Gender Typicality and Sexual Minority Labor Market Differentials

Ian Burn[*]
Department of Economics
University of Liverpool

Michael E. Martell[†]
Department of Economics
Bard College

Abstract

We leverage data on personality and behaviors in the National Study of Adolescent to Adult Health (Add Health) to estimate the relationship between labor market outcomes and masculinity/femininity using a continuous measure of gender typicality. Our measure of gender typicality contributes to empirical methodologies that move social science research beyond the gender binary. We utilize this measure to test if labor market differentials sexual minorities experience reflect, in part, unobserved differences in masculinity and femininity in the sexual minority population. Gender typicality appears to be an important determinant of labor market outcomes in general. Men and women who are more masculine experience more favorable outcomes on average. However, controlling for gender typicality does not affect the size or significance of labor market differentials for gay, lesbian or bisexual individuals. The impact of gender typicality does not differ by sexual orientation either. Therefore, even though the importance of gender typicality is clear, it is not a driver of relative differences for sexual minorities in the labor market.

[*]Email: ian.burn@liverpool.ac.uk.

[†]Corresponding author, Email: mmartell@bard.edu. Acknowledgements: We are grateful for feedback and comments from Kitt Carpenter, Jeff Frank, Marina Gorsuch, Samuel Mann, participants at the 2020 AEA meetings, the AEA Committee on the Status of LGBTQ+ Individuals in the Economics Profession Virtual Research Seminar and the Conference in Economics of Sexual Orientation at Linnaeus University.

1 Introduction

Gay and bisexual men have significantly worse labor market outcomes relative to heterosexual men. In contrast, the evidence for lesbian women and bisexual women suggests their outcomes are often as good as those for heterosexual women. The asymmetry of these labor market differentials stirred a long debate as to the mechanisms that generate sexual orientation based differentials, which goes back to the seminal work by Badgett (1995). Some have argued that these differentials reflect discrimination based on sexual orientation. However, such an outcome could arise from biases towards perceived or real differences in masculinity among sexual minorities (Aksoy et al. 2019; Blandford 2003; Blashill and Powlishta 2009). To our knowledge, this empirical link is under-researched in part because social scientists have done too little to move beyond a gender binary in quantitative research.

The difficulty in testing the effect of gender typicality on wages is exacerbated by the scarcity of high-quality data on characteristics associated with gender typicality and sexual orientation. To overcome these challenges, we utilize data in the National Study of Adolescent to Adult Health (Add Health)[1] and adopt an interdisciplinary method to quantify Add Health respondents' adherence to gender-typical norms following the methods first proposed by Fleming et al. (2017). Our use of gender typicality, which we measure as a continuous characteristic, contributes methodologically to the larger literature on gender differences in the labor market by moving beyond the gender binary. We are one of the first to document evidence of the importance of gender typicality on labor market outcomes of men and women. Our results suggest that gender typicality, in addition to sex, matters in the workplace; preferential treatment of masculinity persists. More gender-typical men have higher earnings and work longer hours, while more gender-typical women work fewer hours. The importance of

[1]This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining Data Files from Add Health should contact Add Health, The University of North Carolina at Chapel Hill, Carolina Population Center, Carolina Square, Suite 210, 123 W. Franklin Street, Chapel Hill, NC 27516 (addhealth_contracts@unc.edu). No direct support was received from grant P01-HD31921 for this analysis.

this characteristic for labor market outcomes bolsters the notion that gender is more than a dummy variable.

We find no evidence that gender typicality explains labor market differentials for sexual minorities in the United States. Gay and bisexual men have hourly wages that are 11% less than heterosexual men, and bisexual men work fewer hours per week and experience a larger differential in annual earnings. We also find that conditional on observable characteristics, lesbian and bisexual women exhibit fewer differences in labor market outcomes relative to heterosexual women, though they earn approximately 5% less than heterosexual women. This differential is only significant for bisexual women. Neither the magnitude nor the statistical significance of estimates of labor market differentials for sexual minority men and women meaningfully declines when we control for gender typicality. While the results indicate that gender typicality is an important factor in labor market outcomes, it does not explain sexual orientation based labor market differentials. Neither does its impact vary by sexual orientation. Therefore, policies aimed at promoting gender equality in the workplace should be broader and aim to protect gender nonconforming individuals regardless of other protected characteristics.

This paper makes two significant contributions to the literature. First, we show that the gender typicality measure developed by Fleming et al. (2017) is predictive of differences in economic outcomes. Measures of gender typicality have primarily been utilized to understand gendered differences in risky behaviors, such as smoking and substance use (Mahalik et al. (2015); Shakya et al. (2019); Wilkinson et al. (2018)). In contrast to the previous literature in public health that examined how one's own gender typicality explained one's own behavior, our analysis is the first to examine how one's gender typicality impacts outcomes determined by others. Our application to the labor market outcomes of sexual minorities highlights the role that gender typicality plays in understanding gender-based gaps more broadly. Second, our results allow for a fuller understanding of labor market outcomes for sexual minorities. The most recent wave of the Add Health data suggests that sexual orientation labor market differentials change little as sexual minorities age. Further, we show that sexual orientation based earnings differentials appear to be independent of any differences in characteristics related to masculinity and femininity.

2 Labor Market Effects of Sexual Orientation

The literature based on Badgett's (1995) early work on the economics of sexual orientation has built a consensus that gay men earn less and lesbian women earn more than their heterosexual counterparts. These differential outcomes are large, typically ranging from 10 to 25 % (Klawitter 2015), and have been documented in a variety of data sets in the US and internationally (e.g., Sweden, Canada, Australia, the Netherlands, the United Kingdom). Some work has shown that disadvantages are larger for bisexual individuals and younger individuals (Bayrakdar and King 2021; Martell 2019; Mize 2016; Sabia 2014, 2015). Recent research in the literature has found evidence that the wage differentials have been getting smaller over time (Carpenter and Eppink 2017; Carpenter 2008; Clarke and Arnold 2018; Jepsen and Jepsen 2021).

There is no consensus as to the source of these sexual orientation based labor market differentials, but most work suggests that discrimination is a key factor at play. The asymmetry of earnings and wage differentials for gay men and lesbian women puzzled researchers. Early efforts to simultaneously explain a lesbian premium and gay penalty led researchers to explore a number of non-discriminatory explanations (see for example Antecol et al. (2008)). Evidence of non-discriminatory explanations is limited, while evidence in favor of discrimination is increasing.[2] The negative differentials persist even when controlling for a wide variety of typically unobserved individual characteristics (Sabia 2015), and the size of the

[2]Non-discriminatory explanations of the earnings effect of sexual orientation have been postulated, but the evidence in favor of these alternative explanations often discussed in the literature (preferences or household specialization) is not very strong. Controlling for labor or supply and individual level heterogeneity does not eliminate the earnings effects of sexual orientation, thus suggesting that preferences can not wholly explain the differences (Cushing-Daniels and Yeung 2009; Elmslie and Tebaldi 2014; Klawitter 2015; Martell 2013b; Sabia 2014). Similarly, evidence in favor of household specialization as an explanation is limited (Daneshvary et al. 2009; Jepsen 2007; Martell and Roncolato 2016). Discrimination as the explanation has been corroborated by resume correspondence studies finding evidence of discrimination for openly gay and lesbian job applicants, (Drydakis 2009; Tilcsik 2011; Weichselbaumer 2003).

penalty declines with geographic and temporal decreases in prejudice (Burn 2020). Moreover, anti-discrimination laws reduce the earnings penalty gay men experience and increase their labor supply (Baumle and Poston Jr. 2011; Burn 2018; Klawitter 2011; Martell 2013a).

It is not straightforward to theorize how discrimination based on sexual orientation plays out because sexual orientation is largely concealable. Sexual minorities could pass as heterosexuals to avoid discrimination. There are three channels through which discrimination could occur. First, many sexual minorities may disclose, at times involuntarily, their sexual orientation. Second, employers may infer sexual orientation based upon observing gender atypical behaviors. Indeed, employers view gay and bisexual men as less gender-conforming than heterosexual men (Steffens et al. 2018), a pattern consistent with self-reports (Lippa 2000). Third, discrimination experienced by sexual minorities may be motivated by their gender typicality and not sexual orientation (see, among others: Blandford (2003); Ahmed et al. (2013); Aksoy et al. (2019)). Gender typicality, because it is comprised of such a wide variety of behaviors and characteristics, may be difficult to conceal in repeated work interactions. Here, labor market differentials would result from discrimination based on gender typicality wherein masculine traits and behaviors are rewarded more than feminine.

Implicit or explicit biases that reward masculine behaviors and characteristics have long been argued to be a source of sex-based differentials more broadly (Bozani 2020; Drydakis et al. 2018). Indeed, the persistence of masculine biases can explain why women's economic progress has stalled despite continued increases in labor force participation as well as education and occupational attainment (Blau and Kahn 2017). Stalled progress could reflect that economic inequality is one of the ways through which the gender hierarchy reproduces itself. The hierarchy is also reproduced by rewarding characteristics regarded as masculine (such as extreme notions of competition and the set of behaviors associated with "toxic masculinity") and penalizing those which are regarded as feminine. Rewarding these characteristics could contribute to the penalties gay men (and heterosexual women) experience and the premia lesbian women (and heterosexual men) experience.[3]

[3]However, the premia for lesbian women may not materialize. Individuals who violate gender norms via gender

The effects of violating gender typicality can vary by sexual orientation. Sexual minorities already violate gender norms via their engagement with same-sex relationships. Therefore, any benefits associated with perceptions of masculinity among sexual minority women may be canceled out by being punished for violating gender typicality. Sexual minority men may be disproportionately punished as "double violators" (Lehavot and Lambert 2007). However, Gorsuch (2019), leveraging experimental manipulation, documents that gender atypicality does not have larger adverse effects on sexual minorities.[4] Gorsuch (2019) found that heterosexual women are penalized for masculine behavior in the labor market, whereas LGBT women are not, with gender conformity having little effect on LGBT men in the US. This is in contrast to Clarke and Arnold (2018), who find a larger role of gender typicality for men. Men are rated less effectual, less respect-worthy, and less hirable in female-typed jobs, but the relationship is weaker for gay men.

3 Quantitative Approaches to Measuring Gender

A large interdisciplinary literature investigating the quantitative measurements of gender, and their applications to economic outcomes, has arisen largely in response to the pioneering work of Bem (1974). A primary contribution of Bem was to measure masculinity and femininity as separate constructs via a standardized survey instrument, the Bem Sex Role Inventory (BSRI). Despite its widespread application, many researchers have highlighted that a fundamental shortcoming of the BSRI is that its construction of gender is essentialist and

atypical behavior (including same-sex relations) threaten the gender hierarchy by engaging in activities outside their socially prescribed roles. For example, women who deviate from their roles by embodying characteristics such as assertiveness and competitiveness, which are typically remunerated among men, are punished in laboratory experiments designed to replicate workplace settings (Bowles et al. 2007; Heilman et al. 2004; Heilman and Chen 2005; Rudman and Glick 2001). This punishment may discourage behavior that threatens the gender hierarchy by breaking the boundaries it sets for women's behavior (Lehavot and Lambert 2007).

[4] Another possible explanation of differential effects of gender typicality for sexual minorities may be that the masculine characteristics of gay men are discounted because they are inconsistent with stereotypes assumed to exist in the workplace. Similarly, lesbian workers violating gender norms may not be punished as much because their greater masculinity is in line with stereotypes. Though it may be the case that the distaste of an employer for gender atypicality outweighs the perceived benefits of greater masculinity, and thus lesbians may be penalized and earn less than similarly situated heterosexual women.

proscribes what characteristics are masculine and feminine (Fleming et al. 2017). The inventory was developed in a single cohort of students at a university, and the items of the inventory do not change depending on the time or place in which they are implemented. This lack of context, coupled with constructing masculinity and femininity as independent factors, does not allow the BSRI to capture the relational aspects of gender or the importance of the context in which gender is measured (Fleming et al. 2017).

A number of new methodologies building on Bem endeavored to relax the essentialist nature of the BSRI measure. Beginning with Lippa and Connelly (1990), these “gender diagnostic techniques” have been used in a number of social science fields, such as psychology and public health. Gender diagnostic measures of gender typicality are based on the theory of gender performance postulated by West and Zimmerman (1987) and later expanded by Butler (1990). The idea of gender performance is that gender is independent of sex and is rooted in how one expresses themselves in interactions with others (a form of signaling). Men and women act in a specific way so as to adhere to the social and cultural norms of male and female behavior in a given society. This theory focuses on an individual’s behavior and how it compares to the behavior of other individuals in society. Thus, that which constitutes ”masculine” and ”feminine” depends on context, the behaviors of individuals in the particular society in which an individual lives. Gender diagnostic techniques were first used on Add Health data by Cleveland et al. (2001). This method was most recently improved upon by Fleming et al. (2017) and has been used in the public health literature (Shakya et al. 2019; Wilkinson et al. 2018).

Methodologically, the crux of these measures is that they utilize survey responses to measure how similar a respondent’s answers are to their same-sex peers. However, the particular survey responses used to measure gender typicality depend on the survey (the time and place) analyzed. Questions on personality characteristics, interests and hobbies, and behaviors are used to determine how an individual behaves. Individuals who are more gender-conforming should give answers that are more predictive of their sex. To measure this, the gender diagnostic methods use a probit regression where a binary indicator of sex is regressed on the survey responses to predict the probability a respondent is male (or female). These predicted probabilities are then used as a measure of conformity because individuals with high predicted values behave in ways that are more typical of their sex.

Gender typicality (as measured using gender diagnostic techniques) is predictive of many different behaviors. Early work using the Cleveland et al. (2001) measure focused on the link between gender conformity and sexual behavior of young adults (Udry and Chantala 2004, 2006). More recent work has focused on the role of gender typicality on health behaviors. Individuals who are the most gender conforming are observed to have the greatest risks of adverse health outcomes (Shakya et al. 2019). High gender conformity among men is correlated with smoking cigarettes, the use of marijuana and recreational drugs, and prescription drug misuse (Cleveland et al. 2001; Lowry et al. 2018; Mahalik et al. 2007, 2015). Men who are more gender typical are 75% more likely to binge drink, and more report a higher frequency of binge drinking compared to less gender typical males (Wilkinson et al. 2018). Women who are more gender typical report a less high frequency of substance use (Wilkinson et al. 2018).

The focus of gender typicality in public health has mainly led researchers to focus on how a desire to conform to perceived norms of behavior influence the decisions of young adults. This focus on risky health behaviors has meant the application of these methods has been tested on a narrow range of outcomes. Thus, our use of gender typicality to understand labor market outcomes represents a significant expansion of the scope of social phenomena gender typicality can explain.

