Three Essays on LGBT Economics and Policy

By

Samuel T. Eppink

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

August 9, 2019

Nashville, Tennessee

Approved:

Christopher S. Carpenter, Ph.D.

Andrew Dustan, Ph.D.

Gilbert Gonzales, Ph.D.

Federico H. Gutierrez, Ph.D.

ACKNOWLEDGEMENTS

First and foremost, I thank my family for their emotional support. I am especially grateful to my mother, Jane Eppink, for her encouragement and for creating many of the opportunities that made pursuing a Ph.D. possible. I thank my friends and fellow students for their feedback and support, particularly Caroline Abraham, Brantly Callaway, Evan Elmore, Andrew Fredrickson, Salama Freed, Aaron Gamino, James Harrison, Emily Lawler, Daniel Mangrum, Paul Niekamp, Francis Ratsimbazafy, Sebastian Tello-Trillo, Joseph Van Matre, and Katie Yewell.

I thank the many professors at Vanderbilt University who offered me guidance throughout my graduate career. In particular I would like to thank my committee members, Andrew Dustan, Gilbert Gonzales, and Federico Gutierrez as well as Kathryn Anderson, Andrew Goodman-Bacon, and Michelle Marcus. They have all been very generous in offering their time, help, and suggestions. I owe a special thank-you to the chair of my committee and my advisor, Christopher Carpenter for his advice and constructive criticism. I could not have asked for a more dedicated mentor.

I would also like to acknowledge the Department of Economics at Vanderbilt University and the Robert Wood Johnson Foundation for their financial support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

This research explores how the socioeconomic and health profiles of lesbian, gay, bisexual, and transgender (LGBT) individuals differ from those of otherwise similar straight cisgender individuals, as well as how those profiles are affected by LGBT related policies. Studying this population is important because, despite being the subject of intense policy debate across the United States, relatively little work has rigorously described the LGBT community or how its members are impacted by policy.

In the first chapter, *Legal Access to Same-Sex Marriage and Sexually Transmitted Diseases*, I use administrative data on disease incidence to estimate a series of dynamic difference-in-differences models to explore the relationship between legal same-sex marriage (SSM) and rates of common sexually transmitted infections. Results provide evidence that legal SSM is associated with a significant decrease in syphilis rates, an especially strong proxy for risky sex between men who have sex with men. This relationship appears to be driven largely by men and young adults. A mechanisms analysis using internet search data suggests reductions in syphilis may be due in part to a marriage induced increase in treatment seeking behavior.

In the second chapter, *Does It Get Better? Recent Estimates of Sexual Orientation and Earnings in the United States* (published in *Southern Economic Journal*), which is joint work with Christopher S. Carpenter, we make use of new information on sexual orientation recently included in the National Health Interview Survey to first reproduce a well-documented finding that self-identified lesbians earn significantly more than comparable heterosexual women. However, these data also show—for the first time in the literature—that self-identified gay men also earn significantly more than comparable heterosexual men, a difference on the order of 10%

of annual earnings. We discuss several possible explanations for the new finding of a gay male earnings premium and suggest that reduced discrimination and changing patterns of household specialization are unlikely to be the primary mechanisms.

In the third chapter, *Transgender Status, Gender Identity, and Socioeconomic Outcomes in the United States* (revised and resubmitted to *Industrial and Labor Relations Review*), which is joint work with Christopher S. Carpenter and Gilbert Gonzales, we provide the first large-scale evidence on transgender status, gender identity, and socioeconomic outcomes in the United States using representative data from 31 states in the Behavioral Risk Factor Surveillance System (BRFSS) that asked identical questions about transgender status and gender identity in at least one year from 2014-2016. Over 1,500 adults ages 18-64 identified as transgender. Individuals who identify as transgender are significantly less likely to be college educated and less likely to identify as heterosexual than individuals who do not identify as transgender. Controlling for these and other observed characteristics, transgender individuals have significantly lower employment rates, lower household incomes, higher poverty rates, and worse self-rated health than otherwise similar men who are not transgender. Differences in household structure account for a substantial share of these differences.

Chapter 1

Legal Access to Same-Sex Marriage and Sexually Transmitted Infections

Samuel T. Eppink*

Abstract

This paper explores the relationship between legal access to same-sex marriage and rates of syphilis, gonorrhea, and chlamydia using data from the CDC's National Notifiable Diseases Surveillance System from 1995 through 2017. By exploiting variation in legal same-sex marriage across states and time, I estimate a series of dynamic difference-in-differences models. Results provide new evidence that legal same-sex marriage is associated with a significant decrease in syphilis rates, an especially strong proxy for risky sex between men. This finding appears to be driven largely by men and young adults. An analysis using internet search data suggests reductions in syphilis may be due in part to a same-sex marriage induced increase in treatment seeking behavior.

*PhD Candidate in the Department of Economics, Vanderbilt University. I thank Christopher Carpenter for his invaluable guidance as my advisor and Andrew Francis, Hugo Mialon, and Handie Peng for generous data sharing. I also thank Marcus Dillender, Andrew Dustan, Andrew Goodman-Bacon, Gilbert Gonzales, Federico Gutierrez, Michelle Marcus, Zach Stark, and seminar participants at Vanderbilt University. All interpretations, errors, and omissions are my own.

1.1 Introduction

Legal recognition for same-sex couples has been, and remains, one of the most hotly debated civil rights issues in decades. Opponents of such recognition frequently claim that the so called "homosexual lifestyle" results in poor health, and that encouraging same-sex relationships, whether through legal recognition or otherwise, presents a risk to public health (see, for example, Family Research Council, 2004). Their argument is essentially that legalizing same-sex marriage will encourage people to live a more actively "homosexual lifestyle," especially with respect to more risky sex (e.g., higher rates of partner exchange and increased frequency of condomless sex), which in turn exacerbates the spread of sexually transmitted infections (STIs). While it is true that men who have sex with men (MSM) make up a disproportionately large share of STI incidence in the United States (CDC, 2018), there are compelling reasons to think that legalizing same-sex marriage could actually reduce the transmission of STIs.[1]

For instance, it is easy to imagine that being married, or even just the opportunity be married in the future, would change sexual minorities' attitudes about monogamy in a way that increases monogamous behavior, reducing the average number of sexual partners and consequently reducing the transmission of STIs. Relatedly, the newly created marriage market for LGB individuals may

