

Understanding the Gay Gap: An Empirical Analysis of Sexual Orientation, Disclosure and Earnings

Erik Plug and Peter Berkhout*

January 2009

Abstract

Gay workers tend to earn less than other men. Does this occur because of discrimination or because of selection? In this paper we address this question and collect new information on workplace disclosure to separate out discrimination effects from selection effects. Using a large sample of recently graduated men in the Netherlands, we find that gay workers earn about 3 to 4 percent less than other men. Our disclosure estimates, however, provide little evidence that the labor market discriminates against gay workers. They rather support the selection story, most prominently observed among undisclosed gay workers who concentrate in lower paid occupations, and earn about 5 to 9 percent less than other men.

*Erik Plug: Amsterdam School of Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands, Tinbergen Institute and IZA. Peter Berkhout: Economisch Instituut voor de Bouwnijverheid, Basisweg 10, 1043 AP Amsterdam, The Netherlands. This research is supported by the Dutch National Science Foundation, VIDI grant no. 452.03.309. For helpful comments, the authors thank Monique de Haan, Virginia Maestri, Jacopo Mazza, Hessel Oosterbeek, Reyn van Ewijk, Mathieu Valdenaire, and conference participants. Copies of the computer programs used to generate the results presented in this paper are available upon request from the authors. E-mail: e.j.s.plug@uva.nl; pberkhout@eib.nl.

1 Introduction

Although it is only recent that economists began to focus on the relationship between sexual orientation and earnings, one observation appears to be quite persistent across a small but growing number of empirical studies: gay men earn less than other men. Over its cause, however, economists do not agree. Some argue that it is a discrimination effect: employers treat observably similar workers with different orientation differently. Others argue that it is a selection effect: workers with different orientation are unobservably different in their productive skills or tastes.¹

The disagreement illustrates the difficulty to empirically separate discrimination effects from selection effects. There are -we think- two information problems that complicate the interpretation of sexual orientation estimates. The first problem is that, when it comes to the workers' sexual orientation, researchers are often better informed than employers. If the workers' sexual orientation is known to us researchers but unknown to some employers, part of the estimated sexual orientation effect cannot be attributed to treatment differentials. The second problem is that, when it comes to the workers' productive skills, employers are usually better informed than researchers. If gay and heterosexual workers look similar to researchers but different to employers, the estimated sexual orientation effect may merely reflect productivity differences that are observable to employers but unobservable to researchers.

In this paper we try to tackle both problems (mainly) by means of collecting new information on the workers' disclosure status. Disclosure information helps us in two ways. First, we distinguish disclosed from undisclosed workers and circumvent the problem that sexual orientation is not always an observable characteristic to employers. And second, we combine disclosure information with a simple discrimination model, and take a first step to uncover the extent of discrimination in sexual orientation estimates. The intuition is as follows. With undisclosed workers, we estimate the relation between sexual orientation and earnings that is arguably driven by productivity differences, and not by employers who discriminate. With disclosed workers, we estimate a similar relationship and attribute the difference in earnings to a combination of employer discrimination and differences in productive traits. If disclosure is exogenous, implying that the workers' productive characteristics do not correlate with disclosure status, our reduced-form estimates enable us to make a distinction between discrimination and selection.

Based on a survey that collects information on Dutch college graduates, including information on the student's sexual orientation and workplace disclosure, our empirical analysis leads us to two main findings. First, we estimate

¹Among the articles that are supportive of the discrimination hypothesis are Badgett (1995); Klawitter and Flatt (1998); Arabsheibani et al. (2005). Articles that raise doubts about the discrimination hypothesis include Berg and Lien (2002); Blandford (2003); Black et al. (2003); Plug and Berkhout (2004).

the association between sexual orientation and male earnings based on a new sample of Dutch college graduates, and find (consistent with previous) evidence that gay men earn less than other men, even in the beginning of a working career. Second, we present new estimates on the effect of workplace disclosure on earnings and find that the difference in earnings between gay and heterosexual men is entirely driven by those gays that work for employers who are unaware about their sexual orientation. If gay workers disclose their identity exogenously, our findings indicate that gay workers experience positive, rather than negative, discrimination.

One important threat is that the disclosure decision is not always exogenous. If gay workers choose to be open about their sexual orientation, our method does not work and produces biased discrimination (and selection) estimates to the extent that disclosed and undisclosed workers are different in their productive traits. We evaluate this possibility and test whether alternative discrimination theories in which gay workers disclose their identity endogenously can serve as credible explanations for our findings. Our results suggest that the gay penalties and disclosure premiums observed in our data are unlikely due to labor market discrimination against gay workers. We are thus led to the conclusion that selection is (almost surely) responsible for the gay gap in earnings.²

One of the advantages of this study is that our discrimination tests do not rely on gender. Previous work challenged the predictions of a discrimination model by showing that lesbian women sometimes earned somewhat more than heterosexual (Berg and Lien 2002; Blandford 2003; Black et al. 2003; Plug and Berkhout 2004). These results require a careful interpretation. As Black et al. (2003) argue, the earnings advantages of lesbian workers over other women are not that informative about the nature of discrimination against homosexual male and female workers if heterosexual women are discriminated against as well. With disclosure information, we test for discrimination using only men and thus avoid potential discriminatory practices against heterosexual and lesbian women that may affect their differences in pay.