4 Data and Methodology

We use the National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994-2018, a longitudinal study of a nationally representative sample of US adolescents in grades 7 through 12 during the 1994-1995 school year. The Add Health cohort was followed into young adulthood with four in-home interviews capturing a wide variety of individual and contextual characteristics over time.

The Wave I in-home survey was administered in 1995 between April and December to a sample of 12,105 students in 132 high schools. Four additional waves of the Add Health follow the Wave I respondents as they transition to adulthood. Wave II was conducted in 1996, approximately one year after the baseline survey. Wave III was administered in 2001 when the respondents were 18 to 26 years old. Wave IV was conducted in 2007 when the respondents were 24 to 32 years old. Wave V was administered in 2016 through 2018 when the subjects were 32 to 42. We utilize Waves III, IV, and V. Due to attrition in the sample, we do not have a perfectly balanced panel. Therefore, we treat the Add Health data as a repeated cross-section of a nationally representative cohort of young Americans.

The detailed questions asked of Add Health respondents allow researchers to estimate sexual orientation differentials for both single and cohabiting sexual minorities. This contrasts with widely used sources of public-use data, such as the American Community Survey or the Current Population Survey, where researchers must infer sexual orientation via family inter-relationships.

However, there are limitations to the Add Health as it applies to the study of the sexual orientation based labor market differentials (Sabia 2014, 2015). First, because the study follows individuals as they enter adulthood, the average age of these individuals is younger than the average American. The relatively young may lead us to over or underestimate wage differentials if income trajectories through adulthood differ by sexual orientation (Martell 2019).

A second drawback of the data, which is common to nearly all research in this area, is that researchers cannot observe whether the respondent has revealed their sexual orientation to their employer or co-workers. Our inability to control for disclosure of sexual orientation may lead to underestimates of the impact of sexual orientation disclosure on earnings. Of course, involuntary disclosure is also possible. Involuntary disclosure may be more likely among sexual minorities who behave in gender-atypical ways that conform to stereotypes. If this is the case, we may be more likely to estimate larger earnings differentials for LGB individuals who are more gender atypical.

4.1 Measuring Sexual Orientation

We classify respondents' sexual orientation based on individual self-reports, which is standard among research utilizing the Add Health data (Sabia 2015). Using Computer-Assisted Self-Interviewing (CASI), Add Health asked respondents in each wave to:

> "Please choose the description that best fits how you think about yourself: 1. 100% heterosexual (straight) 2. Mostly heterosexual (straight) but somewhat attracted to people of your own sex 3. Bisexual, that is, attracted to men and women equally 4. Mostly homosexual (gay), but somewhat attracted to people of the opposite sex 5. 100% homosexual (gay) 6. Not sexually attracted to either males or females."

Those who responded that they were "100% heterosexual" (category 1) were coded as heterosexual.[5] Those who indicated some attraction to both sexes (categories 2, 3, and 4) were coded as bisexual, and those who reported they were "100% homosexual" (category 5) were coded as "gay/lesbian."

As does Sabia (2014), we note that this measurement of sexual orientation conflates two often, but not always, overlapping constructs: sexual attraction with sexual identity.[6] The conflation of identity and attraction leads to some differences in the sample of sexual minorities in the Add Health compared to other nationally representative samples. While the percent of the population that identifies as gay or lesbian is similar to that observed in other surveys (for example, the National Health Interview Survey or General Social Survey (Carpenter and Eppink 2017; Martell and Eschelbach Hansen 2017)), the Add Health includes a higher percent of the population that is bisexual. The larger bisexual population likely reflects

[5]Those not attracted to either sex were coded as their own category ("asexual"), and those preferring not to disclose their sexual orientation are included in the heterosexual category.

[6]Sexual identity is the most relevant construct for the study of labor market outcomes because theories to explain these outcomes revolve around employer perceptions of identity or worker decisions based on their identity (Carpenter 2007; Martell and Eschelbach Hansen 2017).

that the Add Health sample is younger than typical Americans, and, more importantly, that the Add Health allows respondents to report sexual attraction along a continuum. Recording attraction along a continuum better reflects the complex nature of sexuality. We are able to observe individuals with bisexual attractions that may not map into bisexual identities (those who are "mostly heterosexual.") This is most prominent among women for whom bisexual and fluid sexual orientations are more common (Laumann et al. 2000).[7]

In our baselines specifications, we aggregate responses about sexual orientation into two groups, homosexuals and bisexuals, to maximize our sample of sexual minorities and measure the impact of adherence to gender typicality. We combine mostly heterosexual, bisexual, and mostly homosexual into a single category of bisexual. Overall, we classify 8,926 men as heterosexual, 202, men as gay, and 488 men as bisexual. We classify 9,419 women as heterosexual, 102 women as lesbian, and 2020 women as bisexual. As will be discussed in more detail later, our pattern of results is not sensitive to the manner in which we classify respondents' sexual orientation. Aggregating the sexual orientation categories into a single indicator or estimating outcomes separately for each of the five sexual orientation response categories yields similar results.

4.2 Fleming et al. (2017) Gender Typicality Measure

Our measure of adherence to gender typicality is constructed using the gender diagnostic technique proposed by Fleming et al. (2017). A fundamental aspect of the empirically driven approach is that predictors of gender typicality are selected separately for each wave. Thus, the measure - and how it is constructed - varies over time. Because the construction varies over time, the methodology proposed by Fleming et al. (2017) is a multi-step process. The process involves a) identifying predictors of gender typicality b) specifying an empirical

[7]We are not the first to observe that sexual orientation in the Add Health reflects a continuum (e.g., Sabia (2014, 2015)). The distribution within the data is as expected. In Appendix Table A1, we tabulate the distribution of sexual orientation changes across waves. For men, the distribution of sexual orientation is fairly stable and is strongly bimodal, with few men reporting to be purely bisexual. For women, we find sexual orientation is more continuous and fluid, which is well documented in the study of sexuality Laumann et al. (2000); Sabia (2015). The number of women identifying as "mostly heterosexual" grew rapidly between Waves III and IV.

model to predict gender and c) using predictions from the empirical model to construct a continuous measure of gender typicality.

Following Fleming et al. (2017), we first identify the subset of survey questions related to an individual's behavior and preferences in each wave of the Add Health.[8] We exclude any question with more than 300 missing responses. In each wave, we take the subset of questions that remain to specify an empirical model of gender. We calculate the difference between male and female responses to each question and keep the 50 questions with the most significant differences between male and female respondents. We use these 50 variables in a manual backward stepwise logit regression to predict the likelihood a respondent is female. After each iteration, we drop those questions that were insignificant predictors. We then re-estimate the model with the remaining variables. We continue this process until all remaining variables have a p-value of less than 0.0001. After completing the stepwise elimination of insignificant predictors, we are left with an empirical specification we can leverage to generate measures of gender typicality. We regress the selected variables on an indicator variable for being female using a logit regression. Specifically, we estimate:

$$Female_{it} = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n + \epsilon_{it} \tag{1}$$

where $Female_{it}$ is a dummy variable equal to 1 if respondent i surveyed in wave t is female. The independent variables used in the equation are individual i's responses to the question X within each wave. Tables A2 to A4 detail the questions that we use to estimate equation 1. These tables highlight that characteristics typically associated with masculinity and femininity, such as risk-taking (Bem 1974), are important determinants of gender typicality. However, as expected, the characteristics and behaviors associated with gender typicality are much broader. The broad range of predictors differentiates our approach from existing research that focuses on specific characteristics such as preferences for competition

[8]We exclude questions related to one's gender (such as menstruation) because they perfectly predict sex, are demographic because they do not reflect gender, or not referring actively to the respondent (e.g., partner's perceptions of the respondent) because they are not about the respondent.

(Buser et al. 2018). For example, the frequency of crying is the largest predictor of being female in Wave II of the Add Health. Other questions highly predictive of gender in Wave II include frequency of playing sports, getting into a serious physical fight, tanning in the summer or a tanning bed, and frequency of wearing a seat belt. These questions highlight the key theoretical idea that gender is performed through a variety of behaviors and characteristics. These characteristics reinforce commonly held notions of behavioral differences between men and women (for example, in their decision to wear a seat belt or a helmet). While some individual questions may not have a clear link to productivity or characteristics perceived to be relevant in the labor market, these behaviors collectively reflect characteristics and choices individuals make in their behavior that is read by society as ”masculine” or ”feminine.”

We predict the probability a respondent is female using the estimated coefficients.[9] For ease of interpretation, we standardize the probability an individual is female to be mean zero with a standard deviation of one by sex. For men, we multiply the normalized score by negative one so that it reflects the probability of an individual being male. These scores can be understood as a measure of distance. The higher the value of an individual's AGT score, the more gender-typical that individual is relative to the average man or woman in that wave of the Add Health. The mean and standard deviation vary in each wave, so our measure naturally evolves as individuals age, which is a methodological improvement over measures such as the BSRI or Cleveland et al. (2001) who anchor their measures at a single point in time.

The characteristics of our measure of gender typicality suggest that it is a plausible measure of gender typicality. The differences among men and women in the predicted probability of being female are as expected; women are more likely to be predicted female than men (see Figure 1 which shows the distribution of the probability female). The model clearly differentiates between men and women. Importantly, there is also significant variation in gender

[9] In supplementary analyses, we also excluded sexual minorities from the sample used to estimate equation (1) and these coefficients. Results were unchanged.

typicality within each sex.

As shown by the previous literature, Figure 1 clearly shows gay and bisexual men are significantly less gender-typical than heterosexual men (with a significantly lower likelihood of having the lowest probability of being female). Lesbian and bisexual women are also less gender-typical than heterosexual women, though the difference is small.[10] Figure 1 also shows that gay and bisexual men are increasing their gender typicality over time, while we observe a small decreases in gender typicality among lesbian and bisexual women over time.[11]

The characteristics of our measure of gender typicality suggest that it is relevant for the study of labor market outcomes. The AGT measure is strongly correlated with marriage, an outcome related to earnings. In Wave IV data (when respondents are between the ages of 24 and 32), we find gender typicality is correlated with being married among heterosexuals (Table A7).[12] There is a negative correlation among men, indicating that more gender-typical men marry later. On the other hand, a positive correlation exists among women, indicating that gender-typical women marry earlier. In addition to being correlated with factors that affect earnings, our measure of gender typicality is correlated with the gendered perceptions of others. The perception of others is important as it relates to characteristics that an employer may observe, either in an interview or in the workplace. In Wave V, individuals report the extent to which their appearance is perceived as feminine or masculine.[13] Gender typical-

[10] These differences between heterosexual and sexual minorities remain significant even in a regression framework that predicts differences in AGT by sexual orientation conditional on the host of demographic characteristics listed in Equation 2 below (see Appendix Table A5).

[11] In Appendix Table A6, we estimate how AGT changes over time by sexual orientation. Adherence to gender typicality is significantly different for sexual minorities in young adulthood. Sexual minority men respond to the Add Health with answers which are less predictive of them being male, while sexual minority women respond with answers which are more predictive of them being male. As these gay and bisexual men age, these differences lessen over time. Sexual minority men give increasingly similar answers as their heterosexual peers in each wave. This is in contrast to lesbian and bisexual women, who give increasingly different answers as they age.

[12] We restrict this analysis to only heterosexuals because marriage was not legally available to all sexual minorities during Wave IV. The significant correlation persists in a regression framework that predicts marriage among heterosexuals conditional on the host of demographic characteristics listed in Equation 2 below.

[13] Specifically, respondents are asked "A person's appearance, style, or dress may affect the way people think of them. On average, how do you think people would describe your appearance, style, or dress?" Responses range from 1 "very feminine" to 7 " very masculine." The unconditional correlation between AGT and masculine appearance is -0.1325 for women and 0.0899 for men.

ity among men is positively correlated – and gender typicality among women is negatively correlated with – the likelihood of reporting a masculine appearance, style, or dress (see Appendix Table A8). It is important to note, however, that appearance is just a small component of adherence to gender-typical norms. In Appendix Table A8, we find that controlling for appearance and sexual orientation can only explain 11% of the variation in AGT for men and 15% of the variation of AGT for women even though the size of the correlation is large.

4.3 Measuring Labor Market Outcomes and Demographic Characteristics

We combine our novel measure of gender typicality with more standard measures of labor market outcomes (see Sabia (2014).) We estimate differences in income, hourly wages, employment, and hours worked.

To measure employment, we use a respondent's response to the question, "Are you currently working for pay for at least 10 hours a week?" Respondents who answered yes were coded as employed, and respondents who answered no were coded as not employed (we do not differentiate between unemployed and not in the labor force). In Table 1, we find similar employment rates for all groups. Between 80 and 90% of respondents to the Add Health are employed, with gay men having the highest employment rate and heterosexual women having the lowest.

Total earnings from wages are calculated based on subjects' responses to the following question in Wave III, IV, and V, "Now think about your personal earnings. In [the previous year], how much income did you receive from personal earnings before taxes, that is, wages or salaries, including tips, bonuses, and overtime pay, and income from self-employment?" If a respondent replied "do not know" to the earnings question (and in all of Wave V), they were prompted with seven categories of earnings. We follow Sabia (2015) and use the midpoints of each to determine total earnings.[14] Among men, we find that heterosexual men earn

[14]Our results are qualitatively similar if we exclude Wave V due to the categorical coding of income and if we

approximately \$3,000 more than bisexual men and \$4,000 more than gay men per year. The difference in income between lesbian and bisexual women and heterosexual women is smaller. Heterosexual women earn approximately \$2,000 more than lesbian women but about \$1,200 less than bisexual women. The higher earnings of bisexual women are driven by "mostly homosexual" and "mostly heterosexual" women earning much more than bisexual women; this pattern holds among men and on the hourly wages margin as well (see Tables A9 and A10.)

Hours worked per week are based on respondents' answers to the question, "How many hours per week (do/did) you usually work at this job?" We find that heterosexual men work the most hours, while heterosexual women work the least. The differences in hours worked between gay and bisexual men and heterosexual men were approximately 2 hours, similar to the difference between heterosexual and lesbian women. However, the gap between heterosexual and bisexual women was less than 10 minutes.

We calculate hourly wages as total earnings divided by the usual number of hours worker times 50.[15] We find that heterosexual and bisexual men have the highest hourly wages, earning \$21.30 and \$21.47 respectively, per hour, while gay men earned \$19.74. Among women, the gap in hourly wages was larger. Lesbian and bisexual women earned \$16.22 and \$19.32 respectively; heterosexual women earned \$18.45.

Differences by sexual orientation in the characteristics may play a role in determining their labor market differentials. Gay and lesbian, but not bisexual individuals, are more likely to be White than heterosexuals. Sexual minority men are more likely to be Asian than heterosexual men, but sexual minority women are less likely to be Asian than heterosexual women. This pattern holds as well for Hispanic identities.

