

Forthcoming, *Labour Economics*

**Revisiting the Income Penalty for Behaviorally Gay Men:
Evidence from NHANES III**

Christopher S. Carpenter*

*Assistant Professor of Economics/Public Policy
The Paul Merage School of Business, University of California at Irvine
443 GSM, Irvine, CA 92697-3125, USA*

Abstract

Four studies published in the last ten years use the General Social Surveys (GSS) to show that behaviorally gay/bisexual men earn 15-30 percent less than other men. In this paper I use independent data on sexual behavior from the Third National Health and Nutrition Examination Surveys (NHANES III) and find that same-sex behaving men experience a statistically and economically significant income penalty on the order of 23-30 percent. Moreover, the strongest evidence for an income penalty is found for those men most likely to have a gay sexual orientation based on their lifetime sexual behavior. That the penalty for same-sex behaving men is robust across data sources suggests the need for more research into its causes and consequences.

JEL classification: J1; J2

*Tel.: +1-949-824-6112; Fax: +1-949-725-2883.

E-mail address: kittc@uci.edu. Generous financial support from The Paul Merage School of Business at UC Irvine and the Robert Wood Johnson Foundation is gratefully acknowledged. Results do not imply their endorsement.

1. Introduction

No fewer than four recently published studies document a statistically significant earnings penalty associated with same-sex sexual behavior for males in the United States on the order of 15-30 percent of annual earnings (Badgett 1995, Berg and Lien 2002, Black et. al. 2003, Blandford 2003). Notably, however, these large estimated earnings penalties for same-sex behaving men have been thus far limited to data from the General Social Surveys (GSS). One published study using the 1990 Decennial Census finds a much smaller effect – only about 3 percent of annual earnings – for men in same-sex cohabiting relationships (Allegretto and Arthur 2001), while recent studies of men in the Netherlands and California find similarly small effects of a self-reported gay sexual preference or sexual orientation (Plug and Berkhout 2004 and Carpenter 2005, respectively). Consequently, a key question remains: are the previous findings of a large earnings penalty for behaviorally gay and bisexual men a peculiarity of the GSS, or do these results extend across other sources of data on males in the United States? Given the importance of the question and the resulting policy implications if such an effect were true, it is clear that we need further research to either confirm or reject the GSS sexual behavior-based estimates.

In this paper I offer new evidence on the earnings effects of same-sex sexual behavior using a data source that – to my knowledge – has gone previously unexplored by economists for these purposes. Specifically, I employ data from the 1988-1994 waves of the Third Health and Nutrition Examination Survey (NHANES III). NHANES III is well-suited to this research question for a variety of reasons. First, and most importantly, the NHANES III and GSS both ask questions about sexual *behavior*. Previous research

suggests that behavior-based concepts may yield quantitatively and qualitatively different results compared to sexual orientation measures based on household living arrangements or self-reported sexual orientation (see, for example, Allegretto and Arthur 2001 and Carpenter 2005). As such, NHANES III provides an important point of comparison to the four recently published GSS studies. Also, the time period covered (1988-1994) is very similar to the GSS sample frame used in the previous papers (1988-1996). Finally, NHANES III includes similar information on demographic characteristics and human capital accumulation as the GSS.

To preview, I find that similarly specified models using NHANES III data confirm the previously documented GSS penalties for same-sex behaving men. Coefficient estimates on various indicators for same-sex sexual behavior indicate a statistically significant income disadvantage of about 23-30 percent relative to similarly skilled non same-sex behaving men. Moreover, the strongest evidence of an income penalty occurs for the measure of sexual behavior that most likely captures the sample of “openly gay” men: those men who report strictly more lifetime same sex partners than different-sex partners.

