THE EXTENT OF INCOME DISPARITIES BASED ON THE INTERSECTIONS OF SEX AND SEXUAL ORIENTATION: REVISITING A 2003 STUDY BY JOHN M. BLANDFORD

WESLEY J. GROSS

Abstract: Using data from the 1998 to 2012 General Social Survey (GSS), this study uses a broad approach to examine the extent of annual income differences between queer males and queer females and their heterosexual group counterparts. The research primarily revisits a 2003 study by John M. Blandford, but approaches the data using a propensity score matching approach rather than through OLS regression. Overall, while Blandford found that queer males experience a significant income penalty and queer females face an income premium, my research finds that queer status has no statistically significant effect on income. Though the heterosexual and the queer groups are successfully matched, my research finds that queer males and queer females are not systematically receiving higher or lower incomes in the modern labor market. (*JEL J15 J31 J71; Keywords: sexual orientation, queer, income differentials, income disparity)*

Although society continually progresses towards more equity in outcomes for individuals in different demographic groups, economic research continually finds that minority individuals experience worse outcomes relative to their majority group counterparts. Specifically, recent research has found that non-heterosexual individuals often experience differential earnings relative to heterosexuals, with their incomes either being systematically inflated or deflated. These sexuality studies often take a broader approach to quantify income differences, yet the data that researchers utilize all come from 1988-2001. These dated datasets do not capture the recent strides of the gay rights movement, including legalized gay marriage in several states and sexual orientation protection in state employment or all employment in certain states. Therefore, it is crucial to reexamine the experience of non-heterosexual individuals in the modern labor market to understand if differential earnings and the possibility of income discrimination is present. The aim of this study is to holistically examine the extent of income penalties based on the intersections of sex and sexual orientation using an updated dataset and methodology to

assess the existence and magnitude of differential earnings for queer males and queer females (males and females who do not identify as heterosexual).

This research primarily draws upon the previous work by John M. Blandford and his 2003 study "The Nexus of Sexual Orientation and Gender in the Determination of Earnings." Similar to his study, I use data from the General Social Survey (GSS), yet update the dataset using the years 1998-2012. Furthermore, I employ propensity score matching to estimate the treatment effect of being queer on the log of individual annual income, matching the heterosexual and queer groups on several factors that are likely to influence income, such as time spent at work, human capital, family characteristics, location, industry, and occupation. This methodology differs from Blandford's model, which uses OLS regression with 1989-1996 GSS data. Blandford found that gay and bisexual men face a 30.2-31.6% penalty relative to heterosexual men, while lesbian and bisexual women face a 17.4-23.4% premium relative to heterosexual women (2003). Alternatively, using the updated data and model, I find that queer males experience a 5.25% penalty, while queer females observe a 7.19% penalty, yet neither of these results are statistically significant. Therefore, these results note that there are no differential earnings between queer and heterosexual individuals in the modern labor market.

LITERATURE REVIEW

Discrimination studies truly began with Gary Becker's 1971 book The Economics of Discrimination, in which he discussed how the "taste for discrimination" by employers, employees and consumers results in a negative impact on the earnings of minority individuals. In terms of the employer, if he/she has distaste for minority laborers, every dollar paid to the minority worker will result in a monetized disutility for the employer. To minimize this disutility, the employer will offer the minority worker lower wages. If there are enough

nondiscriminatory employers in the labor market, the minority workers will simply relocate to these firms to receive equal wages for equal work. Yet if the number of jobs offered by non-discriminating employers is less than the number of minority job-seekers, then these workers will be forced into a discriminatory environment that offers them lower wages. In this model, not only do the minority laborers lose potential income, but the non-discriminatory employers receive lower profits compared to the employers with a "taste for discrimination" who reduce costs by offering their minority workers lower pay. Thus, in equilibrium, all profit-maximizing firms will pay their minority employees lower wages (1971).

Since 1971, few studies have tested Becker's model by examining the earnings differentials between heterosexual and queer individuals. Badgett's 1995 study is the first to research the effect of sexual orientation in the labor market, finding that the annual income penalty for gay or bisexual men to be 11-27%, while the results for women are not consistently significant. Some of these results can be attributed to occupations, as gay/bisexual men are in the higher-paying jobs but earn less than heterosexual men in these jobs. For women, the model does not adequately control for the unobservable characteristics that distinguish between heterosexuals and non-heterosexuals in the workplace, such as experience and work attachment (Badgett 1995).