The paper continues as follows. Section 2 describes the data on Dutch university graduates, and provides a discussion on the workplace disclosure measures we use in the analysis. Section 3 models the relation between sexual orientation and earnings, focuses on the problem that sexual orientation is not generally an observable characteristic, shows under which conditions we can make a distinc-

²After we sent out our survey –issued in 2004– to collect disclosure information to answer questions about discrimination and sexual orientation, it came to our attention that Comolli (2005) also examined the relationship between gay earnings and disclosure. With data on gay single workers in four US cities, he finds that disclosed workers earn less (and not more) than undisclosed workers. Our study differs from the work of Comolli in at least two directions. First, his research focuses exclusively on gay workers and has little to say about differences in earnings between gay and heterosexual male workers. Second, he uses a switching regression model to allow for endogenous disclosure and relies on instrumental variables/exclusion restrictions that are either statistically weak (age at which respondents first wondered they were gay) or unconvincing (the extent to which respondents feel part of the gay community).

tion between discrimination and selection effects, and presents the parameter estimates. Section 4 attempts to reconcile our findings with various discrimination scenarios that allow for endogenous disclosure. Section 5 concludes.

2 Data, Sample and Measurement

The data used in this paper come from an annual survey of individuals who completed college education in the Netherlands.³ The survey is a questionnaire which contains questions on education, work history since graduation and personal characteristics. The survey focuses on recent graduates. That is, individuals who are interviewed graduated in the academic year two years earlier. Yearly around 30,000 questionnaires are sent out and the response rate is about 30 percent. Now the data contain 10 cohorts of graduates in the beginning of their working career, interviewed between 1997 and 2006.

In this paper we focus our attention on the latter three cohorts that include information on sexual orientation as well as the degree of openness at the workplace.⁴ Of the initial 11,767 male graduates in the 2003/2004, 2004/2005 and 2005/2006 survey, we restrict ourselves to full-time working men (32 and more hours per week) for whom we have complete information on earnings, which reduces our sample to 10,132 observations. We further excluded all workers that were self-employed, worked for temporary employment agencies, earned less than one euro per hour, and for whom data on the various control variables were unavailable. In the end, we are left with a sample of 7,158 male workers, of which 435 are gay (324) or bisexual (111). Descriptive statistics for the most important variables are provided in Table 1.

Measurement of Sexual Orientation and Workplace Disclosure

In the empirical literature on the economics on sexual orientation three sexual orientation measures are currently in use: measures based on (past) sexual experiences; measures based on partnership; and self-reported orientation measures. These measures do not fully overlap and may capture different aspects of sexual orientation. A series of recent papers in economics (Badgett 1995; Black et al. 2003; Plug and Berkhout 2004; Comolli 2005; Carpenter 2005 2007) relate these different aspects to the extent of workplace disclosure and argue that some measures of sexual orientation are more informative to employers (and

³Dutch college education can be divided into two tracks: higher vocational education and university education. Higher vocational education prepares students for specific (categories of) professions, is taught at about 60 special institutes, and its graduates obtain a bachelor degree. University education has a more general, academic character, is provided by 13 universities, and its graduates can obtain bachelor but also master degrees.

⁴Only in our 2003/2004, 2004/2005 and 2005/2006 survey data we collect information on the extent of workplace disclosure.

fellow workers) than others. Within a labor market setting, for example, sexual orientation measures based on partnership are more relevant than measures based on past sexual behavior, simply because it is much easier for employers to obtain information on the gender of the employee's partner, than on whom an employee spent his time in bed with. At the same time, sexual orientation estimates based on partnership ignore all those workers who are gay and single. If finding a partner is somehow related to the worker's productivity, it is possible that the same estimates miss the true impact of sexual orientation on earnings. Variables that measure sexual orientation and the extent of workplace disclosure of all workers, including single and coupled workers, would resolve part of the discussion, but are rarely collected in large systematic data sets.

Not in our data. Regarding sexual orientation, the former college graduates are asked whether they are sexually attracted to men, women or both. Together with the graduates' gender we identify sexual orientation of all single and coupled workers. Regarding workplace disclosure, we ask all identified gay graduates that work at the time of data collection whether their employers/fellow-workers know their sexual orientation, and whether they themselves are open about their sexuality at work. This enables us to construct two disclosure indicators. One measures disclosure directly: employers/colleagues just know the sexual orientation of their employees/fellow-workers. The other one measures disclosure indirectly: employers/colleagues might know the sexual orientation from the open gay lifestyle employees/fellow-workers lead (at work). In our data, we find that the nondisclosure shares among gay workers at the beginning of their working are substantial. Depending on the definition of workplace disclosure, about 40 to 60 percent of all the gay workers in our sample work for employers who are ignorant about their sexual identity.

Note that sexual orientation question as well as the openness questions are part of a special section at the end of the questionnaire that concentrated on general individual and household characteristics. The fact that we infer information on sexual identity and disclosure at the end of the survey, after all other information is gathered, is an additional strength of our data. By doing so, we circumvent potential selectivity in response behavior, when respondents belonging to a sexual minority group would have taken the opportunity to emphasize or even exaggerate problems encountered in relation to their sexual orientation.