2. Previous Literature

The literature examining the effect of sexual orientation on individual earnings has grown substantially since Badgett’s landmark study in 1995 (see Table 1). In that paper, Badgett considered wage discrimination against individuals who were behaviorally gay or lesbian in the General Social Survey (GSS) for the years 1989-1991. She found that behaviorally gay men (defined a number of ways depending on the presence of a

same-sex sex partner) earned between 11% and 27% less than their heterosexual male counterparts, with ambiguous results for lesbians. A family of follow-up studies using more recent waves of the GSS and alternative schemes for coding sexual orientation have confirmed the large earnings penalty for behaviorally gay/bisexual men but have also found earnings premiums for behaviorally lesbian/bisexual women, on the order of 20-30 percent of annual income (see, for example, Blandford (2003), Black et. al. (2003), and Berg and Lien (2002)). This earnings penalty has also been documented internationally using data on self-reported sexual preference in the Netherlands. However, those estimates indicated much smaller effects – around 3 percent (Plug and Berkhout 2004).

Evidence of large and/or significant earnings effects associated with same-sex sexual behavior has garnered a variety of economic explanations for the source of such differences. Included among these are: discrimination against gay men, differential human capital accumulation associated with household specialization theories, and models of gender nonconformity. Notably, the development of models to explain the earnings effects in this literature has been almost explicitly post-hoc theorizing. That is, support for these differing models has been predicated on the existence of large and/or statistically significant earnings effects. In this paper I reconsider whether such evidence exists in the case where the evidence seems strongest: the earnings penalty for behaviorally gay men.

3. Data and Empirical Approach

This study makes use of the Third National Health and Nutrition Examination Survey (NHANES III), which was administered in October 1988 – October 1994 to

almost 40,000 households. NHANES III is well suited for the present research because the data contain questions about sexual behavior that are quite comparable to those fielded in the General Social Surveys.¹

Estimation of the correlation between same-sex sexual behavior and income follows previous papers in this literature as closely as possible (particularly Black et. al. 2003). NHANES respondents are asked at what age they first engaged in sexual intercourse. Men in the subsample reporting any lifetime sexual intercourse are asked how many total sexual intercourse partners they have had.² Of these, the respondent is asked to identify how many were male and how many were female. Notably, NHANES III did not ask about more recent sexual behavior.³

I create three key variables of interest intended to capture sexual orientation using the respondent's self-reported sexual behavior. I first create a variable called ANYSSB equal to one if the individual reports any lifetime same sex sexual behavior. I create

¹ Specifically, NHANES III asks respondents "With how many different partners have you ever had sexual intercourse?". It then asks male respondents "How many partners have been female?" and "How many partners have been male?". The wording of the GSS sexual behavior questions has varied somewhat over the years, but a typical question was "Now thinking about the time since your 18th birthday (including the past 12 months) how many male (female) partners have you had sex with?".

² One may be concerned about the importance of sex miscoding in this case. As Black et al. (2000) point out, the small rate of sexual minorities in the population implies that even a small rate of sex miscoding could result in significant biases when considering measurement schemes that assign sexual orientation on the basis of a respondent's relationship to another person (either through reported sex of one's sex partners or the reported sex of one's unmarried partner). While potentially important, this issue is unlikely to be substantive in this context given the nature of the NHANES interviews. Specifically, NHANES respondents are questioned by trained interviewers on multiple occasions, including a household screener performed either in the respondent's own home or by phone. Respondents also undergo a comprehensive physical examination by a board-certified physician. The questions about the sex of one's sex partners – which were only asked to male respondents under age 60 – were administered in a Mobile Examination Center (MEC) with a trained interviewer in the room. And, there are other sex-specific parts of the NHANES questionnaire that occur prior to questions about sexual behavior, such as sections about birth control use and menstruation. Bias from sex miscoding, therefore, would only occur if one's miscoded sex went uncorrected at multiple stages and through repeated person to person interactions. This is highly unlikely.