Since Badgett's study, several researchers observe gay, lesbian, and bisexual earnings gaps, but find results that conflict with the Badgett's original findings. Blandford and Black et al. both find a gay/bisexual male income penalty of 30.2-31.6% and 14-16% respectively, but find that lesbian/bisexual women observe an earnings premium. Blandford finds that the non-heterosexual female earnings premium to be 17.4-23.4% (2003), while Black et al. find it to be about 20% (2003). Both studies attribute the income gaps to occupational differences. Compared

to heterosexual women, lesbian/bisexual women are less likely to specialize in the home as a caregiver, and focus on labor market activities in higher paying male-dominated occupations. Alternatively, non-heterosexual men concentrate in female-dominated jobs, which depress earnings and returns to human capital accumulation. Although one can also argue that heterosexual men are more likely to specialize in market activities as the traditional provider of the family, the fact that gay/bisexual men earn higher levels of education and are less likely to have dependents would allow them to equally specialize in the labor market (Black et al. 2003, Blandford 2003).

The above studies all utilize data from the General Social Survey (GSS) from about 1989-1996. A 2005 study by Carpenter uses newer GSS data from 1988-2000, and finds that no gay/bisexual male gap exists, yet a lesbian premium does exist of 27%. However, when using data from the California Health Interview Survey, no gay or lesbian earnings gap is found.

A more recent study by Antecol et al. enhanced the research by not only separating lesbian women and gay men from both married and cohabitating heterosexuals, but also examined the causes of the observed earnings penalties and premiums in greater depth. Using data from the 2000 Census, the study finds that lesbians observe a 31.6% premium compared to cohabitating women and a 19.7% premium compared to married women. Gay men also observe an earnings premium compared to cohabitating males of 28.2%, but a penalty of 4.5% compared to married men. The premium for gay men compared to cohabitating men is largely explained by their higher relative education. While there is no strong evidence of occupational segregation influencing the earnings of gay men, occupations do play a large role in the determination of lesbian earnings. Lesbian women experience a larger wage penalty for working in female-dominated occupations, while a smaller wage premium for working in male-dominated

occupations, relative to heterosexual women. However, their higher educational attainment compensates for the negative effects lesbians face within occupations, and primarily explains their premium relative to married women. Thus, educational differences largely account for the gay and lesbian premiums. Alternatively, the earnings penalty observed by gay men relative to married men is likely to arise from unobservable characteristics in the model, such as employer/employee intolerance and poor knowledge of legal protections, as these are more common in lower-paying jobs (Antecol et al. 2008).

Although all of these studies offer strong insight about the relationship between sexual orientation and earnings, I focus on the research by Blandford because his study most connects with my research question. Rather than analyzing the proper measurement of queer status such as Black et al. (2003), or examining the causes of observed differentials such as Antecol et al. (2008), Blandford employs a holistic approach to ascertain the effects of queer status on income. Furthermore, his study offers the most detailed approach to determine the extent and magnitude of income differences for queer males and queer females, and thus the easiest study to reproduce and further improve upon. The documentation has clear variable definitions and precise economic reasoning behind all choices for dependent and independent variables, allowing my study to closely follow his previous research and only diverge to improve a specific aspect of Blandford's approach.

DATA

As with the majority of the sexual orientation income studies, my research uses data from the GSS, a national survey that randomly samples non-institutionalized adults in the United States every two years. The primary purpose for use of this dataset is its inclusion of several questions that inquire about an individual's sexual history, which can be used to proxy for sexual

orientation. Most datasets do not include an individual's sexual orientation, so much of the research uses a proxy for gay, lesbian, and bisexual men and women. Furthermore, other data sources, such as the US Census or the American Community Survey, collect responses based on households, thus any proxy created using that data would exclude any non-heterosexual individual not living in a same-sex household. Furthermore, my research focuses on only those respondents that were involved in full-time or part-time employment in the previous year. I chose to use the years 1998-2012 to allow for the highest amount of queer individuals to enter my dataset without compromising the modernity and relevance of the model.