3 Empirical Testing for Discrimination

In this Section we explore whether we can empirically test for discrimination against gay workers, if sexual orientation is not always observable to employers and we have information on the workers' disclosure status. In our empirical model we concentrate on earnings discrimination and estimate an earnings equation of the form

$$Y_i = \gamma(D_i)H_i + \beta X_i + \epsilon_i, \tag{1}$$

where i indexes the worker, Y is a measure of earnings, H is a measure of sexual orientation that equals 1 for gay workers and 0 otherwise, D is a disclosure index that equals 1 for workers whose sexual orientation is known to employers and 0 otherwise, X is a set of other (observed and unobserved) variables assumed to affect earnings, and ϵ is the remaining error assumed to be uncorrelated with Y (and its determinants H and X). We let parameter $\gamma(D)$ vary with disclosure status

$$\gamma(D_i) = \gamma + \delta D_i, \quad (2)$$

to accommodate that prejudiced employers can only discriminate against gay workers with disclosed identity. That is, the effect of sexual orientation on earnings reflects unobserved productivity differences between gay and heterosexual workers (γ) and possibly discrimination against sexual orientation ($\delta < 0$). The problem of identifying δ empirically, however, is that disclosure is a choice and likely to be correlated with unobserved productivity traits.

To formalize this argument, we let βX_1 be a measure of average productivity for gay workers whose orientation is observable to employers; βX_2 be a measure of the average productivity for gay workers whose orientation is not observable to employers; and βX_0 be a measure of the average productivity for heterosexual workers. If we treat all relevant productivity traits as unobserved, estimate a stripped version of equation (1) and regress Y only on H and HD , we obtain estimates of γ and δ with following properties

$$\text{plim } \hat{\gamma}_{LS} = \beta(X_2 - X_0); \quad \text{plim } \hat{\delta}_{LS} = \delta + \beta(X_2 - X_1). \quad (3)$$

It is easy to see that identification of discrimination against gay workers requires that $\beta X_1 = \beta X_2$. If disclosure is exogenous, we obtain two interesting estimates: one discrimination estimate $\hat{\delta}_{LS}$ that measures the degree of discrimination against gay/bisexual workers, and another selection estimate $\hat{\gamma}_{LS}$ that measures unobserved productivity differences between gay and heterosexual men.

Results

Table 2 presents least squares estimates of the relationship between the (log of) hourly earnings and various measures of sexual orientation on samples of working heterosexual and homosexual men, using varying sets of control variables.

Because much of the empirical literature on the economics on sexual orientation has concentrated on estimating the effect of sexual orientation on earnings, we begin to discuss the estimated effects of sexual orientation on the earnings of all men. In panel A column (i) we regress earnings on a sexual orientation dummy as the single right-hand-side variable and find that gay workers receive 4 percent less in hourly pay than heterosexual workers. In columns (ii) and (iii) we add several other covariates. We find that the inclusion of age, ethnicity, region, school characteristics, and job characteristics (including 15 occupation

indicators) has little effect. The earnings penalty for being gay continues to be statistically significant, negative but small varying between 3 and 4 percent. Our findings are, as such, comparable to those obtained in previous studies: gay men earn (somewhat) less than other men (Badgett 1995; Berg and Lien 2002; Blandford 2003; Black et al. 2003; Plug and Berkhout 2004; Comolli 2005; Carpenter 2005 2007).

To better understand whether this sexual orientation effect is a discrimination effect (or a selection effect for that matter), we must recognize that any discrimination interpretation of the observed earnings penalty depends on the employers' ability to distinguish their employees' sexual orientation. Equations (1) and (2) are designed to address this issue. If we treat the decision to disclose as exogenous, the negative association between sexual orientation and wages, as reported above, has important implications for disclosure and sexual orientation effects. If, for example, the negative correlation is entirely driven by discriminating employers, we expect the disclosure effect to be negative and the sexual orientation effect to be zero ($\hat{\delta}_{LS} < 0$; $\hat{\gamma}_{LS} = 0$). If the negative association just picks up selection effects reflecting a negative relationship between gay workers and their unobserved earnings skills, we expect the disclosure effect to be zero and sexual orientation effect to be negative ($\hat{\delta}_{LS} = 0$; $\hat{\gamma}_{LS} < 0$). And if the penalty captures both discrimination and selection effects, the earnings effects of disclosure and sexual orientation must be negative ($\hat{\delta}_{LS} < 0$; $\hat{\gamma}_{LS} < 0$).

In panel B we estimate the model we previously mentioned in (1) and (2) and regress earnings on a sexual orientation and disclosure dummy using our direct measure of workplace (non)disclosure. Rather surprisingly, we find that none of the implications suggested above matches with our sexual orientation and disclosure estimates. Instead of a negative or zero disclosure coefficient, we find a positive $\hat{\delta}_{LS}$ in every specification. The sexual orientation coefficient $\hat{\gamma}_{LS}$ is always negative. In column (i), for example, our estimates show that gay workers earn on average 9 percent less than other men, and that among gay workers those with disclosed identity earn on average 8 percent more. With several other covariates added results fall but not by much. It is interesting to note that the estimated coefficients $\hat{\gamma}_{LS}$ and $\hat{\delta}_{LS}$ are similar in size but opposite in sign: together they are always statistically insignificant and very close to 0. This suggests that disclosed workers who are gay earn as much as heterosexual workers, and that the earnings penalty for being gay, as reported in Panel A, is primarily driven by those gay workers who did not disclose their sexual orientation to their employers (or fellow employees for that matter). In panel C we estimate the same equation but switch to indirect disclosure measures. With disclosure indicators based on the degree of openness at work, our findings reported in the third panel do not substantially change. The estimates again show that, among all men, only gay workers that are not open and perhaps purposely conceal their sexual identity earn significantly less. Independent of the specification used, the penalty for not being open $\hat{\gamma}_{LS}$ moves around the 5

percent.⁵

If disclosure is exogenous, it is clear that our discrimination model cannot rationalize the negative association between sexual orientation and wages. In particular, our findings indicate that gay workers experience positive discrimination with a premium that varies between 3 to 8 percent. These unconventional numbers, however, should not be taken too literally. They do not make sense. They rather call into question whether the decision to disclose is truly exogenous, an issue to which we shall turn next to.