[1] For the purposes of this paper, the category MSM is inclusive of men who have sex with men and women, which in some other contexts are separately referred to as MSMW.

increase the value of sexual health, inducing those in the market to seek more frequent STI screenings. Because some STIs, such as chlamydia, are frequently asymptomatic and others, such as syphilis, have symptoms that often go unnoticed, increases in screening may identify cases that otherwise would not have been found until later or not at all. If this is the case, data on reported STIs might reflect increases in screening post legalizing same-sex marriage, partially masking decreases in actual STI prevalence. Aside from impacting the attitudes of sexual minorities themselves, the legalization of same-sex marriage may also improve attitudes towards sexual minorities in society at large. Historically, intolerance towards gays and lesbians has forced their interactions to secretive, underground, and socially isolated contexts characterized largely by anonymous encounters with high risk individuals. From an economic perspective, increasing tolerance of lesbian, gay, and bisexual (LGB) individuals decreases the social cost of openly identifying as a sexual minority and enables these individuals to interact in safer and more open environments, thereby reducing the risk of contracting an STI. Marriage also confers a number of benefits that could plausibly impact the incidence of STIs. For example, being married can allow access to the employer sponsored health insurance of one's spouse. If same-sex marriage increases health insurance coverage for sexual minorities, effectively reducing the cost of seeing a doctor, then it may increase STI screening and reduce the spread of STIs.

Despite the numerous channels through which the legalization of same-sex marriage could plausibly impact STI rates, there has been no credible prior research on this question that has spanned the entire rollout of legal same-sex marriage in the United States. By leveraging recent variation in access to same-sex marriage across states and time, I test the competing hypotheses that legal same-sex marriage either increased or decreased the incidence of STIs in a quasi-experimental framework. Specifically, in this paper I estimate a series of dynamic difference-in-differences models using weekly provisional data and finalized annual data on state-level rates of new syphilis, gonorrhea, and chlamydia cases from the Centers for Disease Control and Prevention's (CDC) National Notifiable Diseases Surveillance System (NNDSS).[2] Understanding the relationship between same-sex marriage and STIs matters, not only because STIs are costly and we should care about them independently, but also because STIs are a novel and potentially important measure of investments in personal health. Moreover, the incidence of STIs, as I explain below, varies substantially across populations that should have been differentially affected by the legalization of same-sex marriage.

[2] In the United States, a notifiable disease is an infectious disease or condition for which the Council of State and Territorial Epidemiologists (CSTE), in consultation with the CDC, has determined routine reporting of individual cases is necessary for the prevention and control of the disease.

This paper is most closely related to Dee (2008), who found that laws recognizing same-sex partnerships in Europe were associated with significant reductions in syphilis. My paper differs from Dee's in that I evaluate the effect of same-sex marriage on STIs as opposed to the effect of policies more akin to civil unions or domestic partnerships; look at the context of the United States as opposed to Europe; include chlamydia, the most common STI in the United States, in the analysis in addition to syphilis and gonorrhea; and conduct a novel mechanisms analysis using internet search data. More broadly, this paper contributes to the literature evaluating legal same-sex marriage in the United States by being the first to examine STIs, and to the literature examining the policy determinants of STIs by being the first to evaluate the rollout of legal same-sex marriage across the United States. I report several core findings from this research. First, I find clear evidence that the legalization of same-sex marriage significantly reduced syphilis rates, an exceptionally strong proxy of risky sex between men, in the United States. This finding is robust to a number of alternative specifications and sample restrictions. Second, reductions in syphilis are driven by men and young adults. Third, an analysis using data on internet searches from Google Trends suggests that reductions in syphilis rates may be attributable, in part, to a same-sex marriage induced increase in treatment seeking behavior. While there is some evidence that legal same-sex marriage reduced rates of chlamydia, those estimates are less consistent, possibly because MSM comprise a much smaller portion of chlamydia cases than they do syphilis cases. Overall my results provide the first credible evidence that the legalization of same-sex marriage reduced rates of syphilis in the United States.

The remainder of the paper is as follows. Section 2 describes this paper's place in the existing literature and further discusses the channels through which legalized same-sex marriage may relate to the incidence of certain STIs. Section 3 describes the data and empirical methods used in the analysis. Section 4 presents the results, and section 5 concludes.

1.2 Literature and Mechanisms

This paper contributes to a small literature on LGB related policies and STIs. Although the suggestion that legalizing same-sex marriage might reduce the incidence of STIs has been around for some time (Posner, 1992, p. 311; Philipson and Posner, 1993, pp. 179–80; Eskridge, 1996, p. 120; Müller, 2002; Rauch, 2004, p. 79), there has not been sufficient variation in access to same-sex marriage to rigorously test this hypothesis in the context of the United States until recently.

Dee (2008) uses country-year level panel data from 1980 through 2003 to examine whether STI rates in Europe were impacted by legally recognizing same-sex partnerships. Dee estimates that laws extending marriage-like legal status to same-sex couples significantly reduced rates of syphilis by approximately 43% but did not affect the incidence of gonorrhea or HIV. He attributes these findings to changes in the relative returns of having multiple sex partners brought about by legal recognition. Allowing same-sex couples to share in the benefits afforded by legal recognition increases the returns to monogamy, therefore reducing the relative returns to having multiple sex partners thereby decreasing the spread of STIs. Relatedly, it may be the case that the ability to guarantee sexual health is valued in the new LGB marriage market, resulting in more frequent STI screenings among those, presently partnered or not, with the intention of pursuing marriage.