³ This is an important drawback of NHANES III data relative to the more recent GSS data, which do contain information on the gender of the respondent's sex partners in the past twelve months and the past 5 years.

another variable called ATLEAST equal to one if the man reports at least as many lifetime male sex partners as female sex partners. Finally, MORESSB equals one if the man reports strictly more lifetime male sex partners than female sex partners. The latter two measures are intended to capture those same-sex behaving men who are “most likely” to be openly gay, following Black et al. (2003). Some form of each of these measures has been used in each of the previous GSS studies.⁴

A key limitation of the NHANES III data relative to the GSS is that income in NHANES III is only ascertained at the family level. NHANES III defines the “family” as those individuals in the household living quarters that are related by marriage, blood, or adoption. Absent detailed information on the human capital and labor force characteristics of the other family members, I restrict the sample to men who report their family size is equal to one.⁵ This is the group for whom family income should most closely correspond to individual income. Note that this excludes all married men whose spouses are present at the time of the survey. Because of the rigid definition of “family” in the NHANES III data, it is also unclear how the family size sample selection rule affects cohabiting partnered individuals (both behaviorally gay and straight). I assume that respondents reporting a family size equal to one – whether or not they are partnered – include only their own sources of income when responding to the “family income” question. In some models I also drop from the sample any man who concurrently reports being married and having a family size equal to one, since it is possible that these men

⁴ The relevant sample sizes of men in the ANYSSB, ATLEAST, and MORESSB groups for the main analysis are: 44, 32, and 27, respectively (note that each is a subset of the previous). While these sample sizes are very small, they are quite comparable to those in the published literature using the GSS. See, for example, Badgett (1995). Creating a group of men who report exclusively same-sex sexual behavior results in sample sizes too small for meaningful analysis (10).

⁵ This approach has been used previously where similar “family” or “household” income problems arise. See, for example, Carpenter (2003).

include their spouse's income in their own family income calculation (even if she is not physically present in the household).

Another drawback to the NHANES III data is that it does not allow direct identification of labor earnings. Instead, respondents are asked about family income, as well as income from other sources, including: social security, pension income, AFDC, public assistance, and/or veteran's benefits. In addition to estimating models of family income on the main analysis sample, I also consider the sample of individuals who report receiving no income from any of the aforementioned sources. This approach – combined with the family structure restriction – will also serve to isolate that group of men whose family income most closely approximates individual labor earnings.

Like the GSS, NHANES reports family income only in broad ranges. Moreover, the highest family income bracket is relatively low – \$50,000 per year. I follow Black et al. (2003) and estimate interval regressions which explicitly take into account the censored nature of the data. I adjust the family income interval endpoints by using the Consumer Price Index for all urban consumers and by taking the natural logarithm.⁶

The relevant econometric model can be given by:

$$(1) \text{Log (Annual Income)} = \alpha + \beta_1 X + \beta_2 (\text{Gay Indicator}) + \varepsilon$$

where X is a vector of demographic variables including age and its square, education (5 categories), a dummy variable indicating the respondent is married, race (3 categories), and a dummy variable for urban residence. I use ANYSSB, ATLEAST, and MORESSB

⁶ The actual survey date is not available in the public use files of NHANES III. However, the data do identify the version of the survey employed, from which it is possible to differentiate between responses from 1988-1991, 1992, and 1993-1994. I use the average of the CPI over the relevant years for deflating income to 1984 dollars.

as the gay indicator variables. ε is assumed to be a well-behaved error term. All regressions are weighted by the NHANES III final exam weight.

4. Results

In Columns 1 and 2 of Table 2 we present descriptive statistics from the NHANES III sample for males who did and did not report any lifetime same-sex sexual behavior (among the sample of men reporting family size equals one). A handful of patterns merit mention. First, the level of average income for men reporting same-sex sexual behavior is slightly higher than for other men. Notably, same-sex behaving men are much more highly educated than the balance of the sample: same-sex behaving men are less likely to report having less than a high school education than men who do not report such behavior, and they are much more likely to report a college degree.⁷ The same-sex behaving sample is also much more likely to be white and more likely to live in an urban area compared to other men. Finally, no man reporting lifetime same-sex sexual behavior also reports being currently married.⁸

In Columns 3 and 4 of Table 2 we also report the demographic characteristics of the associated 1988-1996 GSS sample of men reporting a household size equal to one. We do so to provide a direct comparison of characteristics across these two data from similarly defined samples. Although the means for the GSS data differ compared to the NHANES, we find that the overall pattern of observable characteristics is identical across the two data. In the GSS, as in the NHANES, same-sex behaving men report lower

⁷ This educational pattern has also been shown to exist in the GSS, the Decennial Census (using information on same-sex cohabitation), and the 2001 California Health Interview Survey (using self-reported sexual orientation).

incomes and higher education levels than other men. Behaviorally gay men are also more likely to be white, more likely to live in an SMSA, and are older than other men, just as was found in the NHANES. These patterns suggest that the NHANES data provide a reasonable comparison sample for detecting the effects previously found in the GSS.