A major flaw in using the GSS is that I generally am only able to observe a respondent's sexual behavior, rather than their sexual orientation. Since sexual behavior does not explicitly dictate one's sexual orientation, I choose to use the term "queer" to describe non-heterosexual individuals in my research, since "queer" is an umbrella term for all people that do not identify as heterosexual. Thus, my research does not distinguish between gay and bisexual men or lesbian and bisexual women as in previous studies.

Finding the proxy most correlated with sexual orientation is key in this research because for differential earnings to arise, employers, coworkers, and customers must know the worker's sexual orientation. Thus, the worker must choose to disclose his/her sexual orientation. Research by Ragins and Singh finds that the fear to disclose one's queer identity at work negatively affects their psychological well-being, causing fewer promotions, less positive career attitudes, and more physical stress-related symptoms. Thus, the choice to disclose is a balance between the psychological well-being of being open about one's identity, and the negative fear-based consequences that inhibit it (2007).

The 2008 to 2012 GSS datasets do ask respondents to self-report their sexual orientation. Although the 2008 to 2012 samples are too small to produce meaningful results, the data can be used to test proxies for sexual orientation. In the end, my research cannot assuredly state that those individuals who self-report their queer status will disclose their sexual identity in the workplace. Thus, I am assuming that those who self-report their queer status will be in favor of their psychological well-being, and will disclose their queer identity at work, at least partially.

Several proxies have been proposed throughout the literature, including having a same-sex sex partner since age 18, having more same-sex than opposite-sex sex partners since age 18, having exclusively same-sex sex partners or having both male and female sex partners in the past year, and having exclusively same-sex sex partners or having both male and female sex partners in the past five years (Badgett 1995, Black et al. 2003). Additionally, Blandford identifies a fifth proxy which uses both the one year and five year behavioral proxies. He identifies respondents as queer if he/she had exclusively same-sex sex partners or had both male and female sex partners in the past year. If the respondent has not had sex in the past year, the previous five years of sexual history are used (2003). After creating proxies for sexual orientation using all five of these measures, I use confusion matrices to determine how accurately these proxies predict a respondent's self-identified sexual orientation in the 2008 to 2012 samples. Of the proxies, sexual behavior over the past five years most accurately predicted sexual orientation with a success rate of 98.22%.[1] Thus, in my research, a respondent is identified as queer if he/she has had exclusively same-sex sex partners or both male and female sex partners in the past five years, and is identified as heterosexual otherwise.

[1] Confusion matrices are not shown in this research.

Table 1. Variable Means and Percentages for Male Workers, by Sexual Orientation

	Male	
	Heterosexual	Queer
N	3959	175
Percent of Male Workers	95.77%	4.23%
Weeks Worked Per Year	48.42	47.94
Hours Worked Per Week	45.83	43.34
Annual Income	$45,560	$39,890
Education:		
Less Than High School	9.27%	5.71%
High School	50.47%	45.14%
Associate	8.59%	8.57%
Bachelor	21.00%	24.58%
Graduate	10.67%	16.00%
Potential Experience	21.36	20.89
Nonwhite	19.20%	20.00%
Age	41.30	41.58
Married	55.97%	20.00%
Lives in Large Metropolitan Area	14.60%	29.14%
Occupation:		
Management, Professional, & Related Occupations	37.86%	52.39%
Service Occupations	10.68%	12.57%
Sales & Office Occupations	14.82%	19.53%
Natural Resources, Construction, & Maintenance Occupations	15.68%	4.57%
Production, Transportation, & Material Moving Occupations	20.96%	10.94%

Notes: Variable means and percentages for respondents working part-time or full-time, based on author's calculations. Education is grouped as the highest educational degree attained by the respondent. Potential experience is a proxy for a worker's job experience, calculated as age minus the sum of years of education and six. Respondents with a negative level of potential experience are omitted from analysis.

Table 1 provides the basic summary statistics for the male sample. As expected, queer male workers represent a small sample of the total male population, with only 175 individuals and 4.23% of the male population. Queer males earn a much lower income on average, about $6,000 less than their heterosexual counterparts, despite being more concentrated in the higher educational degree categories. However, the lower average income could be attributed to the fact that heterosexual men work more weeks in the year, more hours in the week, and have higher levels of potential experience on average.