While it is easy to imagine that being married or seeking marriage would change the behavior of individuals in same-sex relationships, it is likely that legalizing same-sex marriage alters the behavior of LGB individuals more generally. Francis and Mialon (2010) use a state-year panel dataset spanning from the mid-1970s to mid-1990s to explore the relationship between LGB tolerance and HIV rates. They find that increasing tolerance towards gays in the United States, as expressed through the General Social Survey (GSS), was associated with a significant decrease in HIV rates. Francis and Mialon suggest that this is due in part to increased tolerance inducing gay men to substitute away from more risky sexual behaviors. An intolerant environment increases the costs of being openly gay, encouraging more secrecy and thus driving same-sex sexual behavior to underground contexts characterized by anonymous interactions with high risk individuals in socially disconnected venues (e.g., "cruising" or searching for strangers with whom to have anonymous one-time sex).[3] Conversely, legalizing same-sex marriage may serve as a signal of acceptance that induces LGB individuals to interact in more open, socially mediated, settings.[4]

In a similar vein, Francis et al. (2012), uses a state-year panel from 1981 to 2008 to estimate the effects of banning same-sex marriage or civil unions on STI rates. Because same-sex marriage was already illegal before states adopted same-sex marriage bans, the act of banning same-sex marriage does not constitute a change in access to marriage, but rather simply signals the degree of intolerance held towards LGB individuals. Consistent with the results of Francis and Mialon (2010), that higher levels of tolerance decreased STI rates, they find that laws banning same-sex marriage, a signal of lower levels of tolerance, increased syphilis rates. However, their estimates were smaller and less statistically significant when excluding California, the state with the largest LGB population, from the analysis. They also found that laws allowing same-sex marriage or civil

---

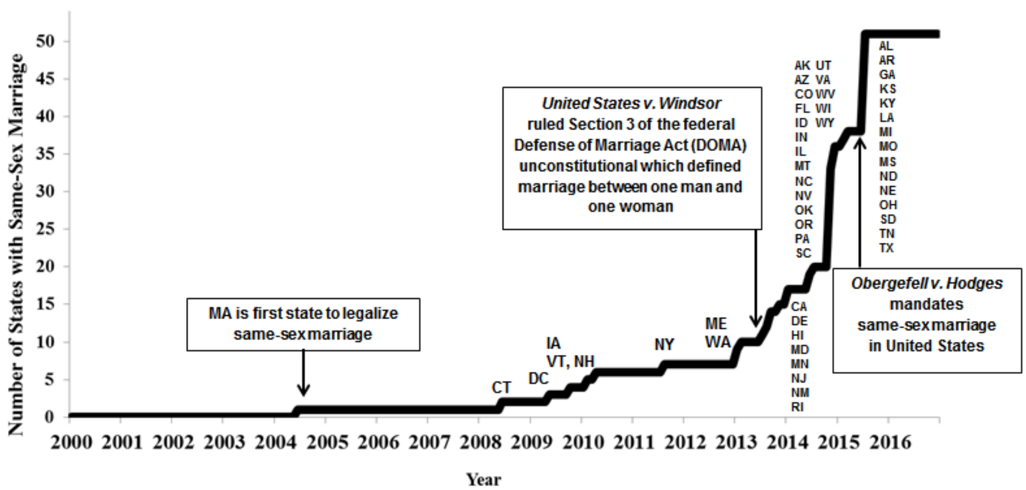
[3] It may also be the case, as noted by Müller (2002), that intolerance increases the incentives for gay men to move to urban areas, thereby reducing the search cost for sexual partners and increasing the spread of STIs.

[4] Legal same-sex marriage may correlate with improved attitudes towards LGB individuals even within those states that were forced to legalize same-sex marriage by federal court order. A number of recent studies have indicated that laws themselves can influence attitudes. For example, Aksoy et al. (2018) finds that laws recognizing same-sex relationships significantly increased favorable attitudes towards sexual minorities. (see also: Alesina and Fuchs-Schündeln, 2007; Barclay and Flores, 2017; Fong et al., 2006; Gallus et al., 2006; Kotsadam and Jakobsson, 2011; Soss and Schram, 2007; Tang et al., 2003; and Svallfors, 2010).

unions were unrelated to syphilis rates, though they are explicit in stating that at the time of their writing there was not enough variation in such laws to yield precise estimates.

This paper furthers the literature on LGB policy and STIs in several key ways. First, I am able to take advantage of much more policy variation than has been possible for the existing literature. For example, by 2008, the end of the period during which Francis et al. (2012) was looking at same-sex marriage bans, only two states (Massachusetts and Connecticut) had legalized same-sex marriage. From 2009 through mid 2013 same-sex marriage was gradually legalized in six more states and DC. In June 2013 the U.S. Supreme Court decision in United States v. Windsor struck down the Defense of Marriage Act (DOMA), thereby extending federal recognition and benefits to those same-sex couples to whom states would grant marriage licenses. Same-sex marriage was rapidly legalized for 29 more states in the following two years. In June 2015, the U.S. Supreme Court issued their decision in Obergefell v. Hodges, which held that the Fourteenth Amendment to the United States Constitution guaranteed same-sex couples the right to marry, legalizing same-sex marriage for the remaining 13 states (see Figure 1.1). This plausibly exogenous policy variation allows me, unlike previous work, to exploit differences in the availability of same-sex marriage through time across all 50 states and DC. Second, by assembling a weekly level panel of state STI rates I am able to identify treatment, in this case legal same-sex marriage, with much more precision than has been possible in previous work, most of which had been at the annual level.

Figure 1.1: Timeline of Legalization of Same-Sex Marriage



Source: National Conference of State Legislatures, the Human Rights Campaign, and various sources.

And third, by collecting information on internet searches I am able to conduct a novel analysis exploring potential channels through which same-sex marriage may have impacted rates of STIs.

This paper also fits into broader literatures studying the effects of same-sex marriage on health outcomes and access to care. A large body of evidence suggests that legal recognition improves mental health outcomes for sexual minorities (see, for example, Hert and Kertzner, 2006; Hatzenbuehler et al., 2010; Riggle et al., 2010; and Wight et al., 2013), but there is mounting evidence that it improves the healthcare access and physical health of sexual minorities as well. Hatzenbuehler et al. (2012) use prospective data from 1,211 sexual minority male patients at a community-based health center in Massachusetts to explore whether health care use and expenditure were reduced after Massachusetts' ban on same-sex marriage was struck down in 2003.[5] They find that in the year following the legalization of same-sex marriage, sexual minority men had a significant decrease in medical care and mental health visits compared to the previous year. They interpret these findings as evidence that same-sex marriage improved the health of sexual minority men, although their results were similar for both partnered and nonpartnered men.