Table 3 reports coefficients and standard errors associated with estimation of equation (1) on the NHANES III data for each of the three gay indicator variables. The top panel (Panel A) reports estimates from the main sample of men who report family size equal to one, while the bottom panel goes further by restricting the sample to men who are not currently married (Panel B). Panel A of Column 1, Table 3 indicates that men who report any lifetime same-sex sexual behavior experience 23 percent lower income than other similarly situated men, and this estimate is marginally significant at the ten percent level.⁹ The other gay indicators that arguably capture samples increasingly more likely to be “correctly” measured as openly gay men also suggest large and statistically significant penalties associated with same-sex sexual behavior on the order of 30 to 37 percent.¹⁰ Notably, the overall pattern of the estimates in Panel B – the sample

⁸ Recall that the sample has already been limited to those reporting family size equal to one, which explains the low marriage rate among those not reporting same-sex sexual behavior.

⁹ The full set of coefficient estimates from this model is reported in Table A1. All other control variables had the predicted sign and significance. OLS models estimated on the midpoints of the income ranges resulted in very similar estimates. To check whether the NHANES results were truly confirming previous GSS estimates, I also re-estimated GSS models from Column 1 of Table 7 in Carpenter (2005) using the ANYSSB indicator but imposing the sample restriction from Table 2 that the respondent reported a household size equal to exactly one. The coefficient on ANYSSB in the full sample GSS was -.243, while in the restricted sample it was -.257. Both were statistically significant at the 5 percent level or better. These estimates are very close to the NHANES estimate of -.235 in Table 3, suggesting that the NHANES data are, in fact, confirming previous GSS results.

¹⁰ Note that one can explicitly account for the potential sex miscoding issue described above in Footnote 2 by eliminating observations that report only having had sex with men, on the observation that many gay men have had sex with women but almost no heterosexual women will have had same-sex sex (Black et al. 2000). Note that this approach has the drawback that it eliminates some “truly” gay men who have not had sex with a woman in their lifetime. Adjusting the sample in this way results in coefficient estimates that

excluding married men – is qualitatively and quantitatively identical to the main analysis sample: same-sex behaving men have significantly lower incomes than other men with similar characteristics.

The bottom panels of Table 3 (Panels C and D) perform the same exercises as the top panel but restrict the sample to men who report no income from other sources such as social security or AFDC. The concern is that if there exists a systematic correlation between same-sex behavior and the likelihood of receiving such benefits, then the lower incomes for same-sex behaving men in Table 3 may have little to do with labor market outcomes per se. As in Panels A and B, I consider both the sample of men reporting family size equal to one (Panel C), as well as the subset of those men who also report they are not currently married (Panel D). The qualitative patterns evident in the top panels remain in the bottom panels with the restricted sample: same-sex behaving men are estimated to earn substantially less than similarly skilled men who do not report such behavior, and the magnitude of these penalties increases with more refined definitions of the “gay” indicator (i.e. moving from Column 1 to Column 3). The sample restrictions on other sources of income, however, do appear to matter substantively, as the point estimates are everywhere smaller than the associated estimates from the full sample in the upper panels.¹¹ Moreover, the ANYSSB specification is no longer statistically significant at even the ten percent level. This sensitivity to the income restriction suggests that the

were very similar to the reported estimates, suggesting that sex miscoding is not substantively important in this context.

¹¹ Though the NHANES III public use files do not identify the date of interview, I did investigate the possible effects of time effects by including controls for the various versions of the survey (see previous footnote). Including these controls did not change the overall pattern of results, though the effects were less precisely estimated. The MORESSB variable remained negative and marginally significant, however. Because the previous GSS studies did not include year effects, I exclude them here for comparison.

other sources of income constitute a larger share of family income for non same-sex behaving men.