Table 2. Variable Means and Percentages for Female Workers, by Sexual Orientation

	Female	
	Heterosexual	Queer
N	3731	188
Percent of Female Workers	95.20%	4.80%
Weeks Worked Per Year	47.16	47.27
Hours Worked Per Week	38.76	41.84
Annual Income	$30,810	$33,710
Education:		
Less Than High School	5.98%	6.38%
High School	52.16%	46.81%
Associate	10.27%	11.70%
Bachelor	21.96%	19.15%
Graduate	9.63%	15.96%
Potential Experience	20.53	16.98
Nonwhite	23.24%	17.55%
Age	40.63	37.36
Married	52.56%	25.00%
Lives in Large Metropolitan Area	14.53%	19.15%
Occupation:		
Management, Professional, & Related Occupations	44.41%	43.62%
Service Occupations	18.25%	17.55%
Sales & Office Occupations	30.18%	23.94%
Natural Resources, Construction, & Maintenance Occupations	0.78%	5.32%
Production, Transportation, & Material Moving Occupations	6.38%	9.57%

Notes: Variable means and percentages for respondents working part-time or full-time, based on author's calculations. Education is grouped as the highest educational degree attained by the respondent. Potential experience is a proxy for a worker's job experience, calculated as age minus the sum of years of education and six. Respondents with a negative level of potential experience are omitted from analysis.

In Table 2, which shows the variables means and percentages for the female population, we find results that differ significantly relative to the male population. The queer female sample earns more on average compared to heterosexual females, as their annual income is about $3,000 greater. Educational differences are unlikely to be the cause of this difference, as the concentrations in each educational degree category do not greatly differ between the female groups. However, time at work does vary slightly, as queer females on average work more. The heterosexual female population does have a higher level of potential experience, yet this may

stem from the fact that they are on average an older population. Again, the queer population is a small percentage of the female group, with only 188 observations.

The occupational distribution by sexual orientation is also interesting. Queer females are more concentrated in traditionally male dominated occupations, which may also cause their higher average income since male dominated occupations are generally higher paying (Gauchat et al. 2012). For queer males, the data provides important further insight on the topic of workplace disclosure. In line with previous research, I do find that queer males are highly concentrated in more female-dominated occupations, such as management, professional, service, sales, office, and related occupations. These occupations are more likely to be accepting of queer males, either because of the workplace environment or the clearly instituted guidelines prohibiting discrimination based on sexual orientation. Thus, the queer males may choose these occupations because disclosure of sexual orientation is less damaging. The previous assumption that queer individuals are disclosing their sexual orientation is therefore supported by this finding, as queer males are concentrated in occupations that are most conducive to disclosure. The possibility that certain occupations or industries are more or less accepting is a factor that may bias my final estimates, yet I am unable to remove the effects of queer acceptance in my research (Drydakis, 2009).

MODEL

A major issue with using the GSS is that last year's income is reported in bracketed categories. Therefore, I choose the midpoint of each category for those below the highest income value, and choose $120,000 for the highest income category to proxy for each respondent's annual income. This calculated income value is used for all analyses in this study. Thus, the income variable is not continuous. Blandford's usage of OLS regression is not the most

appropriate, since the method assumes that dependent variable is continuous. Although OLS regression can be done with discrete dependent variables, alternative models for analysis would be better suited for the data, such as propensity score matching. This method assigns scores to individuals based on specified characteristics other than the treatment effect of interest, which is queer in this study. Matching methods provides a different way to control for variables that are likely to affect the earnings outcome by creating two groups. The non-treated group, heterosexual, is able to have the same characteristics of the treated group, queer, except for sexual orientation. Thus, I can better compare the incomes of heterosexual and queer individuals if the groups are adequately matched. Rather than simply removing the effects of specific factors by controlling for independent variables in an OLS model, I am able to examine the sole effect that being queer has on income compared to a highly similar heterosexual as all other observable characteristics are the same.