A number of studies have found associations between the legal recognition of same-sex couples and insurance coverage among sexual minorities. Buchmueller and Carpenter (2012) found that legalizing same-sex domestic partnerships or civil unions improved insurance coverage for lesbian women in California. This finding was later supported by Dillender (2015), which made use of Current Population Survey data to find that legal recognition for same-sex couples increased the likelihood that women in same-sex couples were insured

---

[5] Hatzenbuehler et al. (2012) consider same-sex marriage legal in Massachusetts beginning in 2003, the year during which the Massachusetts Supreme Judicial Court ruled that the state's same-sex marriage ban was unconstitutional. However, the court stayed its ruling for 180 days, meaning same-sex marriage licenses did not begin being issued until May 17, 2004. For the purposes of this paper, the timing of same-sex marriage is coded with respect to when same-sex couples were able to begin getting married (the so-called effective date) as opposed to when the decision to legalize same-sex marriage was made (the so-called ruling date).

though their partner's employer. Using data from the American Community Surveys, Gonzales (2015) found that legalizing same-sex marriage in New York increased the likelihood that a sexual minority had employer sponsored health insurance. It is natural to expect that gaining health insurance would reduce the personal cost of healthcare and consequently increase its utilization. Indeed, Carpenter et al. (2018) makes use of national cross sectional data from the Behavioral Risk Factor Surveillance System spanning the rollout of same-sex marriage across the entire United States to find that in addition to significantly increasing insurance coverage, access to legal same-sex marriage increased the likelihood that men in same-sex households reported having a usual source of care or a checkup within the last year. If an increase in checkups corresponds with an increase in screening for STIs, then after the legalization of same-sex marriage gay men who have contracted an STI are on average likely aware of the fact sooner than they would have been otherwise, thereby shortening the window during which they might be unwittingly spreading the STI. This would be especially true for STIs that are often asymptomatic, such as chlamydia, or slow to manifest noticeable symptoms, such as syphilis. Empirically, this would likely result in a short-run increase in reported STIs, reflecting increases in screening rather than changes in sex risk behavior or actual disease incidence, followed by a decrease as the window for transmission between becoming contagious and receiving treatment shortens.

1.3 Data Description and Empirical Approach

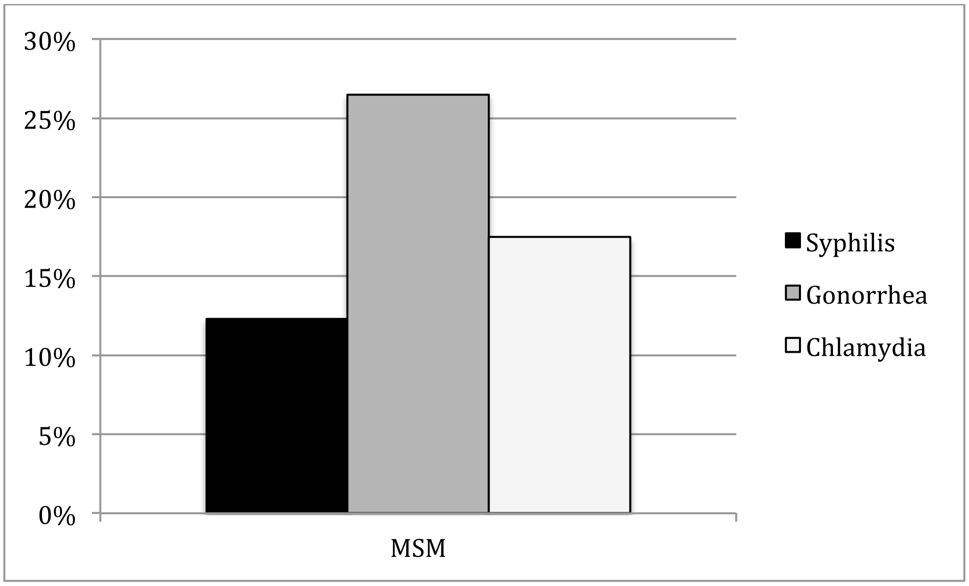
1.3.1 Data on rates of sexually transmitted infections

While there is now sufficient variation in access to same-sex marriage to begin testing its impact, sexual orientation is identified in very few datasets. Moreover, of those in which sexual orientation is included, few if any contained that information prior to the legalization of same-sex marriage. This makes it impossible to test the impact of legal access to same-sex marriage on STI rates amongst sexual minorities directly. However, based on what information is available about STI rates with respect to sexual partners, it is possible to make predictions about which STIs should be most responsive to the legalization of same-sex marriage if indeed same-sex marriage impacts the sexual behavior of MSM.

Data on new STI cases come from the CDC's National Notifiable Diseases Surveillance System (NNDSS). The NNDSS is a system of monitoring and reporting that enables all levels of public health to share information on diseases and conditions for which the Council of State and Territorial Epidemiologists (CSTE), in consultation with the CDC, has determined that regular and frequent information about individual cases is necessary for the prevention and control of the disease. Chlamydia and gonorrhea are the two most common notifiable diseases in the United States, with rates of 528.8 and 171.9 cases per 100,000 population respectively in 2017. Unfortunately, the share of those cases attributable to MSM and the frequency of cases among MSM are unknown, as most jurisdictions do not report the sex of sexual partners for cases of chlamydia or gonorrhea. There is, however, a collaboration of 10 state, county, and city health departments that have agreed to collect and report enhanced clinical and behavioral information about patients within their jurisdictions as part of the STD Surveillance Network (SSuN).6 Although not representative of the United States as a whole and comprised of data from only 30 STD clinics, SSuN data contains some information on chlamydia and gonorrhea cases with respect to sex and the sex of sexual partners. Among patients in SSuN jurisdictions, 17.5% of MSM tested positive for chlamydia and 26.5% for gonorrhea. Syphilis is less common in the general population, with an average of 9.5 cases per 100,000 population in 2017. Within SSuN jurisdictions, 12.3% of MSM tested positive for syphilis, making it less common among MSM than either chlamydia or gonorrhea (see Figure 1.2). However, MSM constitute a much larger share of total syphilis cases than chlamydia or gonorrhea.

[6] The 10 state, county, and city health departments fully participating in SSuN are Baltimore City (Maryland), California (excluding San Francisco County), Florida, Massachusetts, Minnesota, Multnomah County (Oregon), Philadelphia City (Pennsylvania), New York City (New York), San Francisco County (California), and Washington State.