4 Understanding Gay Penalties and Disclosure Premiums

Three main observations emerge from the previous section. Gay workers tend to earn somewhat less than other men. The gay penalty we observe, however, varies with disclosure status. Gay workers with undisclosed identities earn significantly less than heterosexual workers, whereas gay workers with disclosed identities earn as much as heterosexual workers. In a naive discrimination model where gay workers disclose their identity exogenously, these estimates would mean that gay workers experience positive, rather than negative, discrimination in the firms they work for. Since we do not believe that positive discrimination serves as a credible explanation for our findings, we must (regrettably) conclude that the identification method we propose fails in separating discrimination effects from selection effects, and that our discrimination estimates as discussed above remain biased because of endogenous disclosure. In particular, the bias we observe suggests that among gay workers disclosed workers are abler workers, and therefore earn higher wages, even in a labor market that discriminates.

In this Section we will argue that our biased discrimination estimates can still be informative about (the absence of) labor market discrimination. We proceed as follows. We first explore all possible scenarios in which gay penalties and disclosure premiums as observed in our data are the result of a discriminating labor market and endogenous disclosure. Three such scenarios exist: the ones we consider require either productivity advantages of disclosed gay workers, classification error in disclosure status, or undisclosed costs. We then look for some associative evidence and test whether the three hypothesized scenarios will likely hold or not.⁶

⁵In the empirical literature on sexual orientation and earnings it is common to pool gay and bisexual workers. In this paper we follow this procedure. If we would re-estimate our specifications on samples of gay and heterosexual workers, we get estimates that are very similar in sign, magnitude and significance level. According to the specification, as estimated in Table 2 column (iii), the point estimates, in case we exclude all bisexual workers from our sample, are found to be -0.030 [0.010]; -0.071 [0.20] and 0.054 [0.023]; -0.043 [0.014] and 0.26 [0.20] in Panels A, B and C, respectively (with standard errors between brackets).

⁶An alternative procedure to identify discrimination effects without assuming exogenous disclosure would be an instrumental variable approach. However, with the data at hand it is

Discrimination Scenario I: Productivity Advantages

Our first scenario which considers discrimination against disclosed gay workers requires productivity advantages experienced by the same disclosed gay workers. The argument is simple. Our observation that disclosed gay workers earn as much as heterosexual workers can only signal labor market discrimination against gay workers if openly gay workers enjoy substantial productivity advantages over heterosexual workers.

The key question is whether disclosed gays workers indeed have more marketable skills than heterosexual workers. We have our doubts. In fact, there is some non-experimental evidence to suggest the opposite is true; heterosexual workers are at least as productive as disclosed gay workers. Our answer relies on two economic theories that are (possibly) informative about labor market outcomes, productivity traits and sexual orientation.

The first theory we have in mind is a family specialization model of Becker (1981). According to Becker, gay couples are less likely to have children, gain less from specialization in market production, and as a consequence accumulate fewer, rather than more, marketable skills. One of the testable specialization predictions is that gay couples have a lesser extensive division of work than men and women in heterosexual couples. In Table 3 we explore the possible role of labor supply differences between gay and heterosexual workers and regress various division of work outcomes on our set of sexual orientation measures and the usual controls, including age, ethnicity, and region. In column (i) our sexual orientation estimate indicates that gay workers work somewhat less than heterosexual male workers. As Becker argues, specialization should primarily be observed in couples with children. In columns (ii) and (iii) we therefore restrict our sample to workers with partners and find that, compared to heterosexual men, gay workers work somewhat less, while their partners work considerably more. Since almost all men in our sample are young and childless, we have also asked for hypothetical labor supply responses in case our respondents would live in families with young children.⁷ In column (iv) we observe that (hypothetical) children lead to a much stronger reduction in the weekly working hours for gay workers than for heterosexual male workers. In the bottom panels of Table 3 we estimate previous labor supply models, add a disclosure dummy, and distinguish disclosed from undisclosed gay workers. Compared to heterosexual male workers, our estimates also show that disclosed gay workers (as singles and couples) work fewer hours, have partners that work more hours, and they report to work much less in case they would live in families with children. Our data thus (weakly) favor specialization theory.

The second theory we consider is one of sorting. It is again Becker's (1957)

impossible for us to come up with an instrument that is credibly valid.

⁷We assume that gay men are able (and as able as heterosexual men) to visualize their labor supply response to variations in family composition that are hypothetical but not unrealistic. In the Netherlands, for example, it is legal for gay couples to adopt children.

insight on labor market discrimination that helps us to interpret our earnings estimates in Table 2 as productivity differences between heterosexual and disclosed gay workers. If the extent to which firms discriminate varies across firms, and rational gay workers will only disclose their identity if they expect little or no discrimination, sorting predicts that disclosed workers work in non-discriminating firms and undisclosed workers work in discriminating firms. With sorting our estimates cannot detect the extent of earnings discrimination either. This is by assumption. There are still firms with discriminating intentions, but these are not the firms where the disclosed gay workers in our sample end up working. This means that we may set the discrimination coefficient δ in equation (2) to zero, and that because of that our estimates of $\hat{\delta}_{LS}$ and $\hat{\gamma}_{LS}$ together measure (unobserved) productivity differences between disclosed gay and heterosexual men. If we assume sorting to hold, these estimates do not show any productivity advantages experienced by gay workers who are out.