We also find that among men, sexual minorities are more educated than heterosexuals in the Add Health sample. Sexual minority men are more likely to obtain a bachelor's degree or a graduate degree and less likely to receive only a high school diploma. There are no

compare OLS to interval regressions within the Wave V data. Results available upon request.

[15]As does Sabia (2014), we trim hourly wages. Wages are bottom coded at \$2.13 per hour (which is equal to the tipped federal minimum wage since 2007) and top coded at \$500 per hour.

significant differences among women in education. Sexual minority women and heterosexual women obtain college and graduate degrees at very similar rates.

4.4 Regression Framework

We estimate differences in labor market outcomes between sexual minorities and heterosexuals in specifications that replicate the differentials previously observed in the Add Health (Sabia 2014, 2015), with the addition of the Wave V data. We use Wave III, IV, and V data to construct a repeated cross-section of individuals. We estimate differentials in log annual income, log hourly wages, employment status, and hours worked per week. Our wage outcome is bottom coded at \$2.13 per hour (the tipped federal minimum wage throughout Waves III to V) and top coded at \$500 per hour. We include controls for individuals having top- or bottom-coded wages. Approximately 4.9% of employed respondents had hourly wages below \$2.13 per hour, and 0.04% had hourly earnings greater than \$500.

Here, we illustrate the regression models using a single indicator variable for sexual minorities. Sexual minorities can be grouped together in a number of different ways. The effect of changing the groupings of sexual minorities to be larger or smaller is discussed in more depth in the results section.

$$Y_{ist} = \alpha + \theta LGB_{it} + \delta \mathbf{X}_{it} + \sigma_s + \eta_t + \epsilon_{ist} \tag{2}$$

We estimate the models for men and women separately. We begin with a simple specification that predicts labor market differentials for sexual minorities (θ) conditional on demographic characteristics $\mathbf{X}_i$ and high school (σ_t) and wave fixed effects (η_t). $\mathbf{X}_i$ contains controls for age and age squared as well as indicators for race (White, Black, Asian), Hispanic ethnicity, cohabitation status, educational attainment (High School Diploma, Bachelors Degree, Graduate Degree) as well as current enrollment in school, and occupational attainment.

Following Sabia (2014), we include individual-level controls for cognitive ability, physical appearance, physical health, and religiosity.[16] The fixed-effects for high schools (σ_s) capture unobserved differences in high schools and proxying for unobserved community-level differences.[17] Following Sabia (2014), we use the unweighted data from the Add Health. Standard errors are clustered at the school-wave level.

We then add our measure of AGT to test if its inclusion leads to attenuation in estimated labor market differentials. We conclude by estimating equation (3) which also includes an interaction between sexual orientation and AGT to test if heterogeneous effects of AGT by sexual orientation lead to an attenuation of sexual orientation based differentials.

$$\begin{aligned} Y_{ist} = \alpha + \beta AGT_{it-1} + \theta LGB_{it} + \omega(AGT_{it-1} \times LGB_{it}) \\ + \delta \mathbf{X}_{it} + \sigma_s + \eta_t + \epsilon_{ist} \end{aligned} \tag{3}$$

Our first primary parameter of interest is the indicator variable for our sexual minority group. We identify the average gap between sexual minorities and heterosexuals with θ_1, the effect of being lesbian, gay, or bisexual on outcomes conditional on observable characteristics.

Our second parameter of interest is β, the effect of a one standard deviation increase in AGT on labor market outcomes. We lag the AGT measure by one wave to avoid reverse causality between labor market outcomes (such as income) and the activities in which an individual engages.[18] We include an interaction between the lagged gender typicality measure

[16] In addition to the above controls, Sabia (2014) includes controls for engaging in risky behavior, personality, and mental health. We do not use these controls because many of them appear in the measure of adherence to gender-typical norms from Fleming et al. (2017). Their inclusion would potentially be collinear with the gender typicality measure. If we include both these measures in the wage regression, the results are very similar, but the coefficients on AGT are slightly attenuated. Our results are also qualitatively similar if we omit the individual-level controls included in Sabia (2014)

[17] In some specifications discussed below, we also add in controls for sexual minority specific state laws on marriage, discrimination adoption, and hate crime laws. These are only observable in Wave III and IV.

[18] Including these AGT measures introduces the possibility that these estimates may overstate statistical significance due to our use of a generated regressor (Murphy and Topel 2002). Therefore, our results are the upper bound of the effect that AGT has on the sexual orientation wage gap. This bias does not pose a large problem for us because we largely find null results for the effect of gender typicality. The null result of gender typicality does not appear

and the sexual orientation indicator to allow the effect of gender typicality to vary for sexual minorities and heterosexual individuals (ω).

The degree to which including AGT in our equation results in an attenuation of estimated labor market differentials for sexual minorities corresponds to the importance of AGT in contributing to unequal outcomes. In this case, AGT should have a positive effect ($\beta > 0$) on the wages of men and a negative effect ($\beta < 0$) on the wages of women.

5 Results

We begin by presenting results based on Equations 2 and 3, estimating average labor market outcomes without controlling for AGT in Table 2. The top panel reports the results for men, and the bottom panel reports the results for women. In both panels, columns 1, 4, 7, and 10 report our baseline estimates of differences in labor market outcomes for sexual minorities. We then add the control for AGT (columns 2, 5, 8, 11) and its interaction with an indicator for a lesbian/gay/bisexual identity (columns 3, 6, 9, 12).[19]

The labor market differentials presented in Table 2 are consistent with the existing literature. Gay and bisexual men have lower annual incomes than heterosexual men even though the difference for bisexual men (26% and statistically significant) is much larger than that for gay men (8% less, see column 1). Differences in hourly wages are smaller than annual income. Gay men experience 13% lower hourly wages and bisexual men experience 10% lower hourly wages than comparable heterosexual men (column 4). The wage gap for bisexual men being smaller than the income gap is not due to differences in labor market participation.[20] There are no meaningful differences in the likelihood of being employed

to arise due to over controlling. Our pattern of results is also present in alternative specifications that only include a limited vector of basic demographic controls as well as those that exclude controls for occupational attainment.

[19]We find the results are similar when limiting our definition of a sexual minority as those who identified as such throughout Wave III to Wave V (Table A11). Those who consistently identify as LGB may be more likely to differ in their AGT if sexual minorities invest in their identity by rejecting that which is typical for their gender. They may also be more likely to have disclosed their identity. However, results based on this specification are qualitatively similar to those discussed above.

[20]In supplementary analyses, available upon request, we find that this finding is driven by changes among full-time

for gay and bisexual men (column 7),[21] Gay and bisexual men do, however, appear to work fewer hours than heterosexual men work. This differential is much larger and statistically significant for bisexual men (2.12 fewer hours per week) than gay men (1.19 hours per week less than heterosexual men.)

We show in columns 2, 5, 8, and 11 that AGT is strongly correlated with labor market outcomes for men. Men who have AGT scores one standard deviation above the mean earn 3% more annually and 2% more per hour than their less gender typical counterparts. Even though they do not have a differential likelihood of being employed, they also work approximately one quarter more hours per week. This is not surprising. Time spent in paid (household) labor is one way through which masculine (feminine) identities are expressed (Goldin 2014; Bertrand et al. 2015), and labor supply patterns are correlated with local gender norms (Fortin 2005). This pattern of results is consistent with masculinity, a characteristic likely more common among gender typical men, being rewarded in the labor market.

Even though gay men and bisexual men are less gender typical than heterosexual men controlling for AGT does not meaningfully affect the size and significance of the labor market differentials they experience. A minor exception is the statistically insignificant difference in hours worked for gay men falls when controlling for AGT (column 11.) However, this differential increases and becomes statistically significant when we allow the impact of AGT to vary by sexual orientation (column 12.) There is no other evidence of a differential impact of AGT by sexual orientation.[22] More fundamentally, estimated labor market differentials are largely unchanged. This pattern of results is not sensitive to the manner in which we classify respondents' sexual orientation. We show in Table 3 that when we combine homosexuals and bisexuals into a single group or disaggregate bisexuals into specific groups (separating "mostly heterosexual", "bisexual", and "mostly homosexual"), the results across all four

workers (35 hours or more per week) along the intensive margin of work.

[21]One might expect to find differences in employment due to gay and bisexual men obtaining more years of schooling. We control for years of education and being enrolled in school, which would account for any differences due to educational obtainment.

[22]The negative differential effect of AGT for bisexuals on hours worked is driven almost entirely by the mostly heterosexual sample, with AGT impacting the hours worked of bisexuals and mostly homosexuals positively.

margins are very similar.

Moving to Panel B of Table 2, we find no statistically significant labor market differentials for lesbian women. We do, however, find larger and statistically significant differences for bisexual women. Bisexual women earn approximately six percent less annually (column 1) and 5 percent hourly (column 3.) The lack of a wage penalty for lesbian women and the small penalty for bisexual women is in line with the previous literature. Bisexual, but not lesbian, women are also approximately 1 percent less likely to be employed (column 7)[23] Unlike men, there are no significant differences in hours worked per week among women (column 10).

AGT appears to play a more limited role in determining labor market outcomes for women than men. The correlation of AGT to annual income (column 2), hourly wages (column 4),[24] and the likelihood of being employed is small and not robustly significant. The fact that AGT is not correlated with annual incomes for women contrasts with the income effects observed for men. This may be because any gains which arise from being more masculine (which was rewarded by employers when observed in men) are offset by the penalties for gender nonconformity. AGT is negatively correlated with hours worked. A one standard deviation increase in AGT is correlated with a 0.26 hour per week decrease in hours worked, indicating that women who behave less like the average woman work more. AGT is clearly correlated with, and it may be an important determinant of, hours worked for women. Indeed, this relationship is similar in size to that among men on the same margin.[25]

Similar to the results for men, AGT cannot explain the sexual orientation based differ-

[23]Given this significant difference in employment rates, we have run selection corrected estimates of columns 1 through 6 (results available upon request). Because the difference is only 1%, the effect of correcting for selection was negligible.

[24]For all women, we do observe a significant effect of AGT on wages. A one standard deviation increase in AGT decreases wages by 1%. Allowing the effect of AGT to differ by sexual orientation (column 6) does not change the wage penalty but reduces its significance. These results suggest that adhering more to male gender typicality and less to female gender typicality is associated with higher wages for women. Thus, it appears that employers in the United States value masculine norms of behavior in the workplace and reward individuals who more strongly adhere to these norms with higher wages, though the benefits of adhering to these norms for women are slightly smaller than for men.

[25]Unlike men, however, this relationship is driven by changes among part-time workers (less than 35 hours per week) along the intensive margin of work.

ences in hours worked, employment, annual income or hourly earnings. These differentials, largely observed for bisexual women, remain the same size when controlling for AGT. Moreover, the impact of AGT does not appear to vary by sexual orientation (columns 3, 6, 9 and 12.) Similar to that of men, Table 4 shows that when we combine homosexuals and bisexuals into a single group or disaggregate bisexuals into specific groups, our pattern of results is very similar.

The results clearly show that gender typicality is an important explanatory variable for differences in outcomes within sexes but that it does not explain the differential experienced by sexual minorities. For example, take the wage gap we observe for gay and bisexual men in Table 2, gay men earn 10% less and bisexual men earn 11% less. The average AGT for gay and bisexual men is -0.46. If one were to shift the average gay and bisexual men to the maximum AGT score for heterosexual men (which is 1.16), the increase in wages for gay and bisexual men would only be 3.24%. Even such an extreme increase in AGT among sexual minority men leaves gay and bisexual men with sizable disadvantages. Similar exercises yield similar results for differences in annual income and hours worked for gay and bisexual men. For women, there is no significant effect of AGT on wages and income, and the signs of the coefficients suggest that an increase in AGT would decrease wages and income, again allowing one to rule out AGT as an important cause of labor market income differentials for women observed in the data.

5.1 Robustness Analyses

As discussed above, approaches to measure gender typicality continue to evolve. Further, the AGT measure proposed by Fleming et al. (2017) has never been used for economic analyses. Therefore, we implement a number of robustness checks to verify that our pattern of results is not an artifact of the methodological choices embedded in our baseline estimation strategy. Given that we observe no significant differences in our results when we aggregate homosexuals and bisexuals into a single group, we use this highly aggregated grouping in our analyses going forward. The increased sample size provides greater statistical power in the tests for robustness and heterogeneous effects. It also maximizes the likelihood of observing significant differential effects of AGT. We consider three robustness checks: allowing for nonlinear effects of AGT, allowing the effect of AGT to vary across waves, and varying how we construct and control for gender typicality.

First, we turn our attention to the possibility of nonlinear effects of AGT. Our baseline estimates control for AGT linearly. AGT could have a nonlinear effect if larger deviations from average are disproportionately punished or rewarded. We find no evidence that AGT significantly impacts labor market outcomes in nonlinear ways.Only for hours worked for men do we find any evidence that the effect of AGT may by nonlinear. Our results suggest the effect of AGT on hours worked for men is a third-order polynomial, while the effect of AGT on the other three outcomes for men and all outcomes for women are linear.[26] In Appendix Table A12, we show that the results for sexual minority labor market differentials based on specifications that allow for nonlinearities. Each row reports the average differential between LGB individuals and heterosexual individuals from specifications that include the full set of controls listed in equation 3 with varying forms of AGT. The first row of each panel in Appendix Table A12 reprints the average differentials from columns 1, 4, 8, and 12 in Table 2, which uses a first-order polynomial. The next three rows increase the order of the polynomial (second to fourth-order).

These results in Appendix Table A12 show that the limited role of AGT in determining differentials for sexual minorities is not driven by our decision to control for AGT linearly. Appendix Table A12 provides some evidence that there may be nonlinearities of AGT on the hours of work margin (column 4), which may lead to nonlinearities in income (column 1) for men and women.[27] However, the nonlinearity on the hours worked margin suggests our estimates may be conservative. The point estimates of hours differentials get larger as

[26] These results are available upon request.

[27] For men, the income gaps shrink (from 21% to a statistically insignificant but meaningfully large 14%) as we increase the order of the polynomial. For women, the income gap increases from 6% to 10% and becomes more statistically significant as the polynomial of AGT increases.

we increase the polynomial. In all cases, it is important to note that the point estimates from specifications of different polynomials are not statistically significantly different. Moreover, estimates of wage and employment differentials (columns 2 and 3) are largely unchanged. None of these results suggest that AGT explains these sexual orientation based differentials.