Therefore, the major difference my research has with Blandford's original study is the model for econometric analysis of the GSS dataset. To do propensity score matching, I first ran the following logit model:

$$Queer_i = \beta_0 + \beta_i X_i + \varepsilon_i$$

where *queer* is a dummy variable equaling one if the respondent is identified as queer and zero otherwise, and ε_i represents the random error for each respondent *i*, capturing all unobserved variables. The observable factors to be matched on are in X_i, which is a vector of independent variables for each respondent *i*, that includes: weeks worked last year, hours worked last week, four dummies for highest educational degree attained, potential experience and its square, marital status, size of city respondent *i* resides in, three dummies for geographic region, eleven dummies

for industry, and four dummies for occupation.[2] This estimable model controls for the majority of factors that have been shown in previous research to impact an individual's earnings, such as human capital accumulation, industrial and occupational placement, family structure, and location, so that the queer and heterosexual groups can be as similar as possible. The fitted values from this logit model are my propensity scores.

Another key difference between my research and the previous research by Blandford are the variables that I am controlling. Rather than using years of education, I employ educational degree attainment dummies, since the degree an individual holds is likely more influential on earnings compared to simply having one more or one less year of education. I also include industry dummies to allow for more variation in employment characteristics, such that a queer and heterosexual individual matched on both their industry and their occupation have highly similar employment placement. However, the most important independent variables included are weeks worked last year and hours worked last week. The majority of research on earnings differences based on race and sex use hourly wage instead of annual income to eliminate the possibility of income premiums stemming from simply working more, or penalties arising for groups that work less. Since I do not have wage data, it is integral that I match respondents with similar time spent at work to not overestimate or underestimate the effect that queer status has on income. All other control variables were present in Blandford's OLS model.

The final aspect of the model is to examine the treatment effect of queer status on the log of income. By taking the log of income, I normalize my income data and reduce the wide spread

[2] I omit the "Less Than High School" degree category, the "Northeast" region category, the "Personal Services" industry category, and the "Management, Professional, and Related Occupations" occupation category to prevent perfect collinearity. I use the dummy variable "metropolitan" to control for the size of city respondent *i* resides in. Metropolitan takes the value of 1 if the respondent lives in a large metropolitan area with a population of 250,000 or more. Industry is coded using U.S. Bureau of the Census three-digit industrial classification codes for 1998 to 2010 and the 2007 NAICS/Census Industrial Classification Codes for 2012. Occupation is coded using the 1980 U.S. Census three-digit occupational codes for 1998 to 2010 and the 2010 U.S. Census three-digit occupational codes for 2012.

Table 3. Propensity Score Matching Results for the Male and the Female Workers

	Male	Female
Queer Estimate, β	-0.05397 (0.07427)	-0.07460 (0.08918)
Number of Matched Observations	175	188

Notes: Statistical significance is denoted with no asterisk ($p > 0.10$), a single asterisk ($p < 0.10$), a double asterisk ($p < 0.05$) or a triple asterisk ($p < 0.01$).

of data points. The multivariate matching was done with replacement to decrease bias, and used one-to-five exact matching. Thus, each queer individual was matched to five heterosexual individuals with the exact same propensity score, and the exact same characteristics. Through this process, I could nearly ensure that my groups would be closely matched. By using one-to-five matching, I eliminate the possibility that strange outliers within the heterosexual group are driving any observable results, as the incomes of the five matches are averaged prior to comparing with the income of the queer individual. Finally, by testing if balance exists in my matched dataset, I can determine if the propensity score matching process results in successfully matched observed groups. Notably, I complete this process twice for both the male sample and the female sample to generate results for both groups.

RESULTS

The estimated effect of queer status on the log of income is presented in Table 3, for both the male and female sample. The queer estimates are only approximates of the true value. To better interpret the extent of differential earnings for queer individuals, I employ the formula:

$$\%\Delta\hat{y} = 100(\exp(\beta) - 1)$$

Through this method, I find that queer males have an income penalty of 5.25%, while females have an income penalty of 7.19%. Although the coefficients for queer are negative for both the male and the female sample, neither estimate is statistically significant. Thus, I cannot reject the null hypothesis that the income penalty is equal to zero, and there is no systematic difference