In sum, we are not convinced that disclosed gay workers in our sample experience substantial productivity advantages over heterosexual male workers, for two reasons. First, we observe that the disclosed workers in our sample show a mild tendency to behave as family specialists, suggesting that they have probably less marketable skills than heterosexual workers. Second, we also consider an alternative possibility that the disclosed workers in our sample have successfully sorted themselves into non-discriminating firms. If they do, which seems not unreasonable with only 3 to 4 percent of all full-time workers being openly gay, disclosed gay and heterosexual workers turn out to be equally productive. If Becker somehow got it right, whether it is on specialization or on sorting, we must conclude that the labor market does not discriminate against sexual orientation, at least not against disclosed gay workers.

Discrimination Scenario II: Classification Error

Our second scenario in which undisclosed gay workers are discriminated against relies on classification error in disclosure status. If some gay workers who work for prejudiced employers misclassify themselves as undisclosed, it is possible that these undisclosed workers are in fact treated differently by employers who know their gay identity. Under this scenario the negative effect we estimate for being undisclosed and gay ($\hat{\gamma}$) reflects in part a discrimination effect. We consider two sources of such misclassification. First, there is classification error among gay workers who are out: they did not actively inform their employer and, as such, report that their employer does not know. And second, there is classification error among gay workers who are not out (at work): they are unaware that their prejudiced employer might know their identity and wrongfully report that their employer is unaware.

The first type of classification error is -we think- not that serious. A close inspection of Table 2 leads to this conclusion. In there we present results with two distinctive disclosure measures: one measures disclosure directly by asking

gay workers whether their employer explicitly knows their gay identity, and the other measures disclosure indirectly by asking gay workers whether they are openly gay at work. Since the latter measure lacks an explicit reference to the employers' knowledge about the workers' sexual identity, it is arguably more sensitive to the classification error we hypothesize among gay workers who are out. If we would regress sexual orientation and disclosure status on earnings and switch from a direct to an indirect disclosure measure, we likely amplify misclassification bias. That is, we expect the estimated penalty for being undisclosed and gay to increase because it captures the discriminatory effect of the increased share of gay workers whose orientation is known to employers. This is clearly not what we observe.⁸

The second type of classification error is perhaps a bigger concern. There is some indirect evidence to suggest that employers likely know their workers' sexual identity without being notified. Psychologists report that sexual orientation is quite accurately perceived, even at very brief encounters (Ambady, Hallahan and Conner 1999; Rule and Ambady 2008). Psychology students make judgments based on photographs. Out of every 10 pictures of gay men, they show that about 6 to 7 are correctly classified in as little as 50 milliseconds. If their experimental estimates apply to the undisclosed gay workers in our sample, it is indeed possible that the negative effect we estimate for being undisclosed and gay captures a discrimination effect because of classification error. Their numbers, however, are obtained using gays (and lesbians) who are out. If the employers' perceptions are somehow related to the disclosure status, the same estimates will clearly exaggerate the importance of classification error among undisclosed gay workers. Recall that of all gay workers in our sample about 60 percent claim that their employer knows their identity, which is very close to the percentages found by these psychologists. In fact, if there exist a one-to-one relationship between perception and disclosure, our sexual orientation and disclosure estimates do not suffer from missclassification bias at all.

To check whether the earnings penalty experienced by undisclosed gay workers is driven by those gay workers who purposely conceal their gay identity yet work for employers who accurately perceive it, we would need employer information. Because our data does not gather information on the workers' identity through employers, it is impossible to test how serious this misclassification bias actually is.

Discrimination Scenario III: Undisclosure Costs

The third and last scenario we consider is that undisclosed workers are discriminated against if they work for prejudiced employers, fear exposure and because

⁸To find out whether classification bias matters, we interpret differences in earnings of two differently defined groups of (un)disclosed workers in terms of differences in the degree of discriminatory effects and ignore the complication that the same differences in earnings might be tainted by their unobserved differences in productivity.

of that spend valuable time and effort concealing their sexual identity. In this scenario it is the fear of discrimination that is costly and depresses productivity and subsequently earnings.

A similar point about undisclosed workers earning less is put forward by Escoffier (1975) who argues that passing as a heterosexual may require some effort that reduces productivity. Undisclosed workers, he continues, may avoid potentially awkward social interactions that hinder their advancement with other workers. In our survey we can test whether undisclosed workers tend to shy away from social activities at work. One of the questions we have asked workers is to evaluate how much they enjoy social events at work (including diner parties, day trips and informal receptions). Respondents could answer on a 5 points scale, where 1 represents very unattractive and 5 represents very attractive. As our discrimination test we regress this answer on the usual set of sexual orientation measures and controls. The estimates which are reported in Table 3 column (v) reveal that undisclosed workers assess social events favorably. Gay workers who (try to) hide their identity are more inclined to attend social events than gay workers who are out, suggesting that (fear of) discrimination cannot account for the gay penalties and disclosure premiums.

This is as far as the data allow us to go, but we think it is far enough to be confident enough to conclude that the gay gap in earnings as observed in our data is not the consequence of labor market discrimination, at least not against gay college graduates who are young, who work and live in one of the more tolerant countries towards sexual minorities.