Second, we turn our attention to the time-varying effects of AGT. Appendix Table A13 illustrates how the effect of AGT differs across waves. For men, we find evidence that the effect of AGT declines over time, with the income and wage effects decreasing in magnitude wit each wave. The effect of AGT on labor supply, both in terms of employment and hours worked, also appear to vary across waves, decreasing from Wave III to Wave IV, but increasing from Wave IV to Wave V. For women, we find no significant changes in the effect of AGT on income, wages, or employment across waves, but a significant increase in the magnitude of the effect of AGT on hours work when we move from Wave III to Wave IV.

We find that allowing the effect of AGT to vary over time by interacting wave effects with AGT does not affect the labor market differentials observed in the baseline results. We find that income differentials for gay and bisexual men decline as they age due in part to the increases in differences in hours worked, but we do not find a consistent pattern in the changes in wages and employment. For women, we find that labor market differentials do not change as individuals age. For both men and women, there is no evidence that allowing for more flexible forms of controlling for AGT impacts the labor market differentials observed for sexual minorities.

Finally, we turn our attention to how we construct AGT. We find the inability of gender typicality to explain labor market differentials for sexual minorities is not due to how we measure AGT. Our results are also unchanged if we construct AGT (estimate equation 1) on the sample of heterosexual individuals and exclude sexual minorities. Our results are also robust to abandoning the Fleming (2017) methodology that combines multiple survey instruments into a single measure. In Appendix Table A14, we replace AGT with the series of variables used to construct it (listed in Appendix Tables A2 - A4) for each wave. Again, the size of estimated differentials is largely unchanged.

6 Heterogeneous Effects of AGT

Given the robustness of the null effects to alternative specifications of our model, we now turn our attention to the potential for, possibly countervailing, heterogeneous effects of AGT. Heterogeneous effects may arise if demographic characteristics are associated with the incidence and costs of gender nonconformity. We explore if more flexible empirical specifications that allow the effect of AGT to vary by characteristics of the individuals as well as their work. We focus on heterogeneity by work characteristics because these characteristics will better proxy the attitudes of those who observe gender typicality and contexts in which it may be perceived to be productive. These include the gender composition of the occupation they work in, educational obtainment of the worker, and whether the worker lives in a state with an employment nondiscrimination law that protects LGB individuals. We also note that any heterogeneous effect by individual demographic characteristics does not explain sexual minority labor market differentials.[28]

First, we turn our attention to occupations. Even though we control for occupations, we may expect the impact of AGT to vary across occupations, because the disclosure of sexual orientation and the impact of AGT may also depend on the work environment. The environment may matter if the effect of AGT depends on workplace gender norms.[29] These norms likely depend on the sex composition of the occupation because occupations with more men are likely to favor masculine characteristics more than feminine. This bias towards in-group characteristics is a likely contributor to cultures of "toxic masculinity" in some male-dominated occupations. We address this possibility by controlling for the percent of individuals in each occupation that is female. We allow the effect of AGT to vary by sexual orientation as well as indicators for occupations with a high concentration of females (66% or

[28] In the appendix, we present two additional analyses focusing on heterogeneity by race and ethnicity and cohabitation status. We find some evidence of differences in the effect of AGT for individuals of color, but the small sample sizes mean these results are not very robust for smaller groups (Appendix Table A16). We find little evidence of differences in AGT by cohabitation status (Appendix Table A17).

[29] We find some evidence that occupation matters for AGT. If one does not control for occupation in Table 2, we find weaker effects of AGT on labor market outcomes. The qualitative patterns observed in the baseline analyses are unaffected. These results are available upon request.

higher female), gender-neutral (33% to 65% female) or low concentration of females (33% or less female).[30] For men, there is significant evidence that gender typicality (masculinity) is more highly rewarded in male-dominated occupations. For example, the first column of Table 5 shows that the return to gender typicality is (insignificantly) negative in female-dominated jobs but twice as large and significantly positive in male-dominated jobs; the income effect of AGT is 4% for a one standard deviation increase in AGT. The effect on the hours margin is similar. Thus, male-dominated occupations appear to disproportionately reward masculinity - another instance of gender biases meaningfully affecting outcomes for all individuals. However, allowing for heterogeneity by occupation does not explain the labor market differentials gay and bisexual men experience, nor does the effect of AGT vary by sexual orientation. We find no evidence of a differential effect of AGT for gay and bisexual men across occupations.

For women, we find a similar pattern of results. There are significant differences in the effect of AGT across occupations. The negative association between gender typicality and hours worked observed in the baseline estimates is stronger in occupations with more women. For lesbian and bisexual women, we find large income and wage effects in male-dominated occupations. In male documented occupations, lesbian and bisexual women who have lower levels of AGT (and thus are more gender nonconforming and likely perceived as more masculine) earn higher wages than heterosexual women. Thus, masculine characteristics, regardless of sex or sexual orientation, are rewarded in male-dominated occupations. The evidence once again highlights how gender biases meaningfully affect labor market outcomes for women in general. Even though gender atypical lesbian and bisexual women disproportionately benefit from these biases in male-dominated occupations, the differential returns to AGT do not explain average labor market differentials, which remain large and significant.

Next, we consider heterogeneous effects by educational obtainment in Table 6. A sexual

[30] In alternative specifications, we also found qualitatively similar results when we allow the effect of AGT to vary by sexual orientation across the full distribution of the percent female in each occupation.

minority identity may motivate sexual minorities to invest more heavily in their education (Burn and Martell 2020). The increased education may reflect efforts by sexual minorities to sort into careers with less discrimination or an increased ability to manage the selective disclosure of their sexual orientation. The different occupations and careers selected by higher educated workers may reward gender typicality differently than the occupations and careers of less-educated workers. We investigate this heterogeneity by augmenting Equation 3 with interaction terms between the LGB indicator and whether or not an individual completed a bachelor's degree.

In Table 6, we find that allowing for heterogeneity by educational attainment has no effect on the labor market outcomes. For men on average, we do not find a differential effect of AGT by educational obtainment. There is some evidence that gay/bisexual men with a bachelor's degree may have a higher return to AGT in terms of wages. A one standard deviation increase in AGT for a gay or bisexual man with a bachelor's degree increases their wages by 14%. We do not find a differential effect for gay men by education in any of the other outcomes for men. For women, we observe similar results. We do not find any evidence of an effect of AGT that differs by educational attainment, nor do we find a differential effect of AGT on lesbian and bisexual women by education (Table 6). The results suggest there is very little heterogeneity by education, and this does not explain the null results found in our baseline estimation.

Finally, we consider heterogeneous effects based on whether individuals live in a state with employment protections based on sexual orientation. Previous work has shown significantly higher wages for gay men in states with employment protections (Klawitter 2011; Martell 2013a; Burn 2018) If employers are not allowed to discriminate based on sexual orientation, they may rely more on gender conformity when rewarding or punishing individuals. We investigate this heterogeneity by augmenting Equation 3 with interaction terms between the LGB indicator and whether or not an individual lives in a state with an employment nondiscrimination act.

In Table 7, we find that allowing for heterogeneity by employment protections has no effect on the labor market differences for gay and bisexual men. The gaps in earnings (column 1), wages (column 2), employment (column 3), and hours worked (column 4) are nearly identical to the baseline results in Table 2. We find only one difference for women. Lesbian and bisexual women in states with no employment protections receive lower wages if they are more gender typical. A one standard deviation increase in gender typicality for these women decreases their wages by 5%. Taken together, the results suggest there is very little heterogeneity in the effect of AGT by state employment protections.

Our estimates indicate that in many contexts, the effects of AGT are similar. This suggests that gender norms are pervasive, and there are few differences in gender norms across contexts. We find some evidence that norms are more important in situations where there are more individuals of the same sex, where the rewards and punishment for not conforming to gender norms may be more salient. However, the evidence in this section does not suggest that heterogeneous effects of AGT confound our baseline estimates for sexual orientation based differentials. The effect of AGT is felt equally regardless of sexual orientation, with few exceptions. The results, therefore, support gender conformity and gender norms being important for all individuals and a separate source of differentials in the labor market.

7 Conclusion

We leverage the detailed data in the Add Health surveys to construct a novel measure of adherence to gender typicality that has almost exclusively been used in public health research. The results provide convincing evidence that gender typicality, in addition to sex, is an important determinant of labor market outcomes in general.

Gender typicality is significantly correlated with labor market outcomes for men. A one standard deviation increase in AGT was associated with a 3% increase in annual income, a 2% increase in hourly wages, and a 0.24 hour per week increase in work. Gender typicality is less significantly correlated with labor market outcomes for women. A one standard deviation increase in AGT was associated with a 1% decrease in hourly wages and a 0.26 hour per week decrease in work. These effects suggest an important role for employment policies promoting equality by gender. The remuneration of gender typicality for men may reflect the impact of implicit biases towards feminine characteristics, which may be more difficult to detect and remedy than outright animus. The evidence that gender-typical men earn more than gender-atypical men suggests that these biases may disadvantage men in addition to women. More evidence is needed on the outcomes of gender-atypical individuals and the discrimination they may face in the labor market.

These findings provide early evidence that the Fleming et al. (2017) measure of gender typicality has valid empirical uses in the social sciences. Our measure of gender typicality reproduces findings from controlled experiments. We show that sexual minorities are more likely to be gender atypical than their heterosexual counterparts, that gay men exhibit more gender atypicality than lesbian women, and that women have a wider range of behavior than men. Importantly, these patterns persist as individuals age into adulthood.

In arriving at these results, we find that the significant differences in the labor market outcomes of sexual minorities previously documented in the Add Health persist into later adulthood. In Waves III to V of the Add Health, bisexual men earn approximately 25% less annually and 11% less hourly than heterosexual men. The smaller hourly wage differential in part reflects that bisexual men work fewer hours than heterosexual men. Gay men experience similar differences in hourly earnings, though the annual earnings differential is smaller. Differences among women are less pronounced. Lesbian and bisexual women earn approximately 5% less annually and per hour than heterosexual women – though this differential is driven by the robustly significant negative experience of bisexuals. These women are slightly (1%) less likely to be employed. The evidence that gender typicality does not explain labor market differentials survived many additional sub-sample analyses, as well as alternative empirical specifications.

Our methodological approach broadens the discussion of discrimination in economics by focusing on gender as opposed to sex when examining labor market outcomes. We show that gendered norms of behavior are enforced in the workplace and that deviating from them can impact labor market outcomes. There appears to be a bias towards male norms of behavior on the part of employers, such that men who conform less to male gendered norms of behavior are penalized while women who conform less to female gendered norms of behavior are rewarded. These associations are strongest among men. Due to the nature of the Add Health data, we are unable to explore why employers may value gender conformity among workers.

Stronger norms of gender conformity reduce diversity in the workplace as individuals who do not conform are penalized. This lack of diversity can impact the allocation of skills across firms and lead to an inefficient allocation of labor. It also can reduce the productivity of firms since diverse teams are more productive and employees have higher job satisfaction (Bourke 2016; Griffith and Dasgupta 2018; Rock and Grant 2016).

Future work on the outcomes of gender-atypical individuals will contribute to a better understanding of the manifestation of the source of gender differentials in general. Our results indicate that adherence to gender-typical norms is associated with positive labor market outcomes for men, more hours worked for women, and possibly, but not robustly observed, higher wages. These patterns suggest a potentially unexplored explanation that will contribute to understanding variation in the size of gender gaps observed in different populations. Future work should investigate other margins where gender typicality may be more salient for sexual minorities, many of which will evolve over time as the Add Health cohort ages. They may include the impact of gender typicality on educational outcomes, cohabitation, occupational attainment, promotions, and wage trajectories.

References

Ahmed, Ali M, Lina Andersson, and Mats Hammarstedt, "Are gay men and lesbians discriminated against in the hiring process?," *Southern Economic Journal*, 2013, *79* (3), 565–585.

Aksoy, Cevat Giray, Christopher S. Carpenter, Jeff Frank, and Matt L. Huffman, "Gay glass ceilings: Sexual orientation and workplace authority in the UK," *Journal of Economic Behavior & Organization*, 2019, *159*, 167 – 180.

Antecol, Heather, Anneke Jong, and Michael Steinberger, "The sexual orientation wage gap: The role of occupational sorting and human capital," *Industrial and Labor Relations Review*, 2008, *61* (4), 518–543.

Badgett, M V Lee, "The wage effects of sexual orientation discrimination," *Industrial and Labor Relations Review*, June 1995, *48* (4), 726–739.

Baumle, Amanda K. and Dudley L. Poston Jr., "The economic cost of homosexuality: Multilevel analyses," *Social Forces*, 2011, *89* (3), 1005–1031.

Bayrakdar, Sait and Andrew King, "Job Satisfaction and Sexual Orientation in Britain," *Work, Employment and Society*, 2021, p. 0950017020980997.

Bem, Sandra L., "The measurement of psychological androgyny.," *Journal of Consulting and Clinical Psychology*, 1974, *42* (2), 155–162.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan, "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 2015, *130* (2), 571–614.

Blandford, John M., "The Nexus of Sexual Orientation and Gender in the Determination of Earnings," *Industrial and labor relations review*, 2003, *56* (4), 622–643.

Blashill, Aaron J and Kimberly K Powlishta, "The impact of sexual orientation and gender role on evaluations of men.," *Psychology of Men & Masculinity*, 2009, *10* (2), 160.

Blau, Francine and Lawrence Kahn, "The gender wage gap: Extent, trends, and explanations," *Journal of Economic Literature*, 2017, *55* (3).

Bourke, J., *Which Two Heads Are Better Than One?: How Diverse Teams Create Breakthrough Ideas and Make Smarter Decisions*, Australian Institute of Company Directors, 2016.

Bowles, Hannah Riley, Linda Babcock, and Lei Lai, "Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask," *Organizational Behavior and human decision Processes*, 2007, *103* (1), 84–103.

Bozani, Vasiliki, "Masculinity, Femininity, and Workplace Outcomes," *Handbook of Labor, Human Resources and Population Economics*, 2020, pp. 1–27.

Burn, Ian, "Not all laws are created equal: legal differences in state non-discrimination laws and the impact of LGBT employment protections," *Journal of Labor Research*, 2018, *39* (4), 462–497.

—, "The relationship between prejudice and wage penalties for gay men in the United States," *ILR Review*, 2020, *73* (3), 650–675.

— **and Michael E. Martell**, "The role of work values and characteristics in the human capital investment of gays and lesbians," *Education Economics*, 2020, *0* (0), 1–19.

Buser, Thomas, Lydia Geijtenbeek, and Erik Plug, "Sexual orientation, competitiveness and income," *Journal of Economic Behavior and Organization*, 2018, *151*, 191–198.

Butler, Judith, "Gender Trouble (London and New York," 1990.

Carpenter, Christopher S., "Revisiting the income penalty for behaviorally gay men: Evidence from NHANES III," *Labour Economics*, January 2007, *14* (1), 25–34.