5. Conclusion

This paper has revisited the earnings penalty for same-sex behaving men in the United States, showing that the large earnings penalty experienced by behaviorally gay men (as reported in four recently published studies) is confirmed in an independent source of data. Specifically, I estimate that same-sex behaving men in the 1988-1994 Third National Health and Nutrition Examination Survey reported 23-30 percent lower incomes than similarly situated observationally heterosexual men. Notably, the size of the estimated reduction in income associated with same-sex sexual behavior falls within the range of previous estimates using similar techniques and GSS data. That the behaviorally gay earnings penalty is robust to choice of data suggests the need for much further research into its causes and consequences, particularly as they relate to the possible role of discrimination on the basis of sexual orientation in labor markets. Moreover, more research is needed as to why behavior based definitions of sexuality yield qualitatively different patterns than classifications based on other methods.

A handful of limitations – largely due to the nature of the NHANES III data – should be noted. First, because the NHANES III data are primarily used for health related purposes, the included labor market information is limited. For example, I have no information on key job and workplace characteristics such as work effort, occupation, industry, and job tenure. Previous GSS studies have been able to control for broad occupational categories, though the behaviorally gay earnings penalty was generally

robust to such controls. Also, the NHANES III data do not permit controls for space and time, as the survey year and state of residence are not available in the public use data. Indeed, it is possible that changing social attitudes toward gay men and lesbians over this time period may be important in determining the behaviorally gay earnings penalty.

Despite the data limitations, however, these results provide a much needed independent verification of the published GSS results that same-sex behaving men experienced significantly lower incomes than other men in the 1988-1994 period. Indeed, building a body of evidence of the economic effects of sexual orientation with different data sources, sexual orientation concepts, and patterns of results is an important first step in understanding how and why sexual orientation matters for economic well being.

References

- Allegretto, Sylvia and Michelle Arthur (2001). "An Empirical Analysis of Homosexual/Heterosexual Male Earnings Differentials: Unmarried and Unequal?," *Industrial and Labor Relations Review*, 54(3): 631-646.
- Badgett, M. V. (2001). Money, Myths, and Change: The Economic Lives of Lesbians and Gay Men, Chicago: University of Chicago Press.
- . (1995). "The Wage Effects of Sexual-Orientation Discrimination," *Industrial and Labor Relations Review*, 48(4): 726-739.
- Berg, Nathan and Donald Lien (2002). "Measuring the Effect of Sexual Orientation on Income: Evidence of Discrimination?," *Contemporary Economic Policy*, 20(4): 394-414.
- 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(2): 139-154.
- Black, Dan, Hoda Makar, Seth Sanders, and Lowell Taylor (2003). "The Earnings Effects of Sexual Orientation," *Industrial and Labor Relations Review*, 56(3): 449-469.
- Blandford, John (2003). "The Nexus of Sexual Orientation and Gender in the Determination of Earnings," *Industrial and Labor Relations Review*, 56(4): 622-642.
- Carpenter, Christopher (2005). "Self-Reported Sexual Orientation and Earnings: Evidence from California" *Industrial and Labor Relations Review*, 58(2): 258-273.
- (2003). "New Evidence on Gay and Lesbian Household Incomes," *Contemporary Economic Policy*, 22: 78-94.
- Laumann, Edward, John Gagnon, Robert Michael, and Stuart Michaels (1994). The Social Organization of Sexuality: Sexual practices in the United States, Chicago, University of Chicago Press.
- Plug, Eric and Peter Berkhout (2004). "Effects of Sexual Preference on Earnings in the Netherlands," *Journal of Population Economics*, 17(1): 117-131.

Table 1. Studies using the GSS to document earnings penalties for same-sex behaving males.