5 Concluding Remarks

In 1995 Badgett mentioned that a variable measuring the extent of workplace disclosure of gay identity would be more appropriate to include in the wage equation, since disclosure is necessary for workplace discrimination to occur. One decade later, we follow up on her suggestion, collect new information on workplace disclosure, and develop a simple (and possibly valuable) strategy to assess the role of discrimination in sexual orientation estimates. Two messages emerge from our analysis.

The first message is a rather pessimistic one: information on workplace disclosure alone is not sufficient to distinguish discrimination effects from selection effects. Identification strongly relies on disclosure being exogenous. If the decision to disclose is exogenous, implying that the productive characteristics do not correlate with disclosure status, we show that information on disclosure is sufficient to identify discrimination effects. Our data, perhaps not surprisingly, reject exogenous disclosure. We find that disclosed and undisclosed workers are different, and that because of these differences the method we propose fails in separating discrimination effects from selection effects: our discrimination estimates as discussed are biased, and the bias we observe suggests that among gay

workers disclosed workers are abler workers, and therefore earn higher wages, even in a labor market that discriminates.

On a brighter note, the second message is that without the assumption of exogenous disclosure we can still detect (the absence of) labor market discrimination against gay men. We argue that our findings can only be the result of a discriminating labor market with endogenous disclosure if (a) disclosed gay workers experience productivity advantages over other male workers; (b) there is classification error in the disclosure measures we use; or (c) undisclosed itself requires time and effort that hinders productivity. Almost all factors appear to be irrelevant. We therefore have good reasons to believe that our findings are difficult to reconcile with the idea of labor market discrimination. Only in case of serious classification error among undisclosed workers, we are unable to dismiss discrimination as one of the underlying mechanisms. Without employer information on the workers' identity, however, we have little indication of how serious this classification error is. This means that we are confident enough to conclude that the gay gap in earnings as observed in our data is not the consequence of labor market discrimination, assuming that only a small fraction of the variability in disclosure is due to measurement error.

Of course, an important question is what is the external validity of our analysis. As we have discussed throughout, the results we find are specific to a group of young working college graduates in the Netherlands. Our observation of absent discrimination against gay workers may, for example, not be relevant for older generations in the labor market. If gay workers experience losses in earnings because they more frequently end up in dead-end jobs or face glass ceilings, it is likely that discrimination estimates based on starters miss these effects. In addition, our results are specific to the Netherlands and may not hold in other societies either. It is not clear, though, that this is a limitation of our study. Since the Netherlands is one of the more tolerant countries towards sexual minorities, we actually believe that studying earnings effects in this particular country adds a potential value to this study.

Appendix: Some Data Description

In this Appendix we describe how non-standard variables used in our analysis are defined. These variables are sexual orientation, disclosure, labor supply responses to hypothetical variation in family composition and preferences for social activities at work. Other variables used in our analysis are defined in a more standard fashion and will not be described.

Information on sexual orientations is derived from the question: "Which gender do you prefer sexually?" Respondents could choose between three alternatives: only men; only women; or both men and women. If respondents are male and report they are attracted to either only men or both men and women, they are assigned one and classified as gay. This is the "gay" variable in our

analysis.

Information on disclosure applies to gay workers in their current job. We ask this selected group of respondents two questions: “Is your sexual preference known at the workplace?” and “Do you actively disseminate your sexual preference at the workplace?” Respondents could either answer yes or no. The variable we call “known to employer/colleagues” equals one for those respondents who answer yes the first question. And similarly, the variable we call “open lifestyle” equals one for those respondents who answer yes the second question.

Information on labor supply pertains to all working respondents (including gay respondents) in our sample. All labor supply measures are defined in a standard fashion and transformed to the number of hours worked per week in the worker’s current job. One of the more unusual question we ask our respondents is: “Suppose you would have a partner and young children. What would be the ideal amount of working days per week for you?” Respondents could indicate their preferred number of days, ranging from zero to seven. We transform these answers to weekly hours through multiplication by eight. This is the variable we label “hours worked in hypothetical of family with children”.

And finally, information on preference for social events is gathered from working respondents. We ask respondents to evaluate various characteristics that are typical to their current occupation. One of these characteristics involve social events. The question we ask goes as follows: “To what extent do the following items concerning your job appeal to you?” We restrict ourselves to the item “social events with colleagues, such as diner parties, day trips, informal receptions et cetera”. Respondent could answer on a five points scale, indicating very unattractive, unattractive, neutral, attractive and very attractive. The answer respondents give ranges from one to five, where one represents very unattractive, is the variable we define as “preference for social activities at work”.

References

- Ambady, Nalini, Mark Hallahan, and Brett Conner. 1999. “Accuracy of Judgments of Sexual Orientation From Thin Slices of Behavior.” *Journal of Personality and Social Psychology*, Vol. 77, No. 3 (September) pp. 538-547.
- 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*, Vol. 54, No. 3 (April), pp. 631-46.
- Arabsheibani, G. Reza, Alan Marin, and Jonathan Wadsworth. 2005. “Gay Pay in the UK.” *Economica*, Vol. 72, No. 286 (May), pp. 333-47.