—, "Sexual orientation, income, and non-pecuniary economic outcomes: New evidence from young lesbians in Australia," *Review of Economics of the Household*, 2008, *6* (4), 391–408.

— **and Samuel T. Eppink**, "Does it get better? Recent estimates of sexual orientation and earnings in the United States," *Southern Economic Journal*, 2017, *84* (2), 426–441.

Clarke, Heather M. and Kara A. Arnold, "The influence of sexual orientation on the perceived fit of male applicants for both male- and female-typed jobs," *Frontiers in Psychology*, may 2018, *9*, 656.

Cleveland, H Harrington, J Richard Udry, and Kim Chantala, "Environmental and genetic influences on sex-typed behaviors and attitudes of male and female adolescents," *Personality and Social Psychology Bulletin*, dec 2001, *27* (12), 1587–1598.

Cushing-Daniels, Brendan and Tsz-Ying Yeung, "Wage penalties and sexual orientation: An update using the General Social Survey," *Contemporary Economic Policy*, April 2009, *27* (2), 164–175.

Daneshvary, Nasser, C. Jeffrey Wassoups, and Bradley S. Wimmer, "Previous Marriage and the Lesbian Wage Premium," *Industrial Relations: A Journal of Economy and Society*, 2009, *48* (3), 432–453.

Drydakis, Nick, "Sexual Orientation Discrimination in the Labour Market," *Labour Economics*, 2009, *16*, 364–372.

—**, Katerina Sidiropoulou, Vasiliki Bozani, Sandra Selmanovic, and Swetketu Patnaik**, "Masculine vs feminine personality traits and womens employment outcomes in Britain: a field experiment," *International Journal of Manpower*, 2018.

Elmslie, Bruce and Edinaldo Tebaldi, "The Wage Gap against Gay Men: The Leveling of the Playing Field," *Kyklos*, 2014, *67* (3), 330–345.

Fleming, Paul J, Kathleen Mullan Harris, and Carolyn Tucker Halpern, "Description and evaluation of a measurement technique for assessment of performing gender," *Sex Roles*, 2017, *76* (11-12), 731–746.

Fortin, Nicole M, "Gender role attitudes and the labour-market outcomes of women across OECD countries," *Oxford Review of Economic Policy*, 2005, *21* (3), 416–438.

Goldin, Claudia, "A pollution theory of discrimination: male and female differences in occupations and earnings," in "Human capital in history: The American record," University of Chicago Press, 2014, pp. 313–348.

Gorsuch, Marina Mileo, "Gender, sexual orientation, and behavioral norms in the labor market," *ILR Review*, 2019, *72* (4), 927–954.

Griffith, Eric E and Nilanjana Dasgupta, "How the demographic composition of academic science and engineering departments influences workplace culture, faculty experience, and retention risk," *Social Sciences*, 2018, *7* (5), 71.

Harris, Kathleen Mullan and J Richard Udry, "National longitudinal study of adolescent to adult health (add health), 1994-2008 [Public Use]," *Ann Arbor, MI: Carolina Population Center, University of North Carolina-Chapel Hill [distributor], Inter-university Consortium for Political and Social Research [distributor]*, 2018, pp. 08–06.

Heilman, Madeline E, Aaron S Wallen, Daniella Fuchs, and Melinda M Tamkins, "Penalties for success: reactions to women who succeed at male gender-typed tasks.," *Journal of applied psychology*, 2004, *89* (3), 416.

— **and Julie J Chen**, "Same behavior, different consequences: reactions to men's and women's altruistic citizenship behavior.," *Journal of Applied Psychology*, 2005, *90* (3), 431.

Jepsen, Christopher and Lisa Jepsen, "Convergence Over Time or Not? US Wages by Sexual Orientation, 2000-2019," *Labour Economics*, 2021, p. 102086.

Jepsen, Lisa K., "Comparing the earnings of cohabiting lesbians, cohabiting heterosexual women, and married women: Evidence from the 2000 Census," *Industrial Relations: A Journal of Economy and Society*, 2007, *46* (4), 699–727.

Klawitter, Marieka, "Multilevel analysis of the effects of antidiscrimination policies on earnings by sexual orientation," *Journal of Policy Analysis and Management*, 2011, *30* (2), 334–358.

Klawitter, Marieka M., "Meta-analysis of the effects of sexual Orientation on earnings," *Industrial Relations*, 2015, *54* (1), 4–32.

Laumann, Edward O, John H Gagnon, Robert T Michael, and Stuart Michaels, *The social organization of sexuality: Sexual practices in the United States*, University of Chicago press, 2000.

Lehavot, Keren and Alan J. Lambert, "Toward a greater understanding of antigay prejudice: On the role of sexual orientation and gender role violation," *Basic and Applied Social Psychology*, aug 2007, *29* (3), 279–292.

Lippa, Richard A, "Gender-related traits in gay men, lesbian women, and heterosexual men and women: The virtual identity of homosexual-heterosexual diagnosticity and gender diagnosticity," *Journal of Personality*, 2000, *68* (5), 899–926.

Lippa, Richard and Sharon Connelly, "Gender diagnosticity: A new Bayesian approach to gender-related individual differences.," *Journal of Personality and Social Psychology*, 1990, *59* (5), 1051.

Lowry, Richard, Michelle M. Johns, Allegra R. Gordon, S. Bryn Austin, Leah E. Robin, and Laura K. Kann, "Nonconforming Gender Expression and Associated Mental Distress and Substance Use Among High School Students," *JAMA Pediatrics*, nov 2018, *172* (11), 1020.

Mahalik, James R, Caitlin McPherran Lombardi, Jacqueline Sims, Rebekah Levine Coley, and Alicia Doyle Lynch, "Gender, male-typicality, and social norms predicting adolescent alcohol intoxication and marijuana use," *Social Science & Medicine*, 2015, *143*, 71–80.

__, Shaun M Burns, and Matthew Syzdek, "Masculinity and perceived normative health behaviors as predictors of men's health behaviors," *Social science & medicine*, 2007, *64* (11), 2201–2209.

Martell, Michael E, "Differences do not matter: Exploring the wage gap for same-sex behaving men," *Eastern Economic Journal*, 2013, *39* (1), 45–71.

__, "Do ENDAs end discrimination for behaviorally gay men?," *Journal of Labor Research*, November 2013, *34*, 147–169.

__, "Age and the new lesbian earnings penalty," *International Journal of Manpower*, 2019.

__ **and Leanne Roncolato**, "The homosexual lifestyle: time use in same-sex households," *Journal of Demographic Economics*, 2016, *82* (4), 365–398.

__ **and Mary Eschelbach Hansen**, "Sexual identity and the lesbian earnings differential in the US," *Review of Social Economy*, 2017, *75* (2), 159–180.

Mize, Trenton D, "Sexual orientation in the labor market," *American Sociological Review*, 2016, *81* (6), 1132–1160.

Murphy, Kevin M and Robert H Topel, "Estimation and inference in two-step econometric models," *Journal of Business & Economic Statistics*, 2002, *20* (1), 88–97.

Rock, David and Heidi Grant, "Why diverse teams are smarter," *Harvard Business Review*, 2016, *4* (4), 2–5.

Rudman, Laurie A and Peter Glick, "Prescriptive gender stereotypes and backlash toward agentic women," *Journal of social issues*, 2001, *57* (4), 743–762.

Sabia, Joseph J., "Sexual orientation and wages in young adulthood: New evidence from AddHealth," *Industrial and Labor Relations Review*, 2014, *67* (1), 239–267.

— , "Fluidity in sexual identity, unmeasured heterogeneity, and the earnings effects of sexual orientation," *Industrial Relations: A Journal of Economy and Society*, 2015, *54* (1), 33–58.

Shakya, Holly B, Ben Domingue, Jason M Nagata, Beniamino Cislaghi, Ann Weber, and Gary L Darmstadt, "Adolescent gender norms and adult health outcomes in the USA: a prospective cohort study," *The Lancet Child & Adolescent Health*, 2019.

Steffens, Melanie C., Claudia Niedlich, Rosa Beschorner, and Maren C. Köhler, "Do positive and negative stereotypes of gay and heterosexual men affect job-related impressions?," *Sex Roles*, sep 2018, pp. 1–17.

Tilcsik, Andras, "Pride and prejudice: Employment discrimination against openly gay men in the United States," *American Journal of Sociology*, September 2011, *117* (2), 586–626.

Udry, J Richard and Kim Chantala, "Masculinity-Femininity Guides Sexual Union Formation in Adolescents," *Personality and Social Psychology Bulletin*, jan 2004, *30* (1), 44–55.

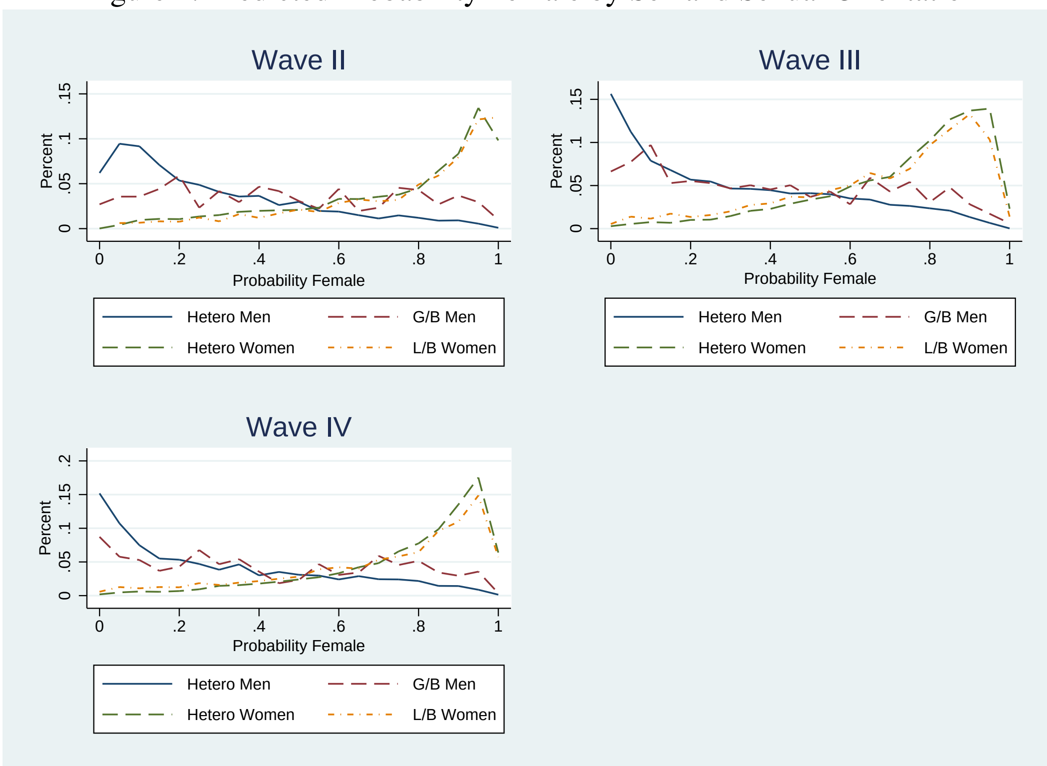
— **and** — , "Masculinity–femininity predicts sexual orientation in men but not in women," *Journal of Biosocial Science*, 2006, *38* (6), 797–809.

Weichselbaumer, Doris, "Sexual Orientation Discrimination in Hiring," *Labour Economics*, 2003, *10*, 629–642.

West, Candace and Don H Zimmerman, "Doing gender," *Gender & society*, 1987, *1* (2), 125–151.

Wilkinson, Andra L, Paul J Fleming, Carolyn Tucker Halpern, Amy H Herring, and Kathleen Mullan Harris, "Adherence to gender-typical behavior and high-frequency substance use from adolescence into young adulthood.," *Psychology of men & masculinity*, 2018, *19* (1), 145.

Figure 1: Predicted Probability Female by Sex and Sexual Orientation



Note: Authors' calculations from Add Health waves III, IV and V (Harris and Udry 2018). For each respondent, predicted probability is based on Fleming et al. (2017). See section 4.2 for further details. Tables A2 to A4 report the variables used in each wave.

Table 1: Descriptive Statistics of the Add Health Sample

	Men			Women		
	Gay	Bisexual	Heterosexual	Lesbian	Bisexual	Heterosexual
Outcomes						
Annual Income	42337.7	43282.2	46568.4	32850.5	36155.0	34918.1
	(37779.4)	(43418.9)	(52000.7)	(24666.7)	(42736.1)	(40631.2)
Hourly wages	19.74	21.47	21.30	16.22	19.32	18.45
	(16.23)	(23.65)	(24.24)	(11.34)	(27.36)	(22.61)
Employed	0.90	0.88	0.86	0.79	0.83**	0.85
	(0.31)	(0.33)	(0.35)	(0.41)	(0.38)	(0.36)
Hours worked per week	41.66**	41.04***	43.73	40.14*	38.63*	38.16
	(10.60)	(12.81)	(12.19)	(9.99)	(11.67)	(11.17)
Demographics						
Age	30.82***	29.99	29.54	31.13***	30.04***	29.53
	(5.76)	(5.89)	(5.82)	(5.42)	(5.73)	(6.01)
HS graduate	0.08***	0.09***	0.15	0.19***	0.09	0.08
	(0.27)	(0.28)	(0.36)	(0.39)	(0.28)	(0.28)
College graduate	0.22	0.25***	0.19	0.18	0.21	0.21
	(0.42)	(0.44)	(0.39)	(0.38)	(0.41)	(0.41)
Graduate school	0.18***	0.14***	0.10	0.12	0.16	0.17
	(0.38)	(0.35)	(0.30)	(0.32)	(0.37)	(0.37)
Peabody score	61.26***	67.45***	54.02	45*	58.65***	50.39
	(27.93)	(28.03)	(28.14)	(30.34)	(28.81)	(29.03)
White	0.60***	0.76***	0.69	0.52***	0.72***	0.64
	(0.49)	(0.43)	(0.46)	(0.50)	(0.45)	(0.48)
Black	0.18	0.13*	0.16	0.32**	0.17***	0.23
	(0.39)	(0.34)	(0.37)	(0.47)	(0.38)	(0.42)
Asian	0.14***	0.05**	0.08	0.04	0.06	0.07
	(0.35)	(0.22)	(0.26)	(0.20)	(0.24)	(0.25)
Other race	0.10	0.08	0.09	0.14**	0.08	0.08
	(0.30)	(0.27)	(0.28)	(0.35)	(0.27)	(0.28)
Hispanic	0.22***	0.15	0.16	0.18	0.13**	0.15
	(0.42)	(0.36)	(0.36)	(0.38)	(0.34)	(0.36)
Observations	202	488	8926	102	2020	9419