Table 4. Matching Balance Checks for the Male and Female Workers

	Male Workers		Female Workers	
	Before Matching	After Matching	Before Matching	After Matching
T-test p-value	$< 2.22e^{-16}$	0.5178	$1.5099e^{-13}$	0.6442
KS Bootstrap p-value	$< 2.22e^{-16}$	0.93	$< 2.22e^{-16}$	1
KS Naïve p-value	$< 2.22e^{-16}$	0.9351	$< 2.22e^{-16}$	0.9999

Notes: The three p-values listed above are only a subset of all the test statistics provided. However, all other test statistics similarly find the same conclusion that the groups are successfully matched, and were thus left out for clarity.

between queer males or queer females and their heterosexual counterparts in terms of earnings. Notably, these results are highly dissimilar to the results found by Blandford's previous study. In his research, gay and bisexual men experienced a large 30.2-31.6% penalty, while females faced a 17.4-23.4% income premium.

After checking the matching balance, I can confirm that the queer groups and the heterosexual groups are successfully balanced. Table 4 shows the several different p-values for the male sample and the female sample before and after matching, which can be used to test the success of the matching process. Results show that the matching was successful. For both the male and female workers, I reject the null hypothesis that there is no difference between the queer group and the heterosexual group before matching, as the T-test p-value, the KS Bootstrap p-value, and the KS Naïve p-value are all very small. However, after matching, I fail to reject this null hypothesis as all the p-values are greater than 0.10. Thus, between these adequately matched groups, no differential earnings arise for the queer males or the queer females.

CONCLUSION

In sum, my research does not find that differential earnings exist for queer males or queer females in the modern labor market. Though the coefficients on these variables were negative when estimating their effect on the log of income, neither were statistically significant. This

result is contrary to the results by Blandford and other econometricians who often find income penalties for queer males and income premiums for queer females.

The cause of these differences may stem from many factors. The lack of results from the more updated dataset could indicate a more equitable labor market with new employment protections or new employer perceptions of the queer community, ending the prior bias against queer males and bias for queer females. Alternatively, by using propensity score matching, I may be better controlling for the factors that influence income compared to an OLS regression method. Through this more appropriate method, I may be better removing their effect, and thus eliminating the prior source of the income differentials. My primary hypothesis, however, is that no prior study controlled for time at work, which is a major determinant of earnings. Without controlling for this factor, previous research may have just been finding that queer females worked more than heterosexual females, while queer males worked less than their heterosexual counterparts, which was observed in the summary statistics. Regardless of the cause, my research uses a new methodology to analyze an updated GSS dataset from 1998-2012, and found that queer males and queer females do not observe income penalties or premiums relative to heterosexuals.

REFERENCES

Antecol, Heather, Anneke Jong, and Michael Steinberger. 2008. "The sexual orientation wage gap: The role of occupational sorting and human capital." *Industrial and Labor Relations Review*, Vol. 61, No. 4 (July), pp. 518-543.

Badgett, Lee M. V. (1995). "The wage effects of sexual orientation discrimination." *Industrial and Labor Relations Review*, Vol. 48, No. 4 (July), pp. 726-739.

Becker, Gary S. 1971. *The economics of discrimination*, 2nd ed. Chicago: University of Chicago Press.

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*, Vol. 56, No. 3 (April), pp. 449-469.

Blandford, John M. 2003. "The nexus of sexual orientation and gender in the determination of earnings." *Industrial and Labor Relations Review*, Vol. 56, No. 4 (July), pp. 622-642.

Carpenter, Christopher S. 2005. "Self-reported sexual orientation and earnings: Evidence from California." *Industrial and Labor Relations Review*, Vol. 58, No. 2 (January), pp. 258-273.

Drydakis, Nick. 2009. "Sexual orientation discrimination in the labour market." *Labour Economics*, Vol. 16, pp. 364-372.

Gauchat, Gordon, Maura Kelly, and Michael Wallace. 2012. "Occupational gender segregation, globalization, and gender earnings inequality in U.S. metropolitan areas." *Gender & Society*, Vol. 26, No. 5 (August), pp. 718-747.

Ragins, Belle R. and Romila Singh. 2007. "Making the invisible visible: Fear and disclosure of sexual orientation at work." *Journal of Applied Psychology*, Vol. 92, No. 4, pp. 1103-1118.

NORC at the University of Chicago. (1998-2012). General Social Survey [Data file and codebook]. Retrieved from http://www3.norc.org/GSS+Website/