

---

MENDELU Working Papers

in Business and Economics

43/2013

What Drives the Gender Gap?  
An Analysis Using Sexual Orientation

Josef Montag

**MENDELU Working Papers in Business and Economics**

Research Centre

Faculty of Business and Economics

Mendel University in Brno

Zemedelska 1, 613 00 Brno

Czech Republic

<http://vyzc.pef.mendelu.cz/en>

+420 545 132 605

**Citation**

Montag, Josef. (2013). What Drives the Gender Gap? An Analysis Using Sexual Orientation. *MENDELU Working Papers in Business and Economics* 43/2013. Mendel University in Brno. Cited from: <http://ideas.repec.org/s/men/wpaper.html>

## **Abstract**

Josef Montag: **What Drives the Gender Gap? An Analysis Using Sexual Orientation**

Gender differences in productivity, if any, that are unobserved to researchers may produce an omitted variable bias in gender gap studies. Finding a subpopulation with less acute differences in unobserved characteristics would allow this concern to be addressed. This paper argues that gays and lesbians are one such interesting group—for the intra-household division of labor and its effects on market productivity cannot be sex-determined in this subpopulation. Indeed, there are substantial intra-household variations in labor market outcomes and other characteristics; the patterns and magnitudes are similar to different-sex households. Simultaneously, the gender wage gap between gays and lesbians is much smaller than in the heterosexual population; in specifications that control for geographic location it is near zero. These findings suggest that the intra-household division of labor is an important factor driving gender differences in labor market outcomes. Such an interpretation is consistent with recent studies that control for productivity.

## **Key words**

Gender gap, sexual division of labor, discrimination, specialization.

**JEL:** J16, D10, J22, J24, J70.

## **Contact**

Faculty of Business and Economics, Mendel University in Brno, Zemedelska 1, 613 00 Brno, Czech Republic, tel.: +420 545 132 340, e-mail: josef.montag@mendelu.cz.

## **Acknowledgements**

For valuable discussions and helpful comments on an early draft of this paper I would like to thank Daniele Bondonio, Dušan Drabik, Libor Dušek, Jaroslav Křivánek, Peter Huber, Peter Mantel, Alfredo Paloyo, Giovanni Ramello, Gregor von Schweinitz, participants at the 2012 Annual Meeting of the American Law and Economics Association at Stanford University, the 2010 Workshop on Perspectives on (Un-) Employment at the Institute for Employment Research, Nuremberg, and participants in seminars at the Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Essen, the Czech Economic Society, Brno, and the University of Economics, Prague. All remaining mistakes and omissions are my responsibility.

## 1 Introduction

The key question with regard to the differences in labor market outcomes between men and women (henceforth the gender gap) is what drives them. Are they a consequence of a market failure that allows for discriminatory preferences to affect the labor market outcomes (Becker 1957), or are they driven by factors that are external to the labor market, such as gender-differences in socialization and child rearing? In particular, increasing returns from specialized human capital create incentives for intra-household division of labor and specialization, which may also lead to gender-related differences in productivity (Becker 1985).

The answer is important, because we need to match the cure with the disease: If markets lack competition, an appropriate policy response would fall in the realm of antitrust policy, lowering entry barriers, and enforcement of anti-discrimination laws. Alternatively, if people are remunerated in accord with their marginal product, an appropriate policy response could be public subsidies to kindergartens, compulsory parental leave for men, or subsidies and quotas for women's employment. However, econometricians face a serious issue when trying to evaluate how alternative factors contribute to the gender gap, namely, that the measures of productivity that are usually available in the data—quite often not much beyond respondents' age and education—are too coarse. It is therefore difficult to control for more fine-grained productivity characteristics of individuals, such as type of education or actual experience, which are readily available to employers.<sup>1</sup>

This paper employs a simple, yet novel, approach to address the issue of unobserved individual-level productivity characteristics: by looking at the wages of gay and lesbian couples, we can compare populations of male and female households. Keeping the households intact has the advantage that any effects of intra-household specialization on market productivity are averaged out within individual households. Put differently, the effects of intra-household division of labor are contained within each gender—so that the differences in labor market outcomes between gays and lesbians can not be driven by intra-household variation in productivity. Thus, we obtain

<sup>1</sup> Two recent exceptions are studies by Bertrand, Goldin, and Katz (2010) and Azmat and Ferrer (2012).

an estimate of the gender gap that is net of the effects of intra-household division of labor on productivity, if any.

Two caveats need to be put forth at this point. First, this approach only allows us to dispense with the effects of intra-household division of labor on productivity on the gender gap. We cannot infer anything about the relative importance of remaining factors. Second, and more importantly, this approach requires us to assume, that gays' wages are not affected by discrimination, or not in an important way.<sup>2</sup> While unable to test this assumption empirically—for the very same difficulties with measuring productivity as in the case of men and women in general—I offer two theoretical arguments in its support: (1) Unlike gender or skin color, sexual orientation is unobserved under most circumstances. It is clearly trickier for a homophobic discrimination to occur—relative to sexist or racist discrimination. (2) To the extent gays are a small minority—compare again with women or ethnic minorities—they may avoid homophobic employers.<sup>3</sup>

I use the 2008 American Community Survey (ACS) to obtain a sample of gays and lesbians and draw a random sample of 30,000 households from the rest of the data. Gay and lesbian couples are identified through a household member's relationship to the householder and their sex. Apart from being recent, the 2008 ACS data is preferable since it includes a fuller sample of same-sex couples than in previous years. I find a substantial variation in same-sex households in labor market outcomes and productivity characteristics, such as working hours, individuals' share of households' income, or years of education. Indeed, the patterns and magnitudes are similar to different-sex households. The estimated raw gender gap in hourly earnings in a subsample of homosexual males and females is approximately 11 log points. That is about one third of the gap between men and women living in heterosexual households and it is similar to the gap among singles. These estimates are insensitive to standard controls for human capital, cohort effects, and two-digit occupations and industries.

<sup>2</sup> See Herek (2000, 2002) for evidence that attitudes towards gays and lesbians differ, to the detriment of the former.

<sup>3</sup> Recently, Acquisti and Fong (2013) have tested whether firms discriminate based on the information available in Facebook profiles of job applicants. They find that firms in conservative states and counties do discriminate against Muslim candidates in favor of Christians. Gays have, however, received the callbacks with equal frequency as straight candidates.

However, the gender gap in hourly earnings between gays and lesbians vanishes once geographic variables are included. This is because gay men tend to live in different places than lesbians—more often in large cities such as New York or Los Angeles with high costs of living as well as high wages. While decisions about where to live are likely to depend on a household's means, the differences in location choices between gays and lesbians may also be driven by factors unrelated to labor markets and wages. Specifically, the lower presence of children in gay households, compared to lesbian, implies differences in valuation of adult- versus child-related amenities; and metropolitan centers are more suitable to child-less adults than to families raising children (Black, Gates, Sanders, and Taylor 2002). To address the concern of reverse causality empirically, I look at whether geography "explains" any of the gender-earnings gap among singles. It does not. At the same time, the presence of children explains about one half of the gender wage gap among singles, but none of the wage differences between gay and lesbian households.

The findings in this paper suggest that the intra-household division of labor and implied specialization of human capital have substantial power to explain the gender-related differences in labor market outcomes. This is not to say that other sources, such as discrimination or gender differences in socialization, are unimportant. They may indeed be the reason why intra-household specialization is gender-specific. However, if the policy goal is to mitigate gender inequality, the results presented herein imply that such programs need to take the intra-household decisions into account.

## **2 Background**

### **2.1 Sexual Orientation and Earnings Literature**

Since its inception by Badgett's (1995) study, the empirical literature looking at the effects of sexual orientation on labor market outcomes has mushroomed.<sup>4</sup> The main challenge that all studies on sexual orientation face is the identification of sexual orientation in the data.

<sup>4</sup> For overviews of the literature on sexual orientation and earnings see Ahmed and Hammarstedt (2010); Baumle (2009); Elmslie and Tebaldi (2007); Schmitt (2008). Black, Sanders, and Taylor (2007b) provide an excellent background article.

First, there is no clear cut definition of what constitutes a sexual orientation.<sup>5</sup> Second, data that contain any possibility of identifying sexual orientation are sparse. Third, respondents' willingness to identify their sexual orientation may not be purely random. Since, as Badgett (1995) points out, sexual orientation is generally unobserved, "coming out" is a decision that has its cost-benefit structure and can be correlated with various factors including earnings; the sign of the correlation being uncertain, however.

Researchers studying sexual orientation and earnings are of course aware of these issues. Indeed, one of the main focuses in the literature is testing the stability of substantive findings in alternative datasets as well as across definitions of sexual orientation. The main structure of the findings is this: gays on average earn less per hour than heterosexual men and lesbians earn more than heterosexual women. These results hold for different sexual orientation definitions, across datasets, and—as the more recent research confirms—also across countries. Specifically, similar results are reported for the U.S.,<sup>6</sup> Canada,<sup>7</sup> the Netherlands,<sup>8</sup> Sweden,<sup>9</sup> and the United Kingdom.<sup>10,11</sup> In the next step, we develop a simple theory that provides a unified framework for understanding these empirical patterns.

## 2.2 A Simple Theory of Intra-Household Specialization

People specialize. We go to the doctor if feeling sick, to an economist—preferably—if we need policy advice, or to a lawyer in the case that we want another person to do something involuntarily. Notwithstanding that a doctor may give a policy opinion and an economist could handle some doctoring, people's physical, intellectual, and memory capacities are limited—often

<sup>5</sup> The three leading methods of defining and identifying sexual orientation are behavioral criteria (gender of sexual partners), self-identification, and household structure.

<sup>6</sup> See Allegretto and Arthur (2001); Berg and Lien (2002); Black, Makar, Sanders, and Taylor (2003); Blandford (2003); Carpenter (2004, 2007); Clain and Leppel (2001); Elmslie and Tebaldi (2007).