Figure 1.2: Proportion of MSM STI Clinic Patients Testing Positive by STI, 2017 SSuN Jurisdictions

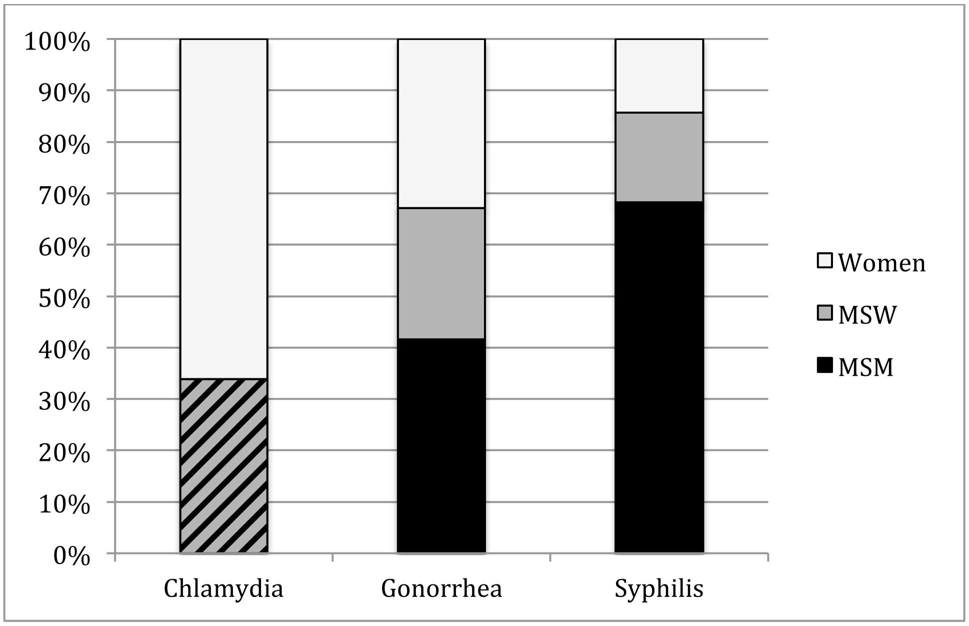


Due to the extreme proportion of syphilis cases attributable to MSM, states do in fact routinely report the sex of sexual partners for those testing positive for syphilis. As of 2017 MSM accounted for 68.2% of all reported syphilis cases among patients with information about sex and the sex of sexual partners, and 79.6% of male cases for which the sex of the partner was known.7 While information about the distribution of chlamydia and gonorrhea cases by sexual behavior is not available for the vast majority of states, all men regardless of sexual partner constitute 33.9% of chlamydia cases and 58% of gonorrhea cases,

[7] Of the 30,644 cases of Primary and Secondary syphilis reported in 2017, men who had sex with exclusively men accounted for 52.1%; men who had sex with men and women (MSMW), 5.8%; men who had sex with women only (MSW), 14.8%; women, 12.1%, men without information about the sex of sex partner, 15%; and unknown sex, 0.1% (CDC 2017).

far less than the 68.2% of syphilis cases attributable to MSM alone (see Figure 1.3).

Figure 1.3: Distribution of STI Cases by Sex and Sexual Behavior, 2017



Notes: The percentage of cases attributable to men (regardless of sexual behavior) and women are calculated using national data. Information about the proportion of chlamydia cases attributable to each MSM and MSW are not available at any level. The proportion of male gonorrhea cases attributable to each MSM and MSW displayed here reflect the proportions reported by clinics in SSuN jurisdictions and are not necessarily nationally representative. The proportion of syphilis cases attributable to each MSM and MSW are calculated using national data.

Despite the fact that there are fewer absolute cases of syphilis amongst MSM than chlamydia or gonorrhea, because MSM account for a much larger share of total syphilis cases than chlamydia or gonorrhea cases, population level syphilis rates should be relatively more responsive to changes in behavior or reporting among

MSM. Put differently, in the absence of MSM specific data on STI rates by state, the rate of syphilis in the general population is a less diluted signal of changes in STI incidence among MSM than the rates of either gonorrhea or chlamydia.

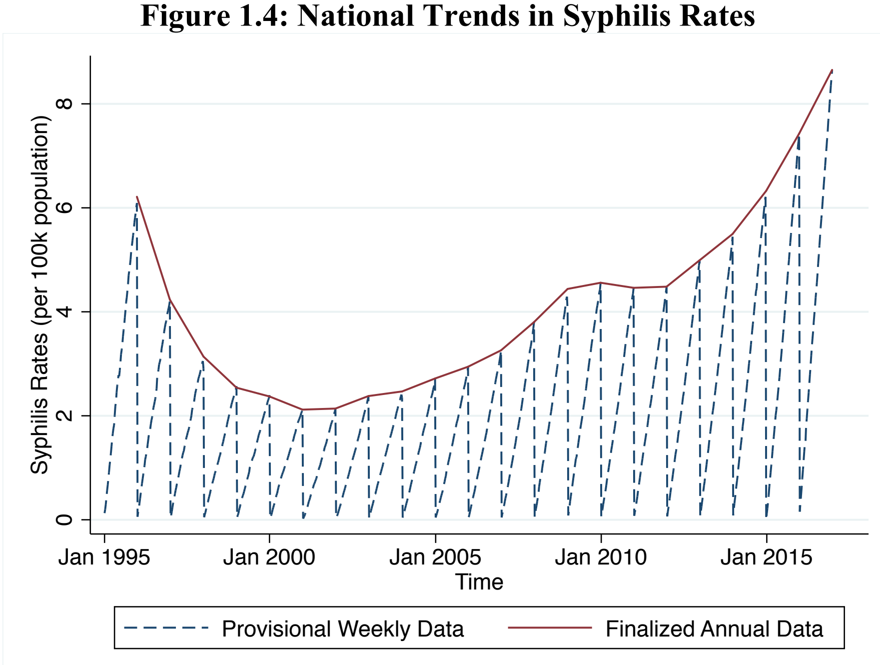
As part of the NNDSS, states' departments of health routinely submit reports on new cases of various diseases, including syphilis, gonorrhea, and chlamydia. Each week the CDC publishes information on new infections in the form of provisional cumulative cases of each disease for each state. These weekly data are provisional in the sense that states may revise their numbers from week to week.[8] For the analysis below these cumulative cases have been extracted and converted to weekly cases for each state-week from 1995 through 2017. Weekly cases were then divided by state population and multiplied by 100,000 to arrive at state-week STI rates per 100,000 population.