Note: Authors' calculations based on Add Health waves III, IV, and V (Harris and Udry 2018). Means and standard deviations are reported in parentheses. See Appendix Table A1 for the breakdown of detailed sexual orientation by wave.
Statistically significant difference relative to heterosexual counterparts at $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Table 2: The Effect of Controlling for AGT on Sexual Orientation Based Labor Market Differentials

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Men	Income	Income	Income	Wages	Wages	Wages	Employed	Employed	Employed	Hours	Hours	Hours
Gay	-0.08	-0.06	-0.13	-0.13**	-0.12**	-0.10*	0.02	0.02	0.04	-1.19	-0.99	-1.85*
	(0.10)	(0.10)	(0.13)	(0.05)	(0.05)	(0.06)	(0.02)	(0.02)	(0.03)	(0.87)	(0.84)	(1.00)
Bisexual	-0.26***	-0.25***	-0.25***	-0.10***	-0.10***	-0.11***	0.00	0.00	0.00	-2.12***	-2.04***	-2.16***
	(0.07)	(0.07)	(0.07)	(0.03)	(0.04)	(0.04)	(0.01)	(0.01)	(0.01)	(0.59)	(0.59)	(0.61)
Lagged AGT		0.03**	0.03*		0.02**	0.02***		-0.00	-0.00		0.25**	0.30**
		(0.01)	(0.01)		(0.01)	(0.01)		(0.00)	(0.00)		(0.12)	(0.13)
Gay × Lagged AGT			-0.09			0.02			0.03			-1.12*
			(0.09)			(0.04)			(0.02)			(0.65)
Bisexual × Lagged AGT			0.03			-0.03			-0.01			-0.41
			(0.07)			(0.03)			(0.01)			(0.48)
Adj. R Squared	0.374	0.375	0.375	0.582	0.582	0.582	0.430	0.430	0.430	0.147	0.147	0.147
N	10630	10630	10630	9616	9616	9616	11257	11257	11257	9616	9616	9616
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Women	Income	Income	Income	Wages	Wages	Wages	Employed	Employed	Employed	Hours	Hours	Hours
Lesbian	0.05	0.04	0.04	-0.03	-0.04	-0.07	-0.01	-0.01	-0.03	1.00	0.83	0.52
	(0.09)	(0.09)	(0.09)	(0.06)	(0.06)	(0.06)	(0.03)	(0.03)	(0.04)	(0.96)	(0.97)	(1.16)
Bisexual	-0.06*	-0.06*	-0.06*	-0.05***	-0.05***	-0.05***	-0.01*	-0.01*	-0.01*	0.18	0.15	0.15
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.33)	(0.33)	(0.33)
Lagged AGT		-0.01	-0.00		-0.01**	-0.01		-0.00	0.00		-0.26**	-0.27**
		(0.01)	(0.01)		(0.01)	(0.01)		(0.00)	(0.00)		(0.11)	(0.13)
Lesbian × Lagged AGT			-0.01			-0.05			-0.03			-0.42
			(0.07)			(0.05)			(0.03)			(0.91)
Bisexual × Lagged AGT			-0.05			-0.02			-0.01			0.08
			(0.03)			(0.02)			(0.01)			(0.31)
Adj. R Squared	0.330	0.330	0.330	0.579	0.579	0.579	0.496	0.496	0.496	0.120	0.120	0.120
N	13171	13171	13171	11541	11541	11541	14514	14514	14514	11541	11541	11541

Note: Authors' calculations based on Add Health waves III, IV, and V (Harris and Udry 2018). Adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects. Standard errors are reported in parentheses and have been clustered at the school level.
* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table 3: Inability of AGT to Explain Differentials is Robust to Estimation on Fully Dis-aggregated Sexual Orientation Responses for Men

	(1) Income	(2) Income	(3) Income	(4) Wages	(5) Wages	(6) Wages	(7) Employed	(8) Employed	(9) Employed	(10) Hours	(11) Hours	(12) Hours
Lagged AGT	0.03*	0.03*	0.03**	0.02***	0.02***	0.02***	0.00	0.00	0.00	0.30**	0.30**	0.31**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.13)	(0.13)	(0.13)
Homosexual/Bisexual	-0.21***			-0.11***			0.01			-2.03***		
	(0.06)			(0.03)			(0.01)			(0.56)		
Homosexual		-0.13			-0.10*			0.04			-1.85*	
		(0.13)			(0.06)			(0.03)			(1.00)	
Bisexual		-0.25***			-0.11***			0.00			-2.16***	
		(0.07)			(0.04)			(0.01)			(0.61)	
Mostly Heterosexual			-0.27***			-0.12***			0.00			-2.11***
			(0.08)			(0.04)			(0.01)			(0.69)
Bisexual			-0.58**			-0.24***			0.00			-2.01
			(0.25)			(0.07)			(0.03)			(1.89)
Mostly Homosexual			0.02			0.06			0.00			-1.38
			(0.15)			(0.09)			(0.04)			(1.73)
Homosexual			-0.13			-0.10*			0.04			-1.85*
			(0.13)			(0.06)			(0.03)			(1.00)
Homosexual/Bisexual × Lagged AGT	-0.02			-0.01			0.00			-0.68*		
	(0.06)			(0.02)			(0.01)			(0.40)		
Homosexual × Lagged AGT		-0.09			0.02			0.03			-1.12*	
		(0.09)			(0.04)			(0.02)			(0.65)	
Bisexual × Lagged AGT		0.03			-0.03			-0.01			-0.41	
		(0.07)			(0.03)			(0.01)			(0.48)	
Mostly Heterosexual × Lagged AGT			0.14			-0.01			-0.02			-1.23**
			(0.10)			(0.03)			(0.01)			(0.54)
Bisexual × Lagged AGT			-0.43**			-0.08			-0.01			2.35
			(0.19)			(0.06)			(0.02)			(1.54)
Mostly Homosexual × Lagged AGT			-0.01			-0.02			0.02			1.03
			(0.10)			(0.07)			(0.03)			(1.62)
Homosexual × Lagged AGT			-0.09			0.02			0.03			-1.13*
			(0.09)			(0.04)			(0.02)			(0.65)
Adj. R Squared	0.37	0.38	0.38	0.58	0.58	0.58	0.43	0.43	0.43	0.147	0.147	0.148
N	10630	10630	10630	9616	9616	9616	11257	11257	11257	9616	9616	9616

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects. Standard errors are reported in parentheses and have been clustered at the school-level.

* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table 4: Inability of AGT to Explain Differentials is Robust to Estimation on Fully Dis-aggregated Sexual Orientation Responses for Women

	(1) Income	(2) Income	(3) Income	(4) Wages	(5) Wages	(6) Wages	(7) Emp	(8) Emp	(9) Emp	(10) Hours	(11) Hours	(12) Hours
Lagged AGT	0.00	0.00	0.00	-0.01	-0.01	-0.01*	0.00	0.00	0.00	-0.27**	-0.27**	-0.28**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.13)	(0.13)	(0.13)
Homosexual/Bisexual	-0.06*			-0.05***			-0.01**			0.18		
	(0.03)			(0.02)			(0.01)			(0.32)		
Homosexual		0.04			-0.07			-0.03			0.52	
		(0.09)			(0.06)			(0.04)			(1.16)	
Bisexual		-0.06*			-0.05***			-0.01*			0.15	
		(0.03)			(0.02)			(0.01)			(0.33)	
Mostly Heterosexual			-0.07**			-0.04**			-0.01			-0.07
			(0.04)			(0.02)			(0.01)			(0.34)
Bisexual			-0.04			-0.16***			-0.05**			1.55
			(0.09)			(0.04)			(0.02)			(0.88)
Mostly Homosexual			0			-0.02			-0.03			0.28
			(0.13)			(0.06)			(0.03)			(1.29)
Homosexual			0.04			-0.07			-0.03			0.5
			(0.09)			(0.06)			(0.04)			(1.16)
Homosexual/Bisexual× Lagged AGT	-0.05*			-0.02			-0.01			0.03		
	(0.03)			(0.02)			(0.01)			(0.30)		
Homosexual × Lagged AGT		-0.01			-0.05			-0.03			-0.42	
		(0.07)			(0.05)			(0.03)			(0.91)	
Bisexual × Lagged AGT		-0.05			-0.02			-0.01			0.08	
		(0.03)			(0.02)			(0.01)			(0.31)	
Mostly Heterosexual × Lagged AGT			-0.04			-0.02			-0.01			0.26
			(0.03)			(0.02)			(0.01)			(0.35)
Bisexual × Lagged AGT			-0.05			0.03			0.00			-0.55
			(0.07)			(0.03)			(0.02)			(0.64)
Mostly Homosexual × Lagged AGT			-0.11			-0.06			-0.03**			-0.39
			(0.09)			(0.04)			(0.02)			(0.82)
Homosexual × Lagged AGT			-0.01			-0.05			-0.03			-0.41
			(0.07)			(0.05)			(0.03)			(0.92)
Adj. R Squared	0.33	0.33	0.33	0.58	0.58	0.58	0.5	0.5	0.5	0.12	0.12	0.12
N	13171	13171	13171	11541	11541	11541	14514	14514	14514	11541	11541	11541

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects. Standard errors are reported in parentheses and have been clustered at the school-level.
* p<0.10 ** p<0.05 *** p<0.01

Table 5: Effect of AGT By Occupation Gender Composition

Men	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual	-0.18***	-0.11***	0.02	-1.75***
	(0.06)	(0.03)	(0.01)	(0.56)
High Female Occ. × Lagged AGT	-0.04	0.01	-0.01	0.20
	(0.03)	(0.02)	(0.01)	(0.25)
Gay/Bisexual × High Female Occ. × Lagged AGT	0.11	0.03	-0.01	0.42
	(0.09)	(0.03)	(0.01)	(0.54)
Gender Neutral Occ. × Lagged AGT	0.04**	0.02	-0.01	0.09
	(0.02)	(0.01)	(0.01)	(0.19)
Gay/Bisexual× Gender Neutral Occ. × Lagged AGT	-0.10	-0.01	-0.00	-0.60
	(0.07)	(0.03)	(0.01)	(0.57)
Low Female Occ. × Lagged AGT	0.04*	0.01	-0.00	0.52***
	(0.02)	(0.01)	(0.00)	(0.19)
Gay/Bisexual × Low Female Occ. × Lagged AGT	0.05	0.01	-0.03	-0.09
	(0.09)	(0.05)	(0.02)	(0.99)
Adj. R Squared	0.348	0.581	0.173	0.148
N	9486	9481	9896	9481

Women	(1) Income	(2) Wages	(3) Employed	(4) Hours
Lesbian/Bisexual	-0.07**	-0.05***	-0.01*	0.15
	(0.03)	(0.02)	(0.01)	(0.32)
High Female Occ. × Lagged AGT	-0.02	-0.01	0.01	-0.46**
	(0.02)	(0.01)	(0.00)	(0.18)
Lesbian/Bisexual × High Female Occ. × Lagged AGT	-0.05	-0.02	-0.02*	0.16
	(0.04)	(0.02)	(0.01)	(0.27)
Gender Neutral Occ. × Lagged AGT	0.02	-0.02	-0.01	-0.02
	(0.02)	(0.01)	(0.01)	(0.17)
Lesbian/Bisexual × Gender Neutral Occ. × Lagged AGT	-0.00	-0.01	-0.00	-0.54
	(0.03)	(0.02)	(0.01)	(0.42)
Low Female Occ. × Lagged AGT	0.04	-0.01	0.01	-0.20
	(0.06)	(0.03)	(0.01)	(0.33)
Lesbian/Bisexual × Low Female Occ. × Lagged AGT	-0.14**	-0.11***	-0.00	-0.63
	(0.06)	(0.04)	(0.02)	(0.73)
Adj. R Squared	0.321	0.579	0.239	0.120
N	11449	11443	12297	11443

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects.

* p<0.10 ** p<0.05 *** p<0.01

Table 6: Effect of AGT By College Education

Men	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual	-0.21***	-0.11***	0.01	-2.03***
	(0.06)	(0.03)	(0.01)	(0.56)
Lagged AGT	0.03*	0.01*	-0.00	0.21
	(0.02)	(0.01)	(0.00)	(0.14)
Gay/Bisexual × Lagged AGT	-0.02	-0.04	-0.00	-0.57
	(0.06)	(0.03)	(0.01)	(0.48)
Lagged AGT × Bachelors	0.00	0.03	-0.01	0.48*
	(0.03)	(0.02)	(0.01)	(0.28)
Gay/Bisexual× Lagged AGT × Bachelors	-0.03	0.13***	-0.01	-0.10
	(0.10)	(0.04)	(0.02)	(0.75)
Adj. R Squared	0.374	0.583	0.430	0.147
N	10630	9616	11257	9616

Women	(1) Income	(2) Wages	(3) Employed	(4) Hours
Lesbian/Bisexual	-0.06*	-0.05***	-0.01**	0.19
	(0.03)	(0.02)	(0.01)	(0.32)
Lagged AGT	0.00	-0.01	-0.00	-0.26*
	(0.01)	(0.01)	(0.00)	(0.14)
Lesbian/Bisexual × Lagged AGT	-0.05	-0.02	-0.01	-0.12
	(0.03)	(0.02)	(0.01)	(0.30)
Lagged AGT × Bachelors	-0.04	-0.02	0.01*	-0.05
	(0.03)	(0.02)	(0.01)	(0.29)
Lesbian/Bisexual × Lagged AGT × Bachelors	-0.05	-0.02	-0.01	0.68
	(0.06)	(0.04)	(0.01)	(0.55)
Adj. R Squared	0.330	0.579	0.496	0.120
N	13171	11541	14514	11541

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects. The models have been fully specified, only coefficients of interest reported in this table for parsimony. Standard errors are reported in parentheses and have been clustered at the school-level.
* p<0.10 ** p<0.05 *** p<0.01

Table 7: Effect of AGT Interacted with Employment Nondiscrimination Acts (ENDA)

Men	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual	-0.25***	-0.13***	0.02	-1.98***
	(0.09)	(0.04)	(0.02)	(0.74)
Lagged AGT	0.09***	0.01	-0.00	0.70***
	(0.03)	(0.01)	(0.01)	(0.23)
Gay/Bisexual × No ENDA × Lagged AGT	-0.10	-0.01	0.02	-1.04
	(0.11)	(0.04)	(0.02)	(0.74)
Gay/Bisexual × ENDA × Lagged AGT	-0.04	-0.00	-0.02	-1.03
	(0.10)	(0.04)	(0.02)	(0.66)
Adj. R Squared	0.317	0.540	0.343	0.171
N	7629	6922	8226	6922

Women	(1) Income	(2) Wages	(3) Employed	(4) Hours
Lesbian/Bisexual	-0.07	-0.05**	-0.02*	0.18
	(0.05)	(0.02)	(0.01)	(0.34)
Lagged AGT	-0.02	-0.00	0.00	-0.00
	(0.03)	(0.01)	(0.01)	(0.20)
Lesbian/Bisexual × No ENDA× Lagged AGT	-0.01	-0.05*	-0.01	-0.30
	(0.06)	(0.03)	(0.01)	(0.55)
Lesbian/Bisexual × ENDA× Lagged AGT	-0.04	0.01	-0.02	0.09
	(0.05)	(0.02)	(0.02)	(0.37)
Adj. R Squared	0.280	0.551	0.351	0.156
N	8839	7964	10108	7964

Note: Authors' calculations based on Add Health waves III and IV (Harris and Udry 2018). The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects. Standard errors are reported in parentheses and have been clustered at the school-level.