<sup>7</sup> See Carpenter (2008).

<sup>8</sup> See Plug and Berkhout (2004).

<sup>9</sup> See Ahmed and Hammarstedt (2010) and Ahmed, Andersson, and Hammarstedt (2013).

<sup>10</sup> See Arabsheibani, Marin, and Wadsworth (2005).

<sup>11</sup> No study I am aware of finds statistically significant results in the opposite directions. Badgett (1995), however, found a wage penalty for both gays and lesbians in pooled 1989–1991 General Social Survey data, but her estimates were not significant in the case of lesbians. Black et al. (2003) analyze the same data and show that her estimates are an artifact of the definition of homosexuality she used (she defined homosexuality as having more same-sex sexual partners since the age of 18). More appropriate definitions of homosexuality (based on recent sexual behavior) yield results in line with the rest of the literature. Thus the first study is also an exception in this respect.

very much so. As a consequence, things have to be sacrificed in order to shift our attention and effort elsewhere. In addition, people often get better at tasks they do repetitively, so that our productivity increases with experience and focus. Specialization, in turn, enhances comparative advantages and creates opportunities for exchange.

This subsection develops a theory linking increasing returns to intra-household specialization and sexual division of labor.<sup>12</sup> It turns out that specialization and related decisions on human capital investments may be an important factor driving the gender differences in labor market outcomes, at least in theory. Moreover, specialization may exacerbate the effects of discrimination—or other sources of disadvantage—on women’s relative wages as well as other labor market outcomes.

### 2.2.1 Increasing Returns and Specialization

To illustrate the effects of specialization on productivity and wages, consider first an economy where two consumption goods,  $c_1$  and  $c_2$ , are produced and can be exchanged one for one on the market. Think also of an agent deciding on the allocation of time between three competing uses:  $t_i$  is the time spent on producing the good  $c_i$ , where  $i = 1, 2$ , and  $l$  is her leisure time. This choice is subject to  $24 = t_1 + t_2 + l$ , since a day has 24 hours. Let  $c_i(t_i) = t_i^a$ , where the parameter  $a$  captures the dependence between the scale of production and productivity: if people get better at a task the more time they devote to it then  $a > 1$ .

Let the agent maximize a utility function  $u(c_1, c_2, l)$ , which has the standard properties,  $u' > 0$ ,  $u'' < 0$ , in all arguments. Since  $c_1$  and  $c_2$  are freely exchangeable via the market—so that the agent’s consumption mix is independent of her production mix—we can write  $u = u(c_1 + c_2, l) = u(t_1^a + t_2^a, l)$ <sup>13</sup> and the problem is to find

$$\arg \max_{t_i} u[t_1^a + t_2^a, 24 - (t_1 + t_2)]. \quad (1)$$

<sup>12</sup> See Becker (1985) for a more thorough discussion of these phenomena.

<sup>13</sup> Note that this utility concept relates to the initial ownership, not the actual consumption achieved after the exchange. The situation is technically analogous to one where  $c_1$  and  $c_2$  are perfect substitutes. Akin to substitutability, exchangeability implies that utility is independent of the composition of inputs.

Then, for any  $l^* \in [0,24]$ , the solution to problem (1) is either  $\{t_1^* = 0, t_2^* = 24 - l^*\}$  or  $\{t_1^* = 24 - l^*, t_2^* = 0\}$ . In other words, she fully specializes in the production of one good and exchanges her 'excess' produce on the market, obtaining her optimum consumption mix. Increasing returns allow her to raise the production of the good she specializes in by a larger amount than she forgoes in production of the other good. Increasing returns and the possibility of market exchange thus generate incentives to specialize.

## 2.2.2 Intra-Household Specialization

Turning our attention to the analysis of household behavior, we need to slightly modify our story. Specifically, some goods and services produced at home are not marketable, nor can they be bought outside. Examples of such household-specific production might be child rearing and a homely atmosphere conducive to it. While it may be possible to hire a babysitter, a cleaner, and order a meal delivery service, these are often imperfect and expensive substitutes to the home-produced alternatives.

One may indeed view a household as a "firm" producing and consuming two types of commodities, denote them again  $c_1$  and  $c_2$ , of which one is household-specific, while the other can be sold on the market and thus exchanged for other goods and services. In other words,  $c_1$  and  $c_2$  can only be "traded" within a household. If production of both types of goods requires specific time inputs and yields increasing returns from each input, a household maximizes its utility if each of its members specializes in production of either commodity and they share the total.

To see this clearly, consider a household with two members represented by utility functions  $u(c_1, c_2, l)$  and  $\mu(c_1, c_2, \lambda)$ , respectively, where  $l$  and  $\lambda$  represent their leisure time and properties  $u' > 0$ ,  $u'' < 0$ ,  $\mu' > 0$ , and  $\mu'' < 0$  hold for all arguments. Assume that both household members are equally capable of producing either type of commodity and let  $t_i$  and  $\tau_i$  represent the household members' time spent in production of the commodity  $i$ , respectively, for  $i = 1, 2$ . The total volume of commodity  $i$  produced by the household is  $C_i = t_i^a + \tau_i^a$  and a general function  $U$  capturing a household's options can be written as

$$U[C_1, C_2, l, \lambda],$$

where  $U' > 0$ ,  $U'' < 0$  in all arguments. The household thus needs to find a vector

$$\arg \max_{t_i, \tau_i} U[t_1^a + \tau_1^a, t_2^a + \tau_2^a, 24 - (t_1 + t_2), 24 - (\tau_1 + \tau_2)]. \quad (2)$$

PROPOSITION 1. If  $a > 1$  then for any  $\{l^*, \lambda^*\} \in [0, 24]$ , at least one household member, the one with more leisure time, fully specializes.

*Proof.* I show this in two steps. First, I show that the full specialization of one household member creates the potential to increase the household's welfare—relative to the situation where no one in the household specialises. In the second step I show that deviating from the point where one household member fully specializes decreases the household's welfare.

Take any  $l^*$  and  $\lambda^*$ , so that  $l^* \geq \lambda^*$ , and assume that  $t_1^* = t_2^*$  and  $\tau_1^* = \tau_2^*$  solve the problem (2). However,  $y \in [0, \tau_1^*]$  exists, such that

$$\begin{aligned} U(t_1^{*a} + \tau_1^{*a}, t_2^{*a} + \tau_2^{*a}, l^*, \lambda^*) &< U[(2t_1^*)^a + (\tau_1^* - y)^a, (\tau_2^* + y)^a, l^*, \lambda^*] \\ &= U[(\tau_1^* + y)^a, (2t_2^*)^a + (\tau_2^* - y)^a, l^*, \lambda^*]. \end{aligned} \quad (3)$$

That is, if the household member with more leisure time fully specializes, (so that  $t_1 = 2t_1^*$  and  $t_2 = 0$ , or vice versa), the other household member may reallocate part of her work effort, denoted  $y$ , to a more productive use. This enables the household to extend the production of at least one commodity, without necessarily changing the amount of time allocated to leisure. This can be seen when we differentiate  $U$ , accounting for the second order effects in production, so

$$\Delta U \approx a \sum_{i=1}^2 \frac{\partial U}{\partial C_i} [t_i^{a-1} \Delta t_i + \frac{1}{2}(a-1)t_i^{a-2} \Delta t_i^2 + \tau_i^{a-1} \Delta \tau_i + \frac{1}{2}(a-1)\tau_i^{a-2} \Delta \tau_i^2].$$

Plugging in  $t_2 = t_1^*$ ,  $\tau_2 = \tau_1^*$ ,  $\Delta t_1 = t_1^*$ ,  $\Delta t_2 = -t_1^*$ ,  $\Delta \tau_1 = -y$ , and  $\Delta \tau_2 = y$ , omitting \*s for clarity, and simplifying yields

$$\begin{aligned} \Delta U \approx a \frac{\partial U}{\partial C_1} [t_1^a + \frac{1}{2}(a-1)t_1^a - \tau_1^{a-1}y + \frac{1}{2}(a-1)\tau_1^{a-2}y^2] \\ + a \frac{\partial U}{\partial C_2} [-t_1^a + \frac{1}{2}(a-1)t_1^a + \tau_1^{a-1}y + \frac{1}{2}(a-1)\tau_1^{a-2}y^2]. \end{aligned} \quad (4)$$

Choosing  $y$  so that the first term in brackets is equal to zero, we can substitute for  $-t_1^a$  in the second term in brackets, in which case

$$\Delta U \approx a(a-1) \frac{\partial U}{\partial C_2} [t_1^a + \tau_1^{a-2} y^2] > 0.$$

Now we need to show that alternatives to specialization are inferior. Specifically, consider a household, where one member fully specializes in production of one commodity, say  $C_1$ , so that  $t_1 = 24 - l$  and  $t_2 = 0$  and the other parameters,  $\lambda$  and  $\tau_i$ s, maximize  $U$ . Then the gain from raising  $t_2$  is

$$\begin{aligned} \frac{\partial U}{\partial t_2} &= \frac{\partial U}{\partial C_2} \frac{\partial C_2}{\partial t_2} + \frac{\partial U}{\partial l} \frac{\partial l}{\partial t_2} \\ &= \frac{\partial U}{\partial C_2} a t_2^{a-1} - \frac{\partial U}{\partial l} \\ &= -\frac{\partial U}{\partial l} < 0. \end{aligned}$$

QED

In the special case when  $l = \lambda$ , complete specialization occurs, so that  $y = \tau_i$ , then we can rewrite inequality (3) as

$$\begin{aligned} U(t_1^{*a} + \tau_1^{*a}, t_2^{*a} + \tau_2^{*a}, l^*, \lambda^*) &< U[(2t_1^*)^a, (2\tau_2^*)^a, l^*, \lambda^*] \\ &= U[(2\tau_1^*)^a, (2t_2^*)^a, l^*, \lambda^*]. \end{aligned} \quad (5)$$