At the end of the calendar year the CDC reconciles and verifies the cumulative weekly cases with data providers in each state in order to produce finalized annual cases for various diseases by state. While the annual data do not

[8] For example, suppose a state reports 10 new cases of syphilis at the end of week 1. Now suppose that state reports 5 new cases of syphilis at the end of week 2 and revises their week 1 number down to 9. The resulting published cumulative cases would then be 10 for week 1 and 14 for week 2. Consequently, differencing the cumulative totals yields 10 for week 1 and 4 for week 2 rather than the true values of 9 for week 1 and 5 for week 2. It is possible for a state to revise a previous week's cases down so much that the week-to-week difference in the published cumulative cases is negative. This occurs for 0.42% of state-week syphilis observations, 0.43% of state-week gonorrhea observations, and 0.40% of state-week chlamydia observations. In such instances, cases are bottom-coded at 0. While this measurement error is an inherent feature of the way provisional data is published, it should not be systematically related to treatment, legal same-sex marriage.

allow for the precise policy coding of the weekly data (because most states enacted same-sex marriage in the middle of the year), they are inclusive of all revisions and therefore free of any measurement error that may result from the provisional nature of the weekly data. Annual data also have the added benefit of being available over a longer time period (1984 through 2016) and, in recent years, are available by demographic groups. Figures 1.4, 1.5, and 1.6 visually compare the weekly and annual data for syphilis, gonorrhea, and chlamydia respectively.

Figure 1.4: National Trends in Syphilis Rates



Notes: The CDC reports provisional weekly data on selected infectious national notifiable diseases in the form of cumulative counts of new cases by state throughout each year. At the beginning of a new year the count begins again at zero.

Figure 1.5: National Trends in Gonorrhea Rates

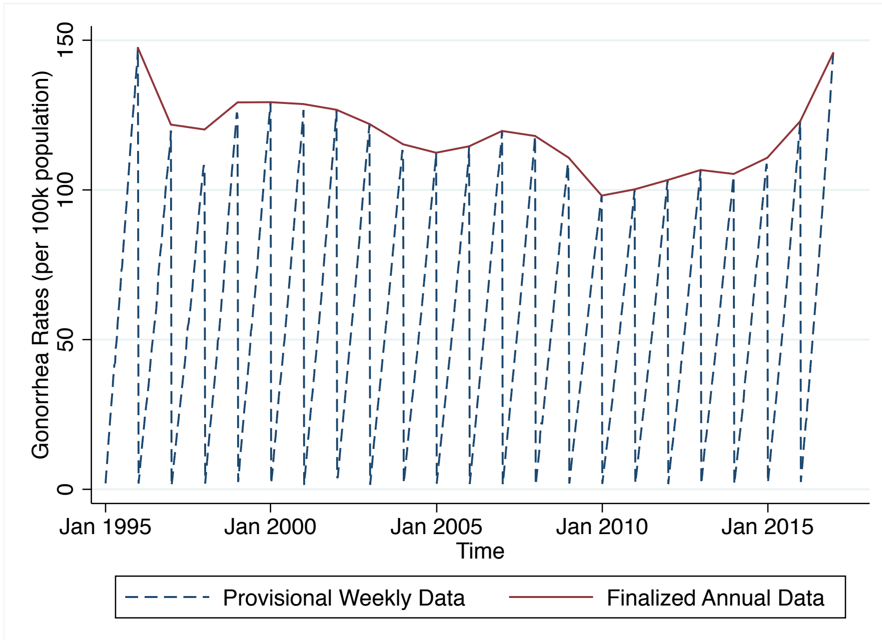
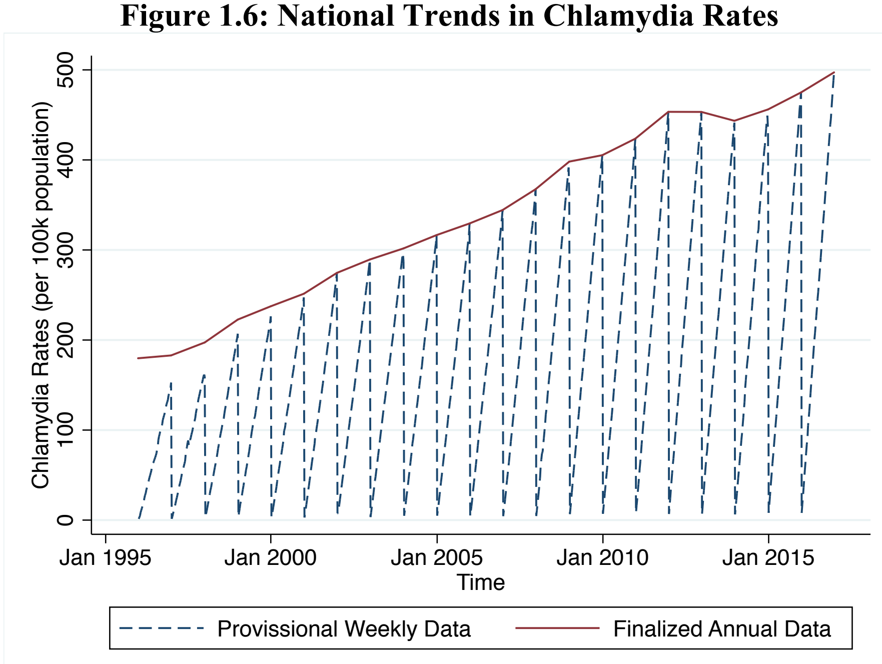


Figure 1.6: National Trends in Chlamydia Rates



In each case, weekly cumulative cases can be seen increasing throughout each year as dashed lines. Annual finalized counts are represented by sold lines. At the end of each year, the cumulative cases in the provisional data fit very closely with the finalized data, suggesting that the provisional data is robust to the CDC's year end process of reconciling and verifying data. This gives some additional confidence in the accuracy of the provisional weekly reports.

1.3.2 Data on internet searches related to sexually transmitted infections

Again, because information concerning sexual orientation is so scarce, it is not possible to explore directly how access to legal same-sex marriage impacts the behavior of LGB individuals with respect to risky sex, attitudes about monogamy, or receipt of treatment for STIs. However, internet search data may reveal changes in the salience of various STIs, interest in testing, and attempts to find treatment in response to the legalization of same-sex marriage. Data on internet searches related to various STIs were collected from Google Trends. Google Trends data are an unbiased sample of the billions of searches Google receives each day that can be pulled from as far back as 2004. Rather than raw search counts, search data is reported as measures of interest over time. These interest over time numbers represent frequency of queries containing given keywords relative to the highest number of such queries for a given geography and time period. Peak popularity for queries containing the keywords of interest for a given geography-time, in this case state-month, are represented by a value of 100. A value of 50 for a given state-month means queries containing the keywords of interest were half as popular. A value of 0 means that there were either no searches containing the keywords of interest or, more likely, that there was insufficient search volume to clear a privacy threshold for the given state-month.

