Z}_i^{'}\phi + Lesbian_i\gamma + IMR_i\eta + \varepsilon_i \qquad (D)$$

where all variables are as previously defined and the IMR is calculated from an auxiliary regression, as specified above.

[43] The IMR is given by the ratio of the probability density function to the cumulative distribution function of a distribution and is also known as the hazard rate.
[44] Work characteristics in **X** are, for obvious reasons, not included in the selection equation. In model (A6), we also exclude occupation and industry controls from the selection equation.

3.4. Results

3.4.1. Base Results

Table A.3 reports the results of the earnings comparison between lesbians and cohabiting heterosexual women based on the 2000 Census. Although we are primarily interested in the coefficient on the lesbian indicator variable, two other features of the results warrant discussion. Firstly, the goodness-of-fit measure, R^2, shows that even the most sparse model is able to explain over half of the variation in logged annual income. Including additional covariates significantly improves the explanatory power of the model, with the full-control model explaining more than sixty-three percent of the variation in logged real income. Second, the estimated coefficients on almost all covariates are both statistically significant and of the expected sign. Importantly, potential experience is estimated to have a positive but diminishing effect on income, representative of the commonly observed arc-shaped earnings profile. Residing in a metropolitan area, being Asian, a U.S. citizen, and fluent in English are all associated with higher income. Interestingly, being of Hispanic origin or black does not adversely affect compensation in the majority of models, while being classified as "other race" or "mixed race" does. As expected, higher educational attainment leads to progressively greater income in all reported models. Finally, having a disability or additional children exerts a negative influence on wage and salary income, all else held constant.[45]

Most importantly for the purposes of this paper, lesbians are predicted to earn significantly higher incomes than cohabiting heterosexual women, a finding that is robust to the inclusion of a range of observable characteristics. Holding only work characteristics constant, for example, lesbians earn a logged real income difference of 0.3006 over cohabiting heterosexuals, representing a premium of 35.07 percent.[46] Controlling further for potential experience, ethnicity, race, English proficiency, U.S. citizenship and disability status reduces the premium to 28.93 percent. The results from Model (A3) demonstrate the critical importance of controlling for education in earnings comparisons, with the addition of the educational-attainment dummies decreasing the estimated earnings advantage to 8.20

[45] Although not reported, the estimated effects of work characteristics conform to prior expectations and are highly significant.

[46] From the specification in Column 1: $\ln(R_i) = \boldsymbol{X}_i'\beta + (\text{Lesbian})\gamma + \varepsilon_i$, the lesbian premium is estimated as $\exp^{(\gamma)} - 1$, or in this case $\exp^{(0.3006)} - 1$ = 35.07 percent. This conversion is used extensively throughout the discussion.

percent.[47] The sequential addition of further controls results in progressively smaller estimates of the lesbian pay gap, with the full-control model suggesting a lesbian premium of 5.55 percent relative to cohabiting heterosexual women.

Table A.4 similarly presents the findings from the 2000 Census earnings comparison between lesbians and married women. Evidently, the results are highly consistent with the cohabiting heterosexuals analysis, with the most noticeable difference (other than differences in the estimated premium) being in the estimated effects of ethnicity and race. Specifically, being Hispanic is now consistently associated with lower earnings, while being black is estimated to have a positive earnings effect in most specifications. Although this contradicts the well-known racial earnings gap, it is likely that selection into marriage causes this seemingly counter-intuitive result. A crucial similarity between the results, however, is that lesbians are once again estimated to earn a premium over otherwise-similar heterosexual women. Model (A1) produces a lesbian premium estimate of 14.43 percent over married women. In contrast to the previous analysis, differences in potential experience, ethnicity, race, English proficiency, U.S. citizenship and disability status exert minimal influence on the income gap, with the premium rising to 14.95 percent upon their inclusion. Holding educational attainment constant again significantly reduces the premium, resulting in an estimated 6.48 percent lesbian advantage. Unlike the cohabiting heterosexuals comparison, the estimated lesbian premium is lowest in (A5), at 1.32 percent, including all controls except industry and occupation. The full-control model suggests occupation and industry asymmetries tend to favour married women, with the lesbian advantage rising to 2.21 percent upon their inclusion.

Finally, we present summarised results for all three samples in Tables 3.3 and 3.4.[48] Given that the focus of this section is on estimating the magnitude of the lesbian pay gap, we omit a detailed discussion of all variables except the lesbian indicator, noting that the important coefficients are significant and of the appropriate sign in almost every case. Examining Table 3.3 reveals a similar pattern across all three samples. Model (A1), which

---

[47] As previously discussed, educational attainment is, to an extent, endogenous. Including it in the regression may thus under-estimate the sexual-orientation earnings effect. In the remainder of this paper we favour over-controlling to generate conservative estimates of the lesbian premium as opposed to omitting key variables and reporting upward-biased estimates.

[48] Detailed regression results for the 2005-2007 and 2008-2010 ACS samples are provided in Appendix A, Tables A.5-A.8. Due to the large number of regressions estimated, brevity considerations allow only the coefficients of interest to be reported in much of this thesis. As this is not ideal, we provide an expanded set of coefficient estimates when it is deemed reasonable to do so.

Table 3.3. Summary of lesbian premium estimates relative to cohabiting heterosexuals

Sample:		Regression model (A1)	(A2)	(A3)	(A4)	(A5)	(A6)
	Lesbian	0.3006***	0.2541***	0.0788***	0.0637***	0.0590***	0.0540***
2000		(0.0069)	(0.0067)	(0.0063)	(0.0063)	(0.0063)	(0.0062)
Census	Observations	176,038	176,038	176,038	176,038	176,038	176,038
	R^2	0.5352	0.5530	0.6097	0.6175	0.6181	0.6363
	Lesbian	0.3396***	0.2725***	0.0856***	0.0728***	0.0706***	0.0657***
2005-2007		(0.0079)	(0.0077)	(0.0072)	(0.0072)	(0.0072)	(0.0070)
ACS	Observations	117,901	117,901	117,901	117,901	117,901	117,901
	R^2	0.5644	0.5864	0.6448	0.6523	0.6525	0.6734
	Lesbian	0.3263***	0.2566***	0.0777***	0.0672***	0.0657***	0.0623***
2008-2010		(0.0076)	(0.0073)	(0.0068)	(0.0068)	(0.0068)	(0.0065)
ACS	Observations	127,303	127,303	127,303	127,303	127,303	127,303
	R^2	0.5740	0.5963	0.6584	0.6659	0.6660	0.6868

Notes: Robust standard errors in parentheses. Models include all variables specified in the text, but only the coefficient on the lesbian indicator is reported. *** represents statistical significance at the 1% level, in a one-tailed test. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

controls only for sexual orientation and work characteristics, produces a lesbian premium of 40.44 and 38.58 percent in the 2005-2007 and 2008-2010 samples, respectively. These estimates are comparable to the 35.07 percent advantage estimated from the 2000 Census. Moreover, the systematic addition of further controls results in a similar reduction in the premium across all three samples. Considering the values from the full-control model (A6), we obtain estimates of the lesbian premium over otherwise-similar cohabiting heterosexuals of 5.55, 6.79 and 6.43 percent in the three respective samples.

Table 3.4 reveals comparable cross-sample consistency in the sexual-orientation pay gap estimates relative to married women. For instance, the magnitude of the lesbian premium differs by no more than two percentage points across samples for each of Models (A3)-(A6). As in the 2000 Census analysis, controlling for personal characteristics other than education and number of children does little to lessen the earnings disparity in the ACS samples. Differences in educational attainment again appear to be one of the primary drivers of the raw compensation gap. Comparing the output for each full-control model yields lesbian-married women adjusted earnings differentials of 2.21, 3.51 and 3.86 percent, respectively. These estimates are all statistically significant and, although they are not the smallest estimates of all

Table 3.4. Summary of lesbian premium estimates relative to married heterosexuals

Sample:		Regression model (A1)	(A2)	(A3)	(A4)	(A5)	(A6)
2000 Census	Lesbian	0.1348***	0.1393***	0.0628***	0.0273***	0.0131**	0.0219***
		(0.0068)	(0.0067)	(0.0063)	(0.0062)	(0.0063)	(0.0062)
	Observations	255,503	255,503	255,503	255,503	255,503	255,503
	R^2	0.5343	0.5461	0.6015	0.6103	0.6106	0.6306
2005-2007 ACS	Lesbian	0.1267***	0.1226***	0.0601***	0.0287***	0.0212***	0.0345***
		(0.0078)	(0.0077)	(0.0071)	(0.0071)	(0.0071)	(0.0070)
	Observations	171,817	171,817	171,817	171,817	171,817	171,817
	R^2	0.5439	0.5591	0.6170	0.6251	0.6252	0.6499
2008-2010 ACS	Lesbian	0.1002***	0.0984***	0.0518***	0.0234***	0.0225***	0.0379***
		(0.0075)	(0.0074)	(0.0067)	(0.0067)	(0.0068)	(0.0066)
	Observations	169,240	169,240	169,240	169,240	169,240	169,240
	R^2	0.5583	0.5737	0.6294	0.6373	0.6373	0.6634

Notes: Robust standard errors in parentheses. Models include all variables specified in the text, but only the coefficient on the lesbian indicator is reported. *** represents statistical significance at the 1% level, in a one-tailed test. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

specifications, represent conservative estimates of the premium due to the inclusion of partially endogenous variables in the equation specification.[49]

3.4.2. Results of Robustness Tests

The previous analysis provides consistent evidence of a lesbian earnings advantage, with a significant premium estimated across a variety of samples and specifications. Ensuring the robustness of this result to the choice of alternative dependent variables will substantially enhance the validity of the findings. Table A.9, Panels A and B, report the results of the comparative real wage analyses with cohabiting and married heterosexuals as the comparison groups, respectively. Although the estimates reported are slightly larger than those from the base specifications, the observed change in the premium as more covariates are added is highly similar for both groups. Model (B6) suggests that lesbians receive 5.92 to 7.62 percent higher real wages than cohabiting heterosexual women, while analogous estimates relative to married women imply a real wage premium of 3.08 to 4.74 percent. Comparing the estimates obtained from Models (A6) and (B6) reveals only minor differences, with the premium

[49] As noted later, differences in unobservable human capital likely augment the premium considerably. Our premium estimates from the full-control model are therefore conservative relative to alternative specifications, but may not be conservative when accounting for differences in human capital accumulation.

estimates being within 1.50 percentage points in all cases. Moreover, the values obtained for all samples are similar in magnitude to findings in prior 2000 Census studies.[50]

Table A.10 presents analogous results for the lesbian income gap among full-time employed women, with no work-characteristic controls. As expected, the failure to control for variation in hours and weeks worked among women results in much larger estimates of the lesbian premium, despite restricting the sample to full-time workers. This specification yields a lesbian advantage over otherwise-similar cohabiting heterosexuals of between 7.93 and 10.25 percent in Models (C5) and (C6), exceeding even the largest comparable estimate from the baseline specification. Examining the lesbian pay gap over married women reveals a similar result. Despite these results being consistent with previous studies that rely on a comparable model specification, such as Baumle and Poston (2011) and Jepsen (2007), they further emphasise that failing to control for hours and weeks of work generates upward-biased estimates of the lesbian premium.

Finally, the results from the Heckman two-stage estimator are provided in Table A.11.[51] For all samples, the inclusion of the selection-correction dramatically reduces the estimated premium relative to cohabiting heterosexual women in Models (D1) and (D2). In all models controlling for educational attainment, however, there is almost no discrepancy between the OLS and two-step estimates. Considering the full-control model, the correction for sample-selectivity appears to have made little difference, with the premium estimate in each sample being within 0.50 percentage points of that implied by OLS.[52] In contrast, the Heckman two-step procedure generates significantly different estimates of the lesbian advantage over married heterosexuals. Notably, the estimates are at least twice as large as those produced by OLS in most cases, including in the full-control model. Taking these results at face value, the estimated lesbian premium relative to married heterosexual women in the respective samples is 7.77, 10.03 and 11.26 percent. This result is in stark contrast to the previous analyses, which persistently display a smaller lesbian advantage over married women, as opposed to cohabiting heterosexuals.

[50] See, for example, Daneshvary et al. (2008, 2009), Gates (2009) and Klawitter (2011) in Table A.1.

[51] As no analogous goodness-of-fit measure is available for the Heckman procedure, such as pseudo-R^2, we do not report on the goodness-of-fit of these models.

[52] Although the selection term appears to have only a minimal effect on the lesbian premium, the selection term is statistically significant in all regressions.

Regardless of the exact estimation method and specification chosen, there is substantial evidence of a lesbian premium in all three samples. For the remainder of the paper, in which we analyse one potential source of the observed premium, we focus primarily on models closely resembling the original specification. We do this for a number of reasons. First, the lack of a continuous weeks-worked variable in the 2008-2010 ACS sample renders the use of hourly wages impracticable.[53] Second, use of hourly wages arguably imposes more constraints on the effect of hours and weeks worked than does including them as controls in the annual income regressions. Third, as previously discussed, the use of logged annual income with no hours or weeks worked controls generates upward biased estimates due to variation in labour-force commitment, even among full-time workers. Finally, despite the Heckman two-step method being able to address sample-selectivity issues, we disregard its use as the primary estimation technique due to concerns surrounding model robustness. Specifically, volatility in the reported estimates and cross-sample inconsistency, potentially due to insufficient or inappropriate exclusion restrictions, may pose problems in assessing possible explanations of the lesbian premium. In unreported regressions, we show that the Heckman Full Information Maximum Likelihood estimator generates results mirroring those obtained via OLS, providing further justification for the primary use of OLS throughout the remainder of the analysis.

4. Maternity Risk and the Premium

4.1. Child-Bearing and Labour-Force Attachment

The results from the previous section strongly suggest the existence of an economically and statistically significant lesbian premium, even when controlling for a wide range of observable characteristics. Specifically, we have shown that lesbians earn approximately 60-70 percent and 30-40 percent more than cohabiting heterosexual women and married women, respectively, not holding other factors constant. When controlling for work and personal characteristics, region of residence, and occupation and industry, we obtain respective lesbian premium estimates of 6-7 percent and 2-4 percent. Considerable effort has been devoted to explaining the adjusted lesbian pay gap in previous studies, but a consensus view on its cause is yet to be reached. For the remainder of this paper we focus on one plausible factor which

[53] It is possible to construct a continuous measure of weeks worked by assigning each individual the value corresponding to the mid-point of their weeks worked category. This method was utilised for comparative purposes in Table 3.1, but its use in generating hourly wages may result in biased estimates.

may contribute to the observed earnings differential. Specifically, we argue that differences in maternity risk, or the likelihood of a woman temporarily leaving the labour market to give birth to a child, may be partially responsible for this differential.

Petit (2007) makes the compelling case that employers may selectively hire employees to reduce the expected costs of maternity leave.[54] That is, holding other factors constant, employers attempt to minimise pecuniary maternity benefits, as well as interruption and replacement costs associated with females leaving the workforce to give birth to or care for children. This reluctance to employ women with greater maternity risk thus results in a compensatory wage reduction to induce the employer to bear the higher associated costs. Moreover, if training and replacement costs are an increasing function of a position's salary, then maternity risk may also manifest itself in discriminatory promotion practices. As lesbians exhibit lower maternity risk than heterosexual women, a sexual-orientation earnings gap may result, even among women with identical levels of human capital.

We are not the first to hypothesise that differential fertility rates among women may exacerbate the lesbian premium. Badgett (2001), for example, argues that lesbians may be perceived as being similar to heterosexual men with respect to labour-force commitment and reluctance to leave their job to have children. Baumle and Poston (2011) and Elmslie and Tebaldi (2007) similarly argue that employers' perceptions regarding sexual-orientation differences in labour-force attachment may be a contributing factor to the observed lesbian premium. A number of other studies also point to differences between lesbians and heterosexuals in terms of labour-force commitment, although these studies are less explicit and tend to focus on differences in human capital accumulation resulting from past labour-market separation.

Despite several authors noting the importance of fertility in wage and salary determination, current research assessing the validity of the maternity-risk hypothesis is severely limited. To our knowledge, all previous empirical studies examining the effects of fertility on earnings and the lesbian pay gap have centred around the earnings effect of existing children, thereby neglecting the consideration of forward-looking child-bearing expectations. As a result, such studies fail to incorporate a vital component of the sexual-

[54] Following a similar method to Petit (2007), Baert (2013) assesses whether sexual-orientation hiring discrimination in Belgium is consistent with differential fertility. He finds weakly significant evidence of higher call-back rates among young lesbians relative to their heterosexual counterparts, consistent with the maternity-risk hypothesis.

orientation maternity differential. For example, a common method used in an attempt to assess whether differences in labour-market commitment affect the lesbian premium is to stratify the analysis by the presence of children.[55] It is argued that independence from family-related career interruptions should result in an appreciably smaller lesbian premium among childless women compared to that among women with children. This method has been used by Jepsen (2007), Blandford (2003) and Ahmed et al. (2013b), among others. In each study, the authors estimate a premium similar in magnitude for both subsamples, providing minimal support for their hypothesis.

The failure to obtain convincing support for this hypothesis is not surprising. Even among women who have not had children, we would expect to see differences in career progression due to discrimination on the basis of perceived maternity risk. This would result in the emergence of an earnings differential among women with otherwise-identical characteristics, despite their never having left the labour force on maternity leave. To further substantiate this claim, we provide results from re-estimating Model (A6) in Table A.12, stratifying the analysis by the presence or absence of children. The results from these supplementary models are mixed. Considering only the lesbian comparison with cohabiting heterosexuals, we observe a substantial decrease in the premium upon restricting the analysis to women without children in all samples, a result consistent with both theories. The comparison with married heterosexuals, however, shows a decidedly larger premium among the subsample of women with no children in two of the samples, contradicting Jepsen's and others' hypothesis. This result is, however, still consistent with discriminatory promotion policy on the basis of perceived maternity risk.

4.2. Forward-Looking Expectations and the Maternity-Risk Hypothesis

This thesis departs from previous studies by assessing whether forward-looking expectations of labour-market separation due to child-birth adversely affect earnings and thus the lesbian premium. As proposed by Phelps (1972) and Arrow (1973), the existence of incomplete information results in employers using perceived group characteristics in assessing the value of potential or current employees, with respect to their expected productivity and costs. As

[55] Other methods include interacting number of children with potential experience (Black et al., 2003), interacting the lesbian indicator dummy with potential experience (Badgett, 1995; Daneshvary, 2008) and allowing the effects of children to differ by sexual orientation (Elmslie and Tebaldi, 2007). A detailed assessment of these methods is beyond the scope of this paper; however, one commonality is their focus on backward-looking effects of motherhood on earnings.

maternity incidence is considerably higher among heterosexual women than lesbians, it is conceivable that employers form their expectations of maternity risk, and thus maternity-leave costs, explicitly taking sexual orientation into consideration. Lesbian women are, to a large extent, thus able to avoid the negative earnings effect of child-birth expectations, giving rise to the observed lesbian premium. To test the validity of this theory, we use within-sample maternity incidence as a measure of employers' perceived maternity risk. Including this perceived risk in earnings regressions will allow a direct assessment of the so-called maternity-risk hypothesis.

4.2.1 Allocated Maternity Risks and Applicability

The data for this section once again consists of the three samples from the 2000 Census, 2005-2007 ACS and 2008-2010 ACS, following an identical sample-selection procedure. Due to the way in which maternity risk is inferred in the present study, the initial analysis necessitates a further restriction of the sample to householders.[56] As the choice of which partner identifies as the householder is likely correlated with income, Table A.13 presents summary statistics for the householder subsamples in each dataset. Evidently, restricting the sample to householders produces a similar pattern to including all householders and residents.[57] This mitigates concerns that the subsequent analysis is not representative of all coupled females. In what follows, a woman is categorised as having had a child in the last year if she is the householder and at least one child under the age of one, within the household, is identified as her "natural born son/daughter".[58] As date of birth is removed for confidentiality purposes, we assume the woman's current age to be the age at which she gave birth to the child.

In order to assign maternity rates, each household type is split into seven age categories, resulting in a total of twenty-one groups. Within each group, actual maternity incidence in the relevant sample period is used as a measure of maternity risk. Table 4.1 presents the maternity risk for each sample and subgroup based on this allocation procedure. Comparing the assigned values suggests that they appropriately capture differential maternity

[56] The ACS data permit an alternative method to identify fertility which can be applied to householders and residents. As a robustness check, we present an analysis of the maternity-risk hypothesis using this identification method to ensure the results are robust to both the inclusion of residents and alternative employer perceptions.

[57] In Table A.14, we re-estimate Model (A6) for householders only. We find very similar estimates of the premium, with the largest divergence in any of the six comparisons being smaller than 0.60 percentage points.

[58] Adoption may also be important to the extent that women take time off work to care for adopted children. We also repeat the analysis including adopted children and obtain highly similar results.

Table 4.1. Allocated maternity risks: by sample and sexual orientation

Sample	Age Group	Lesbians	Cohabiting Heterosexuals	Married Heterosexuals
2000 Census	18-24	0.0828	0.1434	0.1996
	25-31	0.0374	0.0760	0.1267
	32-38	0.0261	0.0393	0.0781
	39-45	0.0086	0.0121	0.0158
	46-52	0.0000	0.0005	0.0011
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000
2005-2007 ACS	18-24	0.0369	0.1437	0.2011
	25-31	0.0285	0.0874	0.1611
	32-38	0.0345	0.0537	0.0908
	39-45	0.0199	0.0147	0.0177
	46-52	0.0048	0.0007	0.0007
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000
2008-2010 ACS	18-24	0.0251	0.1489	0.2029
	25-31	0.0254	0.0901	0.1612
	32-38	0.0194	0.0632	0.0944
	39-45	0.0122	0.0135	0.0171
	46-52	0.0007	0.0007	0.0007
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000

Notes: Values represent number of births per woman. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

risk across age groups and household types. Maternity risk declines with age, being negligible for women above the age of forty-five in all samples. In addition, married women exhibit the greatest degree of maternity risk, followed by cohabiting heterosexuals and then lesbians. Although the relative magnitude of maternity risk across age groups and household types conforms to expectations, the critical importance of the allocated rates to the analysis necessitates further discussion of their applicability.[59]

[59] Although the values discussed here reflect the one-year likelihood of maternity incidence, employers likely base their decisions on maternity costs over the expected lifetime of employment. To the extent that one-year maternity rates reflect relative levels of maternity risk over a longer time horizon, this difference should not pose problems. In unreported robustness checks, we assign maternity risk over three and five-year periods. As these assigned rates approximate linear multiples of one-year maternity risk, it is not surprising that the estimated effect of controlling for maternity risk on the lesbian premium is robust to this choice. As we would expect, the estimated effect of three and five-year maternity risk on earnings is commensurately smaller. For ease of exposition and due to uncertainty regarding the prior relationship and partnership status of individuals, we use one-year rates in the analysis that follows, noting that our control captures the essential elements of near-term (one-to-five year) maternity risk.

The use of an age-fixed-effect maternity risk variable, separated by household type, is motivated by the assumption that employers infer maternity risk based on broad age categories. That is, when assessing an employee's risk of requiring maternity leave, employers form their expectations based only on the individual's age group and perceived household type. This assumption may appear restrictive in that observed maternity rates likely differ by race, marital history and number of pre-existing children, among other factors. However, if information costs are high, it is plausible that employers base their maternity-leave expectations on a limited range of critical determinants, notably sexual orientation, marital status and age group. Moreover, the additional precision gained via further disaggregation may be more than offset by a reduction in accuracy brought about by a decrease in the group sizes from which the estimates are obtained. To reduce concerns that the results generated are heavily conditional upon the allocation procedure implemented, we consider alternative methods in Section 4.3.

Inferring maternity risk from the sample offers several potential advantages over the use of publicly available birth data. Importantly, the present method permits a disaggregation of maternity risk by household type. Published maternity rates do not allow such disaggregation due to the unavailability of information regarding the sexuality of a child's birth mother. Thus, the current method facilitates a direct analysis of the maternity-risk hypothesis that would otherwise be impracticable. Further, absent this issue, inference from the sample provides a more accurate estimate of the maternity risk applicable to women within the research population. Population-level maternity statistics are contaminated by individuals outside of the research population, such as non-partnered individuals and those less likely to be perceived as lesbians. Consequently, the magnitude and age-structure of maternity risk may differ between the two groups. Using sample rates circumvents these issues, enhancing the accuracy of the empirical analysis.[60]

Despite theoretical appeal, several important caveats apply to the use of sample maternity rates. Due to the pooled cross-sectional nature of the data, maternity-related entry and exit from the workforce cannot be captured by the maternity-risk allocation procedure. If

[60] Cross-sample variation in maternity risk among lesbians may lead to concern regarding which rates most accurately represent employers' forward-looking expectations. Given the differences in Census Bureau procedures implemented across samples, the groups are not directly comparable and the perceived likelihood of being a lesbian likely differs between samples. Provided employers take into consideration the perceived likelihood of an employee being a lesbian, rates disaggregated by sample are the most appropriate. In unreported regressions we pool all ACS years and re-run the allocation and estimation procedures. The results do not appear to be heavily dependent on the current methodology.

lesbians are more likely to return to work within a year of giving birth, then the assigned values will understate true differences in maternity risk among employed females. Another problem concerns employers' beliefs surrounding the transition of women between groups. The current allocation procedure assumes employers adopt the naive view that women will not transit between household types in the near-term. However, if employers apply non-zero probability beliefs, perceived maternity risk will differ from actual maternity incidence. In such a case, use of the allocated rates will lead to biased estimates of the earnings effect of maternity risk and its corresponding impact on the lesbian premium. Finally, if employers' perceptions of sexual orientation and partnership status are wildly incorrect, then the analysis herein may be highly inaccurate. The sample restrictions implemented should mitigate these concerns by excluding those individuals whose sexual orientation is unlikely to be accurately assessed by an employer. However, similarly to previous studies on the sexual-orientation pay gap, the following analysis is still heavily dependent on the assumption that employers can accurately perceive sexual orientation and partnership status of individuals in the sample.

4.3. Method

4.3.1. Base Analysis

In order to assess the maternity-risk hypothesis, we perform regression analyses of a similar form to Model (A6).[61] At this point, however, we note a significant concern regarding the analysis of Section 3 that poses difficulties in all comparative wage analyses. Specifically, fewer career interruptions and more life-time hours worked result in lesbians having greater levels of on-the-job experience than heterosexual women for any given level of potential experience. As a result, returns to potential experience should be higher among lesbians. Failing to account for such differences may then lead to a spurious finding whereby lesbians are estimated to earn a premium over heterosexual women, which actually reflects unobservable differences in human capital accumulation.[62] As previously noted, employer

[61] Problems in estimating the pooled-model coefficients for the generalised Blinder-Oaxaca decomposition (caused by unobservable heterogeneity) preclude the use of such decompositions in our analysis. We also reject the use of group-specific coefficients or a weighted matrix as the non-discriminatory earnings structure due to the (unreported) results being relatively sensitive to the chosen weights. A set of decompositions using heterosexual weights is, however, estimated later in the paper for the interested reader to examine.

[62] The same argument can also be applied regarding estimates of the gender pay gap.

perceptions of maternity risk likely result in discriminatory promotion practices, which could further exacerbate the greater return to potential experience received by lesbians.[63]

Failing to consider these differences can result in severe omitted-variables bias when attempting to ascertain the effect of maternity risk on earnings.[64] Specifically, positive bias in the estimated effect of potential experience on heterosexuals' earnings will generate an upward bias in the estimated effect of maternity risk.[65] Conversely, allowing the maternity-risk effect to differ by sexual orientation will generate a large downward bias on the interaction term. In this respect, negative bias in the estimated effect of potential experience on the earnings of lesbians will be captured by lower maternity risk, with large earnings increases being incorrectly attributed to small decreases in maternity risk. In order to obtain unbiased coefficient estimates, one must therefore attempt to account for unobservable differences in human capital accumulation.

A common method for accounting for differences in unobservable human capital is to allow observable human capital to have a differential effect by sexual orientation. However, this results in difficulties in interpretation, which is evident when observing the contrasting treatment of such interaction terms in previous studies. Badgett (1995) includes an interaction term between sexual orientation and the potential experience variable in her regressions. When interpreting the lesbian indicator dummy, she ignores the interaction term, essentially making the assumption that the slope of the interaction term purely captures differences in unobservable productivity. Jepsen (2007), who also includes interactions between educational attainment and the lesbian indicator, follows a similar approach. In contrast, Black et al. (2003) replicate Badgett's (1995) results and evaluate the interaction term at the mean of potential experience, assuming that the interaction term captures other unobservable factors as opposed to differences in human capital. Daneshvary et al. (2008) include interactions between the lesbian indicator and potential experience, its square, and education. They follow a similar approach to Black et al. (2003), extensively incorporating the coefficients of the

---

[63] If discriminatory promotion practices due to perceived maternity risk do in fact contribute to the greater returns to potential experience for lesbians, then this alone would provide sufficient evidence in favour of the maternity-risk hypothesis. This assertion is, unfortunately, not testable with the current data.

[64] As the prior effect of maternity risk is negatively correlated with current, forward-looking maternity risk, failing to control for the prior effect produces upward bias in the estimated maternity risk coefficient. The analysis provided within this section thus captures only the near-term, forward-looking effect of maternity risk on earnings.

[65] This bias will be larger when the bias in the potential experience coefficients is most severe and when heterosexual women comprise a larger proportion of the sample. We thus expect greater bias in the maternity risk coefficient when comparing the earnings of lesbians and married heterosexual women.

interaction terms into their analysis. Evidently, the interpretation of interaction terms is dependent on the effects they truly capture.

As unbiased estimation necessitates the inclusion of several interaction terms, the majority of the analysis in this section focuses on accurately estimating the effect of maternity risk on earnings. Assessing the corresponding effect on the lesbian premium is primarily done on a qualitative basis due to difficulties in interpretation, as explained above, although we initially attempt to make a direct assessment. In what follows, we estimate earnings regressions of the following form:

$$\ln(R_i) = \boldsymbol{X}_i'\beta + \boldsymbol{Z}_i'\phi + Lesbian_i\gamma + \boldsymbol{H}_i'\delta + \varepsilon_i \qquad (A)$$

where i indexes the observation, $\ln(R)$ is the natural log of real annual wage and salary income, and the vector $\boldsymbol{X}$ consists of work characteristics, including usual hours worked per week and five indicator variables for weeks worked in the previous year. $\boldsymbol{Z}$ is a vector of variables controlling for personal characteristics, region of residency, occupation and industry, as described above, with two changes: first, potential experience and its square are centred about their mean values, in line with the method used by Black et al. (2003); second, the full-control model is implemented in all regressions. Although this results in the inclusion of partially endogenous variables, and thus the potential for over-controlling, omitted-variables bias poses a far more serious problem in assessing the maternity-risk hypothesis. In the next section, we relax this constraint to ensure that the results are robust to alterations in the model specification. *Lesbian* is an indicator variable equal to one if the individual is a female cohabiting with a same-sex partner, and zero otherwise. Finally, $\boldsymbol{H}$ is a vector of variables whose inclusion depends on the model to be estimated and contains *Maternity Risk* as well as interactions between the lesbian indicator and the maternity risk variable, potential experience and its square.[66]

To demonstrate the need to include the human capital interaction terms, we run preliminary regressions with the 2000 Census data, noting that the results are broadly representative of all three samples. Model (A1p) represents the base regression which includes the vectors $\boldsymbol{X}$ and $\boldsymbol{Z}$, as well as the lesbian indicator. Model (A2p) adds *Maternity Risk* to the previous specification, while Model (A3p) relaxes the assumption that the effect of

[66] As previously noted, for both ACS samples, we also include two year dummies in all regression models to account for year fixed-effects.

maternity risk is the same across household types by introducing an interaction term, *Maternity Risk* × *Lesbian.* As in previous analyses, all models are estimated separately with cohabiting and married heterosexual women as the comparison group. For comparative purposes, we then estimate Model (A2p), excluding the lesbian indicator, for each of the three household types. These serve to provide evidence that *Maternity Risk* may be picking up the effects of unobservable human capital accumulation or maternity-risk promotion effects in the comparative earnings analyses. Specifically, if the coefficient on *Maternity Risk* is negative in each regression and lesbians receive greater returns to potential experience than heterosexual women, any positive *Maternity Risk* coefficients in the comparative earnings analyses likely arise due to omitted-variables bias.

As a precursor to the results, the above regressions display strong evidence of omitted-variables bias, reaffirming the need to adequately control for human capital differences. The following specifications seek to address this deficiency. Model (A1) is identical to Model (A2p), which includes the vectors ***X*** and ***Z***, in addition to *Lesbian* and *Maternity Risk*. Specification (A2) adds *Potential Experience* × *Lesbian*, while (A3) further includes *Potential Experience Squared* × *Lesbian*. These specifications allow us to examine whether maternity risk negatively influences earnings, but they do not permit an assessment of how the inclusion of *Maternity Risk* affects the lesbian premium. As such, Models (A1) to (A3) are re-estimated, dropping *Maternity Risk* so as to obtain an estimate of the premium both before and after its inclusion.[67]

4.3.2. Robustness Tests: Alternative Specifications

In conducting robustness tests, Model (A3) is taken as the base model as it limits concerns regarding unobservable human capital accumulation influencing the results. To ensure that the findings are not reliant on the choice of dependent variable, we re-estimate Model (A3) restricting the analysis to full-time workers and excluding controls for hours and weeks of work. In addition, we re-run the regression with hourly wages as the dependent variable. As Section 3 highlighted, selection into paid employment may be of importance in obtaining unbiased estimates. We therefore estimate a Heckman two-step model, where the selection

[67] The inclusion of education interaction terms to further control for human capital, and *Maternity Risk* × *Lesbian* to allow maternity risk to have a differential effect across household types, was also assessed. On the basis of F-tests and the Akaike Information Criterion, these specifications were deemed inferior to those included above. The inclusion of education interaction terms is thus relegated to its use in robustness tests, while Appendix E contains a note on the inclusion of a *Maternity Risk* × *Lesbian* interaction term.

equation includes the vectors ***Z*** and ***H***, but omits occupation and industry controls. To identify the system, the selection equation also includes an indicator for the presence of children under the age of six and partner's real income. In unreported regressions, we also estimated models using functional form to identify the selection equation, which produced results similar to those reported.

Following Jepsen (2007), we further control for differences in human capital accumulation by including the lesbian indicator interacted with educational-attainment dummies. As a further robustness test, we then include interactions between *Lesbian* and each of the controls in ***Z***. Provided the original estimates were only biased due to unobservable differences in human capital, the additional interactions should have a minimal effect on the coefficient estimates. As discrimination on the basis of maternity risk may affect occupational sorting, there is a concern that the previous models mask some of the maternity-risk effect due to an over-controlling problem. The next specification addresses this possibility, dropping occupation and industry controls from the base regression. It is also plausible that any observed disadvantage attributed to maternity risk may actually capture the adjustment period associated with new mothers re-entering the workforce, as identified by Anderson et al. (2003). To offset such concerns and isolate the effect of employer perceptions, we re-estimate the base model, excluding all women who gave birth during the previous year. If the maternity-risk disadvantage remains following this restriction, it will provide strong evidence in favour of the maternity-risk hypothesis.

Next, to assess whether the effects of maternity risk differ by employment status, we restrict the sample to full-time workers, which also reduces unobservable heterogeneity. If maternity benefits and replacement costs are more than proportionately higher among full-time workers, we would expect to see a larger negative coefficient on *Maternity Risk* among the sample of full-time employed females.[68] Finally, we assess whether maternity risk has a differential effect on earnings at different points on the income distribution. If maternity costs are largely variable, for example if they are primarily comprised of maternity benefits and training costs, one might expect maternity risk to exert a stronger negative influence at higher

[68] The Family and Medical Leave Act of 1993 applies only to employees meeting certain criteria, including having worked at least 1,250 hours over a twelve month period for a qualified employer (U.S. Department of Labor, n.d.). As many part-time workers in the sample do not meet this criterion, it is reasonable to expect that employers bear greater maternity-leave costs for full-time workers relative to part-time workers.

points on the income distribution.[69] We therefore estimate quantile regressions to examine the effect of maternity risk at different points on the income distribution (10th, 50th, and 90th percentiles).

4.3.3. Robustness Tests: Full Re-estimation

As a final set of robustness checks available for all three samples, we restrict the sample, re-allocate maternity rates, and estimate the base regression model for two separate sub-groups. First, to reduce unobservable heterogeneity and minimise concern regarding differences in racial composition across the twenty-one groups, we re-run the analysis while restricting the sample to white women. Next, we re-estimate Model (A3) having restricted the sample to never-married or currently married individuals.[70] These analyses serve to assess whether disaggregating the maternity-risk allocation process heavily affects the findings. Moreover, women within these subsamples may be more alike with respect to their marital status, extent of family-formation completion, and perceived sexual orientation, enhancing the accuracy of the analysis.