Increasing returns allow the household to achieve a higher productivity and consumption via specialization and exchange (or sharing) of the produce. At the same time, specialization has dramatic effects on productivity, and therefore wages, in specific activities. Put differently, it creates differentials in productivity and wages between people in specific activities. However it is not yet implied that these effects will exhibit any systematic pattern with respect to gender, for instance; as apparent from inequality (5), where the pattern of specialization is indeterminate. Consider now a gender-specific comparative advantage in either activity. Specifically, denote  $\delta \in (0, 1]$  to be a penalty for  $u$  when selling  $c_1$  on the market. This may be due to discrimination,

taste-based or statistical, as well as any other factor diminishing the price of  $u$ 's  $c_1$ s. If  $u$  is an employee, her wage will be  $(1 - \delta)t_1^{a-1} < t_1^{a-1}$ , the right part of the inequality being her marginal product. The inequality (5) then becomes

$$U[(1 - \delta)t_1^{*a} + \tau_1^{*a}, t_2^{*a} + \tau_2^{*a}, l, \lambda] < U[(1 - \delta)(2t_1^*)^a, (2\tau_2^*)^a, l, \lambda] \\ < U[(2\tau_1^*)^a, (2t_2^*)^a, l, \lambda].$$

As result of  $\delta$ ,  $u$  will always specialize in production of  $c_2$  and  $\mu$  in  $c_1$ . Moreover, if only  $c_1$  is produced for the market, comparing wages of  $u$  and  $\mu$  cannot estimate the  $\delta$  as the productivity of the two differs. Put differently,  $u$ 's wage  $(1 - \delta)t_1^{a-1}$  is unobserved in the data.

Now consider households where both members are of type  $u$  as well as some households that have two  $\mu$ s living together—in other words gay and lesbian households. Let  $'$  distinguish the household members, then omitting  $*$ s for clarity we can describe a lesbian household as

$$U[(1 - \delta)t_1^a + (1 - \delta)t_1'^a, t_2^a + t_2'^a, l, \lambda] < U[(1 - \delta)(2t_1)^a, (2t_2')^a, l, \lambda] \\ = U[(1 - \delta)(2t_1')^a, (2t_2)^a, l, \lambda].$$

A gay household is analogously

$$U[\tau_1^a + \tau_1'^a, \tau_2^a + \tau_2'^a, l, \lambda] < U[(2\tau_1)^a, (2\tau_2')^a, l, \lambda] \\ = U[(2\tau_1')^a, (2\tau_2)^a, l, \lambda].$$

Comparing wages across these two types of households then provides a direct estimate of  $\delta$ , that is

$$\frac{(1 - \delta)(2t_1)^{a-1}}{(2\tau_1)^{a-1}} - 1 = \delta,$$

assuming  $a$  is the same for both genders.

### 3 Data and Descriptive Statistics

The data analyzed in this paper come from the 2008 American Community Survey (ACS) Public Use Microdata Sample (PUMS) files. The ACS is a survey of households and groupquarters

conducted continuously by the U.S. Census Bureau. Approximately 3 million housing units are selected each year, and their inhabitants are asked to fill out a questionnaire similar to the decennial census form. A subset of the data, which includes most individual and household characteristics, is made publicly available through the PUMS files. Files have been edited by the Bureau to protect the confidentiality of individuals and households; identifiable geographic areas (Public Use Microdata Areas - PUMAs) have a minimum population of 100,000. Thus, the data are not a pure random sample, however sampling weights are provided and all results reported herein are weighted.

### **3.1 Identification of Sexual Orientation**

The ACS is one of the rare sources of information about the homosexual population. Gays and lesbians can be identified through a household member's relationship to the householder and their sex.<sup>14</sup> If a household member identified him or herself as either husband, wife, or unmarried partner (UP), of the householder, and the two are of the same sex, I label them as a same-sex couple. I label different-sex households those where the householder and his or her partner are of opposite sexes and unpartnered households (singles) those where only the householder is present.

There are three immediate issues with this way of identifying homosexual couples. First, only coupled gays and lesbians may be identified. Second, some homosexual couples may select a different description of their relationship (roomer or nonrelative, for instance). Third, and most important, if one member of a heterosexual couple makes a mistake and marks the wrong gender, the two will end up erroneously counted among same-sex couples. Because the number of same-sex households is quite small (less than 1 percent of partnered households), even a small probability of such mistake in respondents' sex identification may result in a substantial contamination of the data on same-sex households with *de facto* heterosexual couples.

Fortunately, these issues have been known and subject to research since the 1990 U.S. Census allowed homosexual couples to identify themselves as unmarried partners for the first time. I provide a detailed overview of the history of identification of same-sex couples in the U.S.

<sup>14</sup> The relevant part of the questionnaire is shown in Figure A.1 in the Appendix.

Census Bureau’s data and the related literature in the Appendix. The key results relevant to our study are that the vast majority of same-sex couples who identify themselves as unmarried partners are true gay and lesbian couples (Black, Gates, Sanders, and Taylor 2000, 2007a) and their characteristics are similar to those of unpartnered homosexuals (Black et al. 2000; Carpenter 2004).

Recently, the U.S. Census Bureau also reflected the research pointing out a large number of misclassifications among same-sex couples who originally identified as spouses.<sup>15</sup> Starting in 2008, the U.S. Census Bureau introduced an improved ACS questionnaire, as well as new editing rules in order to lower the contamination of data on same-sex spouses by misclassified heterosexuals (U.S. Census Bureau 2009). As a result of that, the estimated number of same-sex spouses declined from 341,000 in 2007 to 150,000 in 2008, while the number of same-sex unmarried partners remained statistically identical with point estimates at 413,00 and 415,000, respectively. Thus, unlike previous datasets, the ACS 2008 data allows the pooling of same-sex couples who state that they are married together with the sample of same-sex unmarried partners.

### 3.2 Estimation Sample and Summary Statistics

To obtain an estimation sample of a workable size, I select all of the 6275 same-sex households, take a random sample of 30,000 households from the remainder of the ACS 2008 data, and adjust weights of the sampled observations accordingly.<sup>16</sup> I restrict the sample to couples who are white, born in the U.S., aged between 25 and 55 years, healthy, who have finished school, have four children at most, and earn between \$1 and \$200 per hour. The relatively narrow age selection criteria is meant to further minimize the effects of classification error described above, since this type of error is more likely to be committed by elderly, non-white, immigrant, and non-English speaking respondents (Black et al. 2007a); and to focus the sample on people

<sup>15</sup> Black et al. (2007a) find that over 40 percent of same-sex unmarried partners were likely misclassified as different sex married couples in the 2000 U.S. Census data. The error was however concentrated among same-sex couples who identified their relationship as marriage, in which cases the U.S. Census Bureau designated the relationship as an unmarried partnership. A similar magnitude of misclassification existed in the 2005-2007 ACS data. As a result, only the data on same-sex unmarried partners was usable. See also Gates and Steinberger (2009).

<sup>16</sup> Denote  $\mathbf{W}$  the vector of sampling weights of length  $N$  and denote  $n$  the size of a random sample drawn from  $N$ . The vector of sampling weights for the subsample  $n$  is  $\mathbf{w} = \mathbf{W} \frac{N}{n}$ .

whose sexual behavior is settled (Black et al. 2003). I check the sensitivity of my results to this procedure and report the results below.

The main outcome of interest, hourly wage, is not measured directly in the ACS data, but can be estimated from yearly earnings and from information about weeks worked and usual weekly hours.<sup>17</sup> While such a concept of wage is noisy because yearly amounts and averages may be mis-estimated by respondents, measurement error would be of concern only if patterns of misreporting were different for men and women. Thus, my main measure of hourly pay is the person's total earned income (wage, salary, or income from self-employment) divided by weeks worked during the past 12 months and usual hours worked per week during the past 12 months. I checked the sensitivity of my results to different wage concepts and did not find appreciable differences.

### **3.2.1 Summary Statistics: Households**

Table 1 presents characteristics of the sample of households in ACS 2008 by household type and marital status; sample sizes and population estimates are reported at the bottom. There were about 400,000 gay and lesbian UP households in the U.S. in 2008 plus an additional 80,000 gay and 90,000 lesbian couples who characterized their relationships as marriages.<sup>18</sup>

The household structure of the data allows us to study the intra-household variation in labor market outcomes as well as other variables across living arrangements; which would be obscured if we only looked at population means. Our discussion in Section 2 suggests that intra-household division of labor may drive intra-household variation in observed labor market outcomes; and, if the mechanism works through increasing returns, the prediction holds for gay and lesbian households as well. Table 1 indeed signals that there is a substantial within-

<sup>17</sup> Weeks worked are measured in 6 intervals (0-14, 14-26, 27-39, 40-47, 48-49, 50-52) and I code them as the middle of respective interval.

<sup>18</sup> These estimates exceed the number of about 35,000 legally married same-sex couples in the U.S. in 2008. There were another 80,000 same-sex partners living in civil unions or domestic partnerships who may have reported themselves as married (U.S. Census Bureau 2009; Gates 2009). Some other same-sex partners may regard their relationship as a marriage and identify as married in the survey, even if there is no legal basis for it. Treatment of respondents' marital status in this paper reflects self-identification and not its legality. While the importance of the classification error in the 2008 ACS data remains to be established, the married same-sex classification can be identified and studied independently.