* $p<0.10$ ** $p<0.05$ *** $p<0.01$

8 Appendix Tables and Figures

Table A1: Sexual Orientation Across Waves

A. Men	Wave III	Wave IV	Wave V
Heterosexual	93.4%	93.0%	91.0%
Mostly Heterosexual	3.0%	3.3%	4.7%
Bisexual	0.7%	0.5%	0.6%
Mostly Homosexual	0.7%	0.9%	0.8%
Homosexual	1.5%	2.0%	2.8%
Asexual	0.2%	0.1%	0.1%
B. Women	Wave III	Wave IV	Wave V
Heterosexual	86.4%	79.9%	79.6%
Mostly Heterosexual	9.9%	15.9%	16.1%
Bisexual	2.1%	2.0%	1.9%
Mostly Homosexual	0.7%	0.8%	1.0%
Homosexual	0.3%	1.1%	1.0%
Asexual	0.3%	0.2%	0.1%

Notes: Authors' calculations based on Add Health Waves III, IV, and V. Subjects are asked to voluntarily report which most accurately reflects their sexual orientation. Not included in this table is those preferring not to disclose. Source: Fleming et al. (2017).

Table A2: Wave II AGT Model Questions

Importance	Question
1	Frequency of crying
2	Frequency of sunbathing in the summer
3	Frequency of playing an active sport
4	How you think of yourself in terms of weight
5	Have you ever driven a car
6	Frequency of doing work around the house
7	Likely to use sunscreen
8	You like to take risks
9	Frequency of poor appetite
10	Difficult problems make you very upset
11	Hours per week playing video/computer games
12	How much do you feel that your friends care about you?
13	Past 12 months, how often get into a serious physical fight
14	You will graduate from college
15	You received testing/treatment for an STI/AIDS in past year
16	Past 12 months, how often deliberately damage property
17	Times used sunlamp or a tanning bed in your life
18	You like yourself just the way you are
19	You live without much thought for the future
20	Number of past thirty days chewed tobacco
21	Frequency of wearing a helmet while cycling
22	You felt you were just as good as other people
23	Frequency wearing a seatbelt in the car
24	How honestly answered questions
25	You felt lonely
26	Frequency of moodiness
27	You are emotional

Notes: Questions are ordered from most important to least important. See section 4.2 for a description of question selection. Source: Fleming et al. (2017).

Table A3: Wave III AGT Model Questions

Importance	Question
1	You were sad, during the past 7 days
2	In past 7 days, how many times doing work around the house
3	What do you think of yourself in terms of weight?
4	In past 7 days, how many times did you participate in gymnastics, weight lifting
5	I can do a good job stretching the truth when I talk to people
6	Hours per week playing video/computer games
7	How many times engage in a hobby (e.g. play cards, arts and crafts, musical, etc.)
8	In past 7 days, how many times did you rollerblade/ski/racquet sports or aerobics ?
9	You like to take risks
10	Have you used legal performance enhancing substances for athletes (i.e. creatine)
11	In past 7 days, how many times did you participate in strenuous team sport
12	Do you own a handgun?
13	Past 12 months, how often deliberately damage property that wasnt yours
14	In past 7 days, how many times hang with friends or talk on the telephone for more than 5 min?
15	Number of past thirty days chewed tobacco
16	Have you ever been expelled from school
17	Past 12 months, how often take part in physical fight where your group against another group
18	Have you ever paid someone to have sex with you?
19	Have you ever played games for money or taken part in another type of gambling for money?
20	In past 7 days, how many times did you walk for exercise
21	In past 7 days, how many times did you bike/skate/dance/skateboard
22	In past 7 days, how many times did you watch TV in the past seven days
23	How important is being faithful is for a successful marriage?

Notes: Questions are ordered from most important to least important. See section 4.2 for a description of question selection. Source: Fleming et al. (2017).

Table A4: Wave IV AGT Model Questions

Importance	Question
1	Have you ever used chewing tobacco at least 20 times in your entire life?
2	Have you ever been arrested?
3	When you go outside on a sunny day for more than one hour, how likely are you to use sunscreen or sunblock?
4	I have a vivid imagination
5	I dont talk a lot
6	I sympathize with others feelings
7	In the past seven days, how many times did you participate in gymnastics, weight lifting, or strength training?
8	Hours per week playing video/computer games
9	In past 7 days, how many times did you walk for exercise
10	I have frequent mood swings
11	How often do you pray privately?
12	During typical summer week, how many hours do you spend in the sun during the day?
13	Have you ever been in the military?
14	In the past 7 days, how many times did you participate in strenuous team sports such as football, soccer, basketball, lacrosse, rugby, field hockey, or ice hockey?
15	In the past 7 days, you felt too tired to do things.
16	I worry about things
17	Compared to other people your age, how intelligent are you?
18	I am not interested in other peoples problems
19	I like to take risks
20	In the past 24 h, have you participated in vigorous activity long enough to work up a sweat, get your heart thumping, or get out of breath?
21	I get stressed out easily
22	I am not really interested in others.

Notes: Questions are ordered from most important to least important. See section 4.2 for a description of question selection. Source Fleming et al. (2017).

Table A5: Effect of Controls on AGT

	(1) Men	(2) Women
LGB	-0.46***	-0.13***
	(0.04)	(0.02)
Age	-0.02	0.11***
	(0.03)	(0.03)
Age × Age	-0.00	-0.00**
	(0.00)	(0.00)
High School	-0.04	0.01
	(0.03)	(0.03)
Bachelors	-0.05	-0.06**
	(0.04)	(0.03)
Post-Bachelors	-0.12***	-0.06**
	(0.04)	(0.03)
Enrolled	-0.04	-0.02
	(0.03)	(0.02)
White	-0.05	0.27***
	(0.07)	(0.06)
Black/AA	0.13*	-0.12*
	(0.08)	(0.06)
Asian	-0.06	0.01
	(0.08)	(0.05)
Other Race	-0.02	0.22***
	(0.10)	(0.07)
Hispanic	-0.06	-0.02
	(0.05)	(0.04)
Currently Cohabiting	-0.01	0.04**
	(0.02)	(0.02)

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects. Standard errors are reported in parentheses and have been clustered at the school-level.

* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table A6: Predicted AGT Over Time

	(1) Men	(2) Women
LGB	-0.77***	0.10*
	(0.08)	(0.06)
Wave IV	-0.00	0.06
	(0.03)	(0.04)
Wave V	-0.05**	0.07**
	(0.03)	(0.03)
LGB × Wave IV	0.35***	-0.31***
	(0.09)	(0.07)
LGB × Wave V	0.41***	-0.34***
	(0.10)	(0.07)
Adj. R Squared	0.017	0.006
N	11257	14514

Note: Authors' calculations based on Add Health waves III, IV, and V (Harris and Udry 2018).The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, highschool fixed effects, and wave fixed effects.

* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table A7: Correlation Between AGT and Marriage Among Heterosexuals

	(1) Men	(2) Women
Lagged AGT	-0.01***	0.02***
	(0.01)	(0.01)
Adjusted R Squared	0.434	0.424
N	4820	5034

Note: Authors' calculations based on Add Health wave IV (Harris and Udry 2018). The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, and high-school fixed effects. Standard errors are reported in parentheses and have been clustered at the school-level.

* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table A8: Correlation Between AGT and Physical Appearance

	(1) Men	(2) Women
LGB	-0.54***	0.35***
	(0.07)	(0.04)
Lagged AGT	0.06***	-0.08***
	(0.01)	(0.02)
Adjusted R Squared	0.111	0.155
N	2955	4314

Note: Authors' calculations based on Add Health Wave V data (Harris and Udry 2018). The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Standard errors are reported in parentheses and have been clustered at the school-level.
* p<0.10 ** p<0.05 *** p<0.01

Table A9: Descriptive Statistics of the Add Health Sample, Men

	(1) Homosexual	(2) Mostly Homosexual	(3) Bisexual	(4) Mostly Heterosexual	(5) Heterosexual
Outcomes					
Annual income	42337.7	47251.1	26922.4**	44911.8	46568.4
	(37779.4)	(44047.6)	(21726.2)	(45311.7)	(52000.7)
Hourly wages	19.74	24.98	13.44**	21.90	21.30
	(16.23)	(28.85)	(9.662)	(23.60)	(24.24)
Employed	0.896	0.831	0.891	0.889	0.857
	(0.306)	(0.377)	(0.315)	(0.315)	(0.351)
Hours worked per week	41.66**	40.77**	39.24***	41.39***	43.73
	(10.60)	(12.50)	(15.68)	(12.40)	(12.19)
Demographics					
Age	30.82	29.45	29.16	30.25**	29.54
	(5.762)	(5.564)	(6.394)	(5.878)	(5.823)
HS graduate	0.0792	0.0843	0.0727	0.0886***	0.148
	(0.271)	(0.280)	(0.262)	(0.285)	(0.356)
College graduate	0.223	0.265*	0.182	0.260***	0.188
	(0.417)	(0.444)	(0.389)	(0.439)	(0.391)
Graduate school	0.178***	0.145	0.0364	0.154***	0.0989
	(0.384)	(0.354)	(0.189)	(0.362)	(0.299)
Peabody score	61.26***	68.02*	56.42	69.05***	54.02
	(27.93)	(27.54)	(32.58)	(27.06)	(28.14)
White	0.599***	0.602*	0.782	0.789***	0.692
	(0.491)	(0.492)	(0.417)	(0.409)	(0.462)
Black	0.183	0.301***	0.109	0.0971***	0.165
	(0.388)	(0.462)	(0.315)	(0.297)	(0.371)
Asian	0.139	0.0482	0	0.0600	0.0752
	(0.346)	(0.215)	(0)	(0.238)	(0.264)
Other race	0.0990***	0.0843	0.127	0.0686	0.0851
	(0.299)	(0.280)	(0.336)	(0.253)	(0.279)
Hispanic	0.223	0.217	0.164	0.137	0.155
	(0.417)	(0.415)	(0.373)	(0.344)	(0.362)
Observations	202	83	55	350	8926

Note: Authors' calculations based on Add Health waves III, IV, and V (Harris and Udry 2018). Means and standard deviations are reported in parentheses. See Appendix Table A1 for the breakdown of detailed sexual orientation by wave.
Statistically significant difference relative to heterosexual counterparts at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10: Descriptive Statistics of the Add Health Sample, Women

	(1) Homosexual	(2) Mostly Homosexual	(3) Bisexual	(4) Mostly Heterosexual	(5) Heterosexual
Outcomes					
Annual income	32850.5	36363.0	28325.3**	37210.0**	34918.1
	(24666.7)	(30021.7)	(26546.8)	(44960.7)	(40631.2)
Hourly wages	16.22	17.84	14.15***	20.10**	18.45
	(11.34)	(12.12)	(14.29)	(29.19)	(22.61)
Employed	0.794	0.849	0.797**	0.831	0.846
	(0.406)	(0.360)	(0.403)	(0.375)	(0.361)
Hours worked per week	40.14	38.88	39.65	38.48	38.16
	(9.998)	(11.90)	(12.45)	(11.54)	(11.17)
Demographics					
Age	31.13***	29.76	29.26	30.16***	29.53
	(5.424)	(5.696)	(6.051)	(5.678)	(6.014)
HS graduate	0.186***	0.0968	0.0823	0.0861	0.0833
	(0.391)	(0.297)	(0.275)	(0.281)	(0.276)
College graduate	0.176	0.226	0.139***	0.216	0.214
	(0.383)	(0.420)	(0.346)	(0.411)	(0.410)
Graduate school	0.118	0.118	0.130	0.166	0.167
	(0.324)	(0.325)	(0.337)	(0.372)	(0.373)
Peabody score	45*	55.30	57.20***	59.03***	50.39
	(30.34)	(29.60)	(31.81)	(28.33)	(29.03)
White	0.520***	0.677	0.684	0.725***	0.643
	(0.502)	(0.470)	(0.466)	(0.447)	(0.479)
Black	0.324**	0.161	0.216	0.167***	0.226
	(0.470)	(0.370)	(0.413)	(0.373)	(0.419)
Asian	0.0392	0.0538	0.0563	0.0619	0.0664
	(0.195)	(0.227)	(0.231)	(0.241)	(0.249)
Other race	0.137***	0.129	0.0909	0.0719	0.0825
	(0.346)	(0.337)	(0.288)	(0.258)	(0.275)
Hispanic	0.176	0.194	0.156	0.124***	0.151
	(0.383)	(0.397)	(0.363)	(0.330)	(0.358)
Observations	102	93	231	1696	9419

Note: Authors' calculations based on Add Health waves III, IV, and V (Harris and Udry 2018). Means and standard deviations are reported in parentheses. See Appendix Table A1 for the breakdown of detailed sexual orientation by wave.