Two additional variables are available in the ACS PUMS which enable further robustness checks to be carried out on the ACS samples. The first indicates if the respondent lived at their current residence twelve months prior to the survey being conducted,[71] while the second captures whether or not a woman gave birth to a child in the previous year. Use of the mobility variable serves to moderate concerns surrounding whether a child born in the previous year was conceived while the women was with her current partner, as well as to restrict the sample to those individuals whose relationship status and sexual orientation are likely to be correctly ascertained by their employers. For example, children born to lesbian couples in the past year may be a result of a previous heterosexual relationship.[72] If neither partner has moved in the prior twelve months, however, then this increases the likelihood that

[69] To further motivate this analysis, we follow Ahmed et al. (2013b) and compute the percentage difference in compensation between lesbians and heterosexual women at each percentile of both the earnings and wage distributions. The resulting figures, which are not reported, display evidence of a larger (raw) premium among higher earning women.
[70] Lack of a suitably large sample precludes similar analyses for non-whites and previously married individuals. A stratification by educational attainment is omitted due to similar concerns.
[71] The 2000 Census also contained a mobility question, however mobility was measured on a five-year basis, rendering it unusable for the current analysis.
[72] Moore and Stambolis-Ruhstorfer (2013), among others, note the importance of prior unions on lesbian child-rearing. Our use of one-year maternity incidence, as opposed to three or five-year incidence, should significantly reduce the possibility that the assigned rates include children from prior unions. The robustness check we implement next serves to further mitigate such concerns.

the child was conceived by the lesbian couple and likely enhances the probability of employers correctly inferring sexual orientation. The fertility variable circumvents the issue of lesbian householders identifying their partner's child as an "own child", resulting in a false positive in the previous analysis.[73] In addition, the previous techniques can be applied to both householders and residents given the availability of fertility information for all women in the sample, thereby enhancing the applicability of results to the coupled-female population.

As a further test of method robustness, we therefore exclude both partners' observations if either partner moved in the last twelve months, and reassign maternity risk across the twenty-one groups. Model (A3) is then re-estimated for each sample relative to both comparison groups. Similarly to the previous analysis, we then drop all females who actually gave birth in the past year and re-run the regressions. Finally, we follow an identical procedure on the full sample, using the fertility variable contained within the PUMS files to assign maternity rates. The results from this analysis also allow us to assess whether the results generalise to both householders and residents.[74]

4.4. Results

4.4.1. Base Analysis

Table A.15 displays the preliminary regression results for the 2000 Census. In the absence of a control for maternity risk, the earnings differential between lesbian and cohabiting heterosexual householders is estimated to be 5.56 percent. When controlling for maternity risk, this estimate drops to 5.16 percent, reducing the lesbian premium by approximately seven percent. The estimates also suggest that a one percentage point increase in maternity risk reduces income by approximately 0.25 percent. These findings are consistent with the notion that maternity risk has a significant adverse effect on income and thus on the lesbian premium. Model (A3p), which relaxes the assumption that the effect of maternity risk is the same across household types, suggests that an increase in maternity risk by one percentage point reduces the earnings of cohabiting heterosexuals and lesbians by 0.51 and 2.65 percent, respectively. This result makes little intuitive sense and suggests *Maternity Risk* and its

[73] Comparing inferred fertility for householders using the two allocation methods (not reported) suggests both false positives and false negatives occur in the base analysis. However, the resulting errors in inferred maternity risk are relatively small and should not significantly bias the base results.
[74] In unreported robustness checks we also assess the effects of stratifying the results by sector and implementing minimum and maximum income restrictions. The results from these specifications were highly similar to those reported.

interaction with *Lesbian* may be picking up the effects of unobservable human capital accumulation, as expected.[75]

Similar estimates relative to married heterosexuals display more persuasive evidence to this end, with the lesbian premium rising from 2.29 to 4.27 percent upon the inclusion of *Maternity Risk*. This outcome is a direct consequence of the estimated positive and statistically significant earnings effect of maternity risk, indicating 0.49 percent higher earnings for each one percentage point increase in maternity risk. Including the interaction term generates wildly different results, with a one percentage point increase in maternity risk implying a reduction in the earnings of married heterosexual and lesbian women by 0.14 and 2.96 percent, respectively. None of these estimates is representative of the coefficients from the individual earnings regressions, suggesting that unobservable heterogeneity and multicollinearity problems may be hampering the accuracy of the analysis.[76] Further reinforcing this possibility is that the estimated effect of potential experience is highly sensitive to the inclusion of *Maternity Risk* and its interaction with the lesbian indicator. Although it is likely that the estimated effect of potential experience implicitly captures some of the maternity-risk effect prior to its explicit inclusion, the significant changes in the coefficients are suggestive of the partial attribution of unobservable effects to maternity risk. Bias in the simple equation specification thus poses potentially severe problems.

Table 4.2 displays the coefficient estimates on *Maternity Risk* in Models (A1) to (A3) across all three samples. As expected, the simple model specification, (A1), does not provide support for the maternity-risk hypothesis. Relative to cohabiting heterosexuals, only the 2000 Census estimate is both statistically significant and of the correct sign. Considering the lesbians versus married heterosexuals comparison yields positive and significant estimates of the earnings effect of maternity risk, in strict disagreement with the maternity-risk hypothesis.

---

[75] This result is also consistent with age-varying accuracy in employers' inferring the sexual orientation of employees. The magnitude of the effect, however, suggests the interaction term is likely capturing differences in unobservable human capital. For a more detailed discussion regarding the inclusion of the interaction term, see Appendix E.

[76] The reader may be concerned by two other features of Table A.15. First, including *Maternity Risk* only marginally improves goodness of fit. Tables A.3-A.8 show that this improvement in R^2 is in line with, or even larger than, the resulting improvement upon controlling for number of children, reducing concern. Second, the coefficient estimates from the individual regressions are much larger among women with lower maternity risk, suggesting the grouped allocation procedure may lead to *Maternity Risk* capturing misspecification in potential experience. In unreported regressions, we alternatively estimate maternity risk by sexual orientation, based on a probit model with a fifth order polynomial in age, to allow a flexible relationship between age and maternity risk. The estimated effect of maternity risk is larger in each case, but the qualitative results are unaffected. We also assess whether including a quartic in potential experience affects the findings. Our results suggest that the comparability of the estimates is enhanced in these specifications, further assuaging concerns.

Table 4.2. Assessing the maternity-risk hypothesis - coefficient on *Maternity Risk*

	Lesbians vs. Cohabs			Lesbians vs. Marrieds		
Sample:	(A1)	(A2)	(A3)	(A1)	(A2)	(A3)
2000 Census	-0.2494** (0.1376)	-0.4777*** (0.1448)	-0.6043*** (0.1462)	0.4889*** (0.1603)	-0.2227 (0.2075)	-0.5211*** (0.2152)
2005-2007 ACS	-0.0021 (0.1644)	-0.3594** (0.1864)	-0.5411*** (0.1936)	0.4721*** (0.1168)	-0.0454 (0.1452)	-0.2752** (0.1534)
2008-2010 ACS	0.1461 (0.1377)	-0.2826** (0.1560)	-0.4492*** (0.1606)	0.5076*** (0.1061)	-0.0881 (0.1325)	-0.3667*** (0.1406)

Notes: Robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficient on the maternity risk variable. ** and *** represent statistical significance at the 5% and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Specifically, across the three samples, an increase in maternity risk by one percentage point is associated with a 0.47 to 0.51 percent increase in earnings based on this specification. As discussed earlier, differences in human capital and discriminatory promotion practices on the basis of perceived maternity risk likely result in steeper earnings profiles for lesbians. The individual regressions in Table A.15 support this notion. As such, these differences must be accounted for when estimating the forward-looking effects of maternity risk.

Model (A2), which includes the interaction between the lesbian indicator and potential experience, demonstrates the importance of such controls, with the coefficient on *Maternity Risk* being of the correct sign in all comparative analyses. Specifically, in the cohabiting heterosexuals comparisons, a one percentage point increase in maternity risk is estimated to reduce earnings by between 0.28 and 0.48 percent. Similar estimates relative to married heterosexuals are statistically insignificant at conventional levels, but also suggest an earnings disadvantage among individuals with higher perceived maternity risk. Finally, the third specification allows unobservable sexual-orientation heterogeneity to be a non-linear function of potential experience. As hypothesised, the results from this specification present significant evidence in favour of the maternity-risk hypothesis, with all coefficient estimates being of the appropriate sign and statistically significant. In the case of the cohabiting heterosexuals comparison, the effect of a one percentage point increase in maternity risk is estimated to reduce income by between 0.45 and 0.60 percent. Comparable estimates relative to married women suggest an effect of 0.28 to 0.52 percent. This indicates that either the penalty associated with an increase in maternity risk is greater among women with a relatively lower level of risk, or that employers do not distinguish between married and cohabiting heterosexuals to the extent implied by our allocation procedure.[77] Regardless, allowing observable human capital to have a differential effect on lesbians appears to permit a more accurate assessment of the partial effect of maternity risk on earnings.

An important consideration regarding the estimates in Table 4.2 is how the lesbian premium is affected by the inclusion of *Maternity Risk*. Table 4.3, therefore, presents the lesbian premium estimates from Models (A1) to (A3) both including and excluding *Maternity Risk* to enable such an assessment. As previously discussed, the results from Model (A1) do not provide support for the maternity-risk hypothesis. In four of the six comparisons, the lesbian premium rises upon the inclusion of *Maternity Risk*, while the change in the premium relative to married heterosexuals represents more than a 50% increase. These observations are counter-intuitive and reflect the positive estimated effect of maternity risk, further demonstrating the importance of including sufficient human capital controls.

Conversely, considering Model (A2) reveals a reduction in the lesbian premium across all six specifications when evaluating the premium at the mean level of potential experience.[78] In the case of cohabiting heterosexuals, the lesbian pay gap drops from 4.48 to 3.43 percent in the 2000 Census sample, upon controlling for maternity risk. This represents a reduction of more than twenty-three percent at the mean level of potential experience. Analogous estimates for the 2005-2007 and 2008-2010 ACS samples imply a reduction in the pay gap of approximately seventeen and twenty-three percent, respectively. The comparison with married women reveals a similar pattern, with the lesbian premium falling from 4.51 to 3.90 percent in the 2000 Census sample, upon the inclusion of *Maternity Risk*. This represents a premium reduction of more than thirteen percent. Comparable estimates from the ACS samples display a reduction in the earnings gap of approximately two and four percent, respectively.

---

[77] Another possible explanation for this finding is that employers incorporate in perceived maternity risk the possibility of transition between categories. Alternatively, the difference may merely reflect statistical imprecision.

[78] This warrants a cautionary note. As previously explained, the literature is divided regarding how to interpret results when including sexual-orientation interaction terms. The discussion herein exclusively evaluates the results at the mean, primarily because this is a traditional convention applied in regression interpretation. This, however, may cause the change in the premium to be of the opposite sign to the maternity risk effect, especially if multicollinearity problems cause significant changes in the interaction term coefficients. Moreover, as much of the observed premium can likely be attributed to real differences, the estimated percentage change in the premium may be biased downwards.

Table 4.3. Assessing the maternity-risk hypothesis - effect on the lesbian premium

		Lesbians vs. Cohabs			Lesbians vs. Marrieds		
Sample:	Specification:	(A1)	(A2)	(A3)	(A1)	(A2)	(A3)
2000 Census	Excluding maternity risk	0.0541***	0.0438***	0.0646***	0.0226***	0.0441***	0.0772***
		(0.0086)	(0.0086)	(0.0104)	(0.0094)	(0.0104)	(0.0127)
	Including maternity risk	0.0503***	0.0337***	0.0556***	0.0418***	0.0383***	0.0690***
		(0.0089)	(0.0091)	(0.0106)	(0.0117)	(0.0117)	(0.0131)
2005-2007 ACS	Excluding maternity risk	0.0698***	0.0579***	0.0724***	0.0396***	0.0625***	0.0912***
		(0.0099)	(0.0103)	(0.0127)	(0.0100)	(0.0107)	(0.0129)
	Including maternity risk	0.0697***	0.0482***	0.0643***	0.0575***	0.0615***	0.0905***
		(0.0102)	(0.0113)	(0.0129)	(0.0110)	(0.0110)	(0.0129)
2008-2010 ACS	Excluding maternity risk	0.0657***	0.0471***	0.0698***	0.0436***	0.0663***	0.1076***
		(0.0092)	(0.0097)	(0.0118)	(0.0095)	(0.0100)	(0.0124)
	Including maternity risk	0.0691***	0.0365***	0.0578***	0.0649***	0.0639***	0.1050***
		(0.0099)	(0.0112)	(0.0125)	(0.0106)	(0.0106)	(0.0124)

Notes: Robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficient on the lesbian indicator. *** represents statistical significance at the 1% level, in a one-tailed test. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Finally, the third specification yields similar results, with all six comparisons displaying a reduction in the lesbian pay gap. First considering the cohabiting heterosexuals comparison for the 2000 Census, we observe a reduction in the premium from 6.67 to 5.72 percent upon controlling for maternity risk. This decrease is economically significant, representing a fall in the pay gap by approximately fourteen percent. As before, similar estimates from the ACS samples imply premium reductions from 7.51 to 6.64 and 7.23 to 5.95 percent, respectively. These specifications suggest that controlling for near-term maternity risk reduces the lesbian pay gap by approximately one percentage point. When considering that real differences in human capital likely augment the estimated premium considerably, these represent economically significant changes in the pay gap.

Despite the earnings effect of maternity risk in Model (A3) being estimated as highly negative and statistically significant in all married heterosexual comparisons (see Table 4.2), the change in the premium provides conflicting evidence regarding its effect on the lesbian premium. Specifically, across the three samples the lesbian premium falls by between one and eleven percent. Thus, despite maternity risk negatively affecting earnings, including *Maternity Risk* does not appear to produce an economically significant reduction in the lesbian premium relative to cohabiting heterosexuals in all samples, when evaluated at the mean of potential experience. Although these findings may appear contradictory, two possible explanations are offered. First, evaluating the premium at the mean of potential experience leaves the results susceptible to changes in the interaction-term coefficients; results are thus not indicative of an analogous change at all levels of potential experience. Second, it may not be overly surprising that the premium changes only marginally with the inclusion of *Maternity Risk* when allowing the effect of potential experience to vary non-linearly by sexual orientation. As maternity incidence is a non-linear, negative function of age, potential experience is likely implicitly controlling for the effect of maternity risk on the lesbian premium. Therefore, when explicitly accounting for differential maternity risk, *Maternity Risk* is observed to have a statistically significant negative earnings effect, but such an effect may have already been partialled out from the premium due to bias in the estimated effect of potential experience. The true effect of maternity risk on the lesbian premium is therefore difficult to ascertain, even in the absence of the other aforementioned issues, and is likely much larger than the discussion here suggests. Moreover, when omitting a control for maternity risk, the compensation effect of potential experience is biased upward due to the reinforcing effect of maternity risk.

4.4.2. Results from Alternative Specifications

Table 4.4 presents the results from assessing the maternity-risk hypothesis with alternative model specifications in the three datasets. The coefficient on *Maternity Risk* is reported for each regression, while the effect of maternity risk on the premium is analysed on a qualitative basis due to previously discussed difficulties. Considering only full-time workers but omitting work-characteristic controls significantly increases the estimated effect of maternity risk on earnings. In the lesbian versus cohabiting heterosexuals comparisons, a one percentage point increase in maternity risk is now associated with an earnings penalty of approximately 0.68 to 1.00 percent, while similar estimates when comparing lesbians with married women indicate a maternity-risk disadvantage of between 0.49 and 0.65 percent. In contrast, use of hourly wages as the dependent variable does not significantly affect the results, with the magnitude of the *Maternity Risk* coefficient increasing slightly in the 2000 Census sample, but falling in the 2005-2007 ACS sample. The level of statistical significance is, however, reduced in the ACS sample. Controlling for selection into paid employment via the Heckman two-step method appears to exert a slightly more pronounced effect on the findings. Within the two earlier samples, cross-cohort consistency in the estimates is significantly enhanced, with the *Maternity Risk* coefficients in the comparative earnings regressions becoming much more

Table 4.4. Coefficient on *Maternity Risk* - alternative specifications

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Married	Cohabs	Married	Cohabs	Married
Base results (Specification A3)	-0.6043***	-0.5211***	-0.5411***	-0.2752**	-0.4492***	-0.3667***
	(0.1462)	(0.2152)	(0.1936)	(0.1534)	(0.1606)	(0.1406)
Full-time, no work characteristic	-1.0001***	-0.6471***	-0.7204***	-0.4852***	-0.6799***	-0.6457***
controls	(0.1369)	(0.2007)	(0.1814)	(0.1427)	(0.1591)	(0.1392)
Log(hourly wages)	-0.6671***	-0.5719***	-0.4148**	-0.2171*	[redacted]	
	(0.1419)	(0.2033)	(0.1829)	(0.1441)		
Heckman two-step method	-0.5689***	-0.5911***	-0.4876***	-0.3769***	-0.1492	-0.4842***
	(0.1467)	(0.2143)	(0.1948)	(0.1538)	(0.1676)	(0.1415)
Including education interactions	-0.5870***	-0.4222**	-0.4763***	-0.2081*	-0.4280***	-0.3405***
	(0.1474)	(0.2177)	(0.1965)	(0.1551)	(0.1631)	(0.1424)
Full interaction model	-0.5880***	-0.4257**	-0.4773***	-0.2027*	-0.4263***	-0.3405***
	(0.1474)	(0.2178)	(0.1966)	(0.1552)	(0.1631)	(0.1424)
Dropping occupation and	-0.7629***	-0.7398***	-0.6661***	-0.4276***	-0.7312***	-0.4602***
industry controls	(0.1490)	(0.2167)	(0.1985)	(0.1580)	(0.1650)	(0.1449)
Excluding women who gave birth in	-0.5757***	-0.4152**	-0.5708***	-0.3083**	-0.3762**	-0.3664***
the previous year	(0.1507)	(0.2194)	(0.1998)	(0.1566)	(0.1649)	(0.1445)
Restricting to full-time employed	-0.9826***	-0.6971***	-0.7476***	-0.5551***	-0.8068***	-0.6481***
	(0.1370)	(0.2009)	(0.1845)	(0.1463)	(0.1618)	(0.1419)
Quantile regression (10th percentile)	-0.6922***	0.1427	-0.5918**	0.0485	-0.5848**	-0.0850
	(0.2659)	(0.3391)	(0.3470)	(0.2582)	(0.3127)	(0.2743)
Quantile regression (50th percentile)	-0.4334***	-0.4191**	-0.3167**	-0.0469	-0.3690**	-0.4147***
	(0.1294)	(0.1945)	(0.1760)	(0.1453)	(0.1620)	(0.1329)
Quantile regression (90th percentile)	-0.7020***	-0.6046**	-0.3378*	-0.6723***	-0.4865**	-0.6784***
	(0.2148)	(0.2614)	(0.2399)	(0.1914)	(0.2342)	(0.1875)

Notes: Robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficient on the maternity risk variable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

alike. However, cross-sample consistency is reduced, with the estimated effect of maternity risk shrinking and becoming statistically insignificant in the 2008-2010 cohabiting heterosexuals comparison.

The addition of the educational-attainment interaction terms marginally reduces the estimated effect of maternity risk across all six regressions. As expected, the inclusion of additional interaction terms with the lesbian indicator has almost no effect on the results. This supports the notion that the bias in the original estimates was due to model misspecification in the form of inadequate human capital controls. Considering the estimates without occupation and industry controls reveals the expected result, with the coefficient on *Maternity Risk* rising in all cases. It appears that women at a greater risk of child-birth may sort into jobs where the penalty associated with temporary labour-market separation is lower. Finally, the results from excluding women who gave birth during the previous year are able to assuage concerns that new mothers returning to the workforce may cause the observed maternity-risk penalty. In all cases, the maternity-risk disadvantage remains statistically significant and the estimated effect is not heavily affected in any of the regression models.

As we would expect, maternity risk is estimated to have a larger adverse effect in each of the six comparative earnings regressions in the full-time employed sample relative to the full sample. Among full-time employed females, a one percentage point increase in maternity risk is associated with a reduction in earnings of between 0.75 and 0.98 percent, when comparing the earnings of lesbians and cohabiting heterosexuals. Analogous estimates relative to married heterosexuals suggest a maternity-risk penalty of approximately 0.56 to 0.70 percent. The increased magnitude of the estimated effects also suggests that perceived maternity rates among full-time employed workers are not grossly over-stated based on the current allocation method, as one would expect the coefficients to diminish in such a case.[79] Finally, the results from the quantile regressions show mixed evidence that employers' maternity costs increase more than proportionately as employees' compensation rises. Considering the cohabiting heterosexuals comparison, maternity risk appears to have the largest relative effect at the bottom of the income distribution, and the smallest effect for those at the median income level. In stark contrast, the married women comparisons suggest

[79] If maternity incidence is lower among full-time employed females, the coefficient on *Maternity Risk* could still increase if maternity-related costs are sufficiently large for full-time workers. This, however, seems unlikely to be the case given that the dependent variable is in log form and the quantile regressions only show slight evidence that maternity costs increase more than proportionately with income.

that maternity risk has no statistically significant effect at the 10th percentile of the income distribution, while the largest effect is exerted at the 90th percentile. Evidently, cross-cohort inconsistency at the 10th percentile of the income distribution is a slight cause for concern, but estimates at other points of the income distribution and cross-sample comparisons display far greater similarity, offsetting this concern to some extent.

4.4.3. Results from Full Re-estimation

Table A.16 presents the allocated maternity risk for each dataset and subgroup following the sample restriction to white women. A comparison with Table 4.1 reveals that this induces a fall in maternity incidence among young lesbians and cohabiting heterosexuals across all samples. For example, among all races, 8.28 percent of 18-24 year old lesbians in the 2000 Census gave birth to a child in the previous year, compared to only 6.25 percent of white women. Table 4.5 contains the results from Model (A3) being re-estimated following the restriction to white women. Evidently, this restriction has varying effects across samples. In the 2000 Census sample, the coefficient on *Maternity Risk* is diminished, resulting in a reduction in statistical significance when comparing lesbians with married women. In contrast, the coefficient on *Maternity Risk* in the 2008-2010 ACS sample increases in absolute value, with the estimated effect becoming almost identical across the two regressions. The 2005-2007 ACS yields slightly volatile estimates, suggesting between a 0.31 and 0.90 percent reduction in earnings for every one percentage point increase in maternity risk.[80] Overall however, the qualitative findings are similar to those from the full sample, with an increase in maternity risk exerting a negative effect on earnings, holding other factors constant.

Table A.17 similarly displays the inferred maternity rates for the sample of never-married and currently married women. No significant differences emerge between the full and current samples, although young never-married women appear to have slightly lower rates of child-birth relative to all women, while the converse applies to never-married women over the age of thirty-one. Table 4.5 presents the results from re-estimating Model (A3) on this restricted sample. Once again, minor changes in the allocated maternity rates, in conjunction with an additional sample restriction, produce significantly different coefficient estimates among the lesbian and cohabiting heterosexuals regressions. Within these regressions, a one percentage point increase in maternity risk is predicted to reduce earnings by 0.75 to 0.94

[80] It is difficult to determine the exact cause of this asymmetric change, but arbitrarily defining maternity risk across the subgroups reveals the results are relatively sensitive to the age-structure of inferred maternity risk.

Table 4.5. Coefficient on *Maternity Risk* - complete subsample re-estimation

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Married	Cohabs	Married	Cohabs	Married
Initial analysis	-0.6043***	-0.5211***	-0.5411***	-0.2752**	-0.4492***	-0.3667***
	(0.1462)	(0.2152)	(0.1936)	(0.1534)	(0.1606)	(0.1406)
Including only white women	-0.5844***	-0.4127*	-0.9040***	-0.3095**	-0.4992***	-0.4899***
	(0.1808)	(0.2528)	(0.2272)	(0.1671)	(0.1881)	(0.1515)
Including only never-married or currently married individuals	-0.8078***	-0.5551***	-0.9421***	-0.2357*	-0.7547***	-0.3802***
	(0.1683)	(0.2161)	(0.2233)	(0.1535)	(0.1832)	(0.1412)

Notes: Robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficient on the maternity risk variable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

percent across the three samples. The lesbians versus married heterosexuals regressions display much smaller changes in the estimates from the base results, suggesting an earnings disadvantage of 0.24 to 0.56 percent for every one percentage point increase in perceived maternity risk. Despite the estimated effect of maternity risk differing based on the choice of heterosexual comparison group, the findings from the subsamples provide significant evidence in favour of the maternity-risk hypothesis.

Due to the potential importance of geographic mobility in employers' assessment of sexual orientation, partnership status, and thus maternity risk, Table A.18 displays the proportion of individuals within each subsample where neither the individual nor their partner moved residence in the past twelve months. As we might expect, married heterosexuals are the least geographically mobile of all partnered women, followed by lesbians and cohabiting heterosexuals. The proportion of non-movers is also dramatically higher among older women, who are more likely to be in stable long-term relationships and settling into permanent residences. Considering the large proportion of young women that have either moved or begun their cohabiting relationship in the past year, it is plausible that inferred sexual orientation and partnership status may be less accurate among these women. Retaining only non-movers thus has the potential to significantly improve the analysis by confining the sample to those individuals whose relationship status is more accurately perceived by employers.

Table A.19 displays the allocated maternity rates after restricting the sample to non-movers. Once again, maternity incidence among married women is relatively unaffected by this change, while rates of child-birth increase among eighteen to thirty-one year old lesbians and cohabiting heterosexual women. Given that we have retained those women likely to be in a longer-term committed relationship, this finding conforms to expectations. Applying these maternity rates and re-estimating Model (A3) yields the estimates contained in the first panel of Table 4.6. Again, the results from the restricted sample are broadly consistent with those from the full sample, with maternity risk estimated to have a statistically significant adverse effect on earnings in all cases. Dropping women who gave birth in the last twelve months slightly affects the results, rendering one of the coefficient estimates statistically insignificant. Across the eight specifications, a one percentage point increase in maternity risk is associated with a 0.17 to 0.60 percent reduction in income, holding other factors constant. Interestingly, the effect of maternity risk is no longer primarily larger among the cohabiting heterosexuals comparisons, as was the case in the majority of prior analyses.

Finally, Table A.20 presents assigned maternity risks for both householders and residents using the fertility variable contained within the ACS PUMS files. To improve the accuracy of the analysis, the sample is again restricted to non-movers, although unreported regressions reveal that the results are not sensitive to this choice. Examining observed maternity incidence from the ACS fertility variable reveals that the age-distribution of maternity risk is skewed much more towards the lower end when considering both householders and residents. This pattern can likely be attributed to younger women only being deemed the householder if they are the primary income-earner in the household, which will generally not be the case if they have recently given birth. Table 4.6 provides the estimated coefficients on *Maternity Risk* for this allocation method and sample. Importantly, the general findings from the previous analyses appear to carry over to use of the ACS fertility variable and the inclusion of residents as well as householders. Moreover, both cross-sample and cross-group consistency is enhanced, with the estimated effect of *Maternity Risk* being contained within a narrow, 0.20 percentage point range. Excluding individuals that gave birth in the past year only marginally affects the results, but reduces the level of statistical significance in the 2008-2010 cohabiting heterosexuals comparison. Over the eight regressions that include both householders and residents, a one percentage point increase in maternity risk is estimated to decrease earnings by 0.22 to 0.46 percent, ceteris paribus, consistent with prior findings.

Evidently, the finding of a negative earnings effect of maternity risk is robust to various model specifications, a restriction to white women, and restricting the sample to

Table 4.6. Coefficient on *Maternity Risk* - ACS-specific robustness tests

	2005-2007 ACS		2008-2010 ACS	
	Cohabs	Married	Cohabs	Married
Householders only				
Dropping both partners' observations if moved in the last twelve months	-0.5800***	-0.3648**	-0.2678*	-0.4796***
	(0.2223)	(0.1616)	(0.1927)	(0.1491)
Dropping from above regressions those who gave birth in last twelve months	-0.6013***	-0.3889***	-0.1731	-0.4412***
	(0.2283)	(0.1649)	(0.1982)	(0.1532)
Householders and residents				
Repeating entire analysis with ACS fertility variable and dropping if moved in last twelve months	-0.4638***	-0.3584***	-0.2802**	-0.4031***
	(0.1641)	(0.0834)	(0.1330)	(0.0793)
Dropping from above regressions those who gave birth in last twelve months	-0.4608***	-0.4115***	-0.2181*	-0.3988***
	(0.1692)	(0.0860)	(0.1367)	(0.0822)

Notes: Robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficient on the maternity risk variable. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

never-married or currently married women. Furthermore, considering only couples where neither partner moved residence in the past twelve months does not appear to affect the conclusions. Neither does excluding women who actually gave birth in the last year. Finally, use of the ACS fertility variable to assign maternity risk to both householders and residents, if anything, improves the consistency and applicability of the findings to the population of all partnered females.

Given the consistent evidence in favour of the maternity-risk hypothesis, the pertinent question regarding how the lesbian premium is affected by the inclusion of *Maternity Risk* re-emerges. Although the results are not reported, the overwhelming majority of specifications display an economically significant reduction in the lesbian premium upon the inclusion of *Maternity Risk*. As previously discussed, however, the magnitude of the decrease is dependent on the centring of the potential experience interaction terms, as well as the extent to which bias in the potential experience coefficients already implicitly partials out the maternity-risk effect from the lesbian premium. Simple quantitative reasoning can, however, aide in understanding the effects of controlling for near-term maternity risk on the lesbian pay gap. Due to the age-patterning of married and cohabiting heterosexual women, cohabiting heterosexual women exhibit the largest average maternity risk, at approximately six percent across the three samples, while married women average five percent. As maternity incidence across lesbians within the sample is approximately 1.5 percent, a maternity gap of 3.5 to 4.5 percentage points emerges. Assuming a true coefficient on *Maternity Risk* of -0.50, this suggests controlling for near-term maternity risk can eliminate approximately 1.75 to 2.25 percentage points of the raw lesbian premium.[81] This adverse earnings effect of maternity risk thus has important ramifications, as maternity-leave policy is likely to have a significant effect on both earnings and the lesbian pay gap.

5. Indirect Tests of the Maternity-Risk Hypothesis

As is evident from the previous section, the ability to directly detect discrimination on the basis of maternity risk is hindered by unobservable heterogeneity amongst females and multicollinearity problems. In order to overcome these problems, we utilise cross-state variation in the lesbian pay gap to further test the maternity-risk hypothesis.[82] Specifically, we assess whether state-level proxies which may affect the perceived maternity-risk gap between lesbians and heterosexual women exert the expected influence on the lesbian premium. For example, state laws granting greater partnership rights to same-sex couples are likely to increase the incidence of child-bearing among lesbian couples; we should thus expect a decrease in the lesbian premium in these states relative to states where no such law is in effect.

5.1. Data Sources and Sample

This section also relies on PUMS data from the 2000 U.S. Decennial Census (5% sample) and 2005 to 2010 American Community Surveys. An identical sample-selection procedure is implemented, although the controls used in the analysis necessitate an additional restriction. This restriction is a consequence of the Americans for Democratic Action (ADA) Voting Records' only being available for the fifty U.S. states. As the state-level ideology measures created by Berry et al. (1998) are computed from ADA scores, their use in the analysis entails

[81] It is important to note, however, that this is not the same as the omitted-variables bias in the adjusted premium resulting from excluding *Maternity Risk*, which was shown earlier to be approximately 0.80 to 1.20 percentage points. A similar conclusion can be obtained via Blinder-Oaxaca decompositions. A detailed discussion of these decompositions is outside the scope of this paper, but we refer the interested reader to Table F.1 for the results from such decompositions, which we provide for comparative purposes only.

[82] As prior effects of maternity risk no longer need to be partialled out, using proxies also allows us to assess the accumulated effects of maternity risk, whereas the analysis in the previous section only permits an assessment of near-term, forward-looking effects. Moreover, we no longer require the assumption that age, sexual orientation and partnership status are the only factors affecting perceived maternity risk.

the exclusion of the District of Columbia from the sample. The use of these controls thus comes at the cost of a reduction in the sample size by approximately 0.17 percent in each sample and the loss of the District of Columbia from the research population.

Several proxies are used in the analysis that follows in an attempt to overcome validity problems that may be associated with any particular proxy. The four proxies used are an indicator variable for states with laws mandating insurance coverage of infertility treatment, an indicator variable for states with laws allowing same-sex domestic partnerships or granting more extensive partnership rights to same-sex couples, an indicator variable for states with state-wide bans on same-sex marriage, and a standardised variable representing the size of the lesbian population as a percentage of the female adult population within each state.[83] Table 5.1 displays the summarised proxy values for each sample. Values for the *Same-Sex Percentage* proxy are displayed prior to standardisation. Between 2000 and 2005, several states enacted laws mandating insurance coverage of infertility treatment or imposed state-wide bans on same-sex marriage. Over time a number of states have also extended partnership, civil union, or marriage rights to same-sex couples. Finally, the data suggests lesbian prevalence increased in the 2005-2007 sample and subsequently dropped; however, differences in questionnaires and procedures followed by the Census Bureau across years likely account for the observed pattern.

As the proxies may be correlated with attitudes towards lesbians and labour-market conditions, unbiased estimation necessitates the inclusion of controls for other state-level factors. In addition to the Berry et al. (1998) citizen and government ideology measures, we include various other controls in an attempt to partial out the effect of maternity risk. These include the percentage of Gross State Product (GSP) from manufacturing, the presence of a state-wide Employment Non-Discrimination Act (ENDA) covering sexual orientation, an indicator for whether the state-level sodomy law was repealed prior to the federal repeal in 2003, and a 2013 same-sex legal equality score. An explanation of each measure and its corresponding source is provided in Appendix D, Table D.2, while a detailed discussion of their use is provided in the following section.

[83] Appendix D, Table D.2 contains a description of each proxy and its corresponding source, while Tables A.21-A.24 provide a detailed list of proxy values by state and year. Finally, Table F.2 provides a brief summary of control values by sample, but we omit a detailed list of values by state and year due to brevity considerations.

Table 5.1. Summary of proxy values by sample period

	2000 Census	2005-2007 ACS	2008-2010 ACS
Percentage of states with laws mandating insurance coverage of infertility treatment (%)	22.00	30.00	30.00
Percentage of states with laws allowing same-sex domestic partnerships or greater rights (%)	2.00	14.67	27.33
Percentage of states with state-wide bans on same-sex marriage (%)	68.00	85.33	83.33
Percentage of the adult female population comprised of coupled lesbians (%)	1.06	1.20	0.98

Sources: A list of sources for each proxy is provided in Table D.2.

5.2. Method

5.2.1. Base Analysis

Our indirect tests of the maternity-risk hypothesis focus on the effects of each proxy on the lesbian premium. For this purpose, in each regression, the level proxy variable is included in case it is correlated with income, as well as an interaction term with the lesbian indicator to allow for a differential effect across household types. Assessing the maternity-risk hypothesis thus entails examining the sign of the lesbian-proxy interaction coefficient estimate.[84] This method overcomes the bias problem created by unobservable human capital as inference is based on the marginal change in the lesbian premium associated with a particular proxy. It also allows us to capture longer-term effects of maternity risk than in the prior analysis. Despite controlling for ideology, factors affecting state-level income, and the rights of homosexuals, the level proxy terms may pick up other minor state-level effects, reducing the informational content of their associated coefficients. However, the interaction terms should

[84] An ideal test of proxy validity would assess whether employers' forward-looking expectations of the maternity gap is affected in the expected manner by the proxy. This test is, for obvious reasons, impracticable. Alternatively, we could assess whether the proxies can explain state-by-state variation in maternity incidence. Due to substantial heterogeneity among women across states, a probit model would be required. Several difficulties plague such an analysis, including specification error, which may bias the coefficients. Some proxies also represent newly formed laws and their effects on maternity rates are likely delayed due to family-formation considerations at the household level. Moreover, uncertainty regarding the continual expansion, or the potential reversal of rights granted to same-sex couples may exacerbate the lag in the maternity-risk effect. As a result, we use alternative techniques to assess proxy validity, as discussed below.

not be heavily affected, given that such unobserved factors are less likely to have a differential effect by household type.

Similarly to the previous analysis, we estimate earnings regressions of the following form to indirectly test the maternity-risk hypothesis:

$$\ln(R_i) = \boldsymbol{X}_i^{'}\beta + \boldsymbol{Z}_i^{'}\phi + Lesbian_i\gamma + \boldsymbol{P}_i^{'}\delta + \varepsilon_i \qquad (A)$$

where $\ln(R)$ is the natural log of real annual wage and salary income, and all other variables excluding $\boldsymbol{P}$ are as defined in Section 4.3. $\boldsymbol{P}$ is a vector of variables whose inclusion is dependent on the model to be estimated. It contains the relevant proxy for differential maternity risk, as well as the citizen and government ideology controls, percentage of GSP from manufacturing and an indicator for the presence of an ENDA covering sexual orientation. Appropriate interaction terms with sexual orientation are also included.[85] To account for the correlation between the error terms of individuals residing in the same state, we estimate cluster-robust standard errors in all models.

In order to assess the effects the proxies exert on the lesbian premium, we adopt a similar step-wise technique to that implemented in Section 3. Specification (A1) adds only the relevant proxy and its interaction with *Lesbian* to the model controlling for work, personal and other characteristics. Model (A2) adds the citizen ideology measure and its interaction with sexual orientation. Specification (A3) further includes *Government Ideology* and its interaction with the lesbian indicator dummy, while (A4) controls for the percentage of GSP from manufacturing and its interaction with *Lesbian*. Finally, (A5) adds an indicator for whether the state in which the individual resides has a state-wide ENDA covering sexual orientation, in addition to its interaction with the lesbian indicator. As these controls may be correlated with cross-state differences in maternity risk, their sequential addition enables the assessment of the proxy effect when controlling for, or alternatively not controlling for, various other state-level factors.[86] The concurrent analysis of the three non-overlapping samples again allows us to assess whether the results are robust to the choice of sample period.

[85] As in our prior analyses of ACS data, we also include year dummies in all regression models to account for year fixed-effects.
[86] We also centre all state-level controls and interactions about their means to aide in interpretation.

5.2.2. Tests of Robustness

As with previous analyses, we take the full-control model (Model (A5)) as the base model from which to conduct tests of robustness due to concerns about omitted-variables bias. As before, we re-estimate (A5) with hourly wages as the dependent variable, as well as restricting the analysis to full-time workers and excluding controls for hours and weeks of work. We also estimate a Heckman two-step model, where the selection equation takes a similar form to those estimated previously. Specifically, the selection equation includes all factors from the earnings equation, excluding work characteristics, and occupation and industry controls. The identifying variables include an indicator for the presence of children under the age of six and partner's real income.

Although the presence of an ENDA captures, to some extent, the legal standing of same-sex couples, it is possible that other state laws may also affect the lesbian premium. To account for this possibility, we consider two additional state-level controls in further tests of robustness. The first model adds an indicator variable for whether the state-level sodomy law was repealed prior to the federal repeal in 2003 to the base specification. The second further includes a 2013 measure of same-sex legal equality in each state. An important characteristic of these measures is that they are time invariant. They therefore serve the purpose of controlling for underlying levels of gay-friendliness in each state as well as potentially capturing the forward-looking expectations of employers regarding the rights of same-sex couples in the future. Provided the 2013 measure of same-sex equality does not significantly capture differential maternity risk in any particular year, these proxies aide in partialling out the maternity-risk effects associated with each proxy.

In a similar vein to the direct analysis, there is a possibility that the proxies may capture cross-state differences in unobservable human capital accumulation. To address this possibility, we estimate additional models in which we interact observable measures of human capital with the lesbian indicator dummy. The first such model interacts sexual orientation with potential experience and its square, while the second further adds indicators of educational attainment interacted with *Lesbian*. The inclusion of these controls should limit the extent to which the proxies capture cross-state variation in the sexual-orientation human capital gap, enhancing the validity of the analysis.

Having tested several supplementary specifications and dependent variable formulations, we then assess whether the results recur when restricting the analysis to specific subsamples. First, we restrict the sample to full-time employed women. If maternity risk has a more negative effect among full-time employed workers, as the direct analysis suggests, we should expect a larger negative coefficient on the lesbian-proxy interaction term when implementing this restriction. To ensure the results are not contingent upon state-level racial composition and to further reduce unobservable heterogeneity, we also re-estimate the base model including only white women. Sample-size considerations preclude an analogous restriction to non-whites. Finally, we re-estimate the base model, stratifying the analysis by age, first including only women aged 40 years or under, and then restricting the sample to women over 40 years of age. If the proxy is truly capturing the effect on the lesbian premium due to marginal changes in the maternity-risk differential, then we would expect the earnings effect to be largest among younger women, where maternity risk is likely to be most significantly affected.[87] The set of estimates these robustness checks produce is thus of crucial importance in assessing the validity of the proxies used.

Another potential concern is whether using OLS with cluster-robust standard errors adequately accounts for the hierarchical nature of the data. In this respect, the data can be thought of as being multilevel, with individuals constituting the first level, and the states they reside in constituting the second level. As previously discussed, only two prior studies analysing the lesbian pay gap implement multilevel estimation techniques, notably Baumle and Poston (2011) and Klawitter (2011). As is evident from these studies, one major disadvantage of such models is that they have difficulty converging when fitting complex models or when the sample size is large. Primo et al. (2007) similarly argue that estimating clustered standard errors is a more practical approach, particularly with large datasets or when including many cross-level interactions. However, to address remaining concerns surrounding the use of OLS, we take a 10% random sample of cohabiting and married heterosexual women, and re-estimate Specification (A5), fitting a multilevel model to the data. We estimate an unstructured variance-covariance matrix in each case to avoid imposing unnecessary restrictions on the covariance between the random slope and random intercept terms.[88] Model

[87] It is important to note, however, that this expectation is not uniform across all proxies. In the case of factors that have likely persisted over time, such as bans on same-sex marriage, or the percentage of the adult female population comprised of same-sex couples, the proxy should also differentially affect older lesbians due to effects on promotion in prior years. We could thus observe a marginal effect on the lesbian premium that is similar in magnitude for both younger and older women. The implications of these age-stratified regressions are discussed in greater detail when considering the results for each proxy.

[88] A detailed discussion of multilevel methods is beyond the scope of this paper. For further information on multilevel models see Hox (2010), or Klawitter (2011) for an applied example. We also estimated models with

(A5) is also re-estimated with the smaller sample to enable a more accurate comparison of the two methods.[89]

5.2.3. Comparison with Gay Males

As discussed, despite controlling for several state-level variables, the maternity-risk proxies may still capture cross-state variation in liberalism, hostility towards homosexuals and labour-market prospects. Another potentially confounding factor is that lesbians may be willing to accept lower income in order to reside in states granting greater rights to same-sex couples. The proxies may then capture selective location effects in addition to the effects of maternity risk. One potential way to address these issues is to re-estimate the models for gay males and examine the estimated proxy-effect on the gay-male pay gap. Assuming that the effects of unobservable factors such as gay sentiment, labour-market prospects, and residency choices are similar for gay males and lesbians, the difference between the effect of the proxy in the male and female regressions can be taken as the partial effect of the proxy on earnings due to its effect on maternity risk.[90] As a final test of the maternity-risk hypothesis, we therefore re-estimate (A5) for males, paying close attention to the difference in the estimated effect of the proxy in the male and female regressions.[91]

5.3. Results

5.3.1. Mandated Insurance Coverage of Infertility Treatment

Table 5.2 provides the base results from the 2000 Census sample for the *Mandated Coverage* proxy, which equals one whenever the individual resides within a state mandating insurance coverage of infertility treatment, and zero otherwise. Within Table 5.2 we display the

no region of residency controls as region arguably constitutes a third data level. These specifications produced qualitatively identical results.

[89] Similarly to Klawitter (2011), we also estimate several other models for each proxy to assess the sensitivity of the results to the specifics of the analysis. The (unreported) results suggest that the findings are moderately sensitive to the sequential exclusion of the largest states from the sample (California, Texas, New York and Florida), but insensitive to the use of a principal component for state-level ideology, choice of controls, income restrictions, and stratification by race. As these robustness checks are considered of secondary importance relative to those included, a detailed analysis of the results is omitted due to brevity considerations.

[90] Family formation among gay males may also increase, for example through adoption, when greater rights are conferred on them, or even when they have improved access to support groups. Although it is not strictly necessary, these parental-leave effects, which are common to both gay males and lesbians, are assumed to be minor in comparison to the direct effects of the proxy on female maternity incidence.

[91] To obtain the male sample, we follow an identical sample-selection procedure to that outlined in Section 3.1 and again use cohabitation to identify different household types. Several other specifications were re-estimated for the male sample, producing similar results to those obtained from the base regressions included here.

Table 5.2. Effect of laws mandating insurance coverage of infertility treatment on the lesbian premium - 2000 Census

	Lesbians versus cohabiting heterosexuals					Lesbians versus married heterosexuals				
	(A1)	(A2)	(A3)	(A4)	(A5)	(A1)	(A2)	(A3)	(A4)	(A5)
Lesbian	0.0400***	0.0396***	0.0390***	0.0379***	0.0415***	0.0168**	0.0102	0.0103	0.0078	0.0099
	(0.0081)	(0.0090)	(0.0094)	(0.0096)	(0.0086)	(0.0089)	(0.0088)	(0.0091)	(0.0092)	(0.0084)
Mandated coverage	0.0381*	0.0060	0.0051	0.0040	-0.0011	0.0412	0.0076	0.0084	0.0065	0.0033
	(0.0224)	(0.0215)	(0.0182)	(0.0180)	(0.0114)	(0.0287)	(0.0229)	(0.0201)	(0.0194)	(0.0111)
Mandated coverage	0.0332**	0.0396**	0.0367**	0.0342**	0.0230*	0.0313*	0.0378**	0.0329**	0.0309**	0.0153
x lesbian	(0.0197)	(0.0211)	(0.0179)	(0.0156)	(0.0149)	(0.0188)	(0.0213)	(0.0169)	(0.0148)	(0.0130)
Citizen ideology		0.0370***	0.0334***	0.0307***	0.0240***		0.0426***	0.0381***	0.0352***	0.0278***
		(0.0070)	(0.0088)	(0.0091)	(0.0063)		(0.0064)	(0.0088)	(0.0091)	(0.0067)
Citizen ideology		-0.0091	-0.0103	-0.0140	-0.0152		-0.0129	-0.0138	-0.0180*	-0.0195*
x lesbian		(0.0114)	(0.0119)	(0.0116)	(0.0118)		(0.0107)	(0.0108)	(0.0105)	(0.0105)
Government ideology			0.0039	0.0043	0.0001			0.0047	0.0053	0.0019
			(0.0034)	(0.0033)	(0.0023)			(0.0041)	(0.0041)	(0.0027)
Government ideology			0.0017	0.0024	0.0003			0.0014	0.0020	-0.0018
x lesbian			(0.0030)	(0.0028)	(0.0027)			(0.0030)	(0.0027)	(0.0026)
% GSP manufacturing				-0.1811	-0.1836*				-0.2009*	-0.1928*
				(0.1202)	(0.1057)				(0.1061)	(0.1046)
% GSP manufacturing				-0.2652**	-0.2634*				-0.2881**	-0.2952**
x lesbian				(0.1306)	(0.1339)				(0.1320)	(0.1328)
State ENDA					0.0792***					0.0760***
					(0.0136)					(0.0166)
State ENDA					0.0196					0.0366**
x lesbian					(0.0198)					(0.0164)
Observations	175,609	175,609	175,609	175,609	175,609	255,185	255,185	255,185	255,185	255,185
R^2	0.6354	0.6368	0.6369	0.6370	0.6379	0.6294	0.6312	0.6314	0.6315	0.6322

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the state-level variables and cross-level interactions. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. The coefficients of interest (*Lesbian* and *Mandated Coverage × Lesbian*) are tested on a one-tailed basis. Source: 2000 Decennial Census PUMS.

coefficient estimates of the other state-level variables to convey the effects they exert upon earnings and the lesbian premium. The reported coefficient on the lesbian indicator represents the premium evaluated at the mean of the state-level controls. Returns to individual-level characteristics, such as potential experience and education, are omitted due to brevity considerations. Furthermore, their estimated effects are relatively stable upon the inclusion of state-level controls, and thus the results are qualitatively similar to those discussed in Section 3.4. Definitions of infertility in U.S. state laws generally entail the individual being unable to conceive or sustain a successful pregnancy after one to two years of unprotected (heterosexual) sexual intercourse. In some cases, the legal definition also requires the individual to be married. Owing to these definitions, mandated insurance coverage of infertility treatment should increase the maternity risk of cohabiting and married heterosexual females relative to lesbians. Thus, if the maternity-risk hypothesis is correct, the lesbian premium should be higher relative to cohabiting and married heterosexuals in states with mandated coverage.

The preceding predictions are borne out by the base results from the 2000 Census, providing evidence in favour of the maternity-risk hypothesis. The comparison with cohabiting heterosexuals reveals the estimated lesbian premium across the five specifications to be between 2.43 and 4.20 percentage points larger in states with mandated insurance coverage of infertility treatment than in states without, when evaluated at the means of other state-level variables.[92] Evidently, controlling for the presence of an ENDA covering sexual orientation is important, resulting in a substantial reduction in the estimated effect of *Mandated Coverage* on the lesbian premium. The results also indicate a significant increase in the lesbian premium over married women in states with mandated coverage. In such states, the premium is estimated to be some 1.56 to 3.89 percentage points larger than in those absent mandated coverage. Excluding the final married women comparison, these estimates are both statistically and economically significant. Holding other factors at their respective means, the estimates represent an increase in the lesbian earnings premium of more than fifty percent vis-à-vis cohabiting heterosexuals and married women. It is questionable whether mandated insurance coverage of infertility treatment could drive this large an earnings differential, given that the total number of children born through infertility treatment worldwide in the past

---

[92] For the specification: $\ln(R_i) = \boldsymbol{X}_i'\beta + \boldsymbol{Z}_i'\phi + Lesbian_i\gamma + Mandated\ Coverage\ \times Lesbian_i\delta_1 + \cdots + \varepsilon_i$, the change in the premium in percentage points is calculated as $\exp^{(\gamma+\delta_1)} - \exp^{(\gamma)}$, or in the case of (A1), $\exp^{(0.0400+0.0332)} - \exp^{(0.0400)} = 3.51$ percentage points. This convention is used throughout the discussion of the proxies.

thirteen years is approximately equal to the annual number of U.S. births (Centres for Disease Control and Prevention (U.S.), 2013; Innes, 2013).

Analysing the coefficients on the controls reveals some interesting results. Citizen ideology appears to be strongly positively correlated with income, with a positive and highly significant coefficient in each case, as might be expected. In all cases, higher citizen ideology results in a smaller lesbian premium, ceteris paribus, but this effect is only statistically significant in two regressions. Including *Government Ideology* adds little explanatory power to the model, with the coefficient on both the level variable and its interaction with *Lesbian* being small in magnitude and statistically insignificant in all regressions. As expected, individuals residing in states with a high percentage of GSP from manufacturing, on average, receive significantly lower incomes, ceteris paribus. The lesbian premium is also reduced in such states. Finally, as identified by Klawitter (2011) and Gates (2009), states with higher average incomes appear to selectively implement ENDA's covering sexual orientation, as evidenced by the statistically significant positive coefficient on the ENDA indicator. Similarly to Gates (2009), we find evidence that labour-market discrimination may have an adverse effect on the earnings of lesbians, with the presence of an ENDA offsetting this to some extent. For the remainder of this section, we focus our attention on the lesbian-proxy interactions, noting that the results discussed here carry over to the other samples and proxies.

Summarised results from the 2000 Census, as well as results for the ACS samples, are provided in Table A.25. The latter results demonstrate the importance of using multiple sample periods in assessing explanations of the lesbian premium. Notably, the 2008 to 2010 sample provides evidence consistent with that from the 2000 Census, with lesbians earning a larger premium over both cohabiting and married heterosexual women in states mandating insurance coverage of infertility treatment. The magnitude of this marginal change is also in line with the results from the 2000 Census. However, the results from the 2005 to 2007 ACS sample provide limited support for the maternity-risk hypothesis. Lesbians are estimated to receive a larger premium over cohabiting heterosexuals of between 1.73 and 2.47 percentage points in states with mandated coverage, although the coefficient is statistically insignificant across all specifications. More troubling, however, is that *Mandated Coverage* is estimated to have no effect on the lesbian premium relative to married women. Although it is plausible that the increase in the premium is greater relative to cohabiting heterosexuals, based on the results from the direct analysis, the finding of no marginal change in the premium relative to married heterosexuals is inconsistent with the maternity-risk hypothesis.

Table A.26 contains the results from the robustness checks for the *Mandated Coverage* proxy. For ease of exposition, within the table we display the percentage point change in the lesbian pay gap associated with the proxy moving from a value of zero to one, when evaluated at the means of all other variables.[93] Evidently, the estimated effect of *Mandated Coverage* on the pay gap is significantly affected by dropping work characteristics and restricting the sample to full-time workers. Specifically, the proxy is now associated with an economically and statistically significant increase in the lesbian premium in the 2005-2007 ACS sample, but the magnitude and significance of the 2008-2010 estimates are reduced. In contrast, using hourly wages as the dependent variable or the Heckman two-step method produces results almost identical to the base specification. The inclusion of additional state-level controls only minimally changes the estimated effect of the proxy, as does the introduction of human capital variables interacted with the lesbian indicator. In contrast, considering only full-time workers dramatically alters the results for the ACS samples. Despite the interaction term being insignificant in the base regression for the 2005-2007 sample, the restriction to full-time employed reveals a significant increase in the lesbian premium relative to both cohabiting (5.14 percentage points) and married heterosexual women (3.98 percentage points) in states mandating coverage. Moreover, the previously significant estimates for the 2008-2010 sample are much smaller and no longer statistically significant. Considering the re-estimation for white women reveals slightly larger estimates of the effect of *Mandated Coverage* on the lesbian premium, resulting in statistical significance in four of the six regressions.

As laws mandating insurance coverage of infertility treatment have generally been implemented in recent years, one would expect a larger increase in the lesbian premium among younger women in states with such laws.[94] This expectation is borne out in both the 2000 and 2008-2010 samples, but the results from the 2005-2007 ACS provide conflicting evidence, with older lesbians receiving a larger increase in the premium than younger lesbians. Analysing the reduced sample produces results similar to the full sample, although the 2005-2007 estimates are now both of the incorrect sign. Somewhat surprisingly, multilevel analysis produces results noticeably different to those generated by OLS, with only

---

[93] The level of statistical significance provided refers to the coefficient on the lesbian-proxy interaction term. In Tables F.3-F.6 we provide supplementary regression results which we use to produce the summary tables in Appendix A.

[94] Although infertility treatments are also available for women over forty years of age, success and uptake rates are significantly lower among this cohort (Society for Assisted Reproductive Technologies, n.d.).

one estimate being both statistically significant and of the appropriate sign. More troubling is that the proxy is now associated with a significant and negative change in the lesbian premium relative to married heterosexuals in the 2005-2007 sample. Finally, the results from the male sample suggest that gay males benefit significantly from living in states mandating insurance coverage for infertility treatment. Calculating the difference between the estimated effects in the male and female regressions reveals sizable negative estimates in all cases. Taking this gap as representing the marginal change in the premium attributable to maternity risk yields a result in stark contrast to prior expectations. A plausible explanation of this phenomenon is that states with less fertile women selectively adopt laws enhancing maternity rates. As such, all else held constant, the sexual-orientation maternity gap may actually be lower in states which have implemented such laws.[95] Absent an accurate method to assess this hypothesis, the results from the *Mandated Coverage* proxy must be deemed inconclusive.

5.3.2. Same-Sex Partnership Laws

Table A.27 contains the base estimates from the ACS samples for the *Legal Partnership* proxy, which equals one whenever the individual resides in a state allowing same-sex partners to enter into a domestic partnership or conferring greater partnership status on same-sex couples, and zero otherwise. In 1999, only Hawaii allowed same-sex domestic partnerships, resulting in minimal variation in the proxy, thus precluding an analysis of the 2000 Census sample. Given that greater legal rights are likely to increase lesbians' willingness to start a family, laws allowing same-sex domestic partnerships are expected to increase future maternity incidence among lesbians. Holding all other factors constant, this should reduce the sexual-orientation maternity-risk gap, exerting a negative effect on the lesbian premium relative to cohabiting and married heterosexual women.

The results in Table A.27 suggest that the lesbian premium is reduced in states allowing same-sex domestic partnerships, consistent with the expectations of the maternity-risk hypothesis. Relative to cohabiting heterosexuals in the 2005-2007 sample, the lesbian premium is lower in these states by between 1.56 and 2.77 percentage points when evaluated at the mean of other state-level variables, although none of the estimates are statistically significant at conventional levels. Similar estimates obtained via a comparative analysis with

[95] Comparing unadjusted fertility rates across states suggests that *Mandated Coverage* exerts the expected effect on maternity risk. If the same is true for adjusted fertility rates, selective adoption of laws cannot explain the observed results. In such a case, it is likely that inadequate controls have been included in the models to account for cross-state heterogeneity, and the proxy is capturing non-maternity-risk related factors.

married women indicate a smaller premium of between 2.04 and 3.75 percentage points. Analysing the 2008-2010 sample yields a similar estimated effect in the full-control model relative to cohabiting heterosexuals (2.64 percentage point reduction), although the estimates diverge considerably across the other specifications. In addition, there exists cross-sample differences in results relative to married women, with the 2008-2010 sample suggesting that partnership laws have no effect on the lesbian premium. Despite this, the base results appear to be broadly consistent with the maternity-risk hypothesis, with all full-control models estimating a negative, albeit sometimes statistically insignificant, effect of *Legal Partnership* on the premium.

Table A.28 presents the robustness checks that serve to further assess the maternity-risk hypothesis in the context of same-sex domestic partnership laws. In contrast to the *Mandated Coverage* proxy, the choice of dependent variable does not substantially alter the findings, although all estimates shrink when restricting to full-time employed females and dropping work characteristics. Use of the Heckman two-step procedure yields similar estimates to the base results in the 2005-2007 sample, and increases the magnitude of the estimated proxy-effect relative to cohabiting heterosexuals in the 2008-2010 period. The base results from both samples emphasise the importance of controlling for cross-state heterogeneity in estimating the partial effect of maternity risk on the premium. Sequentially adding the sodomy law repeal dummy and the same-sex legal equality score only minimally changes the estimated effect of *Legal Partnership* on the pay gap, suggesting that the presence of an ENDA adequately captures the non-partnership-related legal environment faced by same-sex couples. However, including interactions between observable human capital and the lesbian indicator markedly affects the results. That is, the proxy is now associated with a significant reduction in the premium across all specifications in each sample, providing support for the maternity-risk hypothesis. Restricting the sample to full-time employed females reduces the estimated effect of partnership laws, eliminating statistical significance in the 2008-2010 cohabiting heterosexuals comparison. Considering only white women does not appear to significantly affect the results, with the estimates remaining relatively stable across all four regressions.

State laws allowing same-sex domestic partnerships or granting more extensive partnership rights to same-sex couples are a relatively new phenomenon in the U.S. If the maternity-risk hypothesis is correct, we would thus expect partnership laws to result in a greater reduction in the lesbian premium among the younger cohort. Unfortunately, the age-stratified analysis provides conflicting evidence regarding the maternity-risk hypothesis. The 2005-2007 results conform to expectations, with an estimated reduction in the lesbian premium of 4.07 to 4.34 percentage points among women forty years of age or under, while smaller and statistically insignificant effects are observed for women over forty. In contrast, the 2008-2010 sample suggests older lesbians are at a greater disadvantage than younger lesbians when living in states allowing same-sex domestic partnerships, civil unions, or marriages. This is slightly concerning given that we would expect a larger effect among younger women if the proxy truly captures marginal changes in the premium due to the effect on the maternity-risk differential. The multilevel results from the reduced sample reveal that the effect of *Legal Partnership* on the premium is affected by the choice of estimation method, with large discrepancies between the multilevel and OLS results in the 2005-2007 sample. However, the estimates are of the correct sign in each case, so the qualitative conclusions remain unaffected. Finally, the results from the male sample provide further evidence regarding proxy validity, with the gay-male pay gap being largely unaffected by same-sex partnership laws. Differencing the estimated effects in the male and female regressions produces negative values, although this does vary widely across samples. The *Legal Partnership* proxy thus provides relatively strong support for the maternity-risk hypothesis, given the consistent results generated across a variety of robustness tests. However, the results must be interpreted with caution due to the mixed evidence regarding proxy validity obtained via age-stratified regressions.

5.3.3. Same-Sex Marriage Bans

The third proxy used in assessing the maternity-risk hypothesis is *Marriage Bans*, which equals one whenever the individual lives in a state with a law or Constitutional Amendment banning same-sex marriage, and zero otherwise. In the reverse manner to laws allowing same-sex domestic partnerships, bans on same-sex marriage are likely to inhibit family formation among lesbian couples, increasing the maternity divide between heterosexual and lesbian women. As a result, the lesbian premium should be greater in states with such bans in place, ceteris paribus.

The estimates from the initial analysis, presented in Table A.29, display supportive evidence in favour of the maternity-risk hypothesis. Relative to cohabiting and married heterosexual women, the lesbian premium is 1.63 to 3.93 percentage points larger in states banning same-sex marriage, according to the 2000 Census sample. Similar estimates obtained using the 2005-2007 ACS sample suggest an increased premium of 2.93 to 6.05 percentage points, with the majority of estimates being statistically significant.[96] Finally, compared to cohabiting heterosexual women, the 2008-2010 sample suggests a statistically insignificant disadvantage for lesbians living in states banning same-sex marriage, and approximately no effect relative to married heterosexual women. Once again, controlling for other state-level factors appears to be of critical importance in examining the maternity-risk hypothesis, with the coefficients often varying noticeably across specifications.

Results from the *Marriage Bans* robustness checks are reported in Table A.30. Similarly to previous proxies, the choice of estimation method influences the findings to some extent, but the qualitative conclusions remain unchanged. Use of hourly wages, for example, substantially increases the cross-sample consistency of the results, although the estimates in the 2005-2007 sample become statistically insignificant. The results also appear to be relatively invariant to the use of the Heckman two-step procedure, restricting the sample to white women, and the inclusion of additional state and individual-level controls, thus emphasising the robustness of the findings. Restricting the sample to full-time employed females, regardless of whether or not we retain work-characteristic controls, generates large and statistically significant estimates for the proxy-interaction coefficient in the 2000 Census sample, while the ACS sample estimates remain relatively unaffected.

As many of the laws and Constitutional Amendments ruling out same-sex marriage were implemented in the mid-to-late nineties, we would expect to see greater increases in the lesbian premium in states with bans among younger women in the 2000 Census sample. In the later samples, the effects of maternity risk are likely to have filtered through to the premium among older women due to discriminatory promotion practices, substantially weakening the prior expectation regarding the relative age-stratified effects. Examining the results suggests further evidence in favour of the maternity-risk hypothesis, with the effect of *Marriage Bans* on the lesbian premium being greater among the younger cohort in the 2000 Census sample. Moreover, the estimated effect among women aged forty or under is within three-quarters of a percentage point across five of the six specifications. Despite having no strong prior expectation regarding the relative effects across age cohorts in the ACS samples, the 2005-2007 results are slightly concerning, displaying a premium increase twice as large among the older cohort relative to the younger cohort, while the 2008-2010 results are all statistically

[96] Interestingly, holding other state-level factors at their respective means, the 2005-2007 ACS results reveal that there is no statistically significant lesbian premium in states without bans on same-sex marriage.

insignificant. In contrast to the previous proxies, multilevel estimation generates estimates almost identical to those produced by OLS, assuaging concerns about the choice of estimation technique. Again, the results from the gay-male comparison support the maternity-risk hypothesis, with the estimated effect of the proxy on the gay-male pay gap being statistically insignificant in all cases. More importantly, differencing the male and female proxy effects produces positive and relatively stable figures across all comparative analyses. Considering the results as a whole, the *Marriage Bans* proxy thus provides substantial support for the maternity-risk hypothesis.[97]

5.3.4. Prevalence of Same-Sex Couples

The final proxy used in the indirect analysis of the maternity-risk hypothesis is *Same-Sex Percentage*, which is a standardised variable with a mean of zero and standard deviation of one, capturing the size of the lesbian population as a percentage of the adult female population within each state. We believe this proxy to be appropriate for two reasons. First, if relationship status is imperfectly observable but sexual orientation is known, a greater percentage of lesbians within a state results in a higher probability of partnership, enhancing perceived maternity risk. Second, even if relationship status is observable, we posit that a higher concentration of lesbians will lead to enhanced child-bearing rates due to the presence of support groups and greater social acceptance towards lesbian couples raising children. On this basis, it is expected that the lesbian premium will be lower in states characterised by greater prevalence of lesbians as a proportion of the female population.

Table A.31 presents the initial results from the *Same-Sex Percentage* proxy, which display a number of noteworthy features. First, the inclusion of state-level controls significantly influences the results for the 2000 Census sample, with no significant effect of the proxy-interaction term estimated in Specifications (A1) to (A3), and a statistically significant negative effect in three of the four remaining regressions. In contrast, the results from the ACS samples are relatively invariant to the sequential addition of state-level controls. Another important point concerns the magnitude of the effect across samples. Considering the full-control model, results from the 2000 Census suggest that a one standard deviation increase in the lesbian prevalence rate from the mean is associated with a 2.55 and 3.64 percentage point reduction in the lesbian premium relative to cohabiting and married

---

[97] Importantly, the age-stratified and gay-male regressions provide relatively consistent support for proxy validity.

heterosexual women, respectively.[98] Analogous estimates from the ACS samples suggest a smaller effect of between 1.56 and 2.28 percentage points.

Despite being of the correct sign and highly significant, it is unlikely that these estimates solely reflect marginal changes in the premium due to maternity risk, as the results imply an earnings discount for lesbians in states with a high proportion of same-sex couples.[99] This result is not sensible in that it is unlikely that higher probabilities of partnership, the presence of support groups, or greater social acceptance of lesbian family formation would completely offset the maternity-risk differential present between lesbians and heterosexual women. Thus, unless substantial discrimination exists against lesbians, one would not expect an increase in maternity risk to result in a lesbian earnings disadvantage relative to cohabiting and married heterosexual women. It is therefore likely that other factors, such as discrimination, are augmenting the maternity-risk effect. As previously discussed, lesbians may also be willing to accept lower compensation in order to reside in states with greater support networks and tolerance towards lesbians. However, results from the age-stratified analysis and gay-male comparison suggest that both this and the discrimination argument are unlikely to be driving the large negative estimates in the female samples.

Results from the robustness checks for the *Same-Sex Percentage* proxy are provided in Table A.32. Evidently, the sign and statistical significance of the estimates are unaffected in the majority of the supplementary regressions. Using hourly wages as the dependent variable or implementing the Heckman two-step procedure has minimal influence on the results. As with the *Legal Partnership* proxy, restricting the sample to full-time workers and dropping work-characteristic controls does not significantly affect the results. Similarly to the *Marriage Bans* proxy, the robustness of the findings is highlighted by the minimal change in the estimates when restricting the sample to white women, interacting human capital measures with the lesbian indicator, or including further state-level controls. Considering only full-time employed women yields smaller estimated effects in most cases, although the differences are relatively minor.

---

[98] Throughout this section, we evaluate the effects of the *Same-Sex Percentage* proxy by considering a one standard deviation increase from the mean, holding all other factors at their respective means. The effects of larger increases or decreases can be evaluated in an analogous manner.

[99] For example, holding other factors at their respective means, lesbians are predicted to be at an earnings disadvantage relative to married heterosexual women in all samples when residing in states where lesbian prevalence is two standard deviations above the mean. In the 2000 Census sample, lesbians are also at an earnings disadvantage relative to cohabiting heterosexuals when they reside in a state where lesbian prevalence exceeds the state-level mean by 2.2 standard deviations.

Provided the relative concentration of lesbians across states has remained fairly constant over time, there is no obvious prior expectation regarding the age-stratified results.[100] Inspecting the table reveals that in two of the three samples, lesbians aged forty years or under experience the greatest premium reduction when residing in states with a high concentration of same-sex couples. Moreover, the estimates for older women are statistically insignificant in all four comparisons. Considering these results in conjunction with the reverse finding in the 2008-2010 sample suggests that a simple discrimination argument for the proxy's influence is unlikely to suffice, lending strength to the maternity-risk hypothesis. Turning to the estimates corresponding to the reduced sample, the use of multilevel estimation in this case attenuates the estimated effect of maternity risk, but the conclusions remain unaltered.[101] Finally, the results from the gay-male comparison provide further evidence in favour of the maternity-risk hypothesis. The *Same-Sex Percentage* × *Gay Male* coefficient is positive in all cases, suggesting that discrimination and compensation differences due to choice of state-residence are inadequate explanations of the female results. Differencing the male and female coefficients produces negative figures, although these vary widely between the 2000 Census and ACS samples. Given the reservations regarding proxy validity, we take the results as providing tentative support for the maternity-risk hypothesis, noting that the estimated effect is likely augmented by unobservable factors.

6. Discussion

6.1. Summary of Findings

This study examines three non-overlapping samples from the 2000 U.S. Census, 2005-2007 ACS and 2008-2010 ACS to assess whether perceived differences in labour-force attachment contribute to the observed lesbian earnings advantage. No previous empirical studies have analysed the sexual-orientation pay gap using ACS data. As such, in a similar framework to prior studies, Chapter 3 presents an analysis of the lesbian premium for each of the three samples, ignoring both forward-looking expectations of child-bearing and contextual factors that may affect earnings.

[100] The between-sample correlation in state-level lesbian prevalence is approximately 0.85 when using within-sample averages, suggesting that relative lesbian prevalence has remained fairly constant.
[101] The use of a random subsample also appears to affect the conclusions to some extent, although this is not surprising given the significant difference in sample sizes.

Using logged real annual income as the dependent variable, and controlling for hours and weeks of work, we uncover an economically and statistically significant lesbian earnings premium of thirty-five to forty percent relative to cohabiting heterosexual women. Similar estimates compared to married women imply a premium of ten to fifteen percent. The systematic addition of progressively more endogenous controls reduces the premium relative to both comparison groups, although a lesbian pay gap still remains in all three samples. Controlling for work and personal characteristics, region of residence, and occupation and industry, our estimates suggest the premium to be approximately three and six percent relative to married and cohabiting heterosexual women, respectively.

To ensure that the presence of a lesbian pay gap is not conditional upon the dependent variable specification or estimation method, we conduct several robustness tests. Dropping the vector of work characteristics and using the log of real hourly wages as the dependent variable reveals marginally larger estimates of the lesbian premium, with the results being highly similar to previous studies using 2000 Census data. As we would expect, restricting the analysis to full-time employed females and dropping work-characteristic controls generates much larger estimates of the lesbian earnings advantage, due to sexual-orientation differences in annual hours worked among full-time employed females. Finally, use of the Heckman two-step procedure to address sample-selectivity concerns generates similar premium estimates relative to cohabiting heterosexuals, but significantly increases the estimated premium compared to married heterosexuals.

Independent of the chosen specification and estimation technique, we find extensive evidence of an economically and statistically significant lesbian premium. In the remainder of the paper, we examine whether at least some of this premium can be attributed to sexual-orientation differences in maternity risk. To minimise concern associated with omitted-variables bias, all subsequent regression models include the full vector of controls for personal characteristics, region of residence, and occupation and industry. Furthermore, OLS is used as the primary estimation technique due to volatility in the Heckman two-step estimates and difficulties in identifying appropriate exclusion restrictions.

Section 4 presents a direct analysis of the maternity-risk hypothesis. Using intra-household relationships to infer fertility among householders, we allocate maternity rates by sexual orientation and age group using within-sample maternity incidence as an estimate of employers' forward-looking maternity-leave expectations. Preliminary regressions that include allocated maternity risk as an additional regressor display mixed evidence regarding the effect of maternity risk on earnings, and thus the lesbian premium. We argue that prior effects of maternity risk on career progression and unobservable differences in human capital accumulation lead to bias in the relevant coefficient estimates, necessitating the inclusion of additional controls.

To account for sexual-orientation differences in returns to observable human capital that may result, we interact potential experience and its square with the lesbian indicator, and sequentially include them as additional regressors. Results from these specifications provide robust evidence that maternity risk negatively affects earnings, with estimates suggesting a 0.28 to 0.60 percent reduction in earnings for every one percentage point increase in maternity risk. In addition, the penalty appears to be slightly larger when comparing lesbians with cohabiting heterosexual women, indicating that the penalty associated with an increase in maternity risk may be diminishing in the level of maternity risk.

Considering models with some form of potential experience interaction provides consistent evidence that the lesbian premium falls upon controlling for maternity risk. Across the three samples and two comparison groups, these specifications imply that accounting for near-term maternity risk reduces the premium by one to twenty-three percent when evaluated at the mean level of potential experience. As the majority of estimates are clustered within a small range, our best estimates suggest that controlling for the near-term likelihood of labour-market separation to bear children reduces the lesbian pay gap by ten to fifteen percent. These estimates are, however, subject to a considerable degree of uncertainty. Specifically, the estimated change is conditional upon the base premium estimate and the extent to which bias in the effect of potential experience implicitly accounts for differential maternity risk. An important implication is that omitting a control for maternity risk generates upward bias in the estimated earnings effect of potential experience.

In robustness tests, where all models include both potential experience interactions to mitigate the problems associated with unobservable human capital, we again assess alternative specifications and implement the Heckman two-step procedure. Although the level of statistical significance is reduced in three regressions, the general findings remain qualitatively similar across specifications. Further robustness checks, such as including additional sexual-orientation interaction terms, dropping occupation and industry controls, and excluding from the regressions women who gave birth in the previous year, provide robust evidence that maternity risk adversely affects earnings. In particular, the latter result assuages concerns that the maternity-risk earnings disadvantage merely reflects the adjustment period associated with new mothers re-entering the workforce. Restricting to full-time employed females reveals a larger adverse effect of maternity risk. The results from quantile regressions, however, provide mixed evidence that the maternity-risk disadvantage is relatively larger at the upper end of the earnings spectrum.

Repeating the maternity-risk allocation procedure and re-estimating the preferred specification for white women produces qualitatively consistent results, with a one percentage point increase in maternity risk reducing earnings by 0.31 to 0.90 percent. An analogous restriction to never-married or currently married women yields similar results. These results suggest the findings are relatively robust to further disaggregation of the maternity-risk allocation process, however sample-size limitations preclude greater disaggregation within the full sample.

Dropping both partners' observations if either partner moved in the last twelve months, as well as repeating the analysis for householders and residents using the ACS fertility variable, generates results consistent with the maternity-risk hypothesis. The finding of an adverse earnings effect of maternity risk is thus highly robust to various model specifications and sample restrictions. Unreported results again show the reduction in the lesbian premium across the robustness tests to be approximately ten to fifteen percent upon controlling for near-term maternity risk. Using quantitative reasoning (and Blinder-Oaxaca decompositions), we demonstrate that the true maternity-risk effect accounts for approximately 1.75 to 2.25 percentage points of the raw lesbian premium.

In Section 5, we use cross-state variation in the lesbian pay gap to indirectly examine the maternity-risk hypothesis. Specifically, we assess whether proxies that are likely to affect the sexual-orientation maternity gap exert the expected influence on the lesbian premium. The four proxies used are an indicator variable for states with laws mandating insurance coverage of infertility treatment, an indicator variable for states with laws allowing same-sex domestic partnerships or granting more extensive partnership rights to same-sex couples, an indicator variable for states with state-wide bans on same-sex marriage, and a standardised variable representing the size of the lesbian population as a percentage of the female adult population within each state.

Considering each of the four proxies in turn, we estimate regressions in which the proxy and its interaction with the lesbian indicator are added to a base model controlling for work and personal characteristics, region of residency, and occupation and industry. The sequential addition of state-level controls and their interactions with the lesbian indicator then enables an assessment of the proxy-effect when controlling for, or alternatively not controlling for, various other state-level factors. We also run a number of robustness checks, including alternative dependent variable and model specifications, use of multilevel estimation, and restricting the sample to full-time employed and white women. For purposes of evaluating proxy validity, we present an age-stratified analysis and assess the proxy's effect on the gay-male penalty.

Estimates from the *Mandated Coverage* proxy generally suggest an increase in the lesbian premium in states mandating insurance coverage of infertility treatment, as we would expect based on the maternity-risk hypothesis. This finding is relatively robust to choice of dependent variable and the inclusion of additional controls. Moreover, state ideology, the percentage of GSP from manufacturing and the presence of an ENDA covering sexual orientation have the expected effect on income and the lesbian pay gap, implying the model is not grossly misspecified. The gay-male comparison, however, casts doubt on the causal relationship between mandated insurance coverage for infertility treatment, maternity risk, and earnings, as the gay-male disadvantage is significantly reduced in states with such laws in effect. Lacking an accurate method to assess the effect of the proxy on the fertility gap, we deem the results from this proxy inconclusive.

Analogous estimates for the *Legal Partnership* proxy provide significant evidence in favour of the maternity-risk hypothesis. Holding other state-level factors at their respective means, the majority of results suggest that the lesbian premium is at least twenty-five percent lower in states granting partnership rights to same-sex couples. This result also appears to be robust to several model formulations and the use of alternative estimation techniques. Importantly, the results from the male sample reveal the gay-male pay gap is largely unaffected by same-sex partnership laws, suggesting that *Legal Partnership* captures the effect on the lesbian premium due to marginal changes in maternity risk. The age-stratified analysis, however, provides mixed evidence across samples regarding proxy validity.

The results from *Marriage Bans* offer further evidence that differences in maternity risk partially explain the observed lesbian pay gap. Although the estimates vary across samples and comparison groups, the coefficient on the lesbian-proxy interaction term is positive in the majority of regressions, with no estimate being statistically significant and of the incorrect sign. Results from the age-stratified analysis are also broadly consistent with prior expectations. Moreover, differencing the male and female coefficients produces positive and relatively stable estimates across all comparative analyses. Considering the results as a whole suggests that bans on same-sex marriage contribute significantly to the lesbian premium, consistent with the maternity-risk hypothesis.

Finally, we argue that a higher probability of partnership, greater social acceptance towards lesbian child-bearing, and the presence of support networks result in a causal link between higher lesbian prevalence and enhanced perceived maternity risk. Use of the *Same-Sex Percentage* proxy suggests the lesbian premium is negatively affected by greater lesbian prevalence, consistent with this theory. Moreover, results from age-stratified analyses and the gay-male comparison suggest that discrimination or compensation effects associated with residential choice are unlikely to be driving this result. Unfortunately, the magnitude of the estimates suggests that variation in the proxy can more than eliminate the lesbian premium when evaluated at the means of other state-level variables. This result cannot be entirely attributed to maternity-risk effects. We therefore take the estimates as providing tentative evidence in favour of the maternity-risk hypothesis. Further research on the effect of lesbian prevalence on earnings is required to disentangle the maternity-risk effect from unobserved, reinforcing factors.

Overall, the indirect tests of the maternity-risk hypothesis further support the findings from the direct analysis: maternity risk appears to negatively affect income, thereby significantly contributing to the observed lesbian premium. Although there is considerable uncertainty surrounding the estimates, taking the results at face value suggests that longer-term maternity-risk effects may conservatively account for more than twenty-five percent of the remaining lesbian premium over otherwise-similar heterosexual women. These results should be interpreted with caution due to cross-sample inconsistency and our inability to directly assess proxy validity.

6.2. Implications for the Male-Female Pay Gap

Given the importance of maternity risk in determining earnings and thus the lesbian premium, the analysis presented in this study may suggest a hitherto unexplored explanation of the gender pay gap. Specifically, if differential labour-market treatment occurs on the basis of perceived labour-force attachment, then we should expect not only a lesbian premium, but also a general female earnings disadvantage. As such, controlling for differential maternity risk may reduce the estimated male-female earnings disparity.

To assess this possibility, we estimate comparative earnings regressions analogous to Model (A3) from Chapter 4, comparing the earnings of males and females. In particular, the specification controls for work and personal characteristics, region of residency, occupation and industry. Whereas the prior analysis included a sexual-orientation dummy and its interaction with potential experience and its square, the current analysis replaces all sexual-orientation terms with a female indicator to examine differences in income by gender. In order to minimise unobservable heterogeneity and allow for a marriage premium,[102] we run separate regressions for cohabiting heterosexuals and married heterosexuals.[103] The model is estimated both prior to and following the inclusion of *Maternity Risk*, to enable an assessment of the change in the female earnings penalty.

It is important to note that the analysis in this section is subject to greater uncertainty than the direct analysis of the sexual-orientation pay gap due to substantial cross-gender heterogeneity. Moreover, the change in the earnings gap is again dependent on the centring of the potential experience interaction terms and the extent to which bias in the potential experience coefficients implicitly accounts for the effects of maternity risk. Care should therefore be taken when interpreting the results and their inclusion serves only to provide tentative evidence of a potential source of the male-female earnings gap.

Table A.33 displays the results from comparing the earnings of cohabiting heterosexual males and females. In the 2000 Census sample, the female disadvantage is estimated to be 17.88 percent prior to controlling for maternity risk. Including *Maternity Risk* reduces this penalty to 13.85 percent, representing a decrease of approximately twenty-three percent. Similar estimates from the ACS samples display a reduction in the female earnings penalty of roughly nineteen to twenty-five percent. On this basis, controlling for maternity risk appears to significantly reduce the male-female earnings disparity, as hypothesised.

---

[102] Plug and Berkhout (2004) also note the importance of controlling for sexual orientation and marital status in male-female analyses. However, we argue that including dummy variables for sexual orientation and marital status imposes unnecessary restrictions on the coefficient estimates, which may bias the results.

[103] A comparison between gay males and lesbians reveals similar results, although the base pay gap is much smaller. These results are omitted as the included sub-groups predominantly determine the overall gender pay gap.

Results from the gender comparison among married heterosexuals (Table A.34) suggest that controlling for maternity risk has a much smaller effect on the gender pay gap, reducing the female disadvantage from 28.11 percent to 25.49 percent in the 2000 Census sample. Analogous estimates from the ACS sample display a reduction in the female disadvantage of 2.45 and 3.61 percent. This represents less than a one percentage point change in both cases. As maternity risk is predicted to significantly decrease earnings in each sample, bias in the potential experience coefficients likely inhibits an accurate analysis of the true maternity-risk effect. Examining the potential experience interaction terms in both comparisons reveals substantial volatility across specifications. Moreover, particularly in the married heterosexuals comparison, explanatory power is only marginally improved upon the inclusion of *Maternity Risk*. In conjunction with the coefficient estimate being highly statistically significant, this suggests strong confounding with potential experience.

These findings suggest that the inability to adequately control for maternity risk in earnings regressions likely results in upward bias in the adjusted gender pay gap. Moreover, the recurring finding of an adverse maternity-risk earnings effect may be of importance to maternity-leave policy considerations. Specifically, if policymakers can ease the maternity-leave burden borne by employers, employment and promotion prospects among female workers will likely improve, enhancing labour-market equality.

6.3. Improvements and Ideas for Future Research

Although every attempt has been made to ensure the validity and applicability of the findings, future research may facilitate more accurate inference by addressing some of the deficiencies of this study. Alternative datasets with information on actual work experience, for example, could significantly reduce problems associated with human capital accumulation in direct tests of the maternity-risk hypothesis, by reducing unobservable heterogeneity. Moreover, greater returns to actual experience for lesbians would provide evidence of discriminatory promotion practices on the basis of perceived maternity risk, a theory that is not testable with the current data. Obtaining information on job tenure may also be of importance to the extent that employer learning enables a more accurate assessment of partnership status and sexual orientation, and thus maternity risk. Incorporating these factors into the direct analysis would represent a significant improvement on the current methodology.

Another potential improvement on the current analysis may entail a direct assessment of prior maternity-risk effects. Specifically, if differences in unobservable human capital and previous effects of maternity risk can be disentangled (perhaps by examining the difference between the lifetime likelihood and forward-looking expectations of maternity leave), we would have an additional method through which to assess the maternity-risk hypothesis. The ability to explicitly account for prior maternity effects may also enable the maternity-risk allocation process to be disaggregated across a number of additional variables. This would also eliminate the bias in potential experience and maternity risk that lack of such controls generates. Allocating maternity risk differentially by number of pre-existing children or years since last child may capture differences in family-formation completion, while disaggregation by educational attainment could capture delays in child-bearing associated with higher educational attainment.[104] Data limitations, such as the inability to accurately estimate lifetime maternity incidence, unfortunately preclude such an analysis with the current data. We therefore acknowledge that the current method is limited, but provides useful insight into the potential labour-market effects of maternity risk.

With respect to indirect tests of the maternity-risk hypothesis, several potential improvements are apparent. First, the U.S. GSS contains questions on tolerance and attitudes towards homosexual individuals. Access to confidential GSS files would therefore enable the construction of time-varying measures of state-level hostility towards gays. These could then be included in our analysis to further isolate the effect of maternity risk on the lesbian pay gap. This is particularly important, for example, in isolating the causal maternity-risk effect when using the *Same-Sex Percentage* proxy. Second, one major concern regarding the current proxies is that we cannot readily assess their validity in capturing differential maternity risk. Despite having strong prior expectations as to the direction of the proxies' affect on future maternity incidence, many of the laws underlying the proxies have only recently been implemented. It is thus likely that insufficient time has passed for their full fertility and earnings effects to transpire. Similarly to Gates (2009), an extension to the analysis could examine whether time since implementation of relevant state laws is an important factor in determining the magnitude of the proxy-interaction effect.[105] If a longer period since

[104] Due to sample-size considerations, such an allocation process would require a probit model (or similar) to be estimated, to predict fertility based on age and additional factors.

[105] As previously noted, several difficulties prevent us from directly assessing proxy validity. An extension to the analysis could therefore focus on using alternative methods to directly assess the proxies' effects on the lesbian-heterosexual maternity gap.

implementation is associated with greater confidence that the law will not be over-turned, and a correspondingly larger pass-through to maternity incidence, then one would expect to observe smaller effects among states that have recently passed such laws. This hypothesis is not examined herein as developments in U.S. legislation regarding the rights of same-sex couples is a recent phenomenon, thus precluding a more thorough analysis of the maternity-risk effect on the lesbian premium.

As discussed, the use of several proxies can moderate validity concerns associated with any particular proxy.[106] Consideration should therefore be given to alternative proxies which may capture cross-state differences in the sexual-orientation maternity-risk differential. One potentially viable proxy could be based on whether or not state law permits same-sex couples to jointly petition to adopt a child. Although individual homosexuals may petition to adopt in all fifty states and the District of Columbia, joint adoption provides both partners with similar rights and is therefore likely to significantly affect adoption by same-sex couples. Given that adoption, especially of young children, may entail time out of the workforce, same-sex adoption laws can be used to assess the maternity-risk hypothesis. We do not attempt such an analysis due to inconsistencies across sources regarding state-level adoption laws.

Two final proposed proxies, average state-level childcare costs and state extensions to federal maternity-benefit requirements, can be applied in both the lesbian-heterosexual women context and in male-female comparisons. In states with higher average childcare costs, new mothers are more likely to postpone returning to the workforce as the opportunity cost of doing so is greater. In such states, maternity leave is likely to be more costly for employers, resulting in a stronger aversion to maternity risk and a commensurately higher lesbian (or male) earnings premium.[107] Some states, such as California, have also enacted laws which confer greater maternity benefits on women. Benefits include paid maternity leave, additional breaks, and flexible use of sick leave. Holding other factors constant, the sexual-orientation and gender pay gaps should therefore be higher in these states. Despite the availability of current information on childcare costs and maternity benefits at the state-level,

---

[106] For example, the analysis in Section 5 suggests that *Mandated Coverage* may not be an appropriate proxy, while the estimated effect of *Same-Sex Percentage* appears to be augmented by unobserved factors.

[107] Higher childcare costs may also reduce maternity incidence, resulting in the reverse effect. Further research into the relative magnitude of these effects would be required prior to using average childcare costs as a proxy.

time-series measures spanning the length of the three samples are extremely difficult to obtain. As such, an analysis using these proxies is left for future research.

Future research may also focus on the effects of maternity risk on other labour-market outcomes, particularly employment and promotion. Results from the Heckman selection-correction models estimated in Sections 4 and 5 provide evidence that maternity risk adversely affects employment outcomes. Data limitations, however, preclude an assessment of whether the results represent discriminatory hiring practices or if differential employment rates reflect women voluntarily exiting the labour force due to family-formation considerations. Similarly to Frank (2006), use of an alternative dataset may allow us to examine the effect of maternity risk on occupational rank. Finally, it is possible to directly detect hiring discrimination by conducting an audit study similar to Baert (2013), as opposed to using regression analysis. A detailed discussion of the method is outside the scope of this paper, but future research could extend Baert's method by obtaining informal wage offers as in Drydakis (2011). This would provide information on both hiring and wage discrimination. Such an analysis would offer immediate evidence regarding the validity of the maternity-risk hypothesis and could largely avoid the difficulties encountered in our analysis.

6.4. Conclusion

Prior research from the U.S. and abroad reveals a sizable lesbian earnings advantage over otherwise-similar heterosexual women. Using data from the 2000 U.S. Census and 2005-2010 American Community Surveys, we estimate traditional earnings equations and find robust evidence of a lesbian premium, corroborating the findings of previous studies. We then examine the earnings effect of maternity risk to determine whether the perceived likelihood of an employee requiring maternity leave contributes to the observed lesbian premium. The results from numerous direct tests provide considerable evidence that maternity risk negatively affects income. As lesbians exhibit a comparatively smaller degree of maternity risk than their cohabiting and married heterosexual counterparts, our findings indicate that part of the observed premium can be attributed to differences in maternity risk. Although bias in the estimated effect of potential experience and differences in unobservable human capital accumulation prohibit a conclusive assessment, our estimates imply that controlling for near-term maternity risk reduces the lesbian premium by approximately ten to fifteen percent. Further evidence from indirect analyses using several proxy variables suggests that cumulative maternity-risk effects may conservatively explain over twenty-five percent of the lesbian pay gap. As such, the persistent finding of a lesbian premium in previous studies can be attributed, at least in part, to employers' aversion to maternity risk and its associated costs.

These findings are also of critical importance to the general labour-market discrimination literature. Given the adverse earnings effect of maternity risk, our analysis suggests that estimates of the well-established gender earnings disparity are likely to be considerably smaller when incorporating maternity risk into the analysis. Absent the ability to adequately control for maternity risk, strict attention should be paid to potential upward bias in estimated earnings differentials. Moreover, policymakers should consider the broader implications of maternity-leave policy on the labour-market outcomes of females. In this respect, maternity-leave policy will likely influence the hiring and promotion decisions of employers, thereby indirectly affecting sexual-orientation and gender equality in the labour market. However, further research in this area is still required, given the limitations inherent in the direct and indirect analyses. We hope this study will serve to provide the necessary impetus for further empirical work on the maternity-risk hypothesis and its consequential effect on labour-market inequality.

Appendix A: Key Tables

Table A.1. Summary of previous studies examining the earnings differential between lesbians and heterosexual women

Study	Data sources[1] (country)	Sample size	Sexual orientation definition	Estimation technique	Dependent variable(s)	Hours/weeks controls?	Findings: earnings of lesbians relative to heterosexual women
Ahmed, Andersson & Hammarstedt (2013b)	2007 LISA (Sweden)	604,242 females (1,067 lesbians)	Individuals living in a civil union are classified as lesbians	OLS and quantile regressions with a lesbian indicator dummy. Separate analyses for private and public sector, as well as for childless individuals.	Log(annual earnings) and Log(full-time monthly earnings)	No and yes, respectively	4% lower to 3% higher than married heterosexual women, based on OLS and full-time monthly earnings (mostly significant).[2] Full results also show that the lesbian premium is larger when comparing annual earnings, as well as in the public sector and at the upper end of the earnings distribution.
Ahmed & Hammarstedt (2010)	2003 LOUISE (Sweden)	2,015 females (925 lesbians)	Individuals living in a civil union are classified as lesbians	OLS and quantile regressions with a lesbian indicator dummy. Separate analyses for metropolitan and non-metropolitan areas.	Log(annual income)	No	20% lower (significant) to 8% higher than married heterosexual women (based on OLS - quantile regression estimates are similar but more varied). Larger lesbian penalties are observed in non-metropolitan areas and at the lower end of the income distribution.
Antecol, Jong & Steinberger (2008)	2000 Census 5% PUMS (U.S.)	763,977 females (6,205 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	Blinder-Oaxaca decomposition with a lesbian indicator dummy and DiNardo-Fortin-Lemieux decomposition.	Log(hourly wages)	Yes	4% higher than cohabiting heterosexuals but no difference relative to married women, based on Blinder-Oaxaca decomposition. 1% lower to 8% higher than partnered women, based on DiNardo, Fortin, and Lemieux decomposition (at the upper and lower ends of the wage spectrum, respectively).
Arabsheibani, Marin & Wadsworth (2004)	1996 Q1-2001 Q4 LFS (UK)	176,586 females (297 lesbians)	Sex of unmarried partner if currently not married	OLS with a lesbian indicator dummy. Separate analyses for London and elsewhere (by region), as well as by broad age groupings.	Log(hourly wages)	Yes	10% to 13% higher than single women and 3% to 21% higher than partnered women (mostly significant). Larger premiums evident for older lesbians and those not living in London.
Arabsheibani, Marin & Wadsworth (2005)	1996 Q1-2002 Q4 LFS (UK)	204,738 females (359 lesbians)	Sex of unmarried partner if currently not married	Blinder-Oaxaca decomposition with a lesbian indicator dummy.	Log(hourly wages)	Yes	8% higher than coupled heterosexual women and 9% higher than all heterosexuals (both highly significant).

Table A.1 (continued)

Study	Data sources (country)	Sample size	Sexual orientation definition	Estimation technique	Dependent variable(s)	Hours/weeks controls?	Findings: earnings of lesbians relative to heterosexual women
Arabsheibani, Marin & Wadsworth (2007)	2000 Census 5% PUMS (U.S.)	1,668,814 females (≈21,000 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	OLS with a lesbian indicator dummy. Separate analyses by level of education, job sector, region, age, and employment status.	Log(hourly wages)	Yes	3% to 12% higher than partnered women (all highly significant). Lesbian advantage is larger for older women, those employed full-time or in the private sector, and among women with less education.
	1996 Q1-2004 Q4 LFS (UK)	185,266 females (527 lesbians)	Sex of unmarried partner if currently not married	As above.	Log(hourly wages)	Yes	2% lower to 12% higher (larger premiums are significant). Similar pattern to the U.S., except the premium is greater amongst part-time employed and in the UK public sector.
Badgett (1995)	1989-1991 GSS (U.S.)	732 females (34 lesbians/ bisexuals)	At least as many same-sex partners as opposite-sex partners since the age of 18	OLS and Heckman two-step procedure with lesbian indicator dummy and experience × lesbian interaction (imputes income using within occupation medians from CPS).	Log(annual income)	No	11% to 30% lower than unmarried heterosexual women and 24% lower to 15% higher than married women (all insignificant and not evaluated at the mean of potential experience). Estimates would be less negative/more positive if accounting for the interaction term (see Badgett (2006)).[3]
Badgett (2001)	1989-1994 GSS and 1992 NHSLS (U.S.)	1,988 females (53 lesbians/ bisexuals)	At least as many same-sex partners as opposite-sex partners since the age of 18	OLS with a lesbian indicator dummy, using mid-points of income categories.	Log(annual income)	No	3% to 11% higher than currently or previously married heterosexuals, and 2% to 10% higher than those who have never married (all insignificant).
Baumle &Poston (2011)	2000 Census 5% PUMS (U.S.)	1,645,998 females (21,797 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	OLS with a lesbian indicator dummy, as well as HLM with a lesbian indicator dummy and its interaction with state-level variables (compares this to multilevel OLS).	Log(annual income)	Yes	4% higher than married women, based on OLS or HLM (holding the proportion of same-sex couples at zero; the premium falls slightly as this proportion rises). Similarly, 8% to 9% higher than cohabiting heterosexual women (all results highly significant).

Table A.1 (continued)

Study	Data sources (country)	Sample size	Sexual orientation definition	Estimation technique	Dependent variable(s)	Hours/weeks controls?	Findings: earnings of lesbians relative to heterosexual women
Berg & Lien (2002)	1991-1996 GSS (U.S.)	1,310 females (52 lesbians/ bisexuals)	At least one same-sex partner within the last 5 years	Maximum Likelihood Estimation with a lesbian indicator dummy (similar to ordered probit but does not require estimation of income bracket thresholds. Similar results obtained using the equivalent of ordered logit).	Log(annual income)	No	13% to 47% higher than heterosexual women (based on the confidence interval surrounding the point estimate of 30%).
Black, Makar, Sanders & Taylor (2003)	1989-1996 GSS (U.S.)	2,246 females (53 lesbians/ bisexuals)	At least as many same-sex partners as opposite-sex partners since the age of 18	Maximum Likelihood Estimation (interval regressions) with a lesbian indicator dummy (report that OLS produces similar results).	Log(annual income)	No	6% and 9% higher than unmarried and married heterosexual women, respectively (insignificant).
		2,193 females (49 lesbians/ bisexuals)	At least one same-sex partner within the last year	As above.	Log(annual income)	No	Similarly, 22% to 25% higher and 26% to 28% higher than unmarried and married heterosexual women, respectively (significant).
		1,723 females (54 lesbians/ bisexuals)	At least one same-sex partner within the last 5 years	As above.	Log(annual income)	No	27% to 40% higher compared to both unmarried and married heterosexual women (significant).
Blandford (2003)	1989-1996 GSS (U.S.)[4]	2,064 females (<51 open lesbians/ bisexuals)	At least one same-sex partner in last year (five years if no partners in last year) and not currently married.	OLS and Heckman two-step procedure with a lesbian indicator dummy (imputes income within categorical ranges using median annual earnings for each race-gender sub-group from CPS). Separate analysis for women without children.	Log(annual income)	No	15% to 30% higher than unmarried heterosexual women and 17% to 38% higher than married heterosexual women (mostly significant).

Table A.1 (continued)

Study	Data sources (country)	Sample size	Sexual orientation definition	Estimation technique	Dependent variable(s)	Hours/weeks controls?	Findings: earnings of lesbians relative to heterosexual women
Carpenter (2005)	2001 CHIS (U.S.)	8,912 females (179 lesbians)	Self-reported sexual orientation	OLS with a lesbian indicator dummy (reports that including a Heckman selection correction term does not alter the results). Also presents separate analyses by age.	Log(hourly wages)	Yes	6% lower to 4% higher than partnered heterosexual women, based on several specifications (never significant). 3% higher and 4% lower for women 35 and under, and over 35, respectively (insignificant).
	1988-2000 GSS (U.S.)	2,855 females (49 lesbians)	Exclusively same-sex partners within the last 5 years	OLS with a lesbian indicator dummy (imputes income based on categorical range mid-points).	Log(annual income)	No	21% to 31% higher than unmarried heterosexual women (depending on the period analysed, mostly significant).
Carpenter (2008a)	2003-2005 CCHS (Canada)	75,457 females (657 lesbians)	Self-reported sexual orientation	OLS with a lesbian indicator dummy (quantile regressions also estimated but not reported). Separate analyses for partnered and non-partnered individuals.	Log(annual income)	No	16% to 17% higher among all women and 43% higher among partnered women (all highly significant), but only 1% higher when restricted to non-partnered women (not significant).
Carpenter (2008b)	2000 ALSWH (Australia)	7,143 females (69 lesbians)	Self-reported sexual orientation	Interval regression with a lesbian indicator dummy (similar results obtained with ordered logit, not reported).	Log(weekly income)	Yes	24% to 31% lower than heterosexual women (all significant), based on several different model specifications.
Clain & Leppel (2001)	1990 Census 1/1000 PUMS (U.S.)	26,028 female (58 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	Heckman two-step procedure with a lesbian indicator dummy and its interaction with various other controls (based on specification search procedure revolving around statistical significance).	Log(annual income)	No	No premium over coupled females unless living in the Midwest or with dependents (in which case premium is at least 39%). 3% lower to more than 100% higher than women not cohabiting with a partner, depending on age, geographic location and presence of dependents.

Table A.1 (continued)

Study	Data sources (country)	Sample size	Sexual orientation definition	Estimation technique	Dependent variable(s)	Hours/weeks controls?	Findings: earnings of lesbians relative to heterosexual women
Cushing-Daniels & Yeung (2009)	1988-2006 GSS (U.S.)	4,795 females (109 lesbians)	Exclusively same-sex partners in the previous year	OLS and Heckman FIML procedure with a lesbian indicator dummy (imputes income based on categorical range mid-points).	Log(annual income)	No	12% and 9% higher than married and unmarried heterosexual women, respectively, based on OLS. Similarly, 13% lower and 7% higher, using Heckman FIML procedure (all insignificant and based on the entire sample period).
		4,109 females (unknown no. of lesbians)	Exclusively same-sex partners in the previous 5 years	As above.	Log(annual income)	No	11% and 10% higher than married and unmarried heterosexual women, respectively (OLS). Similarly, 14% lower and 5% higher (Heckman) (all insignificant).
Daneshvary, Waddoups & Wimmer (2008)	2000 Census 5% PUMS (U.S.)	98,683 females (6,777 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	OLS with a lesbian indicator dummy and its interaction with education, experience, and experience squared. Also estimates Blinder-Oaxaca decompositions and includes separate analyses by years of education.	Log(hourly wages)	Yes	5% to 6% higher than other women (full sample), 8% to 10% higher than women with less than 16 years of education, and 2% higher than other women with 16 or more years of education (all significant, using Blinder-Oaxaca decomposition). Full results show an earnings disadvantage at low levels of experience and high levels of education and vice versa.
Daneshvary, Waddoups & Wimmer (2009)	2000 Census 5% PUMS (U.S.)	178,794 females (6,785 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	OLS with a lesbian indicator dummy and its interaction with prior marital status. Also estimates Blinder-Oaxaca decompositions, separate analyses by previous marital history and estimates a Tobit earnings model.	Log(hourly wages) and Log(annual income)	Yes	3% to 4% higher, and 2% greater, than cohabiting heterosexual women and single women, respectively (comparing previously married women). Respective lesbian earnings premiums among never-married women are 4% to 7%, and 8%. Relative to married women, never-married lesbians receive a 5% premium, but their previously married counterparts obtain no significant premium (all other cases are significant).

Table A.1 (continued)

Study	Data sources (country)	Sample size	Sexual orientation definition	Estimation technique	Dependent variable(s)	Hours/weeks controls?	Findings: earnings of lesbians relative to heterosexual women
Elmslie & Tebaldi (2007)	2004 CPS (U.S.)	28,626 females (<678 lesbians)	Sex of unmarried partner if both are at least 25 years of age	Heckman two-step procedure with a lesbian indicator dummy. Also includes interaction terms between sexual orientation with other controls, including (but not limited to) race and occupation.	Log(hourly wages) and Log(annual income)[5]	Yes and no, respectively	3% lower to 4% higher than married heterosexual women and 2% to 8% higher (significant) than cohabiting heterosexual women, based on hourly wages (ignoring interaction terms). Regressions based on annual income find a premium of between 1% and 19%.
Frank (2006)	2000-2001 AUT survey of employees at six UK universities	404 females (49 lesbians/ bisexuals)	Self-reported sexual orientation	OLS on data for both genders, including female and female-LGB dummy variables.	Log(annual income)	No	8% higher than heterosexual women among all staff and 5% to 17% higher among academics (all insignificant except 17% premium). Lower earnings gaps are observed when controlling for occupational rank.
Gates (2009)	2000 Census 1% & 5% PUMS (U.S.)	331,026 females (unknown no. of lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	OLS with a lesbian indicator dummy and its interaction with state-level anti-discrimination laws.	Log(hourly wages)	Yes	3% to 6% higher than other women in states without sexual-orientation anti-discrimination laws. Similarly, 4% to 8% higher in states with such laws (all highly significant).
Heineck (2009)	1994 ISSP Data (U.S., Australia, Ireland, Poland and Bulgaria)	2,220 females (32 lesbians)	Exclusively same-sex partners in the previous 5 years	Heckman two-step procedure and Blinder-Oaxaca decomposition with a lesbian indicator dummy (although decomposition not reported for lesbians).	Log(monthly income)	No	11% higher than heterosexual women (but insignificant) in both the gender-pooled and women-only specifications.

Table A.1 (continued)

Study	Data sources (country)	Sample size	Sexual orientation definition	Estimation technique	Dependent variable(s)	Hours/weeks controls?	Findings: earnings of lesbians relative to heterosexual women
Jepsen (2007)	2000 Census 5% PUMS (U.S.)	113,772 females (14,528 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	OLS with a lesbian indicator dummy and its interaction with various controls in different specifications (e.g. education and experience).	Log(annual income)	No[6]	9% to 14% higher than married heterosexual women and 10% to 17% higher than cohabiting heterosexual women (all highly significant) based on specifications with no sexual-orientation interaction terms (interaction-term coefficients are not reported and hence the premium cannot be evaluated at the means).
Klawitter (2011)	2000 Census 5% PUMS (U.S.)	33,077 females (6,356 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	HLM and quantile regression with a lesbian indicator dummy and its interaction with state and local level variables, including presence of an anti-discrimination law. Separate analyses by sector, race and householder status.	Log(annual income) and Log(hourly wages)	No and yes, respectively	3% and 7% higher (significant) than married and cohabiting heterosexual women, respectively, based on hourly earnings. Estimates based on annual income suggest an earnings advantage of at least 16% and 9% (and up to 88%) relative to married and cohabiting heterosexual women, respectively (except in the Government sector and for African Americans).
Klawitter & Flatt (1998)	1990 Census 5% PUMS (U.S.)	17,491 females (3,493 lesbians)	Sex of unmarried partner or spouse as indicated on the housing record	OLS with a lesbian indicator dummy and its interaction with various anti-discrimination and control dummies.	Log(annual income)	No	3% to 16% higher than cohabiting heterosexual women and 11% to 23% higher than married heterosexual women (mostly significant, depending on presence and type of anti-discrimination policies). These differences disappear when restricting the analysis to full-time, full-year workers.
Laurent & Mihoubi (2012)	1996-2007 Employment Survey (France)	116,202 females (327 lesbians)	Same-sex cohabiting "friends" with numerous restrictions	OLS and Heckman two-step procedure with a lesbian indicator dummy. Separate analyses for private and public sector.[7]	Log(monthly income)	Yes	0% to 2% higher than unmarried heterosexual women (mostly insignificant). 1% to 4% higher than married heterosexual women (mostly significant). Lesbian advantage is greatest in the private sector.

Table A.1 (continued)

Study	Data sources (country)	Sample size	Sexual orientation definition	Estimation technique	Dependent variable(s)	Hours/weeks controls?	Findings: earnings of lesbians relative to heterosexual women
Plug & Berkhout (2004)	1998-2000 Survey of Dutch Graduates (Netherlands)	6,437 females (198 lesbians)	Identified based on expressed sexual preferences	OLS on data for both genders, including female and female-lesbian dummy variables.	Log(monthly income) and log(hourly wages)	Yes	1% to 4% higher than heterosexual women (across sixteen different specifications, some significant).

Notes: [1] Sample sizes listed represent the number of females used in the main regression analysis. Where these are unavailable, sample sizes provided refer to the total number of women included in summary statistics tables.

[2] The discussion of significance throughout the table refers to statistical significance.

[3] Results vary only slightly when using alternative definitions.

[4] Blandford also uses 1992 NHSLS data to justify his choice of sexual orientation definition.

[5] Also estimates log(hourly compensation) regressions, the results for which are highly similar to those using log(hourly wages).

[6] Replicating the analysis with log(hourly wages) as the dependent variable did not substantially alter the results (not reported).

[7] For males, further analyses by skill-level, age and time of service are included, as well as a Blinder-Oaxaca type decomposition.

Sources: See table.

Table A.2. Female sample size at each stage of the sequential restriction process

Sample	Stage of the restriction process[1]					
	(1)	(2)	(3)	(4)	(5)	(6)
2000 Census						
Lesbians	32,756	32,756	29,893	28,454	14,298	12,204
Cohabiting heterosexuals	220,847	220,847	215,230	209,312	201,399	163,834
Married heterosexuals	2,821,706	423,256	365,190	355,720	355,720	243,299
No co-residential relationship	4,136,401	620,460	243,775	241,046	239,566	178,276
2005-2007 ACS						
Lesbians	24,346	24,346	21,844	21,631	12,234	10,302
Cohabiting heterosexuals	142,523	142,523	137,064	136,010	134,771	107,599
Married heterosexuals	1,868,847	280,328	237,956	236,824	236,824	161,515
No co-residential relationship	2,520,799	378,119	157,180	156,502	155,589	116,422
2008-2010 ACS						
Lesbians	20,296	20,296	17,620	17,196	12,674	10,653
Cohabiting heterosexuals	162,302	162,302	155,557	153,261	149,635	116,650
Married heterosexuals	1,871,770	280,765	234,861	233,279	232,950	158,587
No co-residential relationship	2,622,410	393,361	165,972	165,041	161,839	117,271

Notes: [1] (1) = All individuals, (2) = (1) + Random sampling, (3) = (2) + Age and GQ restrictions, (4) = (3) + Relationship, age and sex allocation restrictions, (5) = (4) + Marital status allocation restriction, (6) = (5) + Labour-force restrictions.

Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.3. Earnings comparison: lesbians and cohabiting heterosexuals (2000 Census)

	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)
Lesbian	0.3006***	0.2541***	0.0788***	0.0637***	0.0590***	0.0540***
	(0.0069)	(0.0067)	(0.0063)	(0.0063)	(0.0063)	(0.0062)
Potential experience		0.0283***	0.0329***	0.0329***	0.0348***	0.0320***
		(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Potential experience^2		-0.0005***	-0.0005***	-0.0005***	-0.0006***	-0.0005***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hispanic		-0.0327***	0.0363***	-0.0053	0.0021	-0.0049
		(0.0071)	(0.0068)	(0.0068)	(0.0068)	(0.0067)
Black		-0.0727***	-0.0017	-0.0134**	-0.0003	0.0014
		(0.0056)	(0.0053)	(0.0053)	(0.0054)	(0.0053)
Asian		0.2418***	0.1373***	0.0923***	0.0929***	0.0868***
		(0.0124)	(0.0115)	(0.0116)	(0.0115)	(0.0113)
Other race		-0.0893***	-0.0311***	-0.0337***	-0.0251***	-0.0263***
		(0.0084)	(0.0080)	(0.0080)	(0.0080)	(0.0079)
Mixed race		-0.0319***	-0.0171	-0.0372***	-0.0335***	-0.0262***
		(0.0109)	(0.0104)	(0.0103)	(0.0104)	(0.0101)
Speaks English		0.2555***	0.1348***	0.1394***	0.1353***	0.1189***
		(0.0130)	(0.0125)	(0.0125)	(0.0125)	(0.0123)
U.S. citizen		0.0210**	0.0305***	0.0561***	0.0555***	0.0420***
		(0.0097)	(0.0091)	(0.0090)	(0.0090)	(0.0088)
Disabled		-0.1320***	-0.0700***	-0.0695***	-0.0684***	-0.0575***
		(0.0049)	(0.0046)	(0.0046)	(0.0046)	(0.0045)
High school			0.1911***	0.1858***	0.1799***	0.1405***
			(0.0056)	(0.0055)	(0.0055)	(0.0055)
Some college			0.3469***	0.3256***	0.3161***	0.2378***
			(0.0057)	(0.0056)	(0.0057)	(0.0057)
Associate's degree			0.4799***	0.4551***	0.4437***	0.3282***
			(0.0069)	(0.0069)	(0.0069)	(0.0070)
Bachelor's degree			0.7414***	0.6941***	0.6765***	0.5043***
			(0.0065)	(0.0065)	(0.0066)	(0.0069)
Postgraduate			0.9109***	0.8595***	0.8425***	0.6486***
			(0.0084)	(0.0084)	(0.0085)	(0.0090)
Metro. residence				0.1824***	0.1800***	0.1611***
				(0.0035)	(0.0035)	(0.0034)
No. of children					-0.0279***	-0.0260***
					(0.0018)	(0.0017)
Work characteristic controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	-	-	-	-	-	-
Region dummies	No	No	No	Yes	Yes	Yes
Occupation and industry controls	No	No	No	No	No	Yes
Observations	176,038	176,038	176,038	176,038	176,038	176,038
R^2	0.5352	0.5530	0.6097	0.6175	0.6181	0.6363

Notes: Robust standard errors in parantheses. Coefficients for work characteristics, and occupation and industry are omitted. ** and *** represent statistical significance at the 5% and 1% levels, respectively. The coefficient of interest is tested on a one-tailed basis. Source: 2000 Decennial Census PUMS.

Table A.4. Earnings comparison: lesbians and married heterosexuals (2000 Census)

	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)
Lesbian	0.1348***	0.1393***	0.0628***	0.0273***	0.0131**	0.0219***
	(0.0068)	(0.0067)	(0.0063)	(0.0062)	(0.0063)	(0.0062)
Potential experience		0.0238***	0.0227***	0.0218***	0.0232***	0.0207***
		(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Potential experience^2		-0.0005***	-0.0004***	-0.0003***	-0.0004***	-0.0003***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hispanic		-0.0601***	0.0215***	-0.0204***	-0.0165***	-0.0183***
		(0.0065)	(0.0062)	(0.0062)	(0.0062)	(0.0061)
Black		0.0029	0.0551***	0.0374***	0.0407***	0.0466***
		(0.0057)	(0.0053)	(0.0053)	(0.0053)	(0.0052)
Asian		0.2033***	0.1089***	0.0477***	0.0490***	0.0585***
		(0.0083)	(0.0078)	(0.0079)	(0.0079)	(0.0077)
Other race		-0.0788***	-0.0003	-0.0106	-0.0071	-0.0051
		(0.0082)	(0.0079)	(0.0079)	(0.0079)	(0.0077)
Mixed race		-0.0506***	-0.0156	-0.0423***	-0.0414***	-0.0286***
		(0.0116)	(0.0112)	(0.0111)	(0.0111)	(0.0108)
Speaks English		0.2841***	0.1357***	0.1502***	0.1441***	0.1254***
		(0.0100)	(0.0098)	(0.0098)	(0.0098)	(0.0097)
U.S. citizen		0.0914***	0.0723***	0.1019***	0.1020***	0.0801***
		(0.0077)	(0.0074)	(0.0074)	(0.0074)	(0.0072)
Disabled		-0.1217***	-0.0496***	-0.0481***	-0.0485***	-0.0369***
		(0.0044)	(0.0042)	(0.0042)	(0.0042)	(0.0041)
High school			0.1487***	0.1428***	0.1374***	0.0846***
			(0.0055)	(0.0055)	(0.0055)	(0.0054)
Some college			0.2969***	0.2765***	0.2700***	0.1671***
			(0.0057)	(0.0056)	(0.0057)	(0.0058)
Associate's degree			0.4438***	0.4200***	0.4135***	0.2688***
			(0.0066)	(0.0065)	(0.0065)	(0.0067)
Bachelor's degree			0.6558***	0.6151***	0.6063***	0.4143***
			(0.0061)	(0.0061)	(0.0061)	(0.0065)
Postgraduate			0.8775***	0.8285***	0.8181***	0.6045***
			(0.0068)	(0.0068)	(0.0069)	(0.0074)
Metro. residence				0.1848***	0.1841***	0.1691***
				(0.0029)	(0.0029)	(0.0029)
No. of children					-0.0198***	-0.0195***
					(0.0014)	(0.0014)
Work characteristic controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	-	-	-	-	-	-
Region dummies	No	No	No	Yes	Yes	Yes
Occupation and industry controls	No	No	No	No	No	Yes
Observations	255,503	255,503	255,503	255,503	255,503	255,503
R^2	0.5343	0.5461	0.6015	0.6103	0.6106	0.6306

Notes: Robust standard errors in parantheses. Coefficients for work characteristics, and occupation and industry are omitted. ** and *** represent statistical significance at the 5% and 1% levels, respectively. The coefficient of interest is tested on a one-tailed basis. Source: 2000 Decennial Census PUMS.

Table A.5. Earnings comparison: lesbians and cohabiting heterosexuals (05-07 ACS)

	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)
Lesbian	0.3396***	0.2725***	0.0856***	0.0728***	0.0706***	0.0657***
	(0.0079)	(0.0077)	(0.0072)	(0.0072)	(0.0072)	(0.0070)
Potential experience		0.0314***	0.0381***	0.0382***	0.0391***	0.0351***
		(0.0007)	(0.0006)	(0.0006)	(0.0007)	(0.0006)
Potential experience^2		-0.0005***	-0.0006***	-0.0006***	-0.0006***	-0.0006***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hispanic		-0.0398***	0.0322***	-0.0124	-0.0085	-0.0157**
		(0.0083)	(0.0078)	(0.0079)	(0.0079)	(0.0077)
Black		-0.1033***	-0.0301***	-0.0433***	-0.0384***	-0.0315***
		(0.0080)	(0.0073)	(0.0074)	(0.0075)	(0.0072)
Asian		0.2838***	0.1724***	0.1159***	0.1162***	0.1138***
		(0.0141)	(0.0130)	(0.0131)	(0.0131)	(0.0127)
Other race		-0.0748***	-0.0096	-0.0216**	-0.0188**	-0.0200**
		(0.0099)	(0.0094)	(0.0094)	(0.0094)	(0.0092)
Mixed race		-0.0161	0.0006	-0.0225	-0.0218	-0.0096
		(0.0154)	(0.0143)	(0.0143)	(0.0143)	(0.0138)
Speaks English		0.3756***	0.2145***	0.2200***	0.2180***	0.1720***
		(0.0145)	(0.0141)	(0.0140)	(0.0140)	(0.0138)
U.S. citizen		0.0640***	0.0532***	0.0740***	0.0731***	0.0496***
		(0.0114)	(0.0106)	(0.0105)	(0.0105)	(0.0102)
Disabled		-0.2306***	-0.1719***	-0.1636***	-0.1642***	-0.1499***
		(0.0093)	(0.0089)	(0.0089)	(0.0089)	(0.0086)
High school			0.2207***	0.2135***	0.2104***	0.1626***
			(0.0081)	(0.0080)	(0.0080)	(0.0079)
Some college			0.3774***	0.3544***	0.3499***	0.2563***
			(0.0084)	(0.0083)	(0.0083)	(0.0083)
Associate's degree			0.5392***	0.5152***	0.5098***	0.3661***
			(0.0097)	(0.0096)	(0.0097)	(0.0097)
Bachelor's degree			0.7912***	0.7404***	0.7319***	0.5281***
			(0.0090)	(0.0090)	(0.0090)	(0.0094)
Postgraduate			1.0097***	0.9516***	0.9432***	0.7006***
			(0.0106)	(0.0106)	(0.0107)	(0.0113)
Metro. residence				0.1801***	0.1788***	0.1572***
				(0.0045)	(0.0045)	(0.0044)
No. of children					-0.0151***	-0.0157***
					(0.0024)	(0.0023)
Work characteristic controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	No	No	No	Yes	Yes	Yes
Occupation and industry controls	No	No	No	No	No	Yes
Observations	117,901	117,901	117,901	117,901	117,901	117,901
R^2	0.5644	0.5864	0.6448	0.6523	0.6525	0.6734

Notes: Robust standard errors in parantheses. Coefficients for work characteristics, year dummies, and occupation and industry are omitted. ** and *** represent statistical significance at the 5% and 1% levels, respectively. The coefficient of interest is tested on a one-tailed basis. Sources: 2005-2007 American Community Surveys PUMS.

Table A.6. Earnings comparison: lesbians and married heterosexuals (05-07 ACS)

	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)
Lesbian	0.1267***	0.1226***	0.0601***	0.0287***	0.0212***	0.0345***
	(0.0078)	(0.0077)	(0.0071)	(0.0071)	(0.0071)	(0.0070)
Potential experience		0.0246***	0.0243***	0.0236***	0.0243***	0.0214***
		(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Potential experience^2		-0.0005***	-0.0004***	-0.0004***	-0.0004***	-0.0003***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hispanic		-0.0762***	0.0148*	-0.0317***	-0.0296***	-0.0246***
		(0.0081)	(0.0077)	(0.0077)	(0.0077)	(0.0075)
Black		-0.0479***	0.0087	-0.0101	-0.0089	0.0030
		(0.0076)	(0.0070)	(0.0070)	(0.0070)	(0.0068)
Asian		0.2244***	0.1205***	0.0544***	0.0548***	0.0657***
		(0.0094)	(0.0088)	(0.0089)	(0.0089)	(0.0085)
Other race		-0.0903***	-0.0167*	-0.0285***	-0.0273***	-0.0202**
		(0.0102)	(0.0096)	(0.0096)	(0.0096)	(0.0093)
Mixed race		-0.0631***	-0.0405**	-0.0603***	-0.0601***	-0.0549***
		(0.0180)	(0.0167)	(0.0165)	(0.0165)	(0.0162)
Speaks English		0.4060***	0.2178***	0.2334***	0.2300***	0.1776***
		(0.0120)	(0.0117)	(0.0117)	(0.0117)	(0.0115)
U.S. citizen		0.1177***	0.0969***	0.1171***	0.1174***	0.0834***
		(0.0096)	(0.0091)	(0.0090)	(0.0090)	(0.0087)
Disabled		-0.1850***	-0.1220***	-0.1124***	-0.1133***	-0.0988***
		(0.0082)	(0.0078)	(0.0078)	(0.0078)	(0.0075)
High school			0.1558***	0.1521***	0.1489***	0.0848***
			(0.0083)	(0.0082)	(0.0082)	(0.0081)
Some college			0.3082***	0.2900***	0.2867***	0.1656***
			(0.0086)	(0.0085)	(0.0085)	(0.0085)
Associate's degree			0.4724***	0.4541***	0.4506***	0.2783***
			(0.0093)	(0.0092)	(0.0092)	(0.0093)
Bachelor's degree			0.6735***	0.6340***	0.6297***	0.3996***
			(0.0088)	(0.0087)	(0.0088)	(0.0091)
Postgraduate			0.9421***	0.8956***	0.8905***	0.6269***
			(0.0094)	(0.0093)	(0.0093)	(0.0100)
Metro. residence				0.1892***	0.1890***	0.1736***
				(0.0038)	(0.0038)	(0.0036)
No. of children					-0.0114***	-0.0136***
					(0.0019)	(0.0019)
Work characteristic controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	No	No	No	Yes	Yes	Yes
Occupation and industry controls	No	No	No	No	No	Yes
Observations	171,817	171,817	171,817	171,817	171,817	171,817
R^2	0.5439	0.5591	0.6170	0.6251	0.6252	0.6499

Notes: Robust standard errors in parantheses. Coefficients for work characteristics, year dummies, and occupation and industry are omitted. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. The coefficient of interest is tested on a one-tailed basis. Sources: 2005-2007 American Community Surveys PUMS.

Table A.7. Earnings comparison: lesbians and cohabiting heterosexuals (08-10 ACS)

	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)
Lesbian	0.3263***	0.2566***	0.0777***	0.0672***	0.0657***	0.0623***
	(0.0076)	(0.0073)	(0.0068)	(0.0068)	(0.0068)	(0.0065)
Potential experience		0.0301***	0.0377***	0.0377***	0.0385***	0.0344***
		(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Potential experience^2		-0.0005***	-0.0006***	-0.0006***	-0.0006***	-0.0006***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hispanic		-0.0307***	0.0464***	0.0004	0.0034	-0.0049
		(0.0066)	(0.0062)	(0.0062)	(0.0063)	(0.0061)
Black		-0.1166***	-0.0450***	-0.0570***	-0.0532***	-0.0437***
		(0.0070)	(0.0065)	(0.0065)	(0.0066)	(0.0064)
Asian		0.2643***	0.1534***	0.0954***	0.0956***	0.0842***
		(0.0126)	(0.0115)	(0.0115)	(0.0115)	(0.0111)
Other race		-0.0914***	-0.0291***	-0.0414***	-0.0394***	-0.0376***
		(0.0089)	(0.0085)	(0.0085)	(0.0085)	(0.0083)
Mixed race		0.0015	-0.0007	-0.0235**	-0.0228*	-0.0251**
		(0.0128)	(0.0118)	(0.0118)	(0.0118)	(0.0114)
Speaks English		0.3608***	0.1943***	0.2015***	0.1997***	0.1641***
		(0.0135)	(0.0131)	(0.0130)	(0.0130)	(0.0128)
U.S. citizen		0.0683***	0.0524***	0.0715***	0.0703***	0.0474***
		(0.0108)	(0.0101)	(0.0100)	(0.0100)	(0.0096)
Disabled		-0.2410***	-0.1767***	-0.1678***	-0.1683***	-0.1517***
		(0.0095)	(0.0091)	(0.0090)	(0.0090)	(0.0088)
High school			0.1965***	0.1882***	0.1854***	0.1418***
			(0.0077)	(0.0076)	(0.0076)	(0.0076)
Some college			0.3315***	0.3109***	0.3073***	0.2223***
			(0.0078)	(0.0078)	(0.0078)	(0.0078)
Associate's degree			0.5100***	0.4875***	0.4831***	0.3462***
			(0.0090)	(0.0089)	(0.0090)	(0.0090)
Bachelor's degree			0.7586***	0.7080***	0.7010***	0.4946***
			(0.0084)	(0.0084)	(0.0085)	(0.0088)
Postgraduate			1.0016***	0.9424***	0.9356***	0.6882***
			(0.0098)	(0.0098)	(0.0098)	(0.0104)
Metro. residence				0.1620***	0.1609***	0.1447***
				(0.0041)	(0.0041)	(0.0040)
No. of children					-0.0112***	-0.0126***
					(0.0021)	(0.0020)
Work characteristic controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	No	No	No	Yes	Yes	Yes
Occupation and industry controls	No	No	No	No	No	Yes
Observations	127,303	127,303	127,303	127,303	127,303	127,303
R^2	0.5740	0.5963	0.6584	0.6659	0.6660	0.6868

Notes: Robust standard errors in parantheses. Coefficients for work characteristics, year dummies, and occupation and industry are omitted. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. The coefficient of interest is tested on a one-tailed basis. Sources: 2008-2010 American Community Surveys PUMS.

Table A.8. Earnings comparison: lesbians and married heterosexuals (08-10 ACS)

	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)
Lesbian	0.1002***	0.0984***	0.0518***	0.0234***	0.0225***	0.0379***
	(0.0075)	(0.0074)	(0.0067)	(0.0067)	(0.0068)	(0.0066)
Potential experience		0.0248***	0.0242***	0.0235***	0.0236***	0.0211***
		(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Potential experience^2		-0.0005***	-0.0004***	-0.0003***	-0.0004***	-0.0003***
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Hispanic		-0.0737***	0.0129**	-0.0319***	-0.0317***	-0.0261***
		(0.0068)	(0.0064)	(0.0065)	(0.0065)	(0.0063)
Black		-0.0719***	-0.0236***	-0.0383***	-0.0382***	-0.0221***
		(0.0073)	(0.0068)	(0.0068)	(0.0068)	(0.0065)
Asian		0.1958***	0.1041***	0.0402***	0.0402***	0.0504***
		(0.0090)	(0.0085)	(0.0086)	(0.0086)	(0.0082)
Other race		-0.1114***	-0.0255***	-0.0391***	-0.0390***	-0.0322***
		(0.0100)	(0.0095)	(0.0095)	(0.0095)	(0.0092)
Mixed race		-0.0456***	-0.0235	-0.0415***	-0.0415***	-0.0290**
		(0.0158)	(0.0148)	(0.0146)	(0.0146)	(0.0141)
Speaks English		0.4268***	0.2198***	0.2369***	0.2365***	0.1685***
		(0.0117)	(0.0117)	(0.0117)	(0.0117)	(0.0114)
U.S. citizen		0.1042***	0.0877***	0.1046***	0.1047***	0.0777***
		(0.0093)	(0.0089)	(0.0088)	(0.0088)	(0.0086)
Disabled		-0.1856***	-0.1251***	-0.1145***	-0.1146***	-0.0991***
		(0.0090)	(0.0086)	(0.0085)	(0.0085)	(0.0082)
High school			0.1522***	0.1468***	0.1462***	0.0837***
			(0.0083)	(0.0083)	(0.0083)	(0.0082)
Some college			0.2880***	0.2711***	0.2706***	0.1529***
			(0.0085)	(0.0085)	(0.0085)	(0.0085)
Associate's degree			0.4609***	0.4438***	0.4432***	0.2677***
			(0.0093)	(0.0092)	(0.0092)	(0.0093)
Bachelor's degree			0.6499***	0.6129***	0.6123***	0.3772***
			(0.0088)	(0.0087)	(0.0088)	(0.0091)
Postgraduate			0.9117***	0.8640***	0.8633***	0.5917***
			(0.0093)	(0.0092)	(0.0093)	(0.0098)
Metro. residence				0.1733***	0.1733***	0.1599***
				(0.0036)	(0.0036)	(0.0035)
No. of children					-0.0014	-0.0050***
					(0.0018)	(0.0018)
Work characteristic controls	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	No	No	No	Yes	Yes	Yes
Occupation and industry controls	No	No	No	No	No	Yes
Observations	169,240	169,240	169,240	169,240	169,240	169,240
R^2	0.5583	0.5737	0.6294	0.6373	0.6373	0.6634

Notes: Robust standard errors in parantheses. Coefficients for work characteristics, year dummies, and occupation and industry are omitted. ** and *** represent statistical significance at the 5% and 1% levels, respectively. The coefficient of interest is tested on a one-tailed basis. Sources: 2008-2010 American Community Surveys PUMS.

Table A.9. Sexual-orientation differences in wages

Panel A: Lesbians versus cohabiting heterosexuals

Sample:		Regression model (B1)	(B2)	(B3)	(B4)	(B5)	(B6)
2000 Census	Lesbian	0.3319***	0.2724***	0.0848***	0.0696***	0.0645***	0.0575***
		(0.0067)	(0.0065)	(0.0062)	(0.0061)	(0.0061)	(0.0060)
	Observations	176,038	176,038	176,038	176,038	176,038	176,038
	R^2	0.0153	0.0645	0.1960	0.2135	0.2150	0.2523
2005-2007 ACS	Lesbian	0.3859***	0.3014***	0.0965***	0.0835***	0.0811***	0.0734***
		(0.0076)	(0.0074)	(0.0069)	(0.0069)	(0.0069)	(0.0067)
	Observations	117,901	117,901	117,901	117,901	117,901	117,901
	R^2	0.0236	0.0887	0.2415	0.2595	0.2598	0.3094

Panel B: Lesbians versus married heterosexuals

Sample:		Regression model (B1)	(B2)	(B3)	(B4)	(B5)	(B6)
2000 Census	Lesbian	0.1588***	0.1637***	0.0769***	0.0417***	0.0297***	0.0303***
		(0.0067)	(0.0065)	(0.0060)	(0.0060)	(0.0061)	(0.0060)
	Observations	255,503	255,503	255,503	255,503	255,503	255,503
	R^2	0.0024	0.0319	0.1710	0.1911	0.1916	0.2357
2005-2007 ACS	Lesbian	0.1616***	0.1550***	0.0808***	0.0506***	0.0439***	0.0463***
		(0.0075)	(0.0074)	(0.0067)	(0.0067)	(0.0068)	(0.0066)
	Observations	171,817	171,817	171,817	171,817	171,817	171,817
	R^2	0.0029	0.0418	0.1947	0.2134	0.2135	0.2705

Notes: Robust standard errors in parentheses. Models include all variables specified in the text, but only the coefficient on the lesbian indicator is reported. *** represents statistical significance at the 1% level, in a one-tailed test. The dependent variable in each model is the log of real wages. Sources: 2000 Decennial Census and 2005-2007 American Community Surveys PUMS.

Table A.10. Sexual-orientation differences in income among full-time employed women

Panel A: Lesbians versus cohabiting heterosexuals

Sample:		Regression model (C1)	(C2)	(C3)	(C4)	(C5)	(C6)
	Lesbian	0.3536***	0.2958***	0.1022***	0.0858***	0.0806***	0.0763***
2000		(0.0070)	(0.0067)	(0.0062)	(0.0061)	(0.0061)	(0.0060)
Census	Observations	120,282	120,282	120,282	120,282	120,282	120,282
	R^2	0.0250	0.0919	0.2790	0.3040	0.3062	0.3611
	Lesbian	0.3915***	0.3161***	0.1130***	0.0998***	0.0976***	0.0940***
2005-2007		(0.0078)	(0.0076)	(0.0069)	(0.0068)	(0.0069)	(0.0066)
ACS	Observations	79,275	79,275	79,275	79,275	79,275	79,275
	R^2	0.0347	0.1084	0.3109	0.3351	0.3356	0.3999
	Lesbian	0.3877***	0.3097***	0.1014***	0.0884***	0.0861***	0.0838***
2008-2010		(0.0079)	(0.0076)	(0.0069)	(0.0068)	(0.0068)	(0.0065)
ACS	Observations	87,502	87,502	87,502	87,502	87,502	87,502
	R^2	0.0315	0.1028	0.3238	0.3481	0.3487	0.4102

Panel B: Lesbians versus married heterosexuals

Sample:		Regression model (C1)	(C2)	(C3)	(C4)	(C5)	(C6)
	Lesbian	0.2063***	0.2006***	0.1084***	0.0681***	0.0576***	0.0565***
2000		(0.0069)	(0.0067)	(0.0061)	(0.0060)	(0.0061)	(0.0060)
Census	Observations	163,946	163,946	163,946	163,946	163,946	163,946
	R^2	0.0062	0.0543	0.2387	0.2707	0.2713	0.3290
	Lesbian	0.2013***	0.1820***	0.1022***	0.0659***	0.0652***	0.0663***
2005-2007		(0.0078)	(0.0076)	(0.0067)	(0.0067)	(0.0068)	(0.0065)
ACS	Observations	110,803	110,803	110,803	110,803	110,803	110,803
	R^2	0.0067	0.0635	0.2628	0.2933	0.2933	0.3637
	Lesbian	0.1701***	0.1537***	0.0905***	0.0548***	0.0577***	0.0633***
2008-2010		(0.0078)	(0.0076)	(0.0067)	(0.0066)	(0.0067)	(0.0064)
ACS	Observations	114,293	114,293	114,293	114,293	114,293	114,293
	R^2	0.0047	0.0625	0.2605	0.2892	0.2892	0.3591

Notes: Robust standard errors in parentheses. Models include all variables specified in the text, but only the coefficient on the lesbian indicator is reported. *** represents statistical significance at the 1% level, in a one-tailed test. The dependent variable in each model is the log of real annual income and the sample is limited to those working at least 34 hours per week for a minimum of 40 weeks in the previous year. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.11. Heckman two-step estimates of the lesbian earnings premium

Panel A: Lesbians versus cohabiting heterosexuals

Sample:		Regression model (D1)	(D2)	(D3)	(D4)	(D5)	(D6)
2000 Census	Lesbian	0.1834***	0.1893***	0.0728***	0.0583***	0.0581***	0.0532***
		(0.0202)	(0.0110)	(0.0068)	(0.0066)	(0.0067)	(0.0064)
	Censored	39,659	39,659	39,659	39,659	39,659	39,659
	Uncensored	176,038	176,038	176,038	176,038	176,038	176,038
2005-2007 ACS	Lesbian	0.2104***	0.2171***	0.0794***	0.0675***	0.0673***	0.0629***
		(0.0232)	(0.0094)	(0.0072)	(0.0070)	(0.0070)	(0.0068)
	Censored	29,104	29,104	29,104	29,104	29,104	29,104
	Uncensored	117,901	117,901	117,901	117,901	117,901	117,901
2008-2010 ACS	Lesbian	0.1557***	0.1748***	0.0686***	0.0605***	0.0601***	0.0579***
		(0.0211)	(0.0093)	(0.0067)	(0.0065)	(0.0066)	(0.0063)
	Censored	35,006	35,006	35,006	35,006	35,006	35,006
	Uncensored	127,303	127,303	127,303	127,303	127,303	127,303

Panel B: Lesbians versus married heterosexuals

Sample:		Regression model (D1)	(D2)	(D3)	(D4)	(D5)	(D6)
2000 Census	Lesbian	0.0779***	0.2029***	0.1275***	0.0838***	0.0721***	0.0748***
		(0.0125)	(0.0074)	(0.0070)	(0.0068)	(0.0075)	(0.0072)
	Censored	114,515	114,515	114,515	114,515	114,515	114,515
	Uncensored	255,503	255,503	255,503	255,503	255,503	255,503
2005-2007 ACS	Lesbian	0.5777***	0.2433***	0.1478***	0.1093***	0.0908***	0.0956***
		(0.0332)	(0.0092)	(0.0082)	(0.0080)	(0.0087)	(0.0082)
	Censored	77,241	77,241	77,241	77,241	77,241	77,241
	Uncensored	171,817	171,817	171,817	171,817	171,817	171,817
2008-2010 ACS	Lesbian	0.6797***	0.2353***	0.1530***	0.1163***	0.1024***	0.1067***
		(0.0492)	(0.0092)	(0.0082)	(0.0079)	(0.0087)	(0.0081)
	Censored	76,384	76,384	76,384	76,384	76,384	76,384
	Uncensored	169,240	169,240	169,240	169,240	169,240	169,240

Notes: Robust standard errors in parentheses. Models include all variables specified in the text, but only the coefficient on the lesbian indicator is reported. *** represents statistical significance at the 1% level, in a one-tailed test. The dependent variable in each model is the log of real annual income. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.12. Lesbian premium estimates stratified by the presence of children

Sample:		Lesbians vs. Cohabs: Women with children	Lesbians vs. Cohabs: Women with no children	Lesbians vs. Marrieds: Women with children	Lesbians vs. Marrieds: Women with no children
2000 Census	Lesbian	0.0646***	0.0501***	0.0016	0.0438***
		(0.0142)	(0.0068)	(0.0143)	(0.0069)
	Observations	64,239	111,799	131,838	123,665
	R^2	0.6506	0.6051	0.6491	0.5993
2005-2007 ACS	Lesbian	0.0848***	0.0608***	0.0288*	0.0563***
		(0.0160)	(0.0078)	(0.0156)	(0.0078)
	Observations	37,989	79,912	79,524	92,293
	R^2	0.7038	0.6427	0.6773	0.6199
2008-2010 ACS	Lesbian	0.1048***	0.0493***	0.0581***	0.0509***
		(0.0148)	(0.0073)	(0.0143)	(0.0074)
	Observations	42,354	84,949	76,202	93,038
	R^2	0.7076	0.6626	0.6879	0.6397

Notes: Robust standard errors in parentheses. Models include all variables specified in the text, but only the coefficient on the lesbian indicator is reported. * and *** represent statistical significance at the 10% and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.13. Descriptive statistics for householders by sample and sexual orientation

	2000 Census			2005-2007 ACS			2008-2010 ACS		
Variables	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals
Work characteristics									
Annual real income	50,808	32,207	39,284	55,493	32,661	39,979	55,293	32,713	41,024
	(50,687)	(32,324)	(40,218)	(53,671)	(30,989)	(39,913)	(54,078)	(31,640)	(40,449)
Weeks worked	47.12	45.12	45.42	47.22	44.87	45.43	47.22	45.90	46.31
	(10.51)	(12.47)	(12.26)	(10.50)	(12.89)	(12.26)	(9.90)	(11.49)	(11.03)
Hours worked per week	41.92	39.29	38.71	41.81	38.82	37.56	41.09	38.17	37.41
	(10.01)	(9.70)	(10.68)	(10.63)	(10.00)	(11.18)	(10.56)	(10.03)	(11.06)
Personal characteristics									
Age	38.65	34.18	39.50	41.36	35.40	42.74	42.36	35.64	43.57
	(9.95)	(10.50)	(10.47)	(10.50)	(11.56)	(10.72)	(11.04)	(11.66)	(10.77)
Potential experience	18.99	16.00	20.69	21.30	16.91	23.56	22.24	17.01	24.28
	(9.84)	(10.72)	(10.76)	(10.39)	(11.83)	(11.13)	(10.86)	(11.99)	(11.23)
No high school	0.07	0.13	0.10	0.04	0.09	0.05	0.03	0.08	0.05
	(0.26)	(0.34)	(0.30)	(0.19)	(0.29)	(0.23)	(0.16)	(0.27)	(0.22)
High school	0.15	0.30	0.24	0.14	0.28	0.24	0.13	0.25	0.21
	(0.36)	(0.46)	(0.43)	(0.35)	(0.45)	(0.42)	(0.34)	(0.43)	(0.41)
Some college	0.23	0.28	0.24	0.21	0.28	0.22	0.22	0.29	0.23
	(0.42)	(0.45)	(0.42)	(0.41)	(0.45)	(0.42)	(0.41)	(0.45)	(0.42)
Associate	0.08	0.08	0.09	0.08	0.10	0.11	0.09	0.10	0.11
	(0.27)	(0.28)	(0.29)	(0.27)	(0.30)	(0.31)	(0.29)	(0.31)	(0.32)
Bachelor's	0.25	0.15	0.20	0.27	0.18	0.23	0.27	0.19	0.24
	(0.43)	(0.35)	(0.40)	(0.45)	(0.39)	(0.42)	(0.44)	(0.40)	(0.43)
Postgraduate	0.22	0.06	0.13	0.26	0.07	0.15	0.26	0.08	0.16
	(0.41)	(0.23)	(0.34)	(0.44)	(0.26)	(0.35)	(0.44)	(0.28)	(0.36)
Hispanic	0.08	0.10	0.10	0.08	0.12	0.09	0.09	0.13	0.10
	(0.28)	(0.30)	(0.30)	(0.27)	(0.32)	(0.28)	(0.28)	(0.34)	(0.30)
White	0.84	0.75	0.77	0.87	0.80	0.84	0.88	0.80	0.84
	(0.37)	(0.44)	(0.42)	(0.34)	(0.40)	(0.37)	(0.32)	(0.40)	(0.37)
Black	0.08	0.15	0.12	0.05	0.09	0.07	0.05	0.10	0.07
	(0.27)	(0.35)	(0.32)	(0.22)	(0.29)	(0.26)	(0.23)	(0.30)	(0.26)
Asian	0.01	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.04
	(0.11)	(0.13)	(0.17)	(0.14)	(0.14)	(0.18)	(0.13)	(0.15)	(0.19)

Table A.13 (continued)

Variables	2000 Census Lesbians	2000 Census Cohabiting heterosexuals	2000 Census Married heterosexuals	2005-2007 ACS Lesbians	2005-2007 ACS Cohabiting heterosexuals	2005-2007 ACS Married heterosexuals	2008-2010 ACS Lesbians	2008-2010 ACS Cohabiting heterosexuals	2008-2010 ACS Married heterosexuals
Other race	0.05	0.07	0.06	0.04	0.07	0.04	0.03	0.05	0.04
	(0.21)	(0.25)	(0.23)	(0.20)	(0.25)	(0.20)	(0.17)	(0.23)	(0.19)
Mixed race	0.02	0.03	0.02	0.02	0.02	0.01	0.02	0.03	0.02
	(0.15)	(0.16)	(0.15)	(0.14)	(0.14)	(0.11)	(0.14)	(0.16)	(0.12)
Disabled	0.14	0.16	0.17	0.08	0.08	0.07	0.07	0.05	0.05
	(0.34)	(0.37)	(0.37)	(0.28)	(0.27)	(0.25)	(0.25)	(0.23)	(0.22)
English proficient	0.97	0.96	0.94	0.98	0.95	0.96	0.98	0.95	0.96
	(0.17)	(0.19)	(0.23)	(0.13)	(0.21)	(0.20)	(0.12)	(0.21)	(0.21)
U.S. citizen	0.98	0.98	0.97	0.99	0.98	0.98	0.99	0.97	0.98
	(0.12)	(0.13)	(0.17)	(0.10)	(0.16)	(0.15)	(0.09)	(0.16)	(0.16)
No. own children under 18	0.34	0.77	0.96	0.30	0.66	0.90	0.32	0.69	0.91
	(0.78)	(1.07)	(1.13)	(0.71)	(1.01)	(1.09)	(0.76)	(1.03)	(1.11)
U.S. residence									
Northeast	0.22	0.21	0.23	0.20	0.20	0.19	0.20	0.20	0.20
	(0.41)	(0.41)	(0.42)	(0.40)	(0.40)	(0.40)	(0.40)	(0.40)	(0.40)
Midwest	0.18	0.24	0.21	0.19	0.24	0.24	0.19	0.24	0.23
	(0.38)	(0.42)	(0.41)	(0.39)	(0.43)	(0.43)	(0.40)	(0.43)	(0.42)
South	0.30	0.32	0.35	0.32	0.33	0.36	0.34	0.33	0.36
	(0.46)	(0.47)	(0.48)	(0.47)	(0.47)	(0.48)	(0.47)	(0.47)	(0.48)
West	0.31	0.23	0.22	0.29	0.24	0.21	0.27	0.24	0.22
	(0.46)	(0.42)	(0.41)	(0.45)	(0.43)	(0.41)	(0.44)	(0.43)	(0.41)
Metropolitan residence	0.84	0.74	0.74	0.83	0.75	0.72	0.83	0.75	0.71
	(0.37)	(0.44)	(0.44)	(0.37)	(0.44)	(0.45)	(0.37)	(0.43)	(0.45)
Occupation									
Mgt. and professionals	0.51	0.30	0.43	0.56	0.34	0.47	0.55	0.35	0.49
	(0.50)	(0.46)	(0.50)	(0.50)	(0.47)	(0.50)	(0.50)	(0.48)	(0.50)
Service occupations	0.13	0.20	0.15	0.12	0.21	0.13	0.12	0.22	0.14
	(0.34)	(0.40)	(0.35)	(0.32)	(0.41)	(0.34)	(0.33)	(0.42)	(0.34)
Sales	0.24	0.37	0.33	0.23	0.35	0.33	0.24	0.34	0.31
	(0.42)	(0.48)	(0.47)	(0.42)	(0.48)	(0.47)	(0.42)	(0.47)	(0.46)
Construction	0.03	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.01
	(0.18)	(0.10)	(0.09)	(0.16)	(0.09)	(0.08)	(0.15)	(0.09)	(0.07)

Table A.13 (continued)

	2000 Census			2005-2007 ACS			2008-2010 ACS		
Variables	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals	Lesbians	Cohabiting heterosexuals	Married heterosexuals
Production and transport.	0.09	0.11	0.09	0.06	0.08	0.06	0.07	0.07	0.05
	(0.29)	(0.32)	(0.28)	(0.25)	(0.28)	(0.24)	(0.25)	(0.26)	(0.22)
Industry									
Agriculture	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(0.07)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)
Construction	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	(0.15)	(0.12)	(0.13)	(0.15)	(0.13)	(0.14)	(0.13)	(0.12)	(0.13)
Manufacturing	0.10	0.13	0.11	0.09	0.09	0.08	0.08	0.08	0.07
	(0.31)	(0.33)	(0.31)	(0.28)	(0.29)	(0.27)	(0.27)	(0.27)	(0.26)
Wholesale trade	0.03	0.03	0.02	0.03	0.02	0.02	0.02	0.02	0.02
	(0.16)	(0.16)	(0.15)	(0.16)	(0.16)	(0.15)	(0.15)	(0.14)	(0.14)
Retail trade	0.09	0.13	0.10	0.09	0.13	0.10	0.09	0.13	0.09
	(0.29)	(0.34)	(0.30)	(0.28)	(0.34)	(0.30)	(0.29)	(0.34)	(0.29)
Transport. and utilities	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
	(0.20)	(0.17)	(0.17)	(0.18)	(0.17)	(0.17)	(0.18)	(0.17)	(0.17)
Information	0.04	0.04	0.03	0.03	0.03	0.02	0.03	0.02	0.02
	(0.21)	(0.18)	(0.18)	(0.18)	(0.16)	(0.15)	(0.17)	(0.15)	(0.14)
Finance and insurance	0.07	0.08	0.09	0.07	0.09	0.10	0.07	0.08	0.09
	(0.25)	(0.28)	(0.29)	(0.26)	(0.28)	(0.30)	(0.26)	(0.27)	(0.29)
Professional and mgt.	0.11	0.10	0.09	0.11	0.10	0.09	0.12	0.10	0.09
	(0.31)	(0.29)	(0.29)	(0.32)	(0.30)	(0.29)	(0.32)	(0.30)	(0.29)
Education and health	0.31	0.25	0.34	0.34	0.28	0.38	0.34	0.30	0.40
	(0.46)	(0.43)	(0.47)	(0.47)	(0.45)	(0.49)	(0.47)	(0.46)	(0.49)
Arts and entertainment	0.08	0.12	0.07	0.07	0.12	0.06	0.07	0.13	0.06
	(0.27)	(0.33)	(0.25)	(0.25)	(0.33)	(0.23)	(0.25)	(0.33)	(0.23)
Other services	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	(0.20)	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)	(0.18)	(0.19)	(0.19)
Public administration	0.07	0.05	0.06	0.07	0.04	0.06	0.08	0.05	0.06
	(0.25)	(0.21)	(0.23)	(0.25)	(0.20)	(0.23)	(0.27)	(0.21)	(0.23)
Number of observations	6,478	75,056	28,956	5,293	56,090	51,790	5,451	62,325	54,340

Notes: Standard errors in parantheses. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.14. Lesbian premium estimates by householder status

Sample:		Lesbians vs. Cohabs: Householders and residents	Lesbians vs. Cohabs: Householders only	Lesbians vs. Marrieds: Householders and residents	Lesbians vs. Marrieds: Householders only
2000 Census	Lesbian	0.0540***	0.0541***	0.0219***	0.0226***
		(0.0062)	(0.0086)	(0.0062)	(0.0094)
	Observations	176,038	81,534	255,503	35,434
	R^2	0.6363	0.6181	0.6306	0.5981
2005-2007 ACS	Lesbian	0.0657***	0.0698***	0.0345***	0.0396***
		(0.0070)	(0.0099)	(0.0070)	(0.0100)
	Observations	117,901	61,383	171,817	57,083
	R^2	0.6734	0.6614	0.6499	0.6331
2008-2010 ACS	Lesbian	0.0623***	0.0657***	0.0379***	0.0436***
		(0.0065)	(0.0092)	(0.0066)	(0.0095)
	Observations	127,303	67,776	169,240	59,791
	R^2	0.6868	0.6773	0.6634	0.6540

Notes: Robust standard errors in parentheses. Models include all variables specified in the text, but only the coefficient on the lesbian indicator is reported. *** represents statistical significance at the 1% level, in a one-tailed test. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.15. Assessing the maternity-risk hypothesis - preliminary regressions from the 2000 Census

	Comparative earnings regressions						Individual regressions		
	Lesbians vs. Cohabs			Lesbians vs. Marrieds			Lesbians	Cohabs	Marrieds
	(A1p)	(A2p)	(A3p)	(A1p)	(A2p)	(A3p)	(A2p)	(A2p)	(A2p)
Lesbian	0.0541***	0.0503***	0.0962***	0.0226***	0.0418***	0.0797***			
	(0.0086)	(0.0089)	(0.0126)	(0.0094)	(0.0117)	(0.0136)			
Maternity risk		-0.2494**	-0.5067***		0.4889***	-0.1387	-1.3921*	-0.5686***	-0.3911**
		(0.1376)	(0.1434)		(0.1603)	(0.1929)	(0.9989)	(0.1490)	(0.2231)
Maternity risk x lesbian			-2.1445***			-2.8253***			
			(0.3822)			(0.4740)			
Potential experience (centred)	0.0169***	0.0156***	0.0140***	0.0066***	0.0088***	0.0052***	0.0145***	0.0138***	0.0038***
	(0.0003)	(0.0008)	(0.0008)	(0.0004)	(0.0008)	(0.0010)	(0.0024)	(0.0008)	(0.0012)
Potential experience (centred) squared	-0.0006***	-0.0006***	-0.0005***	-0.0005***	-0.0006***	-0.0004***	-0.0007***	-0.0005***	-0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations	81,534	81,534	81,534	35,434	35,434	35,434	6,478	75,056	28,956
R^2	0.6181	0.6182	0.6183	0.5981	0.5982	0.5986	0.5601	0.6165	0.6017

Notes: Robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients most relevant to the analysis. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively. The coefficients of interest (*Lesbian* and *Maternity Risk*) are tested on a one-tailed basis. Source: 2000 Decennial Census PUMS.

Table A.16. Maternity risk by sample and sexual orientation: white women only

Sample	Age Group	Lesbians	Cohabiting Heterosexuals	Married Heterosexuals
2000 Census	18-24	0.0625	0.1255	0.1868
	25-31	0.0279	0.0639	0.1257
	32-38	0.0249	0.0340	0.0810
	39-45	0.0080	0.0110	0.0168
	46-52	0.0000	0.0004	0.0006
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000
2005-2007 ACS	18-24	0.0276	0.1314	0.1974
	25-31	0.0289	0.0771	0.1622
	32-38	0.0372	0.0473	0.0932
	39-45	0.0213	0.0133	0.0177
	46-52	0.0054	0.0007	0.0006
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000
2008-2010 ACS	18-24	0.0187	0.1390	0.2071
	25-31	0.0216	0.0825	0.1639
	32-38	0.0188	0.0588	0.0951
	39-45	0.0121	0.0128	0.0169
	46-52	0.0008	0.0008	0.0006
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000

Notes: Values represent number of births per woman. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.17. Maternity risk by sample and sexual orientation: never-married or currently married women only

Sample	Age Group	Lesbians	Cohabiting Heterosexuals	Married Heterosexuals
2000 Census	18-24	0.0802	0.1408	0.1996
	25-31	0.0362	0.0725	0.1267
	32-38	0.0264	0.0457	0.0781
	39-45	0.0093	0.0160	0.0158
	46-52	0.0000	0.0013	0.0011
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000
2005-2007 ACS	18-24	0.0331	0.1397	0.2011
	25-31	0.0286	0.0840	0.1611
	32-38	0.0414	0.0584	0.0908
	39-45	0.0230	0.0235	0.0177
	46-52	0.0063	0.0011	0.0007
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000
2008-2010 ACS	18-24	0.0262	0.1464	0.2029
	25-31	0.0251	0.0868	0.1612
	32-38	0.0202	0.0669	0.0944
	39-45	0.0074	0.0167	0.0171
	46-52	0.0010	0.0004	0.0007
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000

Notes: Values represent number of births per woman. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.18. Proportion of non-movers by age, sample and sexual orientation

Sample	Age Group	Lesbians	Cohabiting Heterosexuals	Married Heterosexuals
2005-2007 ACS	18-24	0.4038	0.3871	0.5594
	25-31	0.5714	0.5696	0.7722
	32-38	0.7489	0.6888	0.8704
	39-45	0.8449	0.7620	0.9203
	46-52	0.8690	0.8106	0.9394
	53-59	0.9003	0.8564	0.9455
	60-65	0.9003	0.8867	0.9493
2008-2010 ACS	18-24	0.3818	0.4096	0.5694
	25-31	0.5919	0.5951	0.7796
	32-38	0.7388	0.7051	0.8786
	39-45	0.8523	0.7713	0.9271
	46-52	0.8982	0.8208	0.9465
	53-59	0.9168	0.8715	0.9563
	60-65	0.9355	0.8954	0.9607

Sources: 2005-2010 American Community Surveys PUMS.

Table A.19. Maternity risk by sample and sexual orientation: non-moving women only

Sample	Age Group	Lesbians	Cohabiting Heterosexuals	Married Heterosexuals
2005-2007 ACS	18-24	0.0486	0.1595	0.1978
	25-31	0.0447	0.0900	0.1653
	32-38	0.0400	0.0555	0.0903
	39-45	0.0198	0.0143	0.0175
	46-52	0.0056	0.0004	0.0007
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000
2008-2010 ACS	18-24	0.0294	0.1596	0.2111
	25-31	0.0321	0.0937	0.1654
	32-38	0.0227	0.0654	0.0932
	39-45	0.0108	0.0128	0.0173
	46-52	0.0008	0.0007	0.0004
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000

Notes: Values represent number of births per woman. Sources: 2005-2010 American Community Surveys PUMS.

Table A.20. Maternity risk by sample and sexual orientation: ACS fertility variable

Sample	Age Group	Lesbians	Cohabiting Heterosexuals	Married Heterosexuals
2005-2007 ACS	18-24	0.0647	0.1644	0.2343
	25-31	0.0311	0.1032	0.1910
	32-38	0.0251	0.0587	0.1013
	39-45	0.0122	0.0171	0.0199
	46-52	0.0009	0.0021	0.0021
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000
2008-2010 ACS	18-24	0.0480	0.1710	0.2373
	25-31	0.0237	0.1046	0.1935
	32-38	0.0288	0.0727	0.1068
	39-45	0.0094	0.0162	0.0221
	46-52	0.0012	0.0040	0.0042
	53-59	0.0000	0.0000	0.0000
	60-65	0.0000	0.0000	0.0000

Notes: Values represent number of births per woman. Sources: 2005-2010 American Community Surveys PUMS.

Table A.21. Values of *Mandated Coverage* by state and year

	Year						
State	1999	2005	2006	2007	2008	2009	2010
Alabama	0	0	0	0	0	0	0
Alaska	0	0	0	0	0	0	0
Arizona	0	0	0	0	0	0	0
Arkansas	1	1	1	1	1	1	1
California	1	1	1	1	1	1	1
Colorado	0	0	0	0	0	0	0
Connecticut	0	1	1	1	1	1	1
Delaware	0	0	0	0	0	0	0
Florida	0	0	0	0	0	0	0
Georgia	0	0	0	0	0	0	0
Hawaii	1	1	1	1	1	1	1
Idaho	0	0	0	0	0	0	0
Illinois	1	1	1	1	1	1	1
Indiana	0	0	0	0	0	0	0
Iowa	0	0	0	0	0	0	0
Kansas	0	0	0	0	0	0	0
Kentucky	0	0	0	0	0	0	0
Louisiana	0	1	1	1	1	1	1
Maine	0	0	0	0	0	0	0
Maryland	0	1	1	1	1	1	1
Massachusetts	1	1	1	1	1	1	1
Michigan	0	0	0	0	0	0	0
Minnesota	0	0	0	0	0	0	0
Mississippi	0	0	0	0	0	0	0
Missouri	0	0	0	0	0	0	0
Montana	1	1	1	1	1	1	1
Nebraska	0	0	0	0	0	0	0
Nevada	0	0	0	0	0	0	0
New Hampshire	0	0	0	0	0	0	0
New Jersey	0	1	1	1	1	1	1
New Mexico	0	0	0	0	0	0	0
New York	1	1	1	1	1	1	1
North Carolina	0	0	0	0	0	0	0
North Dakota	0	0	0	0	0	0	0
Ohio	1	1	1	1	1	1	1
Oklahoma	0	0	0	0	0	0	0
Oregon	0	0	0	0	0	0	0
Pennsylvania	0	0	0	0	0	0	0
Rhode Island	1	1	1	1	1	1	1
South Carolina	0	0	0	0	0	0	0
South Dakota	0	0	0	0	0	0	0
Tennessee	0	0	0	0	0	0	0
Texas	1	1	1	1	1	1	1
Utah	0	0	0	0	0	0	0
Vermont	0	0	0	0	0	0	0
Virginia	0	0	0	0	0	0	0
Washington	0	0	0	0	0	0	0
West Virginia	1	1	1	1	1	1	1
Wisconsin	0	0	0	0	0	0	0
Wyoming	0	0	0	0	0	0	0

Sources: See Table D.2.

Table A.22. Values of *Legal Partnership* by state and year

State	Year 1999	2005	2006	2007	2008	2009	2010
Alabama	0	0	0	0	0	0	0
Alaska	0	0	0	0	0	0	0
Arizona	0	0	0	0	0	0	0
Arkansas	0	0	0	0	0	0	0
California	0	1	1	1	1	1	1
Colorado	0	0	0	0	0	1	1
Connecticut	0	1	1	1	1	1	1
Delaware	0	0	0	0	0	0	0
Florida	0	0	0	0	0	0	0
Georgia	0	0	0	0	0	0	0
Hawaii	1	1	1	1	1	1	1
Idaho	0	0	0	0	0	0	0
Illinois	0	0	0	0	0	0	0
Indiana	0	0	0	0	0	0	0
Iowa	0	0	0	0	0	1	1
Kansas	0	0	0	0	0	0	0
Kentucky	0	0	0	0	0	0	0
Louisiana	0	0	0	0	0	0	0
Maine	0	1	1	1	1	1	1
Maryland	0	0	0	0	1	1	1
Massachusetts	0	1	1	1	1	1	1
Michigan	0	0	0	0	0	0	0
Minnesota	0	0	0	0	0	0	0
Mississippi	0	0	0	0	0	0	0
Missouri	0	0	0	0	0	0	0
Montana	0	0	0	0	0	0	0
Nebraska	0	0	0	0	0	0	0
Nevada	0	0	0	0	0	1	1
New Hampshire	0	0	0	0	1	1	1
New Jersey	0	1	1	1	1	1	1
New Mexico	0	0	0	0	0	0	0
New York	0	0	0	0	0	0	0
North Carolina	0	0	0	0	0	0	0
North Dakota	0	0	0	0	0	0	0
Ohio	0	0	0	0	0	0	0
Oklahoma	0	0	0	0	0	0	0
Oregon	0	0	0	0	1	1	1
Pennsylvania	0	0	0	0	0	0	0
Rhode Island	0	0	0	0	0	0	0
South Carolina	0	0	0	0	0	0	0
South Dakota	0	0	0	0	0	0	0
Tennessee	0	0	0	0	0	0	0
Texas	0	0	0	0	0	0	0
Utah	0	0	0	0	0	0	0
Vermont	0	1	1	1	1	1	1
Virginia	0	0	0	0	0	0	0
Washington	0	0	0	1	1	1	1
West Virginia	0	0	0	0	0	0	0
Wisconsin	0	0	0	0	0	1	1
Wyoming	0	0	0	0	0	0	0

Sources: See Table D.2.

Table A.23. Values of *Marriage Bans* by state and year

	Year						
State	1999	2005	2006	2007	2008	2009	2010
Alabama	1	1	1	1	1	1	1
Alaska	1	1	1	1	1	1	1
Arizona	1	1	1	1	1	1	1
Arkansas	1	1	1	1	1	1	1
California	1	1	1	1	1	1	1
Colorado	0	1	1	1	1	1	1
Connecticut	1	1	1	1	0	0	0
Delaware	1	1	1	1	1	1	1
Florida	1	1	1	1	1	1	1
Georgia	1	1	1	1	1	1	1
Hawaii	1	1	1	1	1	1	1
Idaho	1	1	1	1	1	1	1
Illinois	1	1	1	1	1	1	1
Indiana	1	1	1	1	1	1	1
Iowa	0	0	0	0	0	0	0
Kansas	1	1	1	1	1	1	1
Kentucky	1	1	1	1	1	1	1
Louisiana	1	1	1	1	1	1	1
Maine	1	1	1	1	1	1	1
Maryland	1	1	1	1	1	1	1
Massachusetts	0	0	0	0	0	0	0
Michigan	1	1	1	1	1	1	1
Minnesota	1	1	1	1	1	1	1
Mississippi	1	1	1	1	1	1	1
Missouri	0	1	1	1	1	1	1
Montana	1	1	1	1	1	1	1
Nebraska	0	1	1	1	1	1	1
Nevada	0	1	1	1	1	1	1
New Hampshire	1	1	1	1	1	1	0
New Jersey	0	0	0	0	0	0	0
New Mexico	0	0	0	0	0	0	0
New York	0	0	0	0	0	0	0
North Carolina	1	1	1	1	1	1	1
North Dakota	1	1	1	1	1	1	1
Ohio	0	1	1	1	1	1	1
Oklahoma	1	1	1	1	1	1	1
Oregon	0	1	1	1	1	1	1
Pennsylvania	1	1	1	1	1	1	1
Rhode Island	0	0	0	0	0	0	0
South Carolina	1	1	1	1	1	1	1
South Dakota	1	1	1	1	1	1	1
Tennessee	1	1	1	1	1	1	1
Texas	1	1	1	1	1	1	1
Utah	1	1	1	1	1	1	1
Vermont	0	0	0	0	0	0	0
Virginia	1	1	1	1	1	1	1
Washington	1	1	1	1	1	1	1
West Virginia	0	1	1	1	1	1	1
Wisconsin	0	0	1	1	1	1	1
Wyoming	0	1	1	1	1	1	1

Notes: California allowed same-sex marriages between 16th June, 2008 and 4th November, 2008. Accounting for this by alternatively coding California as a zero for 2008 does not significantly affect the findings. Sources: See Table D.2.

Table A.24. Values of *Same-Sex Percentage* by state and year - prior to standardisation

	Year						
State	1999	2005	2006	2007	2008	2009	2010
Alabama	0.9758	0.9335	1.2589	0.8488	0.4029	0.7063	0.5519
Alaska	0.9966	1.9512	0.8157	0.6319	0.4796	1.1561	1.1211
Arizona	1.3059	1.3658	1.4699	1.5350	1.0319	1.5159	1.4978
Arkansas	0.6237	0.9664	0.7835	0.6645	0.4489	0.7657	0.6584
California	1.5388	1.5744	1.4606	1.5934	1.2932	1.2207	1.3582
Colorado	1.3264	1.2676	1.3726	1.6105	1.4147	1.4361	1.5821
Connecticut	1.2375	0.9756	1.3593	1.1917	1.1903	1.0868	1.2052
Delaware	1.1135	1.3308	1.8868	1.6393	0.9340	2.2458	0.8217
Florida	1.2363	1.4785	1.3133	1.4127	1.0534	1.1380	1.1696
Georgia	1.2710	1.3159	1.2776	1.2996	0.9576	1.1093	0.8913
Hawaii	1.3341	1.3387	1.3387	1.1693	0.7114	1.5887	1.2302
Idaho	0.6707	0.6390	0.8419	0.7934	0.8177	0.9952	0.5928
Illinois	0.8690	1.0033	1.0884	0.9216	0.8492	0.8051	0.7910
Indiana	0.8713	0.9864	1.2118	1.1223	0.9438	0.9102	0.9834
Iowa	0.6058	0.8661	0.8385	0.5571	0.5300	0.7347	0.8058
Kansas	0.8724	1.2407	1.0157	1.0675	0.5511	0.7016	0.8016
Kentucky	0.8872	1.3864	1.2527	0.9575	0.7870	0.7183	0.9601
Louisiana	1.2884	1.0591	1.0612	1.2837	0.6904	0.6217	0.8072
Maine	1.1620	1.5997	1.7009	1.1752	1.9132	1.2500	1.7076
Maryland	1.2354	1.4539	1.4501	1.4706	1.0813	1.1240	1.0955
Massachusetts	1.7080	1.9254	2.1914	2.3973	2.2413	1.8647	2.3916
Michigan	0.8359	0.8486	1.0087	1.2335	0.7581	0.7507	0.7223
Minnesota	0.8254	1.1832	1.1252	0.9821	0.6778	0.7075	0.7883
Mississippi	1.0752	0.6254	0.4966	0.8999	0.4443	0.5153	0.4475
Missouri	0.8510	1.0754	1.1435	0.9440	0.6176	0.8717	0.7126
Montana	0.7706	0.8722	0.9630	1.0357	0.7183	0.7392	0.5374
Nebraska	0.5690	0.6194	0.4831	0.6717	0.5825	0.5960	0.7375
Nevada	1.2600	1.5283	1.2031	1.0934	0.8283	1.1409	1.0967
New Hampshire	1.4910	2.3056	1.6445	1.3297	0.9904	1.4296	1.2262
New Jersey	1.0966	1.1331	1.3106	1.1902	0.8666	0.9410	1.0885
New Mexico	1.3127	2.0419	1.3517	1.8150	1.5358	1.6235	1.7914
New York	1.2795	1.4029	1.4646	1.4559	1.2433	1.1041	1.2079
North Carolina	1.0603	1.1661	1.0461	1.2587	0.9684	1.0451	1.0452
North Dakota	0.4278	0.7736	0.6479	0.6432	0.4181	0.6882	0.3425
Ohio	0.8544	1.3507	1.2204	1.2338	0.8588	0.9662	0.8450
Oklahoma	0.8038	0.9909	0.7853	0.7191	0.7422	0.8897	0.6557
Oregon	1.5001	1.4007	2.1139	1.8123	1.5413	1.7885	1.2238
Pennsylvania	0.8624	1.0268	1.1193	1.2661	0.6677	0.8718	0.7563
Rhode Island	1.4950	1.9422	1.4948	2.0597	1.3123	0.8763	2.2094
South Carolina	1.1741	1.3778	1.0404	1.1071	0.9379	0.7948	0.8468
South Dakota	0.6458	0.8600	1.3472	1.1536	0.3697	0.1208	0.3688
Tennessee	0.9964	0.9933	1.0826	0.9143	0.8208	0.8461	0.8407
Texas	1.0631	0.9909	1.1766	1.0199	0.8288	0.8994	1.1085
Utah	0.7907	0.7011	0.9743	0.8086	0.7090	0.7391	1.0189
Vermont	1.7651	1.7910	2.0202	0.8915	2.3152	2.5055	1.6937
Virginia	1.0053	1.1552	1.1567	1.1795	0.7214	0.7818	0.7500
Washington	1.4485	1.7905	1.3180	1.5750	1.4906	1.2881	1.2444
West Virginia	0.7780	0.9113	0.9441	1.0174	0.4925	0.5964	0.3814
Wisconsin	0.7610	1.3184	1.0490	0.9080	0.7918	0.7100	0.6815
Wyoming	0.8687	0.3503	1.1854	0.3419	0.1682	0.8258	0.3322

Sources: See Table D.2.

Table A.25. Base results: *Mandated Coverage*

		Lesbians versus cohabiting heterosexuals					Lesbians versus married heterosexuals				
Sample:		(A1)	(A2)	(A3)	(A4)	(A5)	(A1)	(A2)	(A3)	(A4)	(A5)
	Lesbian	0.0400***	0.0396***	0.0390***	0.0379***	0.0415***	0.0168**	0.0102	0.0103	0.0078	0.0099
		(0.0081)	(0.0090)	(0.0094)	(0.0096)	(0.0086)	(0.0089)	(0.0088)	(0.0091)	(0.0092)	(0.0084)
2000	Mandated coverage	0.0332**	0.0396**	0.0367**	0.0342**	0.0230*	0.0313*	0.0378**	0.0329**	0.0309**	0.0153
Census	x lesbian	(0.0197)	(0.0211)	(0.0179)	(0.0156)	(0.0149)	(0.0188)	(0.0213)	(0.0169)	(0.0148)	(0.0130)
	Observations	175,609	175,609	175,609	175,609	175,609	255,185	255,185	255,185	255,185	255,185
	R^2	0.6354	0.6368	0.6369	0.6370	0.6379	0.6294	0.6312	0.6314	0.6315	0.6322
	Lesbian	0.0559***	0.0580***	0.0572***	0.0572***	0.0597***	0.0377***	0.0328***	0.0321***	0.0291**	0.0327***
		(0.0115)	(0.0124)	(0.0123)	(0.0125)	(0.0115)	(0.0109)	(0.0122)	(0.0119)	(0.0121)	(0.0113)
2005 to	Mandated coverage	0.0162	0.0190	0.0216	0.0231	0.0170	0.0026	0.0043	0.0066	0.0088	-0.0050
2007 ACS	x lesbian	(0.0155)	(0.0177)	(0.0178)	(0.0181)	(0.0181)	(0.0164)	(0.0195)	(0.0195)	(0.0197)	(0.0193)
	Observations	117,600	117,600	117,600	117,600	117,600	171,584	171,584	171,584	171,584	171,584
	R^2	0.6723	0.6734	0.6734	0.6741	0.6750	0.6494	0.6507	0.6507	0.6511	0.6517
	Lesbian	0.0519***	0.0511***	0.0527***	0.0525***	0.0534***	0.0389***	0.0299***	0.0319***	0.0307***	0.0305***
		(0.0100)	(0.0106)	(0.0095)	(0.0095)	(0.0096)	(0.0090)	(0.0093)	(0.0087)	(0.0088)	(0.0080)
2008 to	Mandated coverage	0.0239**	0.0357***	0.0321***	0.0334***	0.0289***	0.0117	0.0245**	0.0202*	0.0213**	0.0141
2010 ACS	x lesbian	(0.0132)	(0.0126)	(0.0116)	(0.0114)	(0.0118)	(0.0146)	(0.0132)	(0.0121)	(0.0125)	(0.0111)
	Observations	126,966	126,966	126,966	126,966	126,966	169,034	169,034	169,034	169,034	169,034
	R^2	0.6857	0.6870	0.6870	0.6875	0.6885	0.6626	0.6642	0.6643	0.6645	0.6653

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the lesbian indicator and lesbian-proxy interaction term. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.26. *Mandated Coverage* robustness checks

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Large sample						
Base results	2.43*	1.56	1.82	-0.52	3.09***	1.46
Full-time only, no hours/weeks controls	2.89**	1.83	5.15***	4.62***	1.37*	0.70
Log(hourly wages)	2.35*	1.46	2.06	0.30	[redacted]	
Heckman two-step method	2.44*	1.60	1.68	0.09	2.01*	1.93**
Including sodomy law repeal	2.56*	1.61	1.94	-0.40	3.21***	1.52*
Further including SS legal equality score	2.33*	1.25	2.15	-0.37	3.32***	1.43
Adding potential experience interactions	2.78**	2.09*	1.96	-0.37	3.44***	1.91*
Further including education interactions	2.73**	2.02*	1.98	-0.21	3.39***	1.97*
Including only full-time employed	2.80**	1.97	5.14***	3.98**	1.43	0.61
Including only white women	2.89*	1.89*	2.51	0.22	3.55***	2.17**
Including only women aged 40 or under	3.00**	1.90	0.60	-2.32	3.01**	2.21*
Including only women over 40 years old	0.52	1.25	2.20	1.44	2.96*	1.62
Smaller sample						
Base results	2.38*	2.16	-0.57	-2.15	3.22**	2.02*
Multilevel estimation	1.13	0.76	-1.23	-4.48**	2.78*	1.15
Gay-male comparison						
Female estimates	2.43*	1.56	1.82	-0.52	3.09***	1.46
Male estimates	6.06***	4.44***	7.46***	3.70**	7.39***	2.93*
Difference (Female - Male)	-3.64	-2.88	-5.64	-4.21	-4.30	-1.47

Notes: Models include all variables specified in the text, but we report only the percentage point change in the lesbian pay gap as a result of the proxy moving from a value of zero to one, evaluated at the means of all other variables. *, **, and *** represent statistical significance for the lesbian-proxy interaction term at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.27. Base results: *Legal Partnership*

Sample:		Lesbians versus cohabiting heterosexuals (A1)	(A2)	(A3)	(A4)	(A5)	Lesbians versus married heterosexuals (A1)	(A2)	(A3)	(A4)	(A5)
	Lesbian	0.0670***	0.0691***	0.0692***	0.0697***	0.0719***	0.0398***	0.0358***	0.0359***	0.0341***	0.0364***
		(0.0083)	(0.0087)	(0.0088)	(0.0088)	(0.0084)	(0.0096)	(0.0094)	(0.0095)	(0.0092)	(0.0090)
2005 to	Legal partnership	-0.0162	-0.0158	-0.0160	-0.0147	-0.0261	-0.0255**	-0.0199	-0.0201	-0.0199	-0.0368**
2007 ACS	x lesbian	(0.0211)	(0.0238)	(0.0242)	(0.0238)	(0.0215)	(0.0145)	(0.0170)	(0.0173)	(0.0166)	(0.0176)
	Observations	117,600	117,600	117,600	117,600	117,600	171,584	171,584	171,584	171,584	171,584
	R^2	0.6736	0.6743	0.6743	0.6749	0.6754	0.6507	0.6516	0.6516	0.6519	0.6522
	Lesbian	0.0661***	0.0670***	0.0652***	0.0658***	0.0728***	0.0426***	0.0367***	0.0342***	0.0330***	0.0381***
		(0.0095)	(0.0093)	(0.0083)	(0.0086)	(0.0085)	(0.0088)	(0.0083)	(0.0077)	(0.0074)	(0.0053)
2008 to	Legal partnership	-0.0131	-0.0045	0.0013	0.0002	-0.0249*	-0.0134	0.0020	0.0095	0.0106	-0.0097
2010 ACS	x lesbian	(0.0200)	(0.0209)	(0.0182)	(0.0183)	(0.0178)	(0.0170)	(0.0162)	(0.0136)	(0.0134)	(0.0108)
	Observations	126,966	126,966	126,966	126,966	126,966	169,034	169,034	169,034	169,034	169,034
	R^2	0.6871	0.6878	0.6880	0.6886	0.6888	0.6637	0.6648	0.6649	0.6652	0.6654

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the lesbian indicator and lesbian-proxy interaction term. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2005-2010 American Community Surveys PUMS.

Table A.28. *Legal Partnership* robustness checks

	2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds
Large sample				
Base results	-2.77	-3.75**	-2.64*	-1.00
Full-time only, no hours/weeks controls	-1.75	-3.32**	-1.40	-0.62
Log(hourly wages)	-3.36*	-4.17**	[redacted]	
Heckman two-step method	-2.76	-3.68**	-4.58***	-0.13
Including sodomy law repeal	-2.76	-3.73**	-2.61*	-0.97
Further including SS legal equality score	-1.93	-3.34*	-2.92*	-1.24
Adding potential experience interactions	-3.10*	-4.38***	-3.50**	-2.24**
Further including education interactions	-3.06*	-4.54***	-3.48**	-2.24**
Including only full-time employed	-1.67	-3.08**	-1.29	-0.59
Including only white women	-2.12	-3.15**	-2.91*	-1.63
Including only women aged 40 or under	-4.07**	-4.34*	-0.94	1.15
Including only women over 40 years old	-1.38	-3.97	-6.52***	-4.44**
Smaller sample				
Base results	-5.22**	-3.40*	-1.56	-1.77
Multilevel estimation	-8.24***	-5.77**	-1.56	-1.42
Gay-male comparison				
Female estimate	-2.77	-3.75**	-2.64*	-1.00
Male estimate	3.54	-0.32	-0.40	-0.06
Difference (Female - Male)	-6.31	-3.42	-2.25	-0.94

Notes: Models include all variables specified in the text, but we report only the percentage point change in the lesbian pay gap as a result of the proxy moving from a value of zero to one, evaluated at the means of all other variables. *, **, and *** represent statistical significance for the lesbian-proxy interaction term at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2005-2010 American Community Surveys PUMS.

Table A.29. Base results: *Marriage Bans*

Sample:		Lesbians versus cohabiting heterosexuals (A1)	(A2)	(A3)	(A4)	(A5)	Lesbians versus married heterosexuals (A1)	(A2)	(A3)	(A4)	(A5)
	Lesbian	0.0388**	0.0269*	0.0282**	0.0303**	0.0337**	0.0134	0.0001	0.0017	0.0030	0.0040
		(0.0178)	(0.0190)	(0.0165)	(0.0150)	(0.0161)	(0.0138)	(0.0172)	(0.0170)	(0.0155)	(0.0157)
2000	Marriage bans	0.0229	0.0375*	0.0341*	0.0284*	0.0226	0.0245	0.0333*	0.0293*	0.0229	0.0161
Census	x lesbian	(0.0237)	(0.0252)	(0.0205)	(0.0201)	(0.0202)	(0.0212)	(0.0257)	(0.0210)	(0.0204)	(0.0195)
	Observations	175,609	175,609	175,609	175,609	175,609	255,185	255,185	255,185	255,185	255,185
	R^2	0.6352	0.6368	0.6368	0.6370	0.6379	0.6293	0.6312	0.6313	0.6315	0.6322
	Lesbian	0.0411**	0.0207	0.0210	0.0184	0.0230	0.0168	-0.0076	-0.0071	-0.0167	-0.0139
		(0.0233)	(0.0285)	(0.0292)	(0.0297)	(0.0285)	(0.0241)	(0.0289)	(0.0297)	(0.0286)	(0.0281)
2005 to	Marriage bans	0.0294	0.0543**	0.0541**	0.0577**	0.0511**	0.0284	0.0498*	0.0494*	0.0579**	0.0507*
2007 ACS	x lesbian	(0.0255)	(0.0310)	(0.0316)	(0.0320)	(0.0301)	(0.0265)	(0.0325)	(0.0332)	(0.0321)	(0.0312)
	Observations	117,600	117,600	117,600	117,600	117,600	171,584	171,584	171,584	171,584	171,584
	R^2	0.6719	0.6732	0.6732	0.6739	0.6750	0.6487	0.6504	0.6504	0.6508	0.6517
	Lesbian	0.0636***	0.0814***	0.0832***	0.0869***	0.0907***	0.0224*	0.0319*	0.0352**	0.0331**	0.0363**
		(0.0211)	(0.0246)	(0.0237)	(0.0259)	(0.0254)	(0.0153)	(0.0193)	(0.0184)	(0.0192)	(0.0182)
2008 to	Marriage bans	-0.0002	-0.0170	-0.0191	-0.0231	-0.0290	0.0250*	0.0101	0.0063	0.0080	-0.0002
2010 ACS	x lesbian	(0.0227)	(0.0263)	(0.0253)	(0.0271)	(0.0259)	(0.0176)	(0.0227)	(0.0213)	(0.0222)	(0.0196)
	Observations	126,966	126,966	126,966	126,966	126,966	169,034	169,034	169,034	169,034	169,034
	R^2	0.6854	0.6866	0.6867	0.6871	0.6884	0.6625	0.6640	0.6640	0.6642	0.6651

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the lesbian indicator and lesbian-proxy interaction term. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.30. *Marriage Bans* robustness checks

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Large sample						
Base results	2.36	1.63	5.36**	5.13*	-3.13	-0.02
Full-time only, no hours/weeks controls	6.12***	6.17***	4.63**	4.12*	-1.80	-1.30
Log(hourly wages)	3.80**	3.07**	3.84	2.96	[redacted]	
Heckman two-step method	2.37	1.66	5.41**	4.98*	-3.51	0.29
Including sodomy law repeal	2.50	2.00	5.31**	5.05*	-3.11	-0.03
Further including SS legal equality score	2.91	2.49*	5.17**	4.69*	-3.30	-0.32
Adding potential experience interactions	2.44	1.75	5.51**	5.32**	-3.89	-1.06
Further including education interactions	2.30	1.54	5.52**	5.33*	-3.88	-1.15
Including only full-time employed	5.41***	5.45***	4.05*	4.34*	-1.98	-0.70
Including only white women	2.55	1.58	6.38**	6.90**	-2.68	0.11
Including only women aged 40 or under	3.37*	2.53	3.12	2.76	-0.22	3.05
Including only women over 40 years old	2.41	0.47	7.38**	6.45*	-4.38	-3.58
Smaller sample						
Base results	2.79	3.57*	2.85	6.12**	-1.33	2.53
Multilevel estimation	2.46	3.58*	2.57	6.12**	-1.97	2.09
Gay-male comparison						
Female estimates	2.36	1.63	5.36**	5.13*	-3.13	-0.02
Male estimates	-1.39	-1.71	3.57	2.97	-6.95	-3.83
Difference (Female - Male)	3.75	3.34	1.79	2.16	3.82	3.81

Notes: Models include all variables specified in the text, but we report only the percentage point change in the lesbian pay gap as a result of the proxy moving from a value of zero to one, evaluated at the means of all other variables. *, **, and *** represent statistical significance for the lesbian-proxy interaction term at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.31. Base results: *Same-Sex Percentage*

Sample:		Lesbians versus cohabiting heterosexuals (A1)	(A2)	(A3)	(A4)	(A5)	Lesbians versus married heterosexuals (A1)	(A2)	(A3)	(A4)	(A5)
	Lesbian	0.0493***	0.0521***	0.0522***	0.0523***	0.0530***	0.0202***	0.0192**	0.0195**	0.0196***	0.0184***
		(0.0062)	(0.0060)	(0.0063)	(0.0061)	(0.0061)	(0.0084)	(0.0081)	(0.0082)	(0.0080)	(0.0068)
2000	Same-sex percentage	0.0025	0.0030	-0.0009	-0.0149	-0.0245**	-0.0043	-0.0016	-0.0060	-0.0231**	-0.0364***
Census	x lesbian	(0.0128)	(0.0132)	(0.0114)	(0.0136)	(0.0125)	(0.0116)	(0.0123)	(0.0104)	(0.0122)	(0.0117)
	Observations	175,609	175,609	175,609	175,609	175,609	255,185	255,185	255,185	255,185	255,185
	R^2	0.6359	0.6370	0.6370	0.6370	0.6379	0.6305	0.6317	0.6317	0.6317	0.6323
	Lesbian	0.0631***	0.0672***	0.0672***	0.0686***	0.0683***	0.0331***	0.0330***	0.0331***	0.0324***	0.0303***
		(0.0073)	(0.0072)	(0.0072)	(0.0072)	(0.0075)	(0.0075)	(0.0074)	(0.0074)	(0.0071)	(0.0071)
2005 to	Same-sex percentage	-0.0084	-0.0123*	-0.0125*	-0.0130*	-0.0147**	-0.0111*	-0.0133**	-0.0137**	-0.0169***	-0.0188***
2007 ACS	x lesbian	(0.0075)	(0.0082)	(0.0080)	(0.0078)	(0.0064)	(0.0068)	(0.0071)	(0.0071)	(0.0064)	(0.0057)
	Observations	117,600	117,600	117,600	117,600	117,600	171,584	171,584	171,584	171,584	171,584
	R^2	0.6726	0.6735	0.6735	0.6741	0.6751	0.6496	0.6507	0.6507	0.6510	0.6518
	Lesbian	0.0617***	0.0669***	0.0672***	0.0676***	0.0684***	0.0387***	0.0391***	0.0395***	0.0391***	0.0383***
		(0.0076)	(0.0075)	(0.0061)	(0.0062)	(0.0071)	(0.0075)	(0.0073)	(0.0060)	(0.0061)	(0.0058)
2008 to	Same-sex percentage	-0.0159***	-0.0142***	-0.0150***	-0.0148***	-0.0215***	-0.0171***	-0.0129**	-0.0141***	-0.0143***	-0.0214***
2010 ACS	x lesbian	(0.0058)	(0.0057)	(0.0048)	(0.0050)	(0.0054)	(0.0052)	(0.0058)	(0.0050)	(0.0051)	(0.0066)
	Observations	126,966	126,966	126,966	126,966	126,966	169,034	169,034	169,034	169,034	169,034
	R^2	0.6858	0.6869	0.6869	0.6873	0.6884	0.6626	0.6641	0.6641	0.6643	0.6652

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the lesbian indicator and lesbian-proxy interaction term. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.32. *Same-Sex Percentage* robustness checks

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Large sample						
Base results	-2.55**	-3.65***	-1.67**	-1.92***	-2.28***	-2.20***
Full-time only, no hours/weeks controls	-2.67**	-3.82***	-1.53**	-2.22***	-1.46***	-1.87***
Log(hourly wages)	-2.27**	-3.54***	-1.66**	-2.13***	[redacted]	
Heckman two-step method	-2.61**	-3.51***	-1.56**	-1.84***	-2.49***	-2.09***
Including sodomy law repeal	-2.64**	-3.74***	-1.66**	-2.10***	-2.30***	-2.23***
Further including SS legal equality score	-2.75**	-3.93***	-1.55**	-2.09***	-2.26***	-2.15***
Adding potential experience interactions	-2.72**	-3.95***	-1.73***	-2.21***	-2.40***	-2.39***
Further including education interactions	-2.60**	-3.87***	-1.68***	-2.14***	-2.39***	-2.41***
Including only full-time employed	-2.48**	-3.60***	-1.38**	-1.94***	-1.36***	-1.72***
Including only white women	-3.19***	-4.21***	-2.00***	-2.31***	-2.48***	-2.46***
Including only women aged 40 or under	-2.84***	-4.04***	-3.68***	-3.53***	-1.52**	-1.17*
Including only women over 40 years old	-2.09	-3.33	0.61	-0.43	-3.20***	-3.14***
Smaller sample						
Base results	-2.37**	-4.08***	-1.53**	-0.87	-2.03**	-2.26***
Multilevel estimation	-1.35	-2.95**	-1.39	-0.75	-1.79**	-1.82**
Gay-male comparison						
Female estimates	-2.55**	-3.65**	-1.67**	-1.92***	-2.28***	-2.20***
Male estimates	4.40***	2.43**	2.76**	1.48*	2.12**	0.84
Difference (Female - Male)	-6.96	-6.08	-4.32	-3.40	-4.39	-3.04

Notes: Models include all variables specified in the text, but we report only the percentage point change in the lesbian pay gap as a result of the proxy moving from a value of zero to one (one standard deviation), evaluated at the means of all other variables. *, **, and *** represent statistical significance for the lesbian-proxy interaction term at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.33. Maternity risk and the gender pay gap - cohabiting heterosexuals

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Excluding mat. risk	Including mat. risk	Excluding mat. risk	Including mat. risk	Excluding mat. risk	Including mat. risk
Female	-0.1970***	-0.1491***	-0.1440***	-0.1057***	-0.1382***	-0.1111***
	(0.0046)	(0.0073)	(0.0064)	(0.0101)	(0.0058)	(0.0091)
Maternity risk		-1.1090***		-0.8640***		-0.5800***
		(0.1344)		(0.1770)		(0.1484)
Observations	171,890	171,890	111,893	111,893	122,383	122,383
R^2	0.5584	0.5586	0.6119	0.6121	0.6327	0.6329

Notes: Robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficient on the female indicator and maternity risk variable. *** represents statistical significance at the 1% level, in a one-tailed test. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table A.34. Maternity risk and the gender pay gap - married heterosexuals

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Excluding mat. risk	Including mat. risk	Excluding mat. risk	Including mat. risk	Excluding mat. risk	Including mat. risk
Female	-0.3301***	-0.2942***	-0.2755***	-0.2641***	-0.2706***	-0.2631***
	(0.0060)	(0.0076)	(0.0056)	(0.0062)	(0.0054)	(0.0058)
Maternity risk		-1.4804***		-0.6642***		-0.5360***
		(0.1883)		(0.1417)		(0.1307)
Observations	284,013	284,013	180,024	180,024	175,631	175,631
R^2	0.4603	0.4604	0.5493	0.5494	0.5819	0.5820

Notes: Robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficient on the female indicator and maternity risk variable. *** represents statistical significance at the 1% level, in a one-tailed test. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Appendix B: Previous Evidence Regarding the Maternity-Risk Hypothesis

Academics use several methods to assess whether sexual-orientation differences in labour-force attachment contribute to the observed lesbian premium. As noted, a common feature of these methods is their focus on backward-looking effects of child-bearing on earnings, as opposed to forward-looking expectations of maternity-leave requirements. The current literature does, however, contain results consistent with the maternity-risk hypothesis, as defined in this study. This note briefly highlights the available evidence in the current literature regarding the maternity-risk hypothesis.[108]

Other than assessing the coefficient on interaction terms and the results from age-stratified regressions, perhaps the most convincing evidence regarding the maternity-risk hypothesis is presented by Carpenter (2008a). Carpenter estimates a lesbian premium of approximately 17% among all women. Stratifying his results by partnership status, he finds that this premium is primarily due to a large lesbian advantage among partnered women (43%), while single lesbians receive a premium of only 1%. Carpenter argues that these results provide evidence of differential selection into partnership by sexual orientation. An alternative, or even complementary, explanation is that a larger sexual-orientation maternity gap among partnered women results in a greater lesbian premium.

Ahmed and Hammarstedt (2010) and Ahmed et al. (2013b) also provide suggestive evidence in favour of the maternity-risk hypothesis. Their analyses, which compare the earnings of lesbians in civil unions and married heterosexual women in Sweden, reveal lesbians are at an earnings disadvantage in approximately half of their regressions. Given that Sweden is recognised as one of the most gay-friendly countries in the world (Clarke, 2013), we may expect a larger lesbian premium due to less labour-market discrimination. As argued in Section 5, however, expanded same-sex partnership rights likely enhances family formation among lesbian couples, thereby reducing sexual-orientation differences in maternity incidence.[109] The finding of no consistent lesbian premium thus provides further evidence that employers are averse to bearing the risk of employees taking maternity leave.

[108] As discussed in the text, using an audit study, Baert (2013) finds evidence of hiring discrimination consistent with the maternity-risk hypothesis.

[109] Lesbian couples are also able to adopt children and have equal access to artificial reproduction technologies. Moreover, artificial insemination expenses are covered by the national health care (Ahmed et al., 2013b).

Daneshvary et al. (2009) and Plug and Berkhout (2004) also obtain results that are broadly consistent with the maternity-risk hypothesis, however a causal link is difficult to establish due to the presence of several alternative (and arguably better) explanations.[110] Evidently, the current literature contains some evidence in favour of the maternity-risk hypothesis, but difficulties in empirical estimation have thus far precluded a more direct assessment. In this study, we attempt to provide such an assessment.

[110] Daneshvary et al. (2008) note that their results are consistent with a narrowing of the sexual-orientation gap in labour-force attachment among women with high levels of education. Examining the data, we find no significant reduction in the maternity gap between lesbians and partnered heterosexual women (the comparison group in their study) among women with high educational attainment. This suggests that some other mechanism is likely driving their results.

Appendix C: On the Identification of Lesbians

Research examining the sexual-orientation pay gap is limited due to both sparse representative data and difficulties in defining the meaning of "lesbian" or "gay". Due in part to the increased attention being paid to sexual-orientation discrimination, datasets allowing an assessment of sexual orientation are becoming increasingly available, while the methods used to infer sexual orientation are continually expanding. This note briefly discusses the methods commonly used in previous studies and outlines the advantages and potential problems associated with the identification procedure implemented in this study.[111]

As is evident from Section 2.2, the prior literature defines sexual orientation in three distinct ways. Several studies, particularly those using GSS data, use sexual behaviour to identify lesbians. The definitions used, however, vary with respect to the time horizon over which sexual behaviour is classified and exclusivity of same-sex sexual behaviour. Other studies including Carpenter (2005, 2008a, 2008b) and Frank (2006), use self-reported sexual orientation. The third identification method categorises sexual orientation based on inference from living arrangements. Beginning with the 1990 Census, respondents were able to classify an individual within the household as their "unmarried partner", facilitating the identification of same-sex couples. This so-called cohabitation procedure has also been applied in studies using UK Labour Force Survey and French Employment Survey data.

Although data availability, rather than theoretical considerations, generally determines which identification method is implemented, we consider the cohabitation procedure to be ideal for the purposes of this study. First, one of the major concerns in the literature surrounds the ability of the employer to accurately observe sexual orientation. As sexual orientation is not readily observable, like gender and race, employers must infer sexual orientation based on appearances and mannerisms; affiliations with lesbian, gay, bisexual and transgender advocacy groups; conversations; and information obtained from social networking websites. The accuracy of any study analysing differences in labour-market outcomes by sexual orientation therefore hinges critically on accurate inference on behalf of employers. Jepsen (2007) notes that same-sex female couples represent the group most likely to be correctly identified as lesbians by employers. If two cohabiting females identify as "unmarried

[111] Badgett (2006) provides a thorough discussion of this and a number of other issues pertaining to examining the sexual-orientation pay gap. These issues include endogenous disclosure of sexual orientation and possible explanations for the contrasting findings for gay males and lesbians.

partners", it is likely that their sexuality will be known by the employer due to voluntary or involuntary disclosure. Moreover, use of the cohabitation procedure largely circumvents the ambiguity problem associated with sexual-behaviour based classification systems.

Second, although the sample of cohabiting lesbians does not represent a random sample of all lesbians, the comparison of women in cohabiting partnerships arguably minimises unobservable heterogeneity by considering only those women that can share household responsibilities with a spouse or partner.[112] Third, information pertaining to the respondent's partner may be useful in identifying appropriate exclusion restrictions to address sample-selectivity concerns. Finally, partnered women are significantly more likely than single women to bear children and thus require maternity leave. Given that this study focuses on differential maternity risk as an explanation for the lesbian premium, coupled females (especially those likely to be perceived as such) represent the most appropriate sample for the purposes of this study.

One potential problem associated with the cohabitation procedure, as noted in previous studies, is that of measurement error in classifying same-sex couples. Specifically, in the 2000 Census and subsequent American Community Surveys, the U.S. Census Bureau did not allow same-sex spouse combinations to occur. As a result, whenever the household head identified a resident of the same sex as their "husband/wife", the Census Bureau altered the response to "unmarried partner", provided the sex of either partner had not previously been allocated (U.S. Census Bureau, 2002).[113] To ensure consistency across questions, the Census Bureau also allocated the marital status of the partners in such a case. An unfortunate side-effect of this allocation procedure is that opposite-sex married couples that miscoded the sex of one partner are incorrectly allocated to the same-sex couples group. Given the small size of the lesbian population relative to married heterosexuals, minor misclassification could significantly contaminate the sample of lesbians.[114] Similarly to Black et al. (2007) and Daneshvary et al. (2009), we assuage these concerns by dropping both partners' observations

[112] As partnership rates differ by sexual orientation, selection into partnership may play a role in the observed earnings differentials. Unfortunately, census data cannot be used to address this issue.
[113] As this represents a so-called "logical edit", the relationship status allocation flag in the PUMS files does not capture this change. The Census Bureau also did not allow same-sex spouse combinations to occur in the 1990 Census, but the different allocation procedure implemented was prone to fewer problems.
[114] Incorrect coding of sex among cohabiting heterosexual couples may also contaminate the sample of lesbians, but as sex miscoding is rare and the sample of cohabiting heterosexual couples is relatively small, this is unlikely to heavily affect the results.

when either partner's sex, age, relationship to the householder, or marital status is allocated by the U.S. Census Bureau.[115]

[115] As discussed in Gates and Steinberger (2011), response-mode information in the ACS PUMS can be used to minimise miscoding of same-sex couples while retaining a larger proportion of the sample. As the 2000 Census PUMS files do not contain similar information, we are unable to explore this possibility while maintaining cross-sample consistency.

Appendix D: Definition of Variables

Table D.1. Definition of individual-level variables

Variable	Description
Log(real annual income)	Continuous variable measuring the log of real annual income. Real annual income is measured in 2010 dollars for all samples and is calculated by adjusting nominal income using both the adjustment factor contained within the PUMS files (for ACS data, to account for the timing of survey completion) and accounting for inflation using the CPI. Source: http://www.bls.gov/cpi/cpirsdc.htm, accessed on 20th June, 2013.
Work characteristics:	
Weeks worked:	(Base: Worked less than 14 weeks in the previous year)
Weeks1	Dummy variable, equals 1 if worked 14 to 26 weeks in the previous year, 0 otherwise
Weeks2	Dummy variable, equals 1 if worked 27 to 39 weeks in the previous year, 0 otherwise
Weeks3	Dummy variable, equals 1 if worked 40 to 47 weeks in the previous year, 0 otherwise
Weeks4	Dummy variable, equals 1 if worked 48 to 49 weeks in the previous year, 0 otherwise
Weeks5	Dummy variable, equals 1 if worked 50 to 52 weeks in the previous year, 0 otherwise
Hours worked	Continuous variable measuring the usual hours worked per week in the previous year
Lesbian	Dummy variable, equals 1 if the individual is a female cohabiting with a same-sex partner, 0 otherwise
Potential experience	Continuous variable representing potential work experience, defined as age - years of education - 5
Potential experience^2	Continuous variable representing the square of potential work experience
Hispanic	Dummy variable, equals 1 if the individual is of Hispanic origin, 0 otherwise
Race:	(Base: Individual is exclusively white)
Black	Dummy variable, equals 1 if the individual is exclusively black, 0 otherwise
Asian	Dummy variable, equals 1 if the individual is exclusively Asian, 0 otherwise
Other race	Dummy variable, equals 1 if the individual belongs exclusively to some other racial category, 0 otherwise
Mixed race	Dummy variable, equals 1 if the individual belongs to more than one racial category, 0 otherwise
Speaks English	Dummy variable, equals 1 if English is the respondent's first language, or speaks English "well" or "very well", 0 otherwise
U.S. citizen	Dummy variable, equals 1 if the individual is a U.S. citizen, 0 otherwise

Table D.1 (continued)

Variable	Description
Disabled	Dummy variable, equals 1 if the individual has some form of disability, 0 otherwise[1]
Educational attainment:	(Base: Attained less than a high school diploma)
High school	Dummy variable, equals 1 if highest educational attainment is a high school diploma or equivalent qualification, 0 otherwise
Some college	Dummy variable, equals 1 if attended college but did not obtain any form of degree, 0 otherwise
Associate	Dummy variable, equals 1 if highest educational attainment is an Associate's degree, 0 otherwise
Bachelor	Dummy variable, equals 1 if highest educational attainment is a bachelor's degree, 0 otherwise
Postgraduate	Dummy variable, equals 1 if highest educational attainment is a master's, professional, or doctoral degree, 0 otherwise
Metropolitan residence	Dummy variable, equals 1 if the individual resides in a metropolitan area, 0 otherwise
Region:	(Base: Individual resides in the Northeast of the U.S.)
Midwest	Dummy variable, equals 1 if the individual resides in the Midwest of the U.S., 0 otherwise
South	Dummy variable, equals 1 if the individual resides in the South of the U.S., 0 otherwise
West	Dummy variable, equals 1 if the individual resides in the West of the U.S., 0 otherwise
Number of children	Continuous variable measuring the number of own children under 18 years of age in the household[2]
Occupation:	(Base: Works in a management or professional occupation)
Service occupations	Dummy variable, equals 1 if works in a service occupation, 0 otherwise
Sales	Dummy variable, equals 1 if works in a sales or office occupation, 0 otherwise
Construction	Dummy variable, equals 1 if works in natural resources, construction or maintenance, 0 otherwise
Production and transport.	Dummy variable, equals 1 if works in production, transportation or materials moving, 0 otherwise
Industry:	(Base: Employed in agriculture, forestry, fishing and hunting, or mining)
Construction	Dummy variable, equals 1 if employed in construction, 0 otherwise
Manufacturing	Dummy variable, equals 1 if employed in manufacturing, 0 otherwise
Wholesale trade	Dummy variable, equals 1 if employed in wholesale trade, 0 otherwise

Table D.1 (continued)

Variable	Description
Retail trade	Dummy variable, equals 1 if employed in retail trade, 0 otherwise
Transport. and utilities	Dummy variable, equals 1 if employed in transportation and warehousing or utilities, 0 otherwise
Information	Dummy variable, equals 1 if employed in information, 0 otherwise
Finance and insurance	Dummy variable, equals 1 if employed in finance and insurance, real estate, rental or leasing, 0 otherwise
Professional and mgt.	Dummy variable, equals 1 if employed in professional, scientific, management, administration or waste management services, 0 otherwise
Education and health	Dummy variable, equals 1 if employed in educational services, health care or social assistance, 0 otherwise
Arts and entertainment	Dummy variable, equals 1 if employed in arts, entertainment, recreation, accommodation or food services, 0 otherwise
Other services	Dummy variable, equals 1 if employed in other services (except public administration), 0 otherwise
Public administration	Dummy variable, equals 1 if employed in public administration, 0 otherwise
Year:	(Base: Years 2000, 2005 and 2008 for the three respective samples)
Year 2006	Dummy variable, equals 1 if the individual was surveyed in 2006, 0 otherwise
Year 2007	Dummy variable, equals 1 if the individual was surveyed in 2007, 0 otherwise
Year 2009	Dummy variable, equals 1 if the individual was surveyed in 2009, 0 otherwise
Year 2010	Dummy variable, equals 1 if the individual was surveyed in 2010, 0 otherwise

Notes: [1] The 2000 Census and 2005-2007 ACS PUMS contain a "work-limiting disability" variable. This variable is highly correlated with the general disability variable, and alternative specifications reveal the results are only slightly affected by use of the alternative disability variable.

[2] This is the number of children that the householder identifies as their "natural born son/daughter", which we believe best captures the number of children that fall under the care of the householder and partner. In unreported regressions we recode number of children to include all persons in the household under the age of 18, regardless of relationship to the householder. Again, the results are not significantly affected by this change.

Table D.2. Discussion of state-level proxies and controls

Variable	Description/Source
Maternity risk proxies:	
Mandated coverage	Dummy variable, equals 1 whenever the individual resides within a state mandating insurance coverage of infertility treatment, 0 otherwise.[1] States with mandated insurance coverage of infertility treatment should increase the maternity risk of cohabiting and married heterosexuals relative to lesbians. Thus, if the maternity-risk hypothesis is correct, the lesbian premium should be higher relative to cohabiting and married heterosexuals in states with mandated coverage. Source: http://www.resolve.org/family-building-options/insurance_coverage/state-coverage.html, accessed on 5th July, 2013.
Legal partnership	Dummy variable, equals 1 whenever the individual resides in a state allowing same-sex partners to enter into a domestic partnership or granting greater partnership status to same-sex couples, 0 otherwise. State laws granting greater partnership rights to same-sex couples are likely to increase the incidence of child-bearing among lesbian couples; we thus expect a decrease in the lesbian premium in these states relative to states where no such law is in effect.[2] Sources: http://www.nolo.com/legal-encyclopedia/same-sex-marriage-developments-the-law.html, accessed on 5th July, 2013. Last updated 3rd July, 2013. http://www.ncsl.org/research/human-services/civil-unions-and-domestic-partnership-statutes.aspx, accessed on 5th July, 2013. Last updated 26th June, 2013.
Marriage bans	Dummy variable, equals 1 whenever the individual lives in a state with a law or Constitutional Amendment banning same-sex marriage, 0 otherwise. Bans on same-sex marriage likely inhibit family formation among lesbian couples, increasing the sexual-orientation maternity gap. As a result, the lesbian premium should be greater in states with such bans in place, ceteris paribus. Sources: http://www.nolo.com/legal-encyclopedia/same-sex-marriage-developments-the-law.html, accessed on 5th July, 2013. Last updated 3rd July, 2013. http://www.ngltf.org/downloads/reports/issue_maps/samesex_relationships_5_15_13.pdf, accessed on 5th July, 2013, last updated 15th May, 2013.
Same-sex percentage	Standardised variable representing the size of the coupled lesbian population as a percentage of the female adult population within each state. We believe that the presence of support networks and other factors may increase child-bearing rates among lesbians. In such a case, we would expect a commensurate reduction in the lesbian premium. Due to within-sample variation for each state, several alternative variables were also used to ensure the robustness of the results: The coupled lesbian population as a percentage of the total female population in each state, the within-sample average value for each state (as opposed to the annual value), and the size of the coupled gay (male and female) population as a percentage of the adult population residing in each state. The results were highly consistent with those reported. Source: Own calculations from PUMS files.

Table D.2 (continued)

Variable	Description/Source
State-level controls:	
Citizen ideology	Continuous measure capturing citizen ideology at the state level, based on the "revised 1960-2010 citizen ideology measure" derived by Berry et al. (1998). Citizen ideology is measured annually on a 0-10 scale, with zero reflecting the most conservative and ten representing the most liberal position.[3] Each candidate's ideology score is calculated as the average of interest-group ratings reported by the Committee on Political Education (COPE) and the Americans for Democratic Action (ADA). District-level ideology scores are then computed as a weighted average of the incumbent's and the challenger's (or potential challenger's) ideology score, with weights given by the share of support for each candidate. Finally, state-level ideology is given by an unweighted average of district ideology scores. The authors show that this measure improves on other available measures, particularly because it does not assume citizen ideology is constant over time. Further details surrounding the derivation and reliability of the measure can be found in the original article. Source: http://rcfording.wordpress.com/state-ideology-data/, accessed on 5th July, 2013.
Government ideology	Continuous measure capturing government ideology at the state level, based on the "ADA/COPE measure of state government ideology" derived by Berry et al. (1998). Government ideology is measured annually on a 0-10 scale, with zero again being the most conservative and ten representing the most liberal position. Construction of the measure entails several assumptions regarding the distribution of power among policymakers. Similarly to the citizen ideology measure, ideology values for each policymaker are given by the average of ADA and COPE scores. The final state-level government ideology score is a weighted average of the five ideology scores for the governor, and Democrats and Republicans in a state's lower and upper chambers, with weights given by the share of power held by each individual. Further details can again be found in the original paper. Source: http://rcfording.wordpress.com/state-ideology-data/, accessed on 5th July, 2013.
Percentage of GSP from manufacturing	Continuous measure as used by Baumle and Poston (2011) to capture state-level labour-market conditions which may result in a sexual-orientation earnings differential. Specifically, states with a high proportion of Gross State Product (GSP) from manufacturing may afford residents fewer job opportunities or jobs with lower earnings prospects. If lesbians and heterosexual women exhibit different residency choices on the basis of the percentage of GSP from manufacturing, a lesbian pay gap will emerge. State-level values are calculated by dividing annual GSP from manufacturing by total annual GSP. Source: http://www.bea.gov/regional/, accessed on 17th July, 2013.

Table D.2 (continued)

Variable	Description/Source
State ENDA	Dummy variable, equals 1 if the state in which the individual resides has an Employment Non-Discrimination Act (ENDA) which covers sexual orientation, 0 otherwise. Klawitter and Flatt (1998), Gates (2009) and Klawitter (2011) all use this variable to examine whether the presence of an ENDA covering sexual orientation enhances the earnings of lesbians by reducing sexual-orientation discrimination. To the extent that the presence of an ENDA is correlated with the proxies, controlling for an ENDA may be important in obtaining the marginal effect of maternity risk. Sources: http://www.ngltf.org/downloads/reports/issue_maps/non_discrimination_6_13_color.pdf, accessed on 5th July, 2013. http://www.hrc.org/files/assets/resources/employment_laws_072013.pdf, accessed on 5th July, 2013.
Sodomy repeal	Time-invariant dummy variable, equals 1 if the individual resides in a state where the state-level sodomy law was repealed prior to the 2003 federal repeal, 0 otherwise. We classify the law as repealed only when same-sex sexual intercourse is fully decriminalised. This measure serves to capture underlying attitudes towards homosexuals in each state. Similar measures are used in Klawitter and Flatt (1998) and Baumle and Poston (2011). Source: http://www.sodomy.org/laws/, accessed on 5th July, 2013. http://www.glapn.org/sodomylaws/legal.htm, accessed on 5th July, 2013.
2013 same-sex legal equality score	Time-invariant ordinal variable, ranging from zero to six in half-point increments. This measure captures varying degrees of same-sex legal equality by state based on six categories (hate crimes, non-discrimination, marriage, freedom of gender, youth protection, and adoption), with higher values representing greater equality. The scores are intended to control for underlying levels of gay-friendliness and forward-looking expectations of employers regarding the future rights of same-sex couples. It is possible that including this measure may lead to an over-controlling problem, due to the inclusion of laws affecting maternity risk. Source: http://www.equalitygiving.org/States-of-Equality-and-Gay-Rights-Scorecard, accessed on 5th July, 2013. Last updated 20th June, 2013.

Notes: [1] As with all dummy variable proxies and controls, if a law becomes effective or is repealed during the year, we assign the value associated with the situation following the law change. As there are few law changes throughout the sample period, this coding choice does not appear to significantly affect the results.

[2] We group states with partnership, civil union, or marriage laws due to the limited number of states conferring such rights on same-sex couples. Moreover, individuals and employers likely take into account the probable extension of further rights to same-sex couples in states already offering domestic partnerships or civil unions. The lack of distinction between domestic partnerships conferring spousal rights and civil unions also motivated this choice.

[3] The original measures range from 0-100 but we divide all values by 10 to aide in the presentation of results.

Appendix E: On the Inclusion of a Maternity Risk Interaction Term

The discussion in the text regarding the inclusion of a *Maternity Risk* × *Lesbian* interaction term provides no theoretical foundation as to why we might expect the effect of assigned maternity risk to differ by sexual orientation. This note briefly outlines potential reasons why an interaction term is warranted from an economic perspective, particularly in models that contain potential experience interactions to minimise bias in the maternity-risk coefficient.

As previously discussed, lesbians generally share household responsibilities more equally than heterosexual couples, where household production is primarily undertaken by the female. Lesbians may thus be able to return to the workforce sooner, resulting in lower costs to the employer and a smaller erosion of human capital. On this basis, maternity risk should have a less detrimental effect on the earnings of lesbians than on heterosexuals, all other factors held constant.

There are also several mechanisms by which maternity risk could cause a larger adverse earnings effect among lesbians. As lesbians pose a smaller risk to the employer, they may be placed in positions where training and turnover costs are greater. Furthermore, cohabitation may not be a strong signal of sexual orientation among young women. If employers lack sufficient alternative information to accurately assess sexual orientation, some young lesbians may be incorrectly perceived as being heterosexuals. Similarly, employers may assign a non-trivial probability to young lesbians switching groups in the near-term. The allocated rates used herein may thus understate employers' perceptions, necessitating a correction through a negative coefficient on the interaction term.

Evidently, we cannot unambiguously ascertain whether the effect of maternity risk should be greater for lesbians or heterosexual women. It is likely, however, that the effects of occupational sorting and employer perceptions regarding sexual orientation and transition between groups outweigh household specialisation effects, leading to a negative coefficient on the interaction term. In unreported regressions, we further include *Maternity Risk* × *Lesbian* in all specifications which allow the effect of potential experience to vary by sexual orientation. We generally obtain a negative coefficient on the interaction term, but in each case the coefficient is statistically insignificant at conventional levels. As a result, in the main text we strictly consider models in which the maternity-risk effect does not vary by sexual orientation.

Appendix F: Supplementary Results

Table F.1. Blinder-Oaxaca decompositions - heterosexual coefficients

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Gross differential:	0.4693***	0.2973***	0.5712***	0.3801***	0.5295***	0.3227***
Excluding maternity risk						
Explained:						
Work characteristics	0.1406***	0.1588***	0.1765***	0.2137***	0.1427***	0.1701***
Other controls	0.2733***	0.1177***	0.3250***	0.1312***	0.3189***	0.1117***
Total explained:	0.4139***	0.2765***	0.5015***	0.3449***	0.4615***	0.2818***
Total unexplained:	0.0554***	0.0208**	0.0697***	0.0351***	0.0679***	0.0409***
Including maternity risk						
Explained:						
Work characteristics	0.1405***	0.1588***	0.1764***	0.2137***	0.1426***	0.1701***
Other controls	0.2596***	0.1186***	0.3115***	0.1322***	0.3046***	0.1136***
Maternity risk	0.0215***	0.0145***	0.0200***	0.0061	0.0237***	0.0129***
Total explained:	0.4217***	0.2919***	0.5079***	0.3520***	0.4708***	0.2966***
Total unexplained:	0.0476***	0.0053	0.0633***	0.0280***	0.0586***	0.0262**
Observations	81,534	35,434	61,383	57,083	67,776	59,791

Notes: We assume the coefficients from the heterosexual female regressions represent the non-discriminatory earnings structure, as opposed to a weighted average of the individual regressions. The generalised decomposition is also overlooked due to problems cited in the text regarding unobservable human capital accumulation. Models include all variables specified in the text (Model (A2p)), but we group work characteristics and all other controls for reporting purposes. ** and *** represent statistical significance at the 5% and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table F.2. Summary of control values by sample period

	2000 Census	2005-2007 ACS	2008-2010 ACS
Mean state-level citizen ideology	5.06	5.55	5.47
Mean state-level government ideology	4.63	5.32	5.88
Mean percentage of GSP from manufacturing (%)	14.59	12.50	11.32
Percentage of states with ENDA's covering sexual orientation (%)	22.00	35.33	41.33
Percentage of states which repealed their sodomy law prior to the 2003 federal repeal (%)	72.00	72.00	72.00
Mean 2013 same-sex legal equality score	3.01	3.01	3.01

Sources: A list of sources for each control is provided in Table D.2.

Table F.3. *Mandated Coverage* supplementary regression estimates

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Base results:						
Lesbian	0.0415***	0.0099	0.0597***	0.0327***	0.0534***	0.0305***
	(0.0086)	(0.0084)	(0.0115)	(0.0113)	(0.0096)	(0.0080)
Mandated coverage	0.0230*	0.0153	0.0170	-0.0050	0.0289***	0.0141
x lesbian	(0.0149)	(0.0130)	(0.0181)	(0.0193)	(0.0118)	(0.0111)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6379	0.6322	0.6750	0.6517	0.6885	0.6653
Full-time, no hours/weeks controls:						
Lesbian	0.0654***	0.0466***	0.0765***	0.0459***	0.0827***	0.0599***
	(0.0100)	(0.0098)	(0.0109)	(0.0102)	(0.0090)	(0.0075)
Mandated coverage	0.0267**	0.0173	0.0466***	0.0432***	0.0125*	0.0066
x lesbian	(0.0149)	(0.0168)	(0.0152)	(0.0169)	(0.0097)	(0.0108)
Observations	119,964	163,718	79,050	110,628	87,238	114,129
R^2	0.3658	0.3332	0.4050	0.3677	0.4146	0.3641
Log(hourly wages):						
Lesbian	0.0457***	0.0196**	0.0665***	0.0414***	[redacted]	
	(0.0086)	(0.0086)	(0.0115)	(0.0116)		
Mandated coverage	0.0222*	0.0142	0.0191	0.0029		
x lesbian	(0.0148)	(0.0136)	(0.0160)	(0.0172)		
Observations	175,609	255,185	117,600	171,584		
R^2	0.2556	0.2338	0.3129	0.2740		
Heckman two-step method:						
Lesbian	0.0414***	0.0188**	0.0597***	0.0415***	0.0483***	0.0411***
	(0.0087)	(0.0082)	(0.0115)	(0.0112)	(0.0103)	(0.0078)
Mandated coverage	0.0231*	0.0156	0.0157	0.0009	0.0190*	0.0184**
x lesbian	(0.0152)	(0.0126)	(0.0181)	(0.0198)	(0.0141)	(0.0111)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	-	-	-	-	-	-
Including sodomy law repeal:						
Lesbian	0.0411***	0.0091	0.0591***	0.0315***	0.0530***	0.0300***
	(0.0086)	(0.0087)	(0.0112)	(0.0108)	(0.0094)	(0.0071)
Mandated coverage	0.0243*	0.0158	0.0181	-0.0039	0.0300***	0.0146*
x lesbian	(0.0149)	(0.0129)	(0.0182)	(0.0198)	(0.0120)	(0.0113)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6379	0.6323	0.6751	0.6518	0.6885	0.6653
Further including SS equality score:						
Lesbian	0.0415***	0.0090	0.0579***	0.0306***	0.0525***	0.0305***
	(0.0082)	(0.0078)	(0.0106)	(0.0105)	(0.0097)	(0.0073)
Mandated coverage	0.0221*	0.0123	0.0201	-0.0036	0.0310***	0.0138
x lesbian	(0.0136)	(0.0109)	(0.0190)	(0.0202)	(0.0128)	(0.0118)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6382	0.6324	0.6752	0.6520	0.6887	0.6654

Table F.3 (continued)

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Adding potential exper. interactions:						
Lesbian	0.0663***	0.0450***	0.0872***	0.0727***	0.0826***	0.0736***
	(0.0096)	(0.0088)	(0.0123)	(0.0112)	(0.0127)	(0.0108)
Mandated coverage	0.0257**	0.0198*	0.0178	-0.0034	0.0312***	0.0176*
x lesbian	(0.0147)	(0.0127)	(0.0186)	(0.0202)	(0.0127)	(0.0122)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6380	0.6324	0.6751	0.6521	0.6887	0.6657
Further including educ. interactions:						
Lesbian	0.0758***	0.0586***	0.0924***	0.0813***	0.0883***	0.0773***
	(0.0095)	(0.0084)	(0.0128)	(0.0121)	(0.0125)	(0.0106)
Mandated coverage	0.0250**	0.0189*	0.0179	-0.0019	0.0306***	0.0181*
x lesbian	(0.0149)	(0.0129)	(0.0187)	(0.0198)	(0.0128)	(0.0122)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6380	0.6325	0.6752	0.6521	0.6887	0.6657
Including only full-time employed:						
Lesbian	0.0561***	0.0298***	0.0662***	0.0308***	0.0733***	0.0479***
	(0.0089)	(0.0085)	(0.0105)	(0.0096)	(0.0089)	(0.0071)
Mandated coverage	0.0261**	0.0189	0.0470***	0.0379**	0.0132*	0.0058
x lesbian	(0.0145)	(0.0162)	(0.0150)	(0.0173)	(0.0099)	(0.0106)
Observations	119,964	163,718	79,050	110,628	87,238	114,129
R^2	0.4024	0.3625	0.4344	0.3993	0.4422	0.3943
Including only white women:						
Lesbian	0.0441***	0.0130*	0.0620***	0.0328***	0.0559***	0.0330***
	(0.0085)	(0.0083)	(0.0119)	(0.0121)	(0.0103)	(0.0086)
Mandated coverage	0.0273*	0.0185*	0.0233	0.0021	0.0330***	0.0208**
x lesbian	(0.0170)	(0.0139)	(0.0182)	(0.0196)	(0.0118)	(0.0112)
Observations	137,365	213,342	95,191	144,206	102,969	142,195
R^2	0.6442	0.6403	0.6750	0.6564	0.6902	0.6708
Only women aged 40 or under:						
Lesbian	0.0131	-0.0389***	0.0181*	-0.0346***	-0.0128*	-0.0609***
	(0.0115)	(0.0117)	(0.0127)	(0.0140)	(0.0093)	(0.0086)
Mandated coverage	0.0292**	0.0196	0.0059	-0.0243	0.0300**	0.0232*
x lesbian	(0.0154)	(0.0153)	(0.0269)	(0.0273)	(0.0137)	(0.0166)
Observations	129,424	119,455	77,569	64,706	83,507	60,418
R^2	0.6547	0.6571	0.6941	0.6815	0.7055	0.6965
Only women over 40 years old:						
Lesbian	0.0785***	0.0705***	0.0964***	0.0780***	0.1072***	0.0952***
	(0.0162)	(0.0172)	(0.0175)	(0.0158)	(0.0119)	(0.0112)
Mandated coverage	0.0048	0.0116	0.0198	0.0132	0.0262*	0.0146
x lesbian	(0.0280)	(0.0280)	(0.0227)	(0.0213)	(0.0174)	(0.0187)
Observations	46,185	135,730	40,031	106,878	43,459	108,616
R^2	0.5531	0.6079	0.6099	0.6313	0.6324	0.6478

Table F.3 (continued)

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Smaller sample						
Base results:						
Lesbian	0.0368***	0.0303***	0.0641***	0.0714***	0.0288**	0.0507***
	(0.0101)	(0.0116)	(0.0121)	(0.0144)	(0.0139)	(0.0088)
Mandated coverage	0.0227*	0.0207	-0.0054	-0.0202	0.0308**	0.0190*
x lesbian	(0.0156)	(0.0175)	(0.0153)	(0.0193)	(0.0181)	(0.0136)
Observations	28,419	36,463	20,991	26,344	22,271	26,506
R^2	0.6305	0.6209	0.6583	0.6429	0.6885	0.6614
Multilevel estimation:						
Lesbian	0.0363***	0.0286**	0.0659***	0.0717***	0.0308***	0.0548***
	(0.0113)	(0.0124)	(0.0139)	(0.0144)	(0.0127)	(0.0121)
Mandated coverage	0.0108	0.0074	-0.0116	-0.0426**	0.0266*	0.0108
x lesbian	(0.0194)	(0.0229)	(0.0219)	(0.0230)	(0.0196)	(0.0182)
Observations	28,419	36,463	20,991	26,344	22,271	26,506
R^2	-	-	-	-	-	-
Gay-male comparison						
Female estimates:						
Lesbian	0.0415***	0.0099	0.0597***	0.0327***	0.0534***	0.0305***
	(0.0086)	(0.0084)	(0.0115)	(0.0113)	(0.0096)	(0.0080)
Mandated coverage	0.0230*	0.0153	0.0170	-0.0050	0.0289***	0.0141
x lesbian	(0.0149)	(0.0130)	(0.0181)	(0.0193)	(0.0118)	(0.0111)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6379	0.6322	0.6750	0.6517	0.6885	0.6653
Male estimates:						
Lesbian	-0.0299***	-0.1650***	-0.0004	-0.1240***	0.0122	-0.1058***
	(0.0113)	(0.0115)	(0.0126)	(0.0105)	(0.0132)	(0.0116)
Mandated coverage	0.0606***	0.0510***	0.0720***	0.0410**	0.0705***	0.0321*
x lesbian	(0.0180)	(0.0183)	(0.0242)	(0.0195)	(0.0244)	(0.0213)
Observations	178,186	295,674	120,413	192,527	128,821	187,391
R^2	0.5100	0.4312	0.5607	0.4774	0.5987	0.5145

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the lesbian indicator and lesbian-proxy interaction term. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table F.4. *Legal Partnership* supplementary regression estimates