For this analysis I constructed a dataset of interest in searches containing combinations of the words syphilis, gonorrhea, or chlamydia and keywords like treatment, symptoms, or testing.[9] The resulting values represent the monthly interest in searches for a combination of a given STI and keyword relative to the peak of such searches for each state over the period stretching back to January 2004, the earliest date for which Google Trends data are available. Although the internet search data is fairly noisy, the Google Trends data on the intensity of searches relating to syphilis and the NNDSS data on new cases of syphilis have a correlation of .38. The corresponding correlations for gonorrhea and chlamydia are .306 and .428 respectively.

---

[9] Because Google Trends computes interest over time for searches containing the keywords exactly as specified, I was explicit in including variants of keywords such as common misspellings and plurals. For example, the interest over time values I use relating to syphilis consist of searches containing the words syphilis or the common misspelling, "syphillis." Similarly, any values relating to chlamydia consist of searches containing the words chlamydia or "clamidia."

1.3.3 Empirical Approach

To estimate the effects of same-sex marriage on STI rates and STI-related internet searches, I make use of dynamic difference-in-differences models of the following form:

$$(1.1) \quad Y_{s,t} = \beta_0 + \beta_1 Y_{s,t-1} + \beta_2(\text{LEGAL SAME-SEX MARRIAGE})_{s,t} + \beta_3 Z_{s,t} + \beta_4 S_s + \beta_5 T_t + \beta_6 S_s{*}\text{TREND}_t + \beta_7 S_s{*}\text{TREND}_t^2 + \varepsilon_{s,t}$$

where $Y_{s,t}$ is the rate of a given STI per 100,000 population or the intensity of a given internet search for state *s* at time *t*. Following the literature on STI rates (see, for example, Chesson et al., 2000; Dee, 2008; and Francis et al., 2012), I include $Y_{s,t-1}$ to control for the incidence of a given STI in the previous period for each state. The inclusion of a lagged dependent variable is important in this context because the incidence of an infectious disease in a given period is determined largely by its previous incidence.[10] LEGAL SAME-SEX $\text{MARRIAGE}_{s,t}$ is equal to one for states *s* that have legal same-sex marriage in time *t*. LEGAL SAME-SEX $\text{MARRIAGE}_{s,t}$ is allowed to take fractional values

[10] While the inclusion of a lagged dependent variable can bias estimates, the magnitude of any bias decreases as the number of periods in the panel increases. In fact, for panels with 30 or more periods, least squares fixed effects models with a lagged dependent variable perform just as well or better than many alternatives (Judson and Owen, 1999). The primary analyses using weekly STI data from 1995 through September 2017 consists of 1183 periods, the annual STI data from 1984 through 2016 consists of 33 periods, and the monthly data on internet search intensity from 2004 through 2017 consists of 168 periods.

for cases in which same-sex marriage was legal for a fraction of period *t*.[11] $Z_{s,t}$ is a vector containing time varying state characteristics constructed using the Current Population Survey and consisting of the percentage of those in a state's labor force who are unemployed; the proportion of a state that is black, Asian, Hispanic, or other (non-white) race; the percent of a state's population who's highest level of educational attainment is high school, some college, and college or more; the percent ages 15 to 29 and ages 30 to 44; and the median personal income. $Z_{s,t}$ also contains controls for other aspects of the LGB policy landscape in state *s* at time *t*. These policy controls include: whether a state has banned same-sex marriage, controls for the strength and coverage of domestic partnership and civil union policy, whether a state has a sexual orientation inclusive employment non-discrimination act (ENDA), and whether a state has a religious freedom restoration act (RFRA).[12] As with same-sex marriage, other policy controls were allowed to take fractional values in cases where a policy was in effect for only a fraction of period *t*. S_s are state fixed effects that control for time

[11] Allowing the variable for legal same-sex marriage to take fractional values is especially important for the less precise annual models. For example, a binary treatment variable would assign both Delaware and Hawaii a value of 1 in 2013 despite the fact that same-sex marriage was legal effective July 1, 2013 for Delaware and December 2, 2013 for Hawaii. Using a fractional treatment variable can be thought of as taking into account the intensity of treatment in the first treated period. For Delaware and Hawaii this approach yields values of approximately .504 and .082 respectively for 2013.

[12] Domestic partnership/civil union policies are coded as one of four mutually exclusive categories depending upon a given policy's combination of coverage and strength. Coverage, whether the policy applies to all couples or same-sex couples only, is coded following Dillender (2014). Strength, whether the policy extends all of the state-level rights associated with marriage or a limited set of those rights, is coded following Badgett and Herman (2011).

invariant state characteristics, allowing states to have different baseline levels of a given outcome. T_t is a vector of time indicators to control for national time effects.[13] Models also include state linear and quadratic trends to allow outcomes to evolve differently across states. For annual models, $TREND_t$ is equal to 1 in 1984, 2 in 1985, and so on. For weekly models, $TREND_t$ is equal to 1 for the first week of 1995, 2 for the second week of 1995, and so on. $\varepsilon_{s,t}$ is the error term. β_2 is the coefficient of interest and is identified off of within-state deviations in STI rates from a quadratic trend that are coincident with the timing of the legalization of same-sex marriage. The key identifying assumption here is that conditional on the controls discussed above, STI rates would have evolved the same in states with and without same-sex marriage had it never been legalized. All models are weighted by population and have standard errors clustered by state.

1.4 Results

1.4.1 Evidence on legal same-sex marriage and rates of sexually transmitted infections

Table 1.1 provides direct evidence on the relationship between legal same-sex marriage and rates of syphilis in the top panel, gonorrhea in the middle panel,

[13] For regressions making use of finalized annual STI data, T_t consists of year indicators. For regressions making use of provisional weekly STI data, T_t includes indicators for both the week and year. For regressions using state-month internet search data, T_t includes indicators for the month and year.

and chlamydia in the bottom panel. Specifically, Table 1.1 reports the coefficient on the LEGAL SAME-SEX MARRIAGE variable from a fully saturated model of equation (1.1) that includes the one period lag of the respective STI rate; time varying state characteristics and policies; state, year, and week or month (where applicable) fixed effects; and state-specific linear and quadratic trends.