- Badgett, M.V. Lee. 1995. "The Wage Effects of Sexual Orientation Discrimination." *Industrial and Labor Relations Review*, Vol. 48, No. 4 (July), pp. 726-39.
- Becker, Gary S. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Becker, Gary S. 1981. *A Treatise on the Family*. Cambridge Mass.: Harvard University Press.
- Berg, Nathan C., and Donald Lien. 2002. "Measuring the Effect of Sexual Orientation on Income: Evidence of Discrimination?" *Contemporary Economic Policy*, Vol. 20, No. 4 (October), pp. 394-414.
- Black, Dan A., Hoda R. Makar, Seth G. Sanders, and Lowell J. Taylor. 2003. "The Effects of Sexual Orientation on Earnings." *Industrial and Labor Relations Review*, Vol. 56, No. 3 (April), pp. 449-69.
- Black, Dan A., Seth G. Sanders, and Lowell J. Taylor. 2007. "The Economics of Lesbian and Gay Families." *Journal of Economic Perspectives*, Vol. 21, No. 2 (Spring), pp. 53-70.
- 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. 622-42.
- Carpenter, Christopher S. 2005. "Self-Reported Sexual Orientation and Earnings: Evidence from California." *Industrial and Labor Relations Review*, Vol. 58, No. 2 (January), pp. 258-73.
- Carpenter, Christopher S. 2007. "Revisiting the Income Penalty for Behaviorally Gay Men: Evidence from NHANES III." *Labour Economics*, Vol. 14, No. 1 (January), pp. 24-34.
- Comolli, Renzo. 2005. "The Economics of Sexual Orientation and Racial Perception." Unpublished doctoral dissertation, Yale University.
- Escoffier, Jeffrey. 1975. "Stigmas, Work Environment, and Economic Discrimination against Homosexuals." *Homosexual Counseling Journal*, Vol. 2, No. 1 (January), pp. 8-17.
- Klawitter, Marieka M., and Victor Flatt. 1998. "The Effects of State and Local Anti-Discrimination Policies for Sexual Orientation." *Journal of Policy Analysis and Management*, Vol. 17, No. 4 (Fall), pp. 658-86.
- Plug, Erik, and Peter Berkhout. 2004. "Effects of Sexual Preferences on Earnings in the Netherlands." *Journal of Population Economics*, Vol. 17, No. 1 (February), pp. 117-31.

Rule, Nicholas O., and Nalinin Ambady. 2008. "Brief Exposures: Male Sexual Orientation Is Accurately Perceived at 50 ms." *Journal of Experimental Social Psychology*, Vol. 44, No. 4 (July), pp. 1100-1105.

TABLE 1—SUMMARY STATISTICS FOR SOME SELECTED VARIABLES

	Hetero males (<i>N</i> = 6,723)		G/B males (<i>N</i> = 435)		Disclosed G/B males (<i>N</i> = 267)		Undisclosed G/B males (<i>N</i> = 168)	
LABOR MARKET OUTCOMES:								
hourly wages	9.837	<i>2.411</i>	9.506	<i>2.599</i>	9.814	<i>2.630</i>	8.982	<i>1.822</i>
LABOR SUPPLY OUTCOMES:								
hours worked	38.840	<i>2.939</i>	38.242	<i>2.909</i>	38.221	<i>2.995</i>	38.277	<i>2.765</i>
hours worked partner	27.986	<i>14.93</i>	33.522	<i>12.82</i>	34.981	<i>12.20</i>	27.643	<i>13.72</i>
stated hours worked, in hypothetical family with partner and child	33.245	<i>5.677</i>	30.358	<i>6.122</i>	30.051	<i>6.253</i>	30.940	<i>5.854</i>
INDIVIDUAL CHARACTERISTICS:								
age	26.689	<i>3.072</i>	26.433	<i>2.901</i>	26.521	<i>3.280</i>	26.283	<i>2.102</i>
ethnicity	0.057		0.072		0.035		0.137	
north	0.073		0.054		0.046		0.068	
east	0.186		0.221		0.180		0.291	
south	0.225		0.241		0.263		0.202	
west	0.515		0.484		0.511		0.439	
DISCLOSURE CHARACTERISTICS:								
known to employer/colleagues			0.631		1.000		0.000	
open lifestyle			0.411		0.601		0.086	
VARIOUS HUMAN CAPITAL OUTCOMES:								
GPA (high-school) ^a	6.966	<i>0.600</i>	7.040	<i>0.619</i>	7.034	<i>0.614</i>	7.049	<i>0.629</i>
math grades (high-school) ^b	6.516	<i>1.202</i>	6.393	<i>1.226</i>	6.376	<i>1.244</i>	6.422	<i>1.199</i>
language grades (high-school) ^c	6.834	<i>0.693</i>	6.957	<i>0.749</i>	6.975	<i>0.714</i>	6.925	<i>0.806</i>
MA	0.437		0.445		0.477		0.390	
GPA (college)	7.127	<i>0.539</i>	7.238	<i>0.571</i>	7.249	<i>0.564</i>	7.221	<i>0.586</i>
gender share in college	0.620	<i>0.235</i>	0.493	<i>0.253</i>	0.477	<i>0.253</i>	0.520	<i>0.251</i>
years of parents' schooling	11.698	<i>2.747</i>	11.634	<i>2.574</i>	11.896	<i>2.598</i>	11.186	<i>2.479</i>

^a In high school and college grades range from 1 (worst) to 10 (best).

^b In high school students may matriculate in two types of math, where type B math is considerably more advanced than type A math. The better students enroll in math B and often (40 percent) take math A on the side. Of those who do, the math A grade is on average 1 points higher than the math B grade. We therefore measure math grade as the maximum of the students math B grade and math A grade minus 1.