Study & Years of GSS	Measure	Findings
Badgett (1995) ILRR, 1989-1991	At least as many same-sex as different-sex sex partners since age 18	Significant 31 percent earnings penalty
Berg and Lien (2002) CEP, 1991-1996	Any same-sex sexual behavior in past 5 years	Significant 22 percent earnings penalty
Black et. al. (2003) ILRR, 1988-1996	Any same-sex sexual behavior since age 18	Significant 21 percent earnings penalty
Blandford (2003) ILRR, 1988-1996	Exclusively same-sex sexual behavior in past 12 months (or past 5 years if no behavior in past 12 months), combined with marital status indicators	Significant 31 percent earnings penalty

Table 2. Sample means & standard deviations of key variables.

Variable	ANYSSB=0 NHANES (respondent reports no lifetime same-sex sexual behavior)	ANYSSB=1 NHANES (respondent reports lifetime same-sex sexual behavior)	ANYSSB=0 GSS (respondent reports no lifetime same-sex sexual behavior)	ANYSSB=1 GSS (respondent reports lifetime same-sex sexual behavior)
Income	22782 (16884)	23949 (17240)	34821 (21983)	33810 (21847)
Age	34.6 (11)	36.2 (10.8)	37.5 (10)	39.6 (9.3)
Less than HS	.27 (.44)	.14 (.35)	.10 (.30)	.04 (.19)
HS Degree	.29 (.45)	.27 (.45)	.25 (.43)	.08 (.27)
Some College	.25 (.44)	.27 (.45)	.31 (.46)	.32 (.47)
College Degree	.18 (.39)	.32 (.47)	.34 (.47)	.56 (.50)
White	.53 (.50)	.70 (.46)	.86 (.35)	.96 (.19)
Urban/SMSA	.54 (.50)	.82 (.39)	.20 (.40)	.48 (.50)
Married	.06 (.24)	0	.02 (.15)	.04 (.19)
N	710	44	652	52

Data: 1988-1994 NHANES III and 1988-1996 GSS, author calculations. NHANES III sample includes men reporting family size equal to one. GSS sample includes men who report exactly one person in their household. Income in NHANES is family income; in GSS it is labor income, both in 2000 dollars. Both samples are restricted to men age 18-59. The NHANES asked about urban residence, while the GSS asked about residence in an SMSA.

Table 3. Coefficient estimates on gay indicators in equation (1), NHANES.

Variable	ANYSSB (any lifetime same sex sexual behavior)	ATLEAST (at least as many lifetime same-sex as different-sex partners)	MORESSB (more lifetime same- sex than different-sex partners)
	(1)	(2)	(3)
A. Sample: family size = 1			
Behaviorally Gay Indicator	-.235* (.129)	-.297** (.118)	-.375*** (.130)
N same-sex behaving	44	32	27
N total	731	731	731
B. Sample: family size = 1 & not currently married			
Behaviorally Gay Indicator	-.232* (.129)	-.295** (.118)	-.371*** (.131)
N same-sex behaving	44	32	27
N total	689	689	689
Excluding anyone reporting income from: Social Security, pension, AFDC, public assistance, or veteran's benefits.			
C. Sample: family size = 1			
Behaviorally Gay Indicator	-.184 (.139)	-.220* (.126)	-.306** (.144)
N same-sex behaving	37	26	21
N total	626	626	626
D. Sample: family size = 1			
Behaviorally Gay Indicator	-.183 (.139)	-.220* (.127)	-.305** (.145)
N same-sex behaving	37	26	21
N total	591	591	591

Coefficient estimates and standard errors from interval regression of (1) using NHANES III data. *** indicates significance at 1%, ** at 5%, and * at 10%.

Table A1. Full set of coefficient estimates from Table 3, Column 1, Panel A.

Variable	ANYSSB (any lifetime same sex sexual behavior) (1)
Sample: family size = 1	
Behaviorally Gay Indicator	-.235* (.129)
Age	.163*** (.028)
Age squared	-.002*** (.0004)
Less than high school	-.322** (.135)
Some college	.193** (.092)
College degree	.306*** (.109)
Black race	-.556*** (.073)
Other race	-.585*** (.181)
Hispanic	-.260* (.149)
Urban	.218*** (.079)
Married	.320*** (.128)
N same-sex behaving	44
N total	731

Coefficient estimates and standard errors from interval regression of (1) using NHANES III data. *** indicates significance at 1%, ** at 5%, and * at 10%.