Statistically significant difference relative to heterosexual counterparts at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A11: Effect of Restricting Definition of LGB to Always Identifiers

Men	(1) Income	(2) Income	(3) Income	(4) Wages	(5) Wages	(6) Wages	(7) Employed	(8) Employed	(9) Employed	(10) Hours	(11) Hours	(12) Hours
Always Gay/Bisexual	-0.15*	-0.13*	-0.07	-0.11**	-0.10**	-0.11*	0.04***	0.04**	0.05***	-1.43	-1.23	-2.05**
	(0.08)	(0.08)	(0.09)	(0.05)	(0.05)	(0.06)	(0.01)	(0.01)	(0.02)	(0.88)	(0.87)	(0.94)
Lagged AGT		0.03**	0.03*		0.02***	0.02***		-0.00	-0.00		0.27**	0.31**
		(0.01)	(0.01)		(0.01)	(0.01)		(0.00)	(0.00)		(0.12)	(0.12)
Always Gay/Bisexual × Lagged AGT			0.08			-0.01			0.02**			-1.13*
			(0.08)			(0.03)			(0.01)			(0.63)
Adj. R Squared	0.373	0.374	0.374	0.581	0.582	0.582	0.430	0.430	0.430	0.146	0.146	0.146
N	10630	10630	10630	9616	9616	9616	11257	11257	11257	9616	9616	9616

Women	(1) Income	(2) Income	(3) Income	(4) Wages	(5) Wages	(6) Wages	(7) Employed	(8) Employed	(9) Employed	(10) Hours	(11) Hours	(12) Hours
Always Lesbian/Bisexual	0.03	0.03	0.02	-0.04	-0.04	-0.04	0.01	0.01	0.00	0.69	0.62	0.64
	(0.05)	(0.05)	(0.06)	(0.03)	(0.03)	(0.03)	(0.01)	(0.01)	(0.01)	(0.50)	(0.50)	(0.52)
Lagged AGT		-0.01	-0.01		-0.01**	-0.02**		0.00	0.00		-0.26**	-0.26**
		(0.01)	(0.01)		(0.01)	(0.01)		(0.00)	(0.00)		(0.11)	(0.12)
Always Lesbian/Bisexual × Lagged AGT			-0.03			0.01			-0.02*			0.08
			(0.04)			(0.02)			(0.01)			(0.41)
Adj. R Squared	0.329	0.329	0.329	0.578	0.579	0.579	0.496	0.496	0.496	0.120	0.120	0.120
N	13171	13171	13171	11541	11541	11541	14514	14514	14514	11541	11541	11541

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). LGB individuals are those who recorded an LGB identity in each wave. The adherence to gender typicality (AGT) is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects. Standard errors are reported in parentheses and have been clustered at the school-level.
* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table A12: LGB Differentials After Varying Polynomial of AGT

Men	(1) Income	(2) Wages	(3) Employed	(4) Hours
1^{st} order polynomial of AGT	-0.21***	-0.11***	0.01	-2.03***
	(0.06)	(0.03)	(0.01)	(0.56)
2^{nd} order polynomial of AGT	-0.17**	-0.12***	-0.00	-2.59***
	(0.08)	(0.04)	(0.02)	(0.76)
3^{rd} order polynomial of AGT	-0.15	-0.09**	0.01	-2.84***
	(0.10)	(0.05)	(0.03)	(0.90)
4^{th} order polynomial of AGT	-0.14	-0.09**	0.02	-2.83***
	(0.10)	(0.05)	(0.03)	(0.92)

Women	(1) Income	(2) Wages	(3) Employed	(4) Hours
1^{st} order polynomial of AGT	-0.06*	-0.05***	-0.01**	0.18
	(0.03)	(0.02)	(0.01)	(0.32)
2^{nd} order polynomial of AGT	-0.08**	-0.07***	-0.01	-0.22
	(0.04)	(0.02)	(0.01)	(0.41)
3^{rd} order polynomial of AGT	-0.10*	-0.05*	0.01	-0.89*
	(0.05)	(0.03)	(0.01)	(0.49)
4^{th} order polynomial of AGT	-0.10**	-0.05*	0.01	-0.86*
	(0.05)	(0.03)	(0.02)	(0.49)

Note: Authors' calculations based on Add Health waves III, IV, and V (Harris and Udry 2018). Estimated coefficients report average labor market differentials between sexual minorities and heterosexuals. The first row corresponds to the baseline results. Each subsequent row increases the order of the polynomial of AGT. AGT is measured in the preceding wave using the method developed by Fleming et al. (2017). Drawing on Sabia (2014), controls used in the regression include race, educational obtainment, occupation, physical and health characteristics, religiosity, marital and cohabitation status, high-school fixed effects, and wave fixed effects. Standard errors are reported in parentheses and have been clustered at the school level.

* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table A13: Effect of AGT Across Waves

Men	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual × Wave III	-0.37	-0.09	-0.01	-1.53
	(0.23)	(0.07)	(0.01)	(1.36)
Gay/Bisexual × Wave IV	-0.21**	-0.14***	0.04	-2.07***
	(0.08)	(0.04)	(0.02)	(0.75)
Gay/Bisexual × Wave V	-0.11*	-0.07	-0.02**	-2.21***
	(0.06)	(0.04)	(0.01)	(0.76)
Wave III × Lagged AGT	0.06	0.02	0.00	0.62**
	(0.04)	(0.01)	(0.00)	(0.28)
Wave IV × Lagged AGT	0.03*	0.02**	-0.01*	0.19
	(0.01)	(0.01)	(0.01)	(0.18)
Wave V × Lagged AGT	-0.00	0.01	0.01**	0.27
	(0.02)	(0.01)	(0.00)	(0.25)
Gay/Bisexual × Wave III × Lagged AGT	0.00	-0.01	0.00	-0.91
	(0.16)	(0.06)	(0.01)	(0.90)
Gay/Bisexual × Wave IV × Lagged AGT	-0.05	-0.01	-0.00	-0.35
	(0.06)	(0.04)	(0.03)	(0.56)
Gay/Bisexual × Wave V × Lagged AGT	-0.05	-0.01	-0.01	-0.91
	(0.06)	(0.04)	(0.01)	(0.64)
Adj. R Squared	0.375	0.582	0.430	0.147
N	10630	9616	11257	9616

Women	(1) Income	(2) Wages	(3) Employed	(4) Hours
Lesbian/Bisexual × Wave III	-0.07	-0.04	-0.01	0.07
	(0.13)	(0.04)	(0.00)	(0.70)
Lesbian/Bisexual × Wave IV	-0.05	-0.06***	-0.02	0.17
	(0.04)	(0.02)	(0.01)	(0.37)
Lesbian/Bisexual × Wave V	-0.06	-0.06**	-0.01**	0.25
	(0.04)	(0.03)	(0.00)	(0.59)
Wave III× Lagged AGT	0.02	-0.00	0.00	0.21
	(0.04)	(0.01)	(0.00)	(0.27)
Wave IV × Lagged AGT	-0.01	-0.01	0.00	-0.42***
	(0.02)	(0.01)	(0.01)	(0.15)
Wave V × Lagged AGT	-0.02	-0.01	0.00	-0.40
	(0.02)	(0.01)	(0.00)	(0.26)
Lesbian/Bisexual × Wave III × Lagged AGT	-0.10	0.00	-0.00	-0.28
	(0.10)	(0.04)	(0.00)	(0.67)
Lesbian/Bisexual × Wave IV × Lagged AGT	-0.03	-0.01	-0.01	0.26
	(0.03)	(0.02)	(0.02)	(0.33)
Lesbian/Bisexual × Wave V × Lagged AGT	-0.05	-0.04*	-0.01*	-0.02
Adj. R Squared	0.330	0.579	0.496	0.120
N	13171	11541	14514	11541

Note: Authors' calculations based on Add Health waves III, IV, and V (Harris and Udry 2018). See Table 2 for a description of the data. In this table, we modify Equation 3 by interacting wave fixed effects with AGT.

* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table A14: Effect of Controlling for AGT Components Directly

Wave III	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual Differential	-0.30	-0.06	-0.00	-0.52
	(0.21)	(0.01)	(0.01)	(1.25)
Lesbian/Bisexual Differential	-0.07	-0.02	-0.01	-0.38
	(0.13)	(0.03)	(0.00)	(0.69)

Wave IV	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual Differential	-0.17***	-0.13***	0.04*	-2.21***
	(0.06)	(0.04)	(0.02)	(0.74)
Lesbian/Bisexual Differential	-0.04	-0.04	-0.02	0.04
	(0.03)	(0.02)	(0.02)	(0.37)

Wave V	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual Differential	-0.07	-0.06	-0.01**	-2.30***
	(0.06)	(0.04)	(0.01)	(0.77)
Lesbian/Bisexual Differential	-0.03	-0.05	-0.00	0.15
	(0.05)	(0.03)	(0.00)	(0.60)

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). See Tables A2 through A4 for the list of individuals controls included in each wave. Responses are lagged one wave behind outcomes.

* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table A15: Effect of AGT Estimated Using Panel Data

Men	(1) Income	(2) Wages	(3) Employed	(4) Hours
Lagged AGT	-0.00	-0.01	-0.01	0.08
	(0.02)	(0.01)	(0.00)	(0.18)
Gay/Bisexual × Lagged AGT	-0.04	0.03	0.01	-0.09
	(0.08)	(0.03)	(0.01)	(0.63)

Women	(1) Income	(2) Wages	(3) Employed	(4) Hours
Lagged AGT	-0.02	0.00	0.01**	-0.14
	(0.02)	(0.01)	(0.00)	(0.16)
Lesbian/Bisexual × Lagged AGT	-0.06	-0.01	-0.02*	0.27
	(0.04)	(0.02)	(0.01)	(0.52)

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018) using time varying controls listed in Equation 3.
* p<0.10 ** p<0.05 *** p<0.01

Table A16: Effect of AGT By Race

Men	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual	-0.22***	-0.11***	0.01	-2.09***
	(0.06)	(0.03)	(0.01)	(0.56)
Lagged AGT	0.04**	0.02**	-0.00	0.45***
	(0.02)	(0.01)	(0.00)	(0.17)
Gay/Bisexual × Lagged AGT	-0.00	-0.01	-0.00	-0.73
	(0.06)	(0.03)	(0.01)	(0.53)
Black/AA × Lagged AGT	-0.07**	-0.03**	0.00	-0.89**
	(0.03)	(0.02)	(0.01)	(0.34)
Gay/Bisexual × Black/AA × Lagged AGT	-0.33**	-0.13**	0.01	-0.53
	(0.14)	(0.06)	(0.02)	(1.21)
Asian × Lagged AGT	0.02	0.04*	0.01	-0.51
	(0.06)	(0.02)	(0.01)	(0.40)
Gay/Bisexual × Asian× Lagged AGT	-0.18	0.09	-0.01	-1.12
	(0.15)	(0.06)	(0.03)	(1.50)
Hispanic × Lagged AGT	-0.03	0.02	0.00	0.33
	(0.04)	(0.02)	(0.01)	(0.38)
Gay/Bisexual × Hispanic× Lagged AGT	0.30**	0.07*	0.00	1.32
	(0.12)	(0.04)	(0.02)	(0.95)
Other Race × Lagged AGT	0.01	-0.04	-0.00	-0.33
	(0.05)	(0.03)	(0.01)	(0.48)
Gay/Bisexual × Other Race × Lagged AGT	-0.34*	-0.11*	0.02	-1.89
	(0.17)	(0.06)	(0.02)	(1.24)
Adj. R Squared	0.375	0.583	0.429	0.147
N	10630	9616	11257	9616

Women	(1) Income	(2) Wages	(3) Employed	(4) Hours
Lesbian/Bisexual	-0.06*	-0.05***	-0.01**	0.18
	(0.03)	(0.02)	(0.01)	(0.32)
Lagged AGT	0.00	-0.00	-0.00	-0.32*
	(0.02)	(0.01)	(0.00)	(0.18)
Lesbian/Bisexual × Lagged AGT	-0.06*	-0.02	-0.00	-0.09
	(0.03)	(0.02)	(0.01)	(0.38)
Black/AA × Lagged AGT	0.01	-0.01	0.01	0.34
	(0.03)	(0.02)	(0.01)	(0.29)
Lesbian/Bisexual × Black/AA × Lagged AGT	-0.01	-0.02	0.00	0.43
	(0.07)	(0.03)	(0.02)	(0.52)
Asian × Lagged AGT	0.04	0.03	0.02	0.16
	(0.07)	(0.04)	(0.02)	(0.52)
Lesbian/Bisexual × Asian × Lagged AGT	-0.00	0.09	-0.02	-0.25
	(0.06)	(0.05)	(0.02)	(0.63)
Hispanic × Lagged AGT	-0.15***	-0.03	0.02*	-0.41
	(0.06)	(0.02)	(0.01)	(0.40)
Lesbian/Bisexual × Hispanic × Lagged AGT	0.01	-0.02	-0.00	0.46
	(0.08)	(0.06)	(0.02)	(0.79)
Other Race × Lagged AGT	0.08	-0.01	-0.01	0.18
	(0.07)	(0.03)	(0.02)	(0.45)
Lesbian/Bisexual × Other Race × Lagged AGT	0.06	-0.02	0.03	0.53
	(0.13)	(0.07)	(0.03)	(1.20)
Adj. R Squared	0.330	0.579	0.496	0.120
N	13171	11541	14514	11541

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). Results based on Equation 3 with additional interactions between race, sexual orientation, and AGT. See Table 1 for a description of the data and methodology.
* $p<0.10$ ** $p<0.05$ *** $p<0.01$

Table A17: Effect of AGT by Cohabitation Status

Men	(1) Income	(2) Wages	(3) Employed	(4) Hours
Gay/Bisexual	-0.21***	-0.11***	0.01	-2.03***
	(0.06)	(0.03)	(0.01)	(0.56)
Lagged AGT	0.05*	0.02**	-0.01***	0.60***
	(0.02)	(0.01)	(0.00)	(0.21)
Gay/Bisexual × Lagged AGT	-0.00	0.00	0.01	-0.86*
	(0.07)	(0.03)	(0.01)	(0.45)
Currently Cohabiting × Lagged AGT	-0.04	-0.00	0.02**	-0.53**
	(0.03)	(0.01)	(0.01)	(0.26)
Gay/Bisexual × Currently Cohabiting × Lagged AGT	-0.11	-0.04	-0.01	-0.29
	(0.08)	(0.04)	(0.02)	(0.69)
Adj. R Squared	0.375	0.582	0.430	0.147
N	10630	9616	11257	9616

Women	(1) Income	(2) Wages	(3) Employed	(4) Hours
Lesbian/Bisexual	-0.06*	-0.05***	-0.01**	0.17
	(0.03)	(0.02)	(0.01)	(0.32)
Lagged AGT	0.00	-0.01	-0.00	-0.07
	(0.02)	(0.01)	(0.00)	(0.19)
Lesbian/Bisexual × Lagged AGT	-0.03	-0.02	0.01	0.11
	(0.05)	(0.02)	(0.01)	(0.40)
Currently Cohabiting × Lagged AGT	-0.01	-0.01	0.00	-0.38
	(0.03)	(0.01)	(0.01)	(0.26)
Lesbian/Bisexual × Currently Cohabiting × Lagged AGT	-0.04	-0.00	-0.03***	-0.50
	(0.05)	(0.03)	(0.01)	(0.41)
Adj. R Squared	0.330	0.579	0.496	0.120
N	13171	11541	14514	11541

Note: Authors' calculations based on Add Health waves III, IV and V (Harris and Udry 2018). Results based on Equation 3 with additional interactions between cohabitation status, orientation, and AGT. See Table 1 for a description of the data and methodology.
* p<0.10 ** p<0.05 *** p<0.01