^c In high school all students matriculate in Dutch and English. To measure language grade we take averages.

TABLE 2—LEAST SQUARES ESTIMATES OF SEXUAL ORIENTATION EFFECTS ON HOURLY EARNINGS

	(i)		(ii)		(iii)	
A. SEXUAL ORIENTATION EFFECTS ON LOG EARNINGS.						
Gay ($\hat{\gamma}_D$)	-0.036	<i>0.011***</i>	-0.029	<i>0.009***</i>	-0.035	<i>0.009***</i>
B. SEXUAL ORIENTATION EFFECTS ON LOG EARNINGS FOR (UN)DISCLOSED WORKERS (USING DIRECT DISCLOSURE).						
Gay ($\hat{\gamma}$)	-0.087	<i>0.017***</i>	-0.065	<i>0.015***</i>	-0.071	<i>0.014***</i>
Gay, known to employer/colleagues ($\hat{\delta}$)	0.080	<i>0.022***</i>	0.058	<i>0.019***</i>	0.057	<i>0.018***</i>
$\hat{\delta} + \hat{\gamma}$	-0.006	<i>0.013</i>	-0.007	<i>0.012</i>	-0.014	<i>0.011</i>
C. SEXUAL ORIENTATION EFFECTS ON LOG EARNINGS FOR (UN)DISCLOSED WORKERS (USING INDIRECT DISCLOSURE).						
Gay ($\hat{\gamma}$)	-0.055	<i>0.014***</i>	-0.045	<i>0.012***</i>	-0.048	<i>0.012***</i>
Gay, open lifestyle ($\hat{\delta}$)	0.045	<i>0.021**</i>	0.040	<i>0.018**</i>	0.030	<i>0.017*</i>
$\hat{\delta} + \hat{\gamma}$	-0.009	<i>0.016</i>	-0.005	<i>0.014</i>	-0.018	<i>0.013</i>
CONTROLS:						
Individual, human-capital, region	—		×		×	
Occupation, industry, job characteristics	—		—		×	

NOTE.— Standard errors in italics; ***significant at 1% level; **significant at 5% level; *significant at 10% level; All regressions include controls for survey year.

TABLE 3—DISCRIMINATION TESTS: LEAST SQUARES ESTIMATES THE SEXUAL ORIENTATION EFFECTS ON VARIOUS OTHER LABOR OUTCOMES OF MEN

	(i) Hours Worked (realized) (<i>N</i> =7,153)	(ii) Hours Worked of Gays with Partner (realized) (<i>N</i> =3,236)	(iii) Partners' Hours Worked (realized) (<i>N</i> =3,231)	(iv) Hours Worked in Hypothetical Case of Family with Children (stated) (<i>N</i> =4,485)	(i) Preference for Social Activities at Work (stated) (<i>N</i> =2,302)
A. SEXUAL ORIENTATION EFFECTS ON OTHER LABOR OUTCOMES.					
Gay ($\hat{\gamma}_D$)	-0.612 <i>0.139***</i>	-0.501 <i>0.213**</i>	5.241 <i>1.151***</i>	-2.927 <i>0.347***</i>	0.030 <i>0.063</i>
B. SEXUAL ORIENTATION EFFECTS ON OTHER LABOR OUTCOMES FOR (UN)DISCLOSED WORKERS (USING DIRECT DISCLOSURE).					
Gay ($\hat{\gamma}$)	-0.555 <i>0.223**</i>	-0.000 <i>0.468</i>	0.174 <i>2.529</i>	-2.385 <i>0.581***</i>	0.314 <i>0.102***</i>
Gay, known to employer/colleagues ($\hat{\delta}$) ^a	-0.090 <i>0.278</i>	-0.625 <i>0.520</i>	6.331 <i>2.814**</i>	-0.823 <i>0.709</i>	-0.448 <i>0.127***</i>
$\hat{\delta} + \hat{\gamma}^b$	-0.645 <i>0.172***</i>	-0.625 <i>0.237***</i>	6.504 <i>1.279***</i>	-3.209 <i>0.424***</i>	-0.133 <i>0.078*</i>
C. SEXUAL ORIENTATION EFFECTS ON OTHER LABOR OUTCOMES FOR (UN)DISCLOSED WORKERS (USING INDIRECT DISCLOSURE).					
Gay ($\hat{\gamma}$)	-0.571 <i>0.178***</i>	-0.276 <i>0.285</i>	4.742 <i>1.539***</i>	-2.331 <i>0.451***</i>	0.030 <i>0.077</i>
Gay, open lifestyle ($\hat{\delta}$) ^a	-0.099 <i>0.272</i>	-0.498 <i>0.417</i>	1.101 <i>2.254</i>	-1.401 <i>0.680**</i>	-0.000 <i>0.128</i>
$\hat{\delta} + \hat{\gamma}^b$	-0.669 <i>0.212***</i>	-0.773 <i>0.312**</i>	5.844 <i>1.685***</i>	-3.732 <i>0.523***</i>	-0.030 <i>0.104</i>

NOTE.— Standard errors in italics; ***significant at 1% level; **significant at 5% level; *significant at 10% level. All regressions include additional controls for the worker's the child's age, ethnicity, region of residence, and survey year.

^a This parameter measures the differences in outcomes between disclosed and undisclosed workers.

^b The sum of the parameters δ and γ measures the differences in outcomes between gay workers with disclosed identity and heterosexual workers.