**Table 1**  
**Summary Statistics: Households**

	Same-Sex Couples				Different-Sex Couples		Singles	
	Men		Women		UP	Married	Men	Women
	UP	Married	UP	Married				
Main earner's share of income (%)	64.78	65.31	63.45	67.28	66.86	68.88	100.00	100.00
Main earner's hourly earnings (USD)	41.21	40.85	34.31	32.66	26.73	34.03	24.71	21.41
- difference (USD)	18.40	18.30	13.65	13.40	12.69	16.07	-	-
Weekly hours worked (max.)	47.95	48.22	46.46	46.11	45.79	46.79	45.22	41.62
- difference	7.42	9.57	8.39	10.58	8.42	11.96	-	-
Weeks worked (max.)	50.91	50.77	50.45	50.16	50.29	50.69	48.46	48.20
- difference	3.35	3.59	3.32	6.04	5.88	4.99	-	-
Age (max.)	44.30	43.70	43.91	44.93	39.10	43.79	40.78	41.69
- difference	5.49	3.49	4.53	3.79	4.15	2.95	-	-
Years of education (max.)	16.26	15.77	16.24	16.15	14.37	15.2	14.13	14.41
- difference	1.87	1.80	1.66	1.58	1.63	1.54	-	-
Live in (%):								
- a MSA	89.80	84.26	87.74	83.87	72.66	76.12	79.78	82.72
- a PMSA	48.69	46.78	41.95	38.40	32.28	28.62	32.29	37.51
- S.F., L.A., N.Y., D.C., or Chicago	16.97	11.77	4.44	3.79	2.63	1.62	5.40	5.01
- CA, MA, CT	18.58	27.48	14.93	29.08	10.92	9.44	10.61	14.28
- IA, NH, VT	0.51	0.88	1.54	6.23	2.61	3.66	2.57	1.84
Log of mean costs per bedroom (by PUMAs)	6.33	6.24	6.13	6.09	5.93	5.92	6.00	6.01
<i>N</i> -subsample	663	158	653	176	304	3,472	1,405	1,529
Subpopulation	61,042	15,638	63,227	14,204	1,247,358	12,298,876	6,526,922	6,455,813
<i>N</i> -full sample	2,136	822	2,178	1,119	1,276	15,592	4,558	7,822
Population	194,377	77,517	202,426	89,377	5,644,507	55,427,877	19,657,172	30,828,851

NOTE.—Unit of observation is a household. Rows labeled "max." are means of within-couple maximum values; rows named "difference" represent means of within-couple differences. PUMAs (Public Use Microdata Areas) are geographic areas defined by the U.S. Census Bureau with population of at least 100,000.

household variation in observables among gay and lesbian couples. The main earner's (i.e. the partner with higher total earnings per hour), share of household income is on average close to two thirds among gays and lesbians and slightly higher among heterosexual couples. The absolute differences in hourly earnings are substantial for all groups, but are the largest among gay couples. There is also a substantial intra-household variation in weekly working hours across partnered household types. Differences in weeks worked are higher among heterosexual couples, but remain sizable among gays and lesbians as well. Somewhat surprisingly, the age differences within couples are larger among gays and lesbians than among heterosexuals; this is especially pronounced in the case of gay UP households. Differences in within-couple differences in education across household types are marginally larger in gay households. The similarity of these patterns across household types provides an evidence suggesting that the intra-household division of labor and specialization on market and home production may indeed exist in same-sex households.

The rest of Table 1 reports geographic patterns across our household types. Almost 30 percent of gay and lesbian couples who identified themselves as married live in one of three states in which same-sex marriage was legal at some point in 2008.<sup>19</sup> Compared to different-sex couples, gays as well as lesbians more often live in metropolitan areas; and gays especially are more likely to live in large cities (PMSAs). Gays' higher earnings may thus be partially explained by the fact that they more often live in cities such as San Francisco, Los Angeles, Chicago, Washington D.C., or New York, which are known for their higher cost of living as well as higher wages.<sup>20</sup>

<sup>19</sup> Marriages between same-sex couples have been legal in Massachusetts since 2004 and from June to November 2008 in California, when a constitutional amendment banning same-sex marriages was passed in a referendum. The state Supreme Court in Connecticut also ruled in October 2008 that excluding same-sex marriages was unconstitutional. Iowa and Vermont have allowed same-sex marriages since 2009, and New Hampshire since January 2009.

<sup>20</sup> See also Black et al. (2000, 2002, 2007b). Other top ten cities by their gay population are: Dallas, Philadelphia, Boston, Seattle, and Houston. As many as 34 percent of gay UP households live in these ten cities; for lesbian UPs and married different-sex couples, the proportions are 18 percent and 10 percent, respectively.

### 3.2.2 Summary Statistics: Individuals

Table 2 presents summary statistics for men and women by household types. Gays' hourly wages, as well as yearly incomes, are on average slightly higher compared to men living in heterosexual relationships. Single men's wages are about 20 log points below heterosexual men's and the difference is even more pronounced for yearly earnings and income. When we look at total income of the two partners in a household, it is about 28 log points higher for gays compared to married opposite-sex couples. Shifting attention to women's remuneration, lesbians are the highest earners with hourly earnings about 26 and 22 log points higher than women living in heterosexual relationships and single women, respectively. Lesbians also earn the highest yearly income among women. The combined income of lesbian partners is about 9 log points higher than the average income of heterosexual households, but 19 log points below gays' household income.

On average, gays work about the same number of weeks per year and about one hour less per week than heterosexual men. Men living in heterosexual households work full time with a higher frequency than any other group in the Table. At the same time, partners of heterosexual married men are least likely to be full-time workers. Lesbians (and single women) also work more than women living in heterosexual households, but the differences between lesbians and heterosexual women are larger than the differences among males—more than one week per year and 7.5 hours per week. Consistent with this, lesbians and single women are more likely to be full-time workers.

As in the case with gays, lesbians are more educated than other women. They are also less likely to have children than married women, but married lesbians are about as likely to have kids as single women and women living in UPs. Lesbians and singles are more likely to have served in the military than women with a male partner. Regarding employment, lesbians are slightly less likely to work in a for-profit company than married women and are more likely to own an incorporated business. They are also more likely to work for state or federal government, or to be self-employed. Whereas gays seem to have slightly shorter commuting times than married heterosexual men, lesbians take longer routes compared to other women.

**Table 2**  
**Summary Statistics: Individuals by Household Type**

	Same-Sex Couples		Different-Sex Couples		Singles	
	Men	Women	Men	Women	Men	Women
Log hourly earnings	3.22	3.09	3.18	2.83	2.97	2.87
Log earned income	10.86	10.65	10.84	10.16	10.60	10.41
Log total income	10.89	10.70	10.87	10.19	10.64	10.49
Log household income	11.69	11.50	11.41			
Hours worked per week	44.13	42.02	45.29	36.54	45.22	41.62
Weeks worked per year	49.18	48.47	49.15	47.09	48.46	48.20
Full time workers (%)	87.82	81.73	90.69	63.42	85.41	80.46
Respondent is the main-earner (%)		50	0.68	0.32		100
Respondent's share on income (%)		50	0.62	0.38		100
Age	41.63	41.92	42.61	41.02	40.78	41.69
Years of education	15.22	15.39	14.24	14.46	14.13	14.41
Respondent has children (%) <sup>a</sup>	12.81	24.27	58.18	58.49	11.11	31.98
Number of own children <sup>a</sup>	0.21	0.41	1.07	1.08	0.17	0.53
Works for: (%)						
- a private for profit company	65.88	57.44	68.20	61.61	72.48	66.12
- a private non profit company	9.81	13.03	4.40	12.34	3.45	10.56
- a local government	5.57	11.11	7.27	11.07	4.17	9.14
- a state government	3.06	5.88	3.71	5.13	3.56	6.18
- the federal government	2.69	3.25	3.11	1.11	4.12	2.35
- self-employed, not incorporated	7.18	5.16	6.46	5.49	8.11	4.06
- self-employed, incorporated	5.82	4.14	6.85	3.25	4.11	1.59
Served in the military (%)	6.45	4.00	12.17	1.47	12.04	2.81
Travel time of to work (minutes)	24.21	24.21	25.28	20.07	22.19	22.04
<i>N</i> -subsample	1,642	1,658	3,776	3,776	1,405	1,529
Subpopulation	154,036	155,105	13,568,849	13,447,932	6,526,922	6,455,813
<i>N</i> -full sample	5,928	6,622	16,911	16,911	4,662	8,041
Population	548,117	586,643	61,819,429	60,514,299	20,147,491	31,896,485

<sup>a</sup> Own children in the ACS 2008 data are observed only at the household level. Since only one member of a same-sex couple can be a biological parent of a child, the proportion of gays or lesbians with biological children is obtained by dividing the respective number by two.

Our discussion in Section 2 suggests that, if specialization matters, the differences between gays and lesbians across various labor market outcomes should be smaller relative to differences between heterosexual men and women. This prediction seems to fit the pattern of Table 2. The differences in earnings between gays and lesbians are substantively smaller compared to heterosexual men and women. But, importantly, this is also true for weeks worked per year, weekly working hours, as well as the percentage of full time workers. To the extent that these variables are correlated with job market experience and accumulated labor-market-specific human capital, gays and lesbians may be more comparable across unobserved productivity characteristics. An important, but not surprising, difference between gays and lesbians is in the probability of having children; about a quarter of lesbians have children, which is almost twice as many as do gays. This value is however less than half of the percentage of heterosexuals with children. To the extent children may have a different effect on men's and women's productivity (Bertrand et al. 2010), this should be of smaller concern in the case of gays and lesbians.

## 4 Empirical Model and Results

### 4.1 Model

The base empirical model I employ is the standard Mincer wage equation augmented according to the recommendations of Lemieux (2006). Let  $y_i$  be the log of hourly earnings of a person  $i$ , then

$$y_i = \alpha_0 + \alpha_1 woman_i + \sum_{k=1}^2 \beta_k educ_i^k + \sum_{l=1}^4 \delta_l exper_i^l + \sum_{c(5)} \kappa_c coh_c + \gamma' x_i + e_i, \quad (6)$$

where  $\alpha_0$  is the intercept;  $woman_i$  is a dichotomous variable that equals 1 for women and 0 for men;  $educ_i^k$  is a vector consisting of  $i$ 's years of education and its square;  $exper_i^l$  is a quartic vector in  $i$ 's potential experience, defined as  $exper_i = (age_i - educ_i - 6)$ ;  $coh_c$  is a vector of 5-year cohort indicators;  $x_i$  is a vector of additional control variables; and  $e_i$  is the residual.