Table 1.1: Same-Sex Marriage and STI Rates (per 100,000 population)

	(1) Weekly Data, 1995-2017	(2) Annual Data, 1984-2016
Outcome is Syphilis Rate		
Pre-SSM mean	.094	6.922
SSM legal	-.023** (.009)	-.685* (.379)
R-squared	.367	.907
N	60139	1632
Outcome is Gonorrhea Rate		
Pre-SSM mean	3.019	166.837
SSM legal	.055 (.112)	-5.275 (3.607)
R-squared	.262	.974
N	60233	1630
Outcome is Chlamydia Rate		
Pre-SSM mean	8.579	277.797
SSM legal	-.449* (.244)	1.593 (6.773)
R-squared	.247	.963
N	56983	1516

Notes: * and ** denote statistical significance at 10% and 5%, respectively. All models are weighted by population and include 1 period lags of the respective dependent variable; state characteristics (percent in labor force who are unemployed, percent black, Asian, Hispanic, and other (non-white) race, percent of population who's highest level of educational attainment is high school, some college, and college or more, percent ages 15 to 29 and ages 30 to 44, and the median income); other LGB policies (same-sex marriage bans, the strength and coverage of domestic partnership/civil union policy, sexual orientation inclusive ENDA, and RFRA); state and year fixed effects; and state specific linear and quadratic trends. Weekly models also include week fixed effects. Standard errors are clustered at the state level and are robust to heteroscedasticity.

Column 1 displays results using state-week level data from 1995 through 2017, while column 2 displays results using state-year level data from 1984 through 2016.

Recall that although syphilis is less common among MSM than gonorrhea or chlamydia, because MSM constitute a substantially larger share of total syphilis cases, syphilis rates should be a stronger signal of any behavioral changes among MSM relative to rates of gonorrhea or chlamydia. Indeed, results in Table 1.1 indicate that the legalization of same-sex marriage was associated with large and statistically significant reductions in rates of syphilis but not gonorrhea. Results are more ambiguous for chlamydia, with evidence of reductions appearing in weekly but not annual data. For syphilis, in column 1 of the top panel I estimate that the legalization of same-sex marriage was associated with a significant reduction of .023 new weekly cases of syphilis per 100,000 population, a decrease of 24.5% relative to the pre-reform mean. Similarly, column 2 of the first panel suggests that the legalization of same-sex marriage was associated with a significant reduction of .685 new annual syphilis cases per 100,000 population, a 9.9% decrease relative to the pre-reform mean. Results in the middle panel provide no evidence of a relationship between the legalization of same-sex marriage and gonorrhea rates. For weekly chlamydia rates in column 1 of the bottom panel, legal same-sex marriage was associated with a .449 case reduction in the number of new weekly chlamydia cases per 100,000 population, a decrease of 5.2% relative to the pre-reform mean.

However, as cautioned by Goodman-Bacon (2018), if treatment effects vary over time, then the single-coefficient two-way fixed effects model generating the main difference-in-differences estimates in Table 1.1 may be biased away from the true treatment effect. Figures, 1.7, 1.8, and 1.9 for weekly syphilis, gonorrhea, and chlamydia rates, respectively, show graphically the estimates from event study models identical to equation (1.1) but where the single treatment variable, LEGAL SAME-SEX MARRIAGE, has been replaced by a series of variables representing time relative to the legalization of same-sex marriage in each state.[14] Looking at Figure 1.7, the event study for syphilis, the significance of the event time variables leading up to the legalization of same-sex marriage is suggestive of a pre-trend that may be biasing the results for syphilis in the top panel of Table 1.1. Nonetheless, there is a level shift downward after the legalization of same-sex marriage, which is indicative of an effect of legal same-sex marriage on weekly syphilis rates.

[14] Each state-week observation of syphilis, gonorrhea, or chlamydia rate is coded as falling into one of 12 categories representing the time of that observation relative to the legalization of same-sex marriage in the respective state. Those categories are: 6 or more years before legal same-sex marriage, each year from 1 to 5 years before, each year from 1 to 5 years after, and 6 or more years after. The 10 coefficients for 1 to 5 years before and after legal same-sex marriage are plotted in Figures 1.7, 1.8, and 1.9. The unbalanced endpoints (6 or more years before and 6 or more years after) are excluded from the figures.

Figure 1.7: Effect of Legal Same-Sex Marriage on Weekly Syphilis Rates

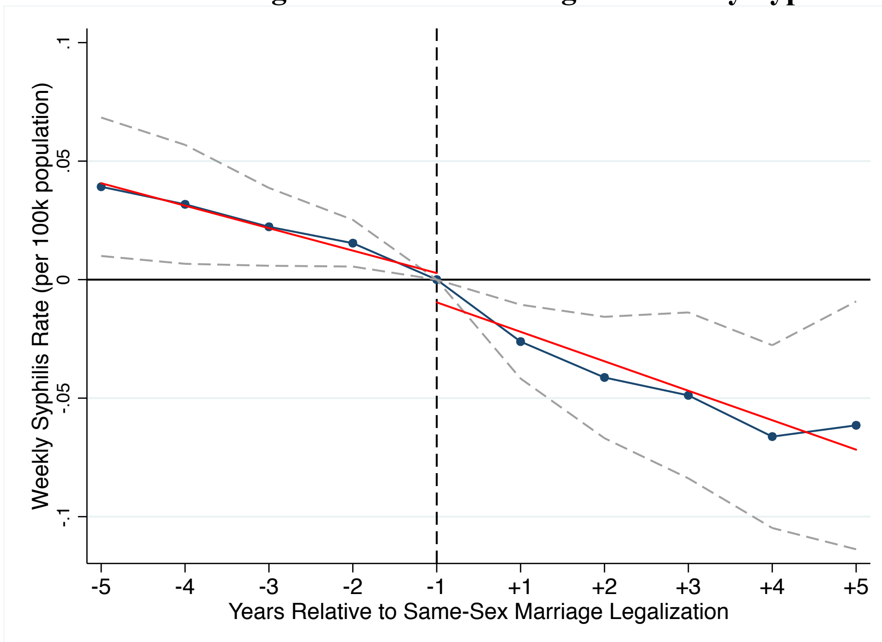


Figure 1.8: Effect of Legal Same-Sex Marriage on Weekly Gonorrhea Rates