	2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds
Base results:				
Lesbian	0.0719***	0.0364***	0.0728***	0.0381***
	(0.0084)	(0.0090)	(0.0085)	(0.0053)
Legal partnership	-0.0261	-0.0368**	-0.0249*	-0.0097
x lesbian	(0.0215)	(0.0176)	(0.0178)	(0.0108)
Observations	117,600	171,584	126,966	169,034
R^2	0.6754	0.6522	0.6888	0.6654
Full-time, no hours/weeks controls:				
Lesbian	0.1001***	0.0700***	0.0912***	0.0626***
	(0.0072)	(0.0072)	(0.0095)	(0.0075)
Legal partnership	-0.0160	-0.0314**	-0.0129	-0.0058
x lesbian	(0.0203)	(0.0174)	(0.0173)	(0.0165)
Observations	79,050	110,628	87,238	114,129
R^2	0.4056	0.3687	0.4153	0.3645
Log(hourly wages):				
Lesbian	0.0808***	0.0496***	[redacted]	
	(0.0080)	(0.0087)		
Legal partnership	-0.0315*	-0.0405**		
x lesbian	(0.0197)	(0.0182)		
Observations	117,600	171,584		
R^2	0.3139	0.2750		
Heckman two-step method:				
Lesbian	0.0713***	0.0477***	0.0690***	0.0482***
	(0.0084)	(0.0091)	(0.0086)	(0.0058)
Legal partnership	-0.0260	-0.0357**	-0.0437***	-0.0012
x lesbian	(0.0215)	(0.0180)	(0.0159)	(0.0114)
Observations	117,600	171,584	126,966	169,034
R^2	-	-	-	-
Including sodomy law repeal:				
Lesbian	0.0718***	0.0358***	0.0727***	0.0378***
	(0.0084)	(0.0090)	(0.0086)	(0.0054)
Legal partnership	-0.0260	-0.0367**	-0.0246*	-0.0094
x lesbian	(0.0216)	(0.0175)	(0.0177)	(0.0108)
Observations	117,600	171,584	126,966	169,034
R^2	0.6754	0.6522	0.6889	0.6654
Further including SS equality score:				
Lesbian	0.0703***	0.0348***	0.0737***	0.0389***
	(0.0079)	(0.0086)	(0.0088)	(0.0056)
Legal partnership	-0.0182	-0.0328*	-0.0275*	-0.0120
x lesbian	(0.0249)	(0.0208)	(0.0176)	(0.0107)
Observations	117,600	171,584	126,966	169,034
R^2	0.6755	0.6522	0.6890	0.6656

Table F.4 (continued)

	2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds
Adding potential exper. interactions:				
Lesbian	0.1004***	0.0784***	0.1051***	0.0862***
	(0.0104)	(0.0106)	(0.0104)	(0.0082)
Legal partnership	-0.0284*	-0.0413***	-0.0320**	-0.0208**
x lesbian	(0.0211)	(0.0167)	(0.0182)	(0.0118)
Observations	117,600	171,584	126,966	169,034
R^2	0.6755	0.6525	0.6890	0.6659
Further including educ. interactions:				
Lesbian	0.1053***	0.0879***	0.1105***	0.0901***
	(0.0108)	(0.0112)	(0.0110)	(0.0081)
Legal partnership	-0.0279*	-0.0425***	-0.0317**	-0.0207**
x lesbian	(0.0215)	(0.0170)	(0.0184)	(0.0120)
Observations	117,600	171,584	126,966	169,034
R^2	0.6755	0.6526	0.6891	0.6659
Including only full-time employed:				
Lesbian	0.0897***	0.0521***	0.0819***	0.0502***
	(0.0070)	(0.0069)	(0.0086)	(0.0065)
Legal partnership	-0.0154	-0.0297**	-0.0120	-0.0056
x lesbian	(0.0200)	(0.0166)	(0.0164)	(0.0145)
Observations	79,050	110,628	87,238	114,129
R^2	0.4353	0.4004	0.4429	0.3947
Including only white women:				
Lesbian	0.0752***	0.0379***	0.0773***	0.0448***
	(0.0078)	(0.0085)	(0.0090)	(0.0061)
Legal partnership	-0.0199	-0.0308**	-0.0273*	-0.0157
x lesbian	(0.0250)	(0.0172)	(0.0211)	(0.0148)
Observations	95,191	144,206	102,969	142,195
R^2	0.6755	0.6570	0.6905	0.6710
Only women aged 40 or under:				
Lesbian	0.0284***	-0.0373***	0.0035	-0.0542***
	(0.0092)	(0.0104)	(0.0083)	(0.0070)
Legal partnership	-0.0404**	-0.0427*	-0.0094	0.0121
x lesbian	(0.0185)	(0.0272)	(0.0225)	(0.0241)
Observations	77,569	64,706	83,507	60,418
R^2	0.6945	0.6817	0.7058	0.6966
Only women over 40 years old:				
Lesbian	0.1073***	0.0890***	0.1353***	0.1114***
	(0.0131)	(0.0119)	(0.0118)	(0.0098)
Legal partnership	-0.0125	-0.0370	-0.0586***	-0.0405**
x lesbian	(0.0397)	(0.0309)	(0.0239)	(0.0188)
Observations	40,031	106,878	43,459	108,616
R^2	0.6102	0.6319	0.6329	0.6480

Table F.4 (continued)

	2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds
Smaller sample				
Base results:				
Lesbian	0.0719***	0.0636***	0.0464***	0.0621***
	(0.0112)	(0.0109)	(0.0103)	(0.0086)
Legal partnership	-0.0498**	-0.0324*	-0.0150	-0.0168
x lesbian	(0.0226)	(0.0237)	(0.0169)	(0.0158)
Observations	20,991	26,344	22,271	26,506
R^2	0.6586	0.6430	0.6885	0.6614
Multilevel estimation:				
Lesbian	0.0726***	0.0618***	0.0427***	0.0619***
	(0.0120)	(0.0118)	(0.0133)	(0.0117)
Legal partnership	-0.0797***	-0.0558**	-0.0151	-0.0134
x lesbian	(0.0289)	(0.0292)	(0.0289)	(0.0252)
Observations	20,991	26,344	22,271	26,506
R^2	-	-	-	-
Gay-male comparison				
Female estimates:				
Lesbian	0.0719***	0.0364***	0.0728***	0.0381***
	(0.0084)	(0.0090)	(0.0085)	(0.0053)
Legal partnership	-0.0261	-0.0368**	-0.0249*	-0.0097
x lesbian	(0.0215)	(0.0176)	(0.0178)	(0.0108)
Observations	117,600	171,584	126,966	169,034
R^2	0.6754	0.6522	0.6888	0.6654
Male estimates:				
Lesbian	0.0229**	-0.1061***	0.0481***	-0.0881***
	(0.0098)	(0.0083)	(0.0171)	(0.0155)
Legal partnership	0.0340	-0.0036	-0.0038	-0.0007
x lesbian	(0.0330)	(0.0284)	(0.0381)	(0.0322)
Observations	120,413	192,527	128,821	187,391
R^2	0.5609	0.4778	0.5983	0.5137

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the lesbian indicator and lesbian-proxy interaction term. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2005-2010 American Community Surveys PUMS.

Table F.5. *Marriage Bans* supplementary regression estimates

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Base results:						
Lesbian	0.0337**	0.0040	0.0230	-0.0139	0.0907***	0.0363**
	(0.0161)	(0.0157)	(0.0285)	(0.0281)	(0.0254)	(0.0182)
Marriage bans	0.0226	0.0161	0.0511**	0.0507*	-0.0290	-0.0002
x lesbian	(0.0202)	(0.0195)	(0.0301)	(0.0312)	(0.0259)	(0.0196)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6379	0.6322	0.6750	0.6517	0.6884	0.6651
Full-time, no hours/weeks controls:						
Lesbian	0.0335**	0.0102	0.0605***	0.0307	0.1020***	0.0725***
	(0.0181)	(0.0176)	(0.0256)	(0.0286)	(0.0229)	(0.0210)
Marriage bans	0.0575***	0.0593***	0.0427**	0.0392*	-0.0164	-0.0122
x lesbian	(0.0210)	(0.0211)	(0.0259)	(0.0287)	(0.0223)	(0.0222)
Observations	119,964	163,718	79,050	110,628	87,238	114,129
R^2	0.3659	0.3332	0.4048	0.3674	0.4139	0.3634
Log(hourly wages):						
Lesbian	0.0276**	0.0032	0.0438*	0.0174	[redacted]	
	(0.0147)	(0.0143)	(0.0301)	(0.0300)		
Marriage bans	0.0363**	0.0301**	0.0361	0.0287		
x lesbian	(0.0188)	(0.0182)	(0.0318)	(0.0328)		
Observations	175,609	255,185	117,600	171,584		
R^2	0.2556	0.2338	0.3130	0.2740		
Heckman two-step method:						
Lesbian	0.0335**	0.0130	0.0220	-0.0006	0.0845***	0.0464***
	(0.0165)	(0.0154)	(0.0285)	(0.0291)	(0.0284)	(0.0172)
Marriage bans	0.0227	0.0163	0.0516**	0.0486*	-0.0328	0.0028
x lesbian	(0.0208)	(0.0188)	(0.0300)	(0.0316)	(0.0290)	(0.0185)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	-	-	-	-	-	-
Including sodomy law repeal:						
Lesbian	0.0328**	0.0006	0.0233	-0.0138	0.0905***	0.0362**
	(0.0172)	(0.0162)	(0.0283)	(0.0277)	(0.0253)	(0.0184)
Marriage bans	0.0239	0.0198	0.0506**	0.0499*	-0.0288	-0.0003
x lesbian	(0.0214)	(0.0203)	(0.0298)	(0.0306)	(0.0258)	(0.0195)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6380	0.6323	0.6751	0.6517	0.6884	0.6651
Further including SS equality score:						
Lesbian	0.0294**	-0.0045	0.0243	-0.0113	0.0921***	0.0387**
	(0.0166)	(0.0078)	(0.0277)	(0.0270)	(0.0254)	(0.0186)
Marriage bans	0.0279	0.0247*	0.0492**	0.0463*	-0.0306	-0.0031
x lesbian	(0.0219)	(0.0182)	(0.0296)	(0.0300)	(0.0258)	(0.0195)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6382	0.6324	0.6753	0.6519	0.6887	0.6654

Table F.5 (continued)

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Adding potential exper. interactions:						
Lesbian	0.0593***	0.0403***	0.0509**	0.0271	0.1259***	0.0891***
	(0.0141)	(0.0146)	(0.0300)	(0.0286)	(0.0281)	(0.0211)
Marriage bans	0.0227	0.0167	0.0510**	0.0505**	-0.0349	-0.0097
x lesbian	(0.0203)	(0.0196)	(0.0295)	(0.0304)	(0.0273)	(0.0219)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6380	0.6325	0.6752	0.6521	0.6886	0.6656
Further including educ. interactions:						
Lesbian	0.0695***	0.0551***	0.0564**	0.0368	0.1324***	0.0939***
	(0.0148)	(0.0147)	(0.0308)	(0.0309)	(0.0276)	(0.0203)
Marriage bans	0.0212	0.0145	0.0509**	0.0501*	-0.0346	-0.0105
x lesbian	(0.0205)	(0.0196)	(0.0297)	(0.0313)	(0.0273)	(0.0221)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6380	0.6325	0.6752	0.6521	0.6886	0.6656
Including only full-time employed:						
Lesbian	0.0286**	-0.0018	0.0546**	0.0109	0.0944***	0.0554***
	(0.0167)	(0.0163)	(0.0244)	(0.0253)	(0.0213)	(0.0202)
Marriage bans	0.0512***	0.0532***	0.0376*	0.0420*	-0.0182	-0.0066
x lesbian	(0.0197)	(0.0198)	(0.0251)	(0.0258)	(0.0209)	(0.0213)
Observations	119,964	163,718	79,050	110,628	87,238	114,129
R^2	0.4024	0.3625	0.4343	0.399	0.4417	0.3937
Including only white women:						
Lesbian	0.0364**	0.0087	0.0195	-0.0258	0.0910***	0.0405**
	(0.0177)	(0.0152)	(0.0299)	(0.0272)	(0.0281)	(0.0212)
Marriage bans	0.0243	0.0155	0.0607**	0.0684**	-0.0248	0.0011
x lesbian	(0.0228)	(0.0196)	(0.0335)	(0.0311)	(0.0297)	(0.0233)
Observations	137,365	213,342	95,191	144,206	102,969	142,195
R^2	0.6442	0.6403	0.6751	0.6564	0.6901	0.6706
Only women aged 40 or under:						
Lesbian	0.0001	-0.0505***	-0.0061	-0.0712***	0.0024	-0.0786***
	(0.0190)	(0.0117)	(0.0305)	(0.0293)	(0.0278)	(0.0086)
Marriage bans	0.0331*	0.0263	0.0309	0.0292	-0.0022	0.0325
x lesbian	(0.0234)	(0.0220)	(0.0353)	(0.0365)	(0.0316)	(0.0285)
Observations	129,424	119,455	77,569	64,706	83,507	60,418
R^2	0.6547	0.6571	0.6942	0.6815	0.7055	0.6964
Only women over 40 years old:						
Lesbian	0.0637***	0.0716***	0.0462*	0.0307	0.1514***	0.1284***
	(0.0208)	(0.0231)	(0.0353)	(0.0359)	(0.0329)	(0.0238)
Marriage bans	0.0224	0.0044	0.0681**	0.0607*	-0.0384	-0.0320
x lesbian	(0.0279)	(0.0296)	(0.0384)	(0.0389)	(0.0326)	(0.0265)
Observations	46,185	135,730	40,031	106,878	43,459	108,616
R^2	0.5531	0.6079	0.6100	0.6313	0.6323	0.6477

Table F.5 (continued)

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Smaller sample						
Base results:						
Lesbian	0.0259	0.0127	0.0387	0.0079	0.0529**	0.0381**
	(0.0206)	(0.0199)	(0.0369)	(0.0323)	(0.0258)	(0.0221)
Marriage bans	0.0268	0.0346*	0.0270	0.0589**	-0.0127	0.0241
x lesbian	(0.0220)	(0.0235)	(0.0360)	(0.0335)	(0.0266)	(0.0222)
Observations	28,419	36,463	20,991	26,344	22,271	26,506
R^2	0.6305	0.6209	0.6584	0.6428	0.6883	0.6612
Multilevel estimation:						
Lesbian	0.0235	0.0062	0.0398*	0.0041	0.0541**	0.0410**
	(0.0192)	(0.0209)	(0.0282)	(0.0284)	(0.0285)	(0.0248)
Marriage bans	0.0237	0.0350*	0.0244	0.0592**	-0.0188	0.0199
x lesbian	(0.0233)	(0.0256)	(0.0307)	(0.0312)	(0.0316)	(0.0272)
Observations	28,419	36,463	20,991	26,344	22,271	26,506
R^2	-	-	-	-	-	-
Gay-male comparison						
Female estimates:						
Lesbian	0.0337**	0.0040	0.0230	-0.0139	0.0907***	0.0363**
	(0.0161)	(0.0157)	(0.0285)	(0.0281)	(0.0254)	(0.0182)
Marriage bans	0.0226	0.0161	0.0511**	0.0507*	-0.0290	-0.0002
x lesbian	(0.0202)	(0.0195)	(0.0301)	(0.0312)	(0.0259)	(0.0196)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6379	0.6322	0.6750	0.6517	0.6884	0.6651
Male estimates:						
Lesbian	0.0054	-0.1290***	0.0026	-0.1343***	0.1006**	-0.0554
	(0.0213)	(0.0223)	(0.0320)	(0.0316)	(0.0476)	(0.0464)
Marriage bans	-0.0139	-0.0196	0.0350	0.0334	-0.0649	-0.0413
x lesbian	(0.0272)	(0.0290)	(0.0382)	(0.0352)	(0.0517)	(0.0492)
Observations	178,186	295,674	120,413	192,527	128,821	187,391
R^2	0.5100	0.4312	0.5607	0.4770	0.5985	0.5137

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the lesbian indicator and lesbian-proxy interaction term. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

Table F.6. *Same-Sex Percentage* supplementary regression estimates

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Base results:						
Lesbian	0.0530***	0.0184***	0.0683***	0.0303***	0.0684***	0.0383***
	(0.0061)	(0.0068)	(0.0075)	(0.0071)	(0.0071)	(0.0058)
Same-sex percentage	-0.0245**	-0.0364***	-0.0147**	-0.0188***	-0.0215***	-0.0214***
x lesbian	(0.0125)	(0.0117)	(0.0064)	(0.0057)	(0.0054)	(0.0066)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6379	0.6323	0.6751	0.6518	0.6884	0.6652
Full-time, no hours/weeks controls:						
Lesbian	0.0783***	0.0561***	0.0985***	0.0649***	0.0893***	0.0628***
	(0.0060)	(0.0066)	(0.0062)	(0.0061)	(0.0077)	(0.0059)
Same-sex percentage	-0.0250**	-0.0368***	-0.0140**	-0.0212***	-0.0134***	-0.0177***
x lesbian	(0.0130)	(0.0127)	(0.0074)	(0.0071)	(0.0050)	(0.0056)
Observations	119,964	163,718	79,050	110,628	87,238	114,129
R^2	0.3658	0.3333	0.4047	0.3677	0.4140	0.3636
Log(hourly wages):						
Lesbian	0.0567***	0.0276***	0.0763***	0.0428***	[redacted]	
	(0.0061)	(0.0064)	(0.0072)	(0.0073)		
Same-sex percentage	-0.0217**	-0.0351***	-0.0155**	-0.0206***		
x lesbian	(0.0130)	(0.0117)	(0.0067)	(0.0058)		
Observations	175,609	255,185	117,600	171,584		
R^2	0.2556	0.2390	0.3131	0.2742		
Heckman two-step method:						
Lesbian	0.0529***	0.0276***	0.0677***	0.0420***	0.0588***	0.0512***
	(0.0062)	(0.0071)	(0.0075)	(0.0071)	(0.0079)	(0.0058)
Same-sex percentage	-0.0250**	-0.0347***	-0.0147**	-0.0178***	-0.0238***	-0.0201***
x lesbian	(0.0125)	(0.0119)	(0.0064)	(0.0058)	(0.0061)	(0.0064)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	-	-	-	-	-	-
Including sodomy law repeal:						
Lesbian	0.0532***	0.0178***	0.0684***	0.0301***	0.0685***	0.0381***
	(0.0060)	(0.0070)	(0.0075)	(0.0071)	(0.0072)	(0.0057)
Same-sex percentage	-0.0253**	-0.0374***	-0.0156**	-0.0206***	-0.0217***	-0.0217***
x lesbian	(0.0124)	(0.0117)	(0.0067)	(0.0060)	(0.0054)	(0.0063)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6379	0.6324	0.6751	0.6518	0.6885	0.6652
Further including SS equality score:						
Lesbian	0.0528***	0.0165**	0.0681***	0.0296***	0.0685***	0.0384***
	(0.0064)	(0.0069)	(0.0073)	(0.0071)	(0.0074)	(0.0058)
Same-sex percentage	-0.0264**	-0.0394***	-0.0146**	-0.0205***	-0.0213***	-0.0209***
x lesbian	(0.0116)	(0.0102)	(0.0069)	(0.0066)	(0.0054)	(0.0060)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6382	0.6325	0.6753	0.6520	0.6887	0.6654

Table F.6 (continued)

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Adding potential exper. interactions:						
Lesbian	0.0789***	0.0554***	0.0964***	0.0716***	0.0984***	0.0829***
	(0.0069)	(0.0083)	(0.0097)	(0.0091)	(0.0102)	(0.0088)
Same-sex percentage	-0.0255**	-0.0381***	-0.0158***	-0.0208***	-0.0220***	-0.0222***
x lesbian	(0.0123)	(0.0113)	(0.0063)	(0.0055)	(0.0058)	(0.0076)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6380	0.6325	0.6752	0.6521	0.6886	0.6656
Further including educ. interactions:						
Lesbian	0.0875***	0.0681***	0.1014***	0.0806***	0.1037***	0.0867***
	(0.0071)	(0.0078)	(0.0102)	(0.0097)	(0.0104)	(0.0086)
Same-sex percentage	-0.0241**	-0.0368***	-0.0153***	-0.0199***	-0.0218***	-0.0223***
x lesbian	(0.0121)	(0.0111)	(0.0062)	(0.0055)	(0.0058)	(0.0077)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6380	0.6326	0.6752	0.6522	0.6887	0.6656
Including only full-time employed:						
Lesbian	0.0686***	0.0395***	0.0879***	0.0471***	0.0802***	0.0506***
	(0.0054)	(0.0062)	(0.0060)	(0.0056)	(0.0071)	(0.0055)
Same-sex percentage	-0.0234**	-0.0352***	-0.0127**	-0.0187***	-0.0126***	-0.0165***
x lesbian	(0.0123)	(0.0121)	(0.0070)	(0.0067)	(0.0050)	(0.0059)
Observations	119,964	163,718	79,050	110,628	87,238	114,129
R^2	0.4024	0.3627	0.4343	0.3993	0.4417	0.3938
Including only white women:						
Lesbian	0.0573***	0.0225***	0.0736***	0.0338***	0.0726***	0.0438***
	(0.0059)	(0.0069)	(0.0078)	(0.0074)	(0.0073)	(0.0062)
Same-sex percentage	-0.0306***	-0.0420***	-0.0188***	-0.0226***	-0.0233***	-0.0238***
x lesbian	(0.0115)	(0.0106)	(0.0070)	(0.0057)	(0.0058)	(0.0067)
Observations	137,365	213,342	95,191	144,206	102,969	142,195
R^2	0.6442	0.6404	0.6751	0.6565	0.6901	0.6707
Only women aged 40 or under:						
Lesbian	0.0274***	-0.0280***	0.0245***	-0.0427***	0.0017	-0.0496***
	(0.0077)	(0.0085)	(0.0088)	(0.0091)	(0.0084)	(0.0080)
Same-sex percentage	-0.0280***	-0.0424***	-0.0366***	-0.0375***	-0.0153**	-0.0124*
x lesbian	(0.0113)	(0.0104)	(0.0069)	(0.0080)	(0.0072)	(0.0088)
Observations	129,424	119,455	77,569	64,706	83,507	60,418
R^2	0.6547	0.6572	0.6943	0.6815	0.7055	0.6964
Only women over 40 years old:						
Lesbian	0.0827***	0.0771***	0.1036***	0.0812***	0.1223***	0.1040***
	(0.0108)	(0.0116)	(0.0118)	(0.0104)	(0.0087)	(0.0087)
Same-sex percentage	-0.0194	-0.0313	0.0055	-0.0040	-0.0287***	-0.0287***
x lesbian	(0.0241)	(0.0249)	(0.0110)	(0.0090)	(0.0071)	(0.0087)
Observations	46,185	135,730	40,031	106,878	43,459	108,616
R^2	0.5532	0.6080	0.6099	0.6314	0.6323	0.6477

Table F.6 (continued)

	2000 Census		2005-2007 ACS		2008-2010 ACS	
	Cohabs	Marrieds	Cohabs	Marrieds	Cohabs	Marrieds
Smaller sample						
Base results:						
Lesbian	0.0474***	0.0397***	0.0624***	0.0586***	0.0440***	0.0601***
	(0.0081)	(0.0098)	(0.0105)	(0.0102)	(0.0090)	(0.0081)
Same-sex percentage	-0.0229**	-0.0400***	-0.0145**	-0.0082	-0.0196**	-0.0215***
x lesbian	(0.0138)	(0.0147)	(0.0086)	(0.0081)	(0.0085)	(0.0079)
Observations	28,419	36,463	20,991	26,344	22,271	26,506
R^2	0.6306	0.6210	0.6582	0.6427	0.6884	0.6612
Multilevel estimation:						
Lesbian	0.0415***	0.0315***	0.0610***	0.0551***	0.0413***	0.0599***
	(0.0090)	(0.0099)	(0.0100)	(0.0100)	(0.0099)	(0.0090)
Same-sex percentage	-0.0130	-0.0290**	-0.0132	-0.0071	-0.0173**	-0.0173**
x lesbian	(0.0127)	(0.0132)	(0.0106)	(0.0103)	(0.0101)	(0.0094)
Observations	28,419	36,463	20,991	26,344	22,271	26,506
R^2	-	-	-	-	-	-
Gay-male comparison						
Female estimates:						
Lesbian	0.0530***	0.0184***	0.0683***	0.0303***	0.0684***	0.0383***
	(0.0061)	(0.0068)	(0.0075)	(0.0071)	(0.0071)	(0.0058)
Same-sex percentage	-0.0245**	-0.0364***	-0.0147**	-0.0188***	-0.0215***	-0.0214***
x lesbian	(0.0125)	(0.0117)	(0.0064)	(0.0057)	(0.0054)	(0.0066)
Observations	175,609	255,185	117,600	171,584	126,966	169,034
R^2	0.6379	0.6323	0.6751	0.6518	0.6884	0.6652
Male estimates:						
Lesbian	-0.0118*	-0.1501***	0.0263***	-0.1119***	0.0414***	-0.0940***
	(0.0080)	(0.0074)	(0.0092)	(0.0085)	(0.0095)	(0.0086)
Same-sex percentage	0.0436***	0.0279**	0.0265**	0.0164*	0.0201**	0.0092
x lesbian	(0.0129)	(0.0145)	(0.0125)	(0.0100)	(0.0116)	(0.0094)
Observations	178,186	295,674	120,413	192,527	128,821	187,391
R^2	0.5100	0.4314	0.5608	0.4773	0.5985	0.5137

Notes: Cluster-robust standard errors in parantheses. Models include all variables specified in the text, but we report only the coefficients on the lesbian indicator and lesbian-proxy interaction term. *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, in a one-tailed test, respectively. Sources: 2000 Decennial Census and 2005-2010 American Community Surveys PUMS.

References

Ahmed, A., Andersson, L., & Hammarstedt, M. (2013a). Are Gay Men and Lesbians Discriminated Against in the Hiring Process? *Southern Economic Journal, 79*(3), 563-585.

Ahmed, A., Andersson, L., & Hammarstedt, M. (2013b). Sexual Orientation and Full-Time Monthly Earnings, by Public and Private Sector: Evidence from Swedish Register Data. *Review of Economics of the Household, 11*(1), 83-108.

Ahmed, A. M., & Hammarstedt, M. (2010). Sexual Orientation and Earnings: A Register Data-Based Approach to Identify Homosexuals. *Journal of Population Economics, 23*(3), 835-849.

Anderson, D. J., Binder, M., & Krause, K. (2003). The Motherhood Wage Penalty Revisited: Experience, Heterogeneity, Work Effort, ad Work-Schedule Flexibility. *Industrial and Labor Relations Review, 56*(2), 273-294.

Antecol, H., Jong, A., & Steinberger, M. (2008). The Sexual Orientation Wage Gap: The Role of Occupational Sorting and Human Capital. *Industrial and Labor Relations Review, 61*(4), 518-543.

Arabsheibani, G. R., Marin, A., & Wadsworth, J. (2004). In the Pink: Homosexual-Heterosexual Wage Differentials in the UK. *International Journal of Manpower, 25*(3), 343-354.

Arabsheibani, G. R., Marin, A., & Wadsworth, J. (2005). Gay Pay in the UK. *Economica, 72*(286), 333-347.

Arabsheibani, G. R., Marin, A., & Wadsworth, J. (2007). Variations in Gay Pay in the USA and the UK. In M. V. L. Badgett & J. Frank (Eds.), *Sexual Orientation Discrimination: An International Perspective* (pp. 44-61). London: Routledge.

Arrow, K. J. (1973). The Theory of Discrimination. In Ashenfelter, O., & Rees, A. (Eds.), *Discrimination in Labor Markets* (pp. 3-33). Princeton, NJ: Princeton University Press.

Babcock, L., & Laschever, S. (2003). *Women Don't Ask: Negotiation and the Gender Divide*. Princeton, NJ: Princeton University Press.

Badgett, M. V. L. (1995). The Wage Effects of Sexual Orientation Discrimination. *Industrial and Labor Relations Review, 48*(4), 726-739.

Badgett, M. V. L. (2001). *Money, Myths, and Change: The Economic Lives of Lesbians and Gay Men*. Chicago, IL: University of Chicago Press.

Badgett, M. V. L. (2006). Discrimination Based on Sexual Orientation: A Review of the Literature in Economics and Beyond. In V. Rogers III (Ed.), *Handbook on the Economics of Discrimination* (pp. 161-186). Cheltenham, UK: Edward Elgar Publishing.

Badgett, M. V. L., Sears, B., Lau, H., & Ho, D. (2009). Bias in the Workplace: Consistent Evidence of Sexual Orientation and Gender Identity Discrimination 1998-2008. *Chicago-Kent Law Review, 84*(2), 559-596.

Baert, S. (2013). Career Lesbians: Getting Hired for Not Having Kids? IZA Discussion Paper No. 7767.

Baumle, A. K., & Poston, D. L. (2011). The Economic Cost of Homosexuality: Multilevel Analyses. *Social Forces, 89*(3), 1005-1031.

Becker, G. S. (1971). *The Economics of Discrimination (2nd ed.)*. Chicago: University of Chicago Press.

Becker, G. S. (1991). *A Treatise on the Family (Enlarged ed.)*. Cambridge, MA: Harvard University Press.

Berg, N., & Lien, D. (2002). Measuring the Effect of Sexual Orientation on Income: Evidence of Discrimination? *Contemporary Economic Policy, 20*(4), 394-414.

Berry, W. D., Ringquist, E. J., Fording, R. C., & Hanson, R. L. (1998). Measuring Citizen and Government Ideology in the American States, 1960-93. *American Journal of Political Science, 42*(1), 327-348.

Black, D., Gates, G., Sanders, S., & Taylor, L. (2000). Demographics of Gay and Lesbian Population in the United States: Evidence from Available Systematic Data Sources. *Demography, 37*(2), 139-154.

Black, D., Makar, H. R., Sanders, S. G., & Taylor, L. J. (2003). The Earnings Effects of Sexual Orientation. *Industrial and Labor Relations Review, 56*(3), 449-469.

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.

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources, 8*(4), 436-455.

Budig, M. J., & England, P. (2001). The Wage Penalty for Motherhood. *American Sociological Review, 66*(2), 204-225.

Carpenter, C. (2004). New Evidence on Gay and Lesbian Household Incomes. *Contemporary Economic Policy, 22*(1), 78-94.

Carpenter, C. (2005). Self-Reported Sexual Orientation and Earnings: Evidence from California. *Industrial and Labor Relations Review, 58*(2), 258-273.

Carpenter, C. (2008a). Sexual Orientation, Work, and Income in Canada. *The Canadian Journal of Economics, 41*(4), 1239-1261.

Carpenter, C. (2008b). Sexual Orientation, Income, and Non-Pecuniary Economic Outcomes: New Evidence from Young Lesbians in Australia. *Review of Economics of the Household, 6*(4), 391-408.

Centres for Disease Control and Prevention (U.S.). (2013). *Births and Natality*. Retrieved 1 December 2013, from http://www.cdc.gov/nchs/fastats/births.htm.

Clain, S. H., & Leppel, K. (2001). An Investigation into Sexual Orientation Discrimination as an Explanation for Wages Differences. *Applied Economics, 33*(1), 37-47.

Clarke, W. (2013, May 17). France Least Tolerant County in Western Europe of Homosexuals. *The Telegraph*. Retrieved 27 November 2013, from http://www.telegraph.co.uk.

Cushing-Daniels, B., & Yeung, T. (2009). Wage Penalties and Sexual Orientation: An Update Using the General Social Survey. *Contemporary Economic Policy, 27*(2), 164-175.

Daneshvary, N., Waddoups, C. J., & Wimmer, B. S. (2008). Educational Attainment and the Lesbian Wage Premium. *Journal of Labor Research, 29*(4), 365-379.

Daneshvary, N., Waddoups, C. J., & Wimmer, B. S. (2009). Previous Marriage and the Lesbian Wage Premium. *Industrial Relations, 48*(3), 432-453.

DiNardo, J., Fortin, N., & Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Analysis. *Econometrica, 64*(5), 1001-1044.

Drydakis, N. (2011). Women's Sexual Orientation and Labor Market Outcomes in Greece. *Feminist Economics, 17*(1), 89-117.

Elmslie, B., & Tebaldi, E. (2007). Sexual Orientation and Labor Market Discrimination. *Journal of Labor Research, 28*(3), 436-453.

Elmslie, B., & Tebaldi, E. (2012). *Discrimination Against Gay Men: The Leveling of the Playing Field.* Manuscript submitted for publication.

Frank, J. (2006). Gay Glass Ceilings. *Economica, 73*(291), 485-508.

Gates, G. J.(2009). The Impact of Sexual Orientation Anti-discrimination Policies on the Wages of Lesbians and Gay Men. UC Los Angeles: California Center for Population Research.

Gates, G. J., & Steinberger, M. D. (2011). *Same-Sex Unmarried Partner Couples in the American Community Survey: The Role of Misreporting, Miscoding, and Misallocation*. Working paper presented at UCLA-California Center for Population Research and UC-Berkeley Population Center Miniconference on Census Microdata and Applications, UCLA, 3 June 2011.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica,47*(1), 153-161.

Heineck, G. (2009). Sexual Orientation and Earnings: Evidence from the ISSP. *Applied Economic Letters, 16*(13), 1351-1354.

Hox, J. (2010). *Multilevel Analysis: Techniques and Applications (2nd ed.).* New York, NY: Routledge.

Innes, E. (2013, October 16). Five Million Babies Have Now Been Born by IVF - and Half Since 2007, New Figures Confirm. *Daily Mail*. Retrieved 15 November 2013, from http://www.dailymail.co.uk.

Jepsen, L. K. (2007). Comparing the Earnings of Cohabiting Lesbians, Cohabiting Heterosexual Women, and Married Women: Evidence from the 2000 Census. *Industrial Relations, 46*(4), 699-727.

Klawitter, M. (2011). Multilevel Analysis of the Effects of Antidiscrimination Policies on Earnings by Sexual Orientation. *Journal of Policy Analysis and Management, 30*(2), 334-358.

Klawitter, M. M., & Flatt, V. (1998). The Effects of State and Local Antidiscrimination Policies on Earnings for Gays and Lesbians. *Journal of Policy Analysis and Management, 17*(4), 658-686.

Korenman, S., & Neumark, D. (1992). Marriage, Motherhood, and Wages. *The Journal of Human Resources, 27*(2), 233-255.

Kurdek, L. A. (1993). The Allocation of Household Labor in Gay, Lesbian, and Heterosexual Married Couples. *Journal of Social Issues, 49*(3), 127-139.

Laurent, T., & Mihoubi, F. (2012). Sexual Orientation and Wage Discrimination in France: The Hidden Side of the Rainbow. *Journal of Labor Research, 33*(4), 487-527.

Moore, M. R., & Stambolis-Ruhstorfer, M. (2013). LGBT Sexuality and Families at the Start of the Twenty-First Century. *Annual Review of Sociology, 39*, 491-507.

Oaxaca, R. L. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review, 14*(3), 693-709.

Peplau, L. A., & Fingerhut, A. (2004). The Paradox of the Lesbian Worker. *Journal of Social Issues, 60*(4), 719-735.

Petit, P. (2007). The Effects of Age and Family Constraints on Gender Hiring Discrimination: A Field Experiment in the French Financial Sector. *Labour Economics, 14*(3), 371-391.

Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *The American Economic Review, 62*(4), 659-661.

Plug, E., & Berkhout, P. (2004). Effects of Sexual Preferences on Earnings in the Netherlands. *Journal of Population Economics, 17*(1), 117-131.

Primo, D. M., Jacobsmeier, M. L., & Milyo, J. (2007). Estimating the Impact of State Policies and Institutions with Mixed-Level Data. *State Politics and Policy Quarterly, 7*(4), 446-459.

Ransom, M., & Oaxaca, R. L. (2005). Intrafirm Mobility and Sex Differences in Pay. *Industrial and Labor Relations Review, 58*(2), 219-237.

Society for Assisted Reproductive Technologies. (n.d.). *SART National Summary*. Retrieved 14 September 2013, from http://www.sart.org/frame/detail.aspx?id=3893.

Tebaldi, E., & Elmslie, B. (2006). Sexual Orientation and Labour Supply. *Applied Economics, 38*(5), 549-562.

U.S. Census Bureau. (2002). *Technical Note on Same-Sex Unmarried Partner Data from the 1990 and 2000 Censuses*. Retrieved 13 May 2013, from http://www.census.gov.

U.S. Census Bureau. (2009). *Changes to the American Community Survey between 2007 and 2008 and their Potential Effect on the Estimates of Same-Sex Couple Households*. Retrieved 19 October 2013, from http://www.census.gov.

U.S. Department of Labor. (n.d.). *Family and Medical Leave Act*. Retrieved 13 August 2013, from http://www.dol.gov/whd/fmla/.

U.S. Equal Employment Opportunity Commission. (n.d.). *The Pregnancy Discrimination Act of 1978*. Retrieved 18 November 2013, from http://www.eeoc.gov/laws/statutes/pregnancy.cfm.

Weichselbaumer, D. (2003). Sexual Orientation Discrimination in Hiring. *Labour Economics, 10*, 629-642.

Weichselbaumer, D. (2013). Testing for Discrimination against Lesbians of Different Marital Status: A Field Experiment. IZA Discussion Paper No. 7425.

Zhang, X. (2009). Earnings of Women With and Without Children. *Perspectives on Labour and Income*, *10*(3), 5-13.