The model is estimated on three subsamples: same-sex couples, heterosexual couples, and singles. The main coefficient of interest is  $\alpha_1$  estimated for the same-sex couples subsample. As argued in Section 2, the potential effects of intra-household division of labor on productivity, if any, do not confound the estimate—because we are in fact comparing male and female households. Letting the intra-household differences in market productivity average out within

each household (and therefore within each gender) allows the identification of the effect of gender on wages.

## 4.2 Base Results

The base set of results is reported in Table 3, which presents estimates of the gender gap in hourly earnings for individuals living in three types of households: same-sex couples, heterosexual couples, and singles. Specifications (1), (5), and (9) present estimates of the raw gap, controlling for marital status only so that UPs (i.e. the modal relationship among gays and lesbians) are the reference category. The raw gender gap in hourly earnings among same-sex couples, captured by the coefficient on the variable *'woman'*, is 11.3 log points, which is about two fifths of the estimated raw gap among heterosexual UPs (27.4 log points) and less than one third of the gender gap among heterosexual married couples (35.4 log points), while it is statistically similar to the gender gap among singles (10.2 log points). The coefficients on interactions of marital status with dummies for gays and lesbians are much smaller than in the case of heterosexual couples. Their interpretation is, however, not clear, as the reporting error discussed in Section 3 may be at work.

Columns (2), (6), and (10) report results of the model defined in equation (6). Adding human capital variables results in a small but statistically insignificant decrease in the same-sex gender wage gap and an increase of the gap among heterosexual couples and singles. At the same time, comparing columns (5) and (6), the estimated difference between gender-earnings gaps among heterosexual UPs and married couples decreases and coefficients on interactions between marriage and dummies for men and women lose most of their value. This suggests that the greater gender wage gap among married heterosexual couples may be explained by higher intra-household variability in observed characteristics.

In specifications (3), (7), and (9) I add dummies for employer types, as listed at the bottom of Table 2, as well as a full set of dummies for two-digit occupation codes and two-digit industry codes. I note that inclusion of these variables is controversial—job characteristics may reflect differences in preferences or household specialization effects, but they can also be affected by discrimination, in which case one should not control for them. At the same time, the measures

**Table 3**  
**Gender–Earnings Gap by Respondents’ Household Type**

	Same-Sex Couples				Different-Sex Couples				Singles			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Woman	-0.113*	-0.105*	-0.097*	-0.056	-0.274*	-0.289*	-0.240*	-0.281*	-0.102*	-0.130*	-0.107*	-0.140*
	(0.036)	(0.031)	(0.028)	(0.028)	(0.052)	(0.052)	(0.050)	(0.052)	(0.029)	(0.027)	(0.030)	(0.026)
Married × man	0.005	0.053	0.031	0.057	0.267*	0.088 <sup>+</sup>	0.047	0.104 <sup>+</sup>	-	-	-	-
	(0.060)	(0.051)	(0.048)	(0.048)	(0.049)	(0.045)	(0.037)	(0.044)				
Married × woman	-0.061	-0.055	-0.071	-0.068	0.187*	0.058	0.034	0.070	-	-	-	-
	(0.053)	(0.048)	(0.046)	(0.045)	(0.044)	(0.038)	(0.037)	(0.036)				
Part time		-0.218*	-0.108*	-0.219*		-0.145*	-0.055 <sup>+</sup>	-0.156*		-0.232*	-0.111*	-0.250*
		(0.040)	(0.039)	(0.039)		(0.024)	(0.023)	(0.024)		(0.043)	(0.042)	(0.042)
Years of education		0.018	-0.014	0.002		0.056 <sup>+</sup>	0.016	0.053 <sup>+</sup>		-0.013	-0.034	-0.007
		(0.066)	(0.048)	(0.059)		(0.028)	(0.028)	(0.027)		(0.032)	(0.031)	(0.030)
Years of education <sup>2</sup> /10		0.030	0.031 <sup>+</sup>	0.031		0.024 <sup>+</sup>	0.028*	0.021 <sup>+</sup>		0.047*	0.045*	0.041*
		(0.021)	(0.015)	(0.019)		(0.009)	(0.010)	(0.009)		(0.011)	(0.011)	(0.010)
Potential experience		0.111	0.128 <sup>+</sup>	0.128 <sup>+</sup>		0.039	0.045	0.046		0.117 <sup>+</sup>	0.126*	0.106 <sup>+</sup>
		(0.058)	(0.058)	(0.056)		(0.031)	(0.031)	(0.030)		(0.046)	(0.040)	(0.045)
Potential experience <sup>2</sup> /10		-0.077	-0.089	-0.091		-0.019	-0.027	-0.025		-0.069	-0.074 <sup>+</sup>	-0.058
		(0.050)	(0.049)	(0.047)		(0.025)	(0.024)	(0.024)		(0.041)	(0.035)	(0.040)
Potential experience <sup>3</sup> /10 <sup>3</sup>		0.212	0.242	0.258		0.044	0.078	0.064		0.179	0.191	0.137
		(0.174)	(0.167)	(0.163)		(0.077)	(0.075)	(0.075)		(0.140)	(0.119)	(0.137)
Potential experience <sup>4</sup> /10 <sup>6</sup>		-2.005	-2.268	-2.548		-0.261	-0.761	-0.510		-1.597	-1.704	-1.091
		(2.107)	(1.994)	(1.967)		(0.814)	(0.790)	(0.793)		(1.630)	(1.397)	(1.603)
Lives in a MSA				0.190*				0.129*				0.149*
				(0.042)				(0.022)				(0.034)
Lives in a PMSA				0.122*				0.126*				0.163*
				(0.030)				(0.021)				(0.033)
Lives in S.F., L.A., N.Y., D.C., Chicago				0.238*				0.172 <sup>+</sup>				0.123
				(0.051)				(0.069)				(0.065)
Lives in CA, MA, CT				0.151*				0.096*				0.014
				(0.037)				(0.031)				(0.042)
5-year cohorts	-	Yes	Yes	Yes	-	Yes	Yes	Yes	-	Yes	Yes	Yes
Job characteristics	-	-	Yes	-	-	-	Yes	-	-	-	Yes	-
<i>N</i>	3,300	3,300	3,282	3,300	7,552	7,552	7,498	7,552	2,934	2,934	2,912	2,934
<i>R</i> <sup>2</sup>	0.009	0.220	0.321	0.274	0.075	0.230	0.334	0.255	0.006	0.201	0.316	0.236

NOTE.—Standard errors clustered on households are in parentheses: <sup>+</sup>  $p < 0.05$ , \*  $p < 0.01$ . The cohort born in 1965–1970 is the omitted category. Job characteristics include dummies for the employer types, as listed at the bottom of Table 2, two-digit occupations, and two-digit industries.

of women's human capital characteristics—notably the potential experience—are probably biased upwards. The results in Table 3 however suggest that human capital and job characteristics do not have large explanatory power with respect to the gender wage gap in any of the three subpopulations. Also note that the estimates are less sensitive in the cases of a gay-lesbian wage gap as well as in the case of singles. This is consistent with the patterns seen in Table 2, where gays and lesbians as well as single men and women are more similar in their labor market outcomes and characteristics, relative to heterosexual men and women.

The last set of models reported in Table 3 include dummies, which capture differences in geographic location noted in Table 1. Controlling for geography results in the gay-lesbian gender-earnings gap losing more than half its value; and it is no longer statistically significant. There is no such effect among heterosexual couples or singles, as can be seen in comparison with specifications (2), (6), and (10). This result reflects the fact discussed in Section 3, that lesbians do not locate in large cities and metropolitan centers, characterised by higher living costs as well as high income levels, to the extent gays do. So, is it legitimate to control for geography?

### **4.3 The Role of Geography**

There are various factors giving gays and lesbians incentives to locate in urban areas. They may avoid discrimination in a more anonymous city environment and the urban population may be more tolerant to non-modal lifestyles. At the same time, gay and lesbian marriage markets are probably better in cities and metropolitan areas than in villages. This is not only because of higher population density, but also due to sexual-minority specific “social infrastructure”, such as gay bars, community-specific social networks, and job opportunities. In addition, since gays and lesbians are less likely to have children, they may place higher value on adult-specific amenities and attach less importance to child-related amenities. Finally, their location choices may be more independent of their parents' address, compared to couples with children or planning to have them.

But why should gays and lesbians locate differently? Black et al. (2002) found that indeed gays' choices of geographic location are explained by the availability of adult-specific amenities

in cities rather than by anti-gay sentiments.<sup>21</sup> At the same time, as reported in Table 2, lesbian couples have children twice as often as gays.<sup>22</sup> To the extent that the presence of children generates differences in the valuation of child-related versus adult-related amenities across households, it can also lead to differences in geographic choices between gay and lesbian couples.<sup>23</sup> It would, however, be naive to state at this point that the choice of geographic location drives the gay-lesbian wage differential. Reverse causality may also be at work: Since lesbians earn on average less than gays, they may not be able to afford to live in the same areas. It would not be surprising then that the gender gap among same-sex couples may be “explained away” by differences in geographic location.

One possible way to address the issue of reverse causality is by looking at singles. Because single men and women, unlike different-sex couples, can also choose their geographic location in ways reflecting their gender-specific means or preferences, they constitute a natural counterfactual to same sex-couples in this respect. To the extent that the geographic differences between gays and lesbians can be driven by the differences in wages, one would expect a similar effect in the case of single men and women who exhibit the same gender wage differential. Unlike same-sex couples, however, the location choices of single men and women are not independent of the choices of the opposite sex. This is because marriage markets would create a natural drag to the geographic sorting of singles; at least for those interested in meeting partners of the opposite sex. As a result, the comparison of same-sex couples with singles is imperfect. But still, if wages determine geographic location, one would expect the estimate of the gender gap among singles to be lower in specifications when controlling for geography.

Table 4 reports the result of alternative models which capture the effects of geography on the gay-lesbian gender wage gap as well as the gap between single men and women. The estimates of the raw gap (controlling for the marital status and the cohort effects only) are virtually the

<sup>21</sup> See also Black et al. (2007b).

<sup>22</sup> The difference for unmarried partners is fourfold: 5.23 percent of gay UPs have children versus 20.92 percent of lesbian UPs with children. Among the married couples, 42.21 percent of gays and 38.81 percent of lesbians have children. Since these are children who are present in the household, there is apparently a substantial number of children living in gay and lesbian households in the United States and the ACS data is potentially an interesting source of information about them.

<sup>23</sup> It also may be possible that there are gender-specific (or gay- and lesbian-specific) preferences about living environment, which could be accommodated by gay and lesbian couples, resulting in geographic sorting.

**Table 4**  
**Gay-Lesbian Earnings Gap and Geography**

	Same-Sex Couples					Singles				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Woman	-0.108*	-0.043	-0.001	0.000	-0.016	-0.111*	-0.125*	-0.120*	-0.115*	-0.116*
	(0.035)	(0.032)	(0.032)	(0.032)	(0.032)	(0.029)	(0.028)	(0.028)	(0.028)	(0.027)
Married × man	-0.000	0.026	0.059	0.053	0.043	-	-	-	-	-
	(0.059)	(0.049)	(0.051)	(0.051)	(0.049)					
Married × woman	-0.073	-0.080	-0.070	-0.075	-0.057	-	-	-	-	-
	(0.053)	(0.049)	(0.049)	(0.048)	(0.048)					
Lives in a MSA		0.282*					0.198*			
		(0.046)					(0.037)			
Lives in a PMSA		0.150*	0.356 <sup>+</sup>				0.188*	0.268		
		(0.034)	(0.177)				(0.036)	(0.190)		
Log of mean costs per bedroom (by PUMAs)					0.460*					0.455*
					(0.030)					(0.032)
S.F., L.A., N.Y., D.C., Chicago	-	Yes	Yes	Yes	-	-	Yes	Yes	Yes	-
CA, MA, CT	-	Yes	Yes	Yes	-	-	Yes	Yes	Yes	-
MSAs & CMSAs	-	-	Yes	Yes	-	-	-	Yes	Yes	-
PMSAs	-	-	-	Yes	-	-	-	-	Yes	-
5-year cohorts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	3,300	3,300	3,300	3,300	3,300	2,934	2,934	2,934	2,934	2,934
<i>R</i> <sup>2</sup>	0.054	0.156	0.228	0.249	0.163	0.030	0.098	0.197	0.229	0.130

NOTE.—Standard errors clustered on households are in parentheses: <sup>+</sup>  $p < 0.05$ , \*  $p < 0.01$ . The cohort born in 1965–1970 is the omitted category. PUMAs (Public Use Microdata Areas) are geographic areas defined by the U.S. Census Bureau with a population of at least 100,000.

same for same-sex couples and singles, 10.8 and 11.1 log points, respectively, as reported in columns (1) and (6). In specifications (2) and (7) I add dummies for living in metropolitan and central metropolitan areas, a set of dummies for living in one of the five cities with the highest presence of gays, and dummies for the states in which same-sex marriage was legal at some point in 2008. The coefficients on geographic location are highly statistically significant and of comparable size for both groups—however controlling for geography has very different effects on the estimates of the gender wage gap.

The estimated gender wage gap between single men and women barely changes, as is apparent when comparing columns (6) and (7), but the point estimate of the gay-lesbian gender wage gap loses about 60 percent of its value, as well as statistical significance. Specifications (3) and (4) allow for an unrestricted set of dummies for metropolitan and central metropolitan statistical areas as well as primary metropolitan statistical areas, and there are 264 MSAs and CMSAs and 74 PMSAs in the data. This results in the gender wage gap between gays and lesbians evaporating entirely; yet there is virtually no impact of geography on the gender wage gap between single men and women, as seen in specifications (8) and (9). Note that PMSAs identify within-city areas and geographic sorting is hypothetically possible within a city-specific marriage market.

However, one may still worry that if wages determine geographic location while marriage markets prevent full geographic sorting of single men and women, our specifications with dummies for metropolitan statistical areas may not be appropriate. In large geographic spaces the two effects may simply cancel out. Finer controlling for geography then might produce different results. The smallest identifiable geographic area in the ACS data are Public Use Microdata Areas (PUMA), each with a population slightly over 100,000. There are approximately 3,000 PUMAs across the United States and 630 in the ACS data. If wages determine location, single men and women may live in different PUMAs, reflecting their means, while remaining on the same marriage market. It is thus conceivable that geographic sorting may take place across these relatively small areas. Controlling for living costs across PUMAs should therefore capture the sorting of single men and women related to wages.

To probe this hypothesis, in specifications (5) and (10) I drop all the geographic dummies and replace them with a log of mean costs per bedroom (rent or mortgage payments) for each PUMA. The coefficients on the cost-of-living variable are, not very surprisingly, highly statistically significant; and the point estimates are virtually the same for same-sex couples and singles. However, the costs of living are apparently not correlated with gender in the case of singles. At the same time, the estimated gender wage gap among gays and lesbians remains near zero. These results do not support the hypothesis that wages drive gender differences in geographic location in an important way.

The fact that the gender-earnings gap among singles is fully resistant to controlling for geographic location, as well as living costs, weakens the concern that the wage differences between gays and lesbians drive the gay-lesbian differences in geographic location. As it is unlikely that marriage markets would completely prevent geographic sorting of single men and women, these results suggest that the gender-earnings gap has different sources in the case of same-sex couples and singles. Taken at their face value, the results in Table 4 suggest that different choices in geographic location between gay and lesbian households explain the entire same-sex gender-earnings gap.

#### **4.4 Robustness Checks**

While gay-lesbian differences in the likelihood of having children may create gay-lesbian differences in choices of location, children may also affect incentives to work hard and drive productivity up. This is because children increase household expenses and consume household resources. Altruism towards one's own children is also likely to increase one's work motivation. At the same time, if intra-household specialization among heterosexuals correlates with the sexual division of labor, one may expect similar effects also to take place in the case of singles with children. In addition, children of separated parents live more often with their mothers, while fathers are responsible for paying alimony. This may further and maintain the sexual division of labor even among singles. Ideally, we would like to control for peoples' lifetime plans regarding children. If, however, having children is correlated with such plans, we can partially control for their effect on productivity by including the interaction between the number of children and male and female dummies.

Table 5 reports the results. Looking at singles first, the estimates confirm that having children affects men's and women's earnings differently, with each child estimated to subtract about 9 percent of women's wages. The estimated effect on men's wages, albeit smaller and not statistically significant, is positive and of a similar magnitude. As a result, the estimated gender-earnings gap among singles loses about one half of its value. The picture is quite different among lesbians; here women earn more with each child, an effect that compares well with that for single men. One should however be cautious when interpreting these coefficients causally; lesbians with children may be positively selected. In which case, the coefficients would be picking up other differences in productivity between lesbians with and without children that are not related to children themselves. There is no apparent effect of children on gays' earnings; however the coefficients on children are not comparable between gays and lesbians, because gays' children are only identified at the household level.

While the estimate of the same-sex gender gap is now systematically negative, we see again that most of it can still be attributed to geographic location, whereas children do not explain any substantial portion of the same-sex gender wage gap. If children positively affect lesbians' productivity, the estimates in Table 5 that control for children might be preferable. On the other hand, lesbians' own children may be picking up the effects of intra-household division of labor on productivity, or as just mentioned, lesbians with children may be positively selected, in which case estimates that do not control for children are appropriate. Here I leave it to the consideration of the reader. In either case, these estimates of gender wage gap are substantially smaller than those usually obtained in population samples.

As a further robustness check, I re-estimate Table 4, relaxing the age restriction to 18–65 years as well as including nonwhites and individuals from households in which only one of the partners works.<sup>24</sup> The only change in specifications is that cohorts are interacted with dummies for ethnic groups, to control for any cohort-race-specific shocks. As can be seen in Table 6 the sample size almost doubles, however the results are similar. Although the estimated same-sex earnings gap in specifications controlling for geographic location is no longer zero, it loses 50 to 75 percent of its value and remains substantively smaller than the standard estimates of the

<sup>24</sup> Remaining restrictions for both partners are thus: born in the U.S., out of school, no disability.

**Table 5**  
**Gay–Lesbian Earnings Gap and Geography, Controlling for Children**

	Same-Sex Couples					Singles				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Woman	-0.123*	-0.056	-0.014	-0.009	-0.030	-0.040	-0.057	-0.061 <sup>+</sup>	-0.058	-0.054
	(0.035)	(0.032)	(0.032)	(0.032)	(0.032)	(0.031)	(0.030)	(0.030)	(0.030)	(0.030)
Married × man	0.034	0.035	0.061	0.030	0.051	-	-	-	-	-
	(0.065)	(0.054)	(0.054)	(0.055)	(0.053)					
Married × woman	-0.082	-0.091	-0.078	-0.084	-0.070	-	-	-	-	-
	(0.053)	(0.049)	(0.049)	(0.048)	(0.049)					
Number of children in household × man <sup>a</sup>	-0.053	-0.014	-0.003	0.039	-0.013	0.052	0.085 <sup>+</sup>	0.061	0.064	0.091*
	(0.039)	(0.034)	(0.039)	(0.040)	(0.033)	(0.037)	(0.037)	(0.037)	(0.038)	(0.035)
Number of own children × woman <sup>a</sup>	0.062	0.066 <sup>+</sup>	0.072 <sup>+</sup>	0.074 <sup>+</sup>	0.077 <sup>+</sup>	-0.117*	-0.099*	-0.092*	-0.088*	-0.086*
	(0.034)	(0.034)	(0.033)	(0.034)	(0.033)	(0.021)	(0.021)	(0.021)	(0.020)	(0.021)
Lives in a MSA		0.282*					0.203*			
		(0.046)					(0.036)			
Lives in a PMSA		0.147*	0.347 <sup>+</sup>				0.185*	0.263		
		(0.034)	(0.171)				(0.036)	(0.185)		
Log of mean costs per bedroom (by PUMAs)					0.462*					0.450*
					(0.030)					(0.032)
S.F., L.A., N.Y., D.C., Chicago	-	Yes	Yes	Yes	-	-	Yes	Yes	Yes	-
CA, MA, CT	-	Yes	Yes	Yes	-	-	Yes	Yes	Yes	-
MSAs & CMSAs	-	-	Yes	Yes	-	-	-	Yes	Yes	-
PMSAs	-	-	-	Yes	-	-	-	-	Yes	-
5-year cohorts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	3,300	3,300	3,300	3,300	3,300	2,934	2,934	2,934	2,934	2,934
<i>R</i> <sup>2</sup>	0.057	0.157	0.230	0.251	0.166	0.043	0.109	0.204	0.236	0.140

NOTE.—Standard errors clustered on households are in parentheses: <sup>+</sup>  $p < 0.05$ , \*  $p < 0.01$ . The cohort born in 1965–1970 is the omitted category. PUMAs (Public Use Microdata Areas) are geographic areas defined by the U.S. Census Bureau with population of at least 100,000.

<sup>a</sup> Number of children refers to own children present in the household. The ACS 2008 data does not identify the father, it however identifies whether a woman has her own children or not. In cases when children are present in a household and a woman has own children I assume that she is their mother and code accordingly.

**Table 6**  
**Gay–Lesbian Earnings Gap and Geography, Broad Sample**

	Same-Sex Couples					Singles				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Woman	-0.108*	-0.055 <sup>+</sup>	-0.038	-0.041	-0.027	-0.129*	-0.136*	-0.138*	-0.136*	-0.126*
	(0.026)	(0.025)	(0.025)	(0.025)	(0.025)	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)
Married × man	0.003	0.034	0.063	0.055	0.053	-	-	-	-	-
	(0.044)	(0.040)	(0.041)	(0.041)	(0.039)					
Married × woman	-0.029	-0.034	-0.019	-0.022	-0.024	-	-	-	-	-
	(0.039)	(0.037)	(0.037)	(0.037)	(0.036)					
Lives in a MSA		0.252*					0.193*			
		(0.037)					(0.028)			
Lives in a PMSA		0.123*	0.262 <sup>+</sup>				0.183*	-0.132		
		(0.026)	(0.130)				(0.026)	(0.195)		
Log of mean costs per bedroom (by PUMAs)					0.435*					0.444*
					(0.023)					(0.023)
S.F., L.A., N.Y., D.C., Chicago	-	Yes	Yes	Yes	-	-	Yes	Yes	Yes	-
CA, MA, CT	-	Yes	Yes	Yes	-	-	Yes	Yes	Yes	-
MSAs & CMSAs	-	-	Yes	Yes	-	-	-	Yes	Yes	-
PMSAs	-	-	-	Yes	-	-	-	-	Yes	-
Races × cohorts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>N</i>	5,592	5,592	5,592	5,592	5,592	5,206	5,206	5,206	5,206	5,206
<i>R</i> <sup>2</sup>	0.134	0.205	0.254	0.264	0.217	0.084	0.136	0.205	0.225	0.170

NOTE.—Standard errors clustered on households are in parentheses: <sup>+</sup>  $p < 0.05$ , \*  $p < 0.01$ . Races × cohorts refer to a full set of interactions between 5-year cohorts and nine ethnic group categories. The cohort of whites who were born in 1965–1970 is the omitted category. PUMAs (Public Use Microdata Areas) are geographic areas defined by the U.S. Census Bureau with population of at least 100,000.

gender-wage gap. Expanding the sample has no discernible effect on estimates of the gender gap among singles and—importantly and consistent with our previous estimates—the estimate on the women’s dummy is fully resistant to controlling for geography.

The results do not change appreciably when the sample is restricted to employees or to full time-full year workers. Omitting same-sex couples who identified themselves as married did not change the results either. The point estimates of the gay-lesbian wage gap rise by an additional 3 log points after the inclusion of immigrants (by 1 log point for singles), and become statistically significant. However, there are complex selection issues that drive immigration and the gender-related differences in productivity characteristics as well as gender roles and stereotypes may be different among immigrants and natives. These concerns are only bolstered if we consider immigration of sexual minorities. What remains robust, however, is the low estimate of the gender wage gap between gays and lesbians (and between single men and women), compared to coupled heterosexuals, and the sensitivity of the estimates to controls for geography, a sensitivity which does not exist in the case of singles.

## **5 Conclusions**

This paper addresses the issue of unmeasured gender-specific productivity differences in the gender gap research. It builds on the simple idea that gender-related differences in productivity characteristics, if any, should be smaller among gays and lesbians compared to heterosexuals. This is because the specialization in market and home production is not determined by gender in gay and lesbian households. Thus, even if partners living in same-sex relationships specialize, the effects of specialization on work effort, human capital investments, and related choices average out within households and consequently within each gender. This approach works under the assumption that discrimination based on sexual orientation does not affect gays and lesbians differentially. Although I cannot probe the validity of this assumption empirically, there are reasons to believe such a differential effect should not be overwhelming. First, sexual orientation—unlike gender or race—is generally unobserved, therefore discriminating against gays is not as easy as discriminating against women or ethnic minorities. Second, since gays and lesbians are small minorities, they may be able to avoid bigoted employers.

I find a substantial intra-household variation in labor market outcomes among gay and lesbian couples in the 2008 American Community Survey. These intra-household differences exhibit similar patterns, as well as comparable magnitudes, to the differences within different-sex households. Consequently, at the aggregate level, gay and lesbian populations exhibit substantially smaller differences than those between heterosexual men and women. Focusing mainly on the gender gap in pay, the raw difference in hourly earnings between coupled gays and lesbians in the 2008 American Community Survey is 11.3 log points and it is insensitive to standard controls (education, potential experience, occupations, and industries). Most of the gap, however, disappears when geographic location is controlled for. The point estimate is zero in some specifications. Geography has no effect on the gap among singles, while children explain about one half of it.

Taken at their face value, results in this paper suggest that the main driver of the gender wage gap, as well as other differences in labor market outcomes, is the household. This finding is consistent with the results of two recent studies that have detailed data on productivity unmatched in the previous literature, namely Bertrand et al. (2010) and Azmat and Ferrer (2012). Both papers find that the gender differences in productivity explain a major portion of the gender differences in pay. In both studies, the presence of children and family orientation seem to be important drivers of the productivity differences between men and women.

Since intra-household division of labor apparently produces important effects on productivity and labor market outcomes of men and women, both in theory as well as empirically, more research of this phenomenon would be not be out of place. A direct study of intra-household division of labor and specialization in same-sex households, possibly one using a time use survey, would be especially valuable. Hopefully, such data will become available in the future.

Finally, a word needs to be said about policy implications: Discrimination as the factor behind gender-related differences in labor market outcomes may be overrated. This may be simply because markets are competitive enough, so that discriminating is too costly. The results in this paper, as well as the research just cited, suggest that policies pursuing gender equality need to be designed bearing in mind their potential effects on intra-household decisions as to the division of labor as well as the implications for specialization in market versus home production

and the related human capital investments. Examples of such policies can be parental leave for fathers or subsidies to kindergartens.

## Appendix

### Who are the Same-Sex Couples in the ACS 2008 Data?

Same-sex couples in the American Community Survey are identified through a household member's relationship to the householder and their sex. All household members are asked the same question regarding their relationship to the householder. Specifically, same-sex couples are those where a household member identifies him/herself as either a husband/wife or an unmarried partner (UP) of the householder and the two are of the same sex. The relevant part of the ACS 2008 questionnaire is shown in Figure A.1. There are several issues to be considered before using same-sex couples from the ACS data as data on the gay and lesbian population at large.

**Person 1**

(Person 1 is the person living or staying here in whose name this house or apartment is owned, being bought, or rented. If there is no such person, start with the name of any adult living or staying here.)

**1 What is Person 1's name?**  
Last Name (Please print) First Name MI

**2 How is this person related to Person 1?**  
 Person 1

**3 What is Person 1's sex? Mark (X) ONE box.**  
 Male  Female

**4 What is Person 1's age and what is Person 1's date of birth?**  
Please report babies as age 0 when the child is less than 1 year old.  
Print numbers in boxes.  
Age (in years) Month Day Year of birth

**Person 2**

**1 What is Person 2's name?**  
Last Name (Please print) First Name MI

**2 How is this person related to Person 1? Mark (X) ONE box.**  
 Husband or wife  Son-in-law or daughter-in-law  
 Biological son or daughter  Other relative  
 Adopted son or daughter  Roomer or boarder  
 Stepson or stepdaughter  Housemate or roommate  
 Brother or sister  Unmarried partner  
 Father or mother  Foster child  
 Grandchild  Other nonrelative  
 Parent-in-law

**3 What is Person 2's sex? Mark (X) ONE box.**  
 Male  Female

**4 What is Person 2's age and what is Person 2's date of birth?**  
Please report babies as age 0 when the child is less than 1 year old.  
Print numbers in boxes.  
Age (in years) Month Day Year of birth

FIG. A.1.—Identification of relationships between household members in the 2008 ACS questionnaire.

First, the ACS does not allow identification of unpartnered gays and lesbians. Some of them will be pooled with singles; some may live in different-sex households. Second, there are other possible relationship descriptions that some same-sex couples may choose instead (roomer, housemate, other nonrelative), and there are alternative theories on how this may be correlated with outcomes, the sign of the correlation being undetermined (see e.g. Badgett (1995) for discussion of these issues). The more serious concern, however, relates to the question of whether same-sex households are truly those of gays and lesbians; because of these couples may in reality be different-sex couples in which one of the partners has been mislabeled.

Specifically, all mistakes in sex identification made by heterosexuals result in the sample of same-sex couples being contaminated by heterosexual couples.<sup>25</sup> Because same-sex couples constitute a small fraction of households (less than 1 percent), even a small number of mistakes in identification of respondents' sex can cause high contamination of the same-sex sample with *de facto* heterosexual couples. To understand the magnitude of the misclassification problem, we need to briefly enter the history of how identification of same-sex couples was tackled in the U.S. Census and the ACS data.

Black et al. (2000) were the first to take on this issue thoroughly by studying demographic characteristics of same-sex couples in the 1990 U.S. Census and comparing them with estimates from other sources in which sexual orientation can be identified.<sup>26</sup> They conclude that misclassification is not a serious issue in the 1990 U.S. Census, and that almost all same-sex couples are identified homosexuals. However, they estimate that only about one third of all same-sex couples identified themselves as UP in that data—only about 0.1 percent of households (145,130) were same-sex couples in the 1990 U.S. Census.<sup>27</sup> Carpenter (2004) analyzes the 1996–2000 data from the Centers for Disease Control, which has individual level information on sexual behavior and family planning and suffers from less severe under-reporting of same-sex couples than the 1990 U.S. Census. He finds that the sexual behavior of same-sex couples systematically differs from married heterosexual couples as well as from different-sex UP couples and that it is consistent with health literature on homosexual men and women. Importantly, he shows that economic outcomes of same-sex couples from the Centers for Disease Control data are similar to results based on the 1990 U.S. Census data and are not an artifact of self-selection.

An important change occurred in between the 1990 and 2000 U.S. Censuses: If a couple identified themselves as same-sex and married in the 1990 U.S. Census, such a response was treated as a logical contradiction and the sex of one of the partners was changed so that these couples were counted as different-sex couples. This led to an undercount of same-sex couples. In 2000 the Census Bureau changed the relationship status to UP instead, so that

<sup>25</sup> I abstract from the unlikely case in which both partners select the wrong sex in the form.

<sup>26</sup> The 1990 U.S. Census was the first in which same-sex couples could identify as unmarried partners.

<sup>27</sup> In the 2008 ACS data it is 0.34–0.49 percent, or 397,650–565,500 households, depending on whether same-sex households who originally identified as husband/wife are included.

these couples were enumerated as same-sex UPs. Black et al. (2007a) find that, because of this new handling of same-sex couples who identified as spouses, over 40 percent of same-sex unmarried partners were likely misclassified as different-sex married couples in the 2000 U.S. Census data. While the change of relationship status itself was not flagged, the Bureau also edited responses to marital status questions, which were flagged so that households in which two same-sex partners had originally indicated themselves as being in a marital relationship could be dropped. The resulting sample of same-sex couples would again contain only those who indicated themselves as UP, a sample of mostly true same-sex couples Black et al. (2000, 2007a).

The situation was analogous in the ACS data from 2000 through 2007. However, in response to Black et al.'s research, the Census Bureau implemented a set of changes in 2008 in order to improve the identification of same-sex couples. The changes introduced between the 2007 and the 2008 ACS were of two types: changes in processing and editing rules and changes in the questionnaire format (see U.S. Census Bureau (2009) for more details). Whereas the latter were expected to decrease the probability of mistakes in responses, the former were introduced to better identify and correct such mistakes. As a result, the estimate of the number of same-sex households dropped from 750,000-780,000 in 2005-2007 ACS data files to 565,000 in 2008 ACS data. The whole decline is, however, attributable to the decrease in the number of unmarried same-sex couples who were originally recorded as same-sex spouses, while the number of same-sex couples who identified themselves as UPs remained statistically the same: 413,000 and 415,000 in 2007 and 2008 ACS data, respectively (U.S. Census Bureau 2009). While the role of misclassification error in the sub-sample of married gays and lesbians remains to be established, these couples can be identified and studied independently in the 2008 ACS data.

## References

- Acquisti, Alessandro and Christina M. Fong (2013), "An experiment in hiring discrimination via online social networks." Working paper, Carnegie Mellon University, Pittsburgh, PA.
- Ahmed, Ali, Lina Andersson, and Mats Hammarstedt (2013), "Sexual orientation and full-time monthly earnings, by public and private sector: Evidence from Swedish register data." *Review of Economics of the Household*, 11, 83–108.
- Ahmed, Ali M. and Mats Hammarstedt (2010), "Sexual orientation and earnings: A register data-based approach to identify homosexuals." *Journal of Population Economics*, 23, 835–849.
- Allegretto, Sylvia A. and Michelle M. Arthur (2001), "An empirical analysis of homosexual/heterosexual male earnings differentials: Unmarried and unequal?" *Industrial and Labor Relations Review*, 54, 631–646.
- Arabsheibani, G. Reza, Alan Marin, and Jonathan Wadsworth (2005), "Gay pay in the UK." *Economica*, 72, 333–347.
- Azmat, Ghazala and Rosa Ferrer (2012), "Gender gaps in performance: Evidence from young lawyers." CEP Discussion Paper no. 1136, London School of Economics and Political Science, London.
- Badgett, M. V. Lee (1995), "The wage effects of sexual orientation discrimination." *Industrial and Labor Relations Review*, 48, 726–739.
- Baumle, Amanda K. (2009), "The cost of parenthood: Unraveling the effects of sexual orientation and gender on income." *Social Science Quarterly*, 90, 983–1002.
- Becker, Gary S. (1957), *The Economics of Discrimination*. University Of Chicago Press, Chicago.
- Becker, Gary S. (1985), "Human capital, effort, and the sexual division of labor." *Journal of Labor Economics*, 3, S33–S58.
- Berg, Nathan and Donald Lien (2002), "Measuring the effect of sexual orientation on income: Evidence of discrimination?" *Contemporary Economic Policy*, 20, 394–414.

- Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz (2010), "Dynamics of the gender gap for young professionals in the financial and corporate sectors." *American Economic Journal: Applied Economics*, 2, 228–255.
- Black, Dan, Gary Gates, Seth Sanders, and Lowell Taylor (2000), "Demographics of the gay and lesbian population in the united states: Evidence from available systematic data sources." *Demography*, 37, 139–154.
- Black, Dan, Gary Gates, Seth Sanders, and Lowell Taylor (2002), "Why do gay men live in San Francisco?" *Journal of Urban Economics*, 51, 54–76.
- Black, Dan, Gary Gates, Seth Sanders, and Lowell Taylor (2007a), "The measurement of same-sex unmarried partner couples in the 2000 US Census." Working Paper no. CCPR-023-07, Center for Population Research, UCLA, Los Angeles.
- Black, Dan A., Hoda R. Makar, Seth G. Sanders, and Lowell J. Taylor (2003), "The earnings effects of sexual orientation." *Industrial and Labor Relations Review*, 56, 449–469.
- Black, Dan A., Seth G. Sanders, and Lowell J. Taylor (2007b), "The economics of lesbian and gay families." *Journal of Economic Perspectives*, 21, 53–70.
- Blandford, John M. (2003), "The nexus of sexual orientation and gender in the determination of earnings." *Industrial and Labor Relations Review*, 56, 622–642.
- Carpenter, Christopher S. (2004), "New evidence on gay and lesbian household incomes." *Contemporary Economic Policy*, 22, 78–94.
- Carpenter, Christopher S. (2007), "Revisiting the income penalty for behaviorally gay men: Evidence from NHANES III." *Labour Economics*, 14, 25–34.
- Carpenter, Christopher S. (2008), "Sexual orientation, work, and income in canada." *Canadian Journal of Economics*, 41, 1239–1261.
- Clain, Suzanne H. and Karen Leppel (2001), "An investigation into sexual orientation discrimination as an explanation for wage differences." *Applied Economics*, 33, 37–47.
- Elmslie, Bruce and Edinaldo Tebaldi (2007), "Sexual orientation and labor market discrimination." *Journal of Labor Research*, 28, 436–453.

- Gates, Gary J. (2009), "Same-sex spouses and unmarried partners in the american community survey, 2008." Technical report, Williams Institute, UCLA, Los Angeles.
- Gates, Gary J. and Michael D. Steinberger (2009), "Same-sex unmarried partner couples in the american community survey: The role of misreporting, miscoding and misallocation." Working paper, presented at the Annual Meeting of the Population Association of America, Detroit, MI.
- Herek, Gregory M. (2000), "Sexual prejudice and gender: Do heterosexuals' attitudes toward lesbians and gay men differ?" *Journal of Social Issues*, 56, 251–266.
- Herek, Gregory M. (2002), "Gender gaps in public opinion about lesbians and gay men." *Public Opinion Quarterly*, 66, 40–66.
- Lemieux, Thomas (2006), "The "mincer equation" thirty years after *Schooling, Experience, and Earnings*." In *Jacob Mincer, A Pioneer of Modern Labor Economics* (Shoshana Grossbard, ed.), 127–145, Springer, New York.
- Plug, Erik and Peter Berkhout (2004), "Effects of sexual preferences on earnings in the netherlands." *Journal of Population Economics*, 17, 117–131.
- Schmitt, Elizabeth D. (2008), "Discrimination versus specialization: A survey of economic studies on sexual orientation, gender and earnings in the united states." *Journal of Lesbian Studies*, 12, 17–30.
- 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." Technical report, U.S. Census Bureau, Suitland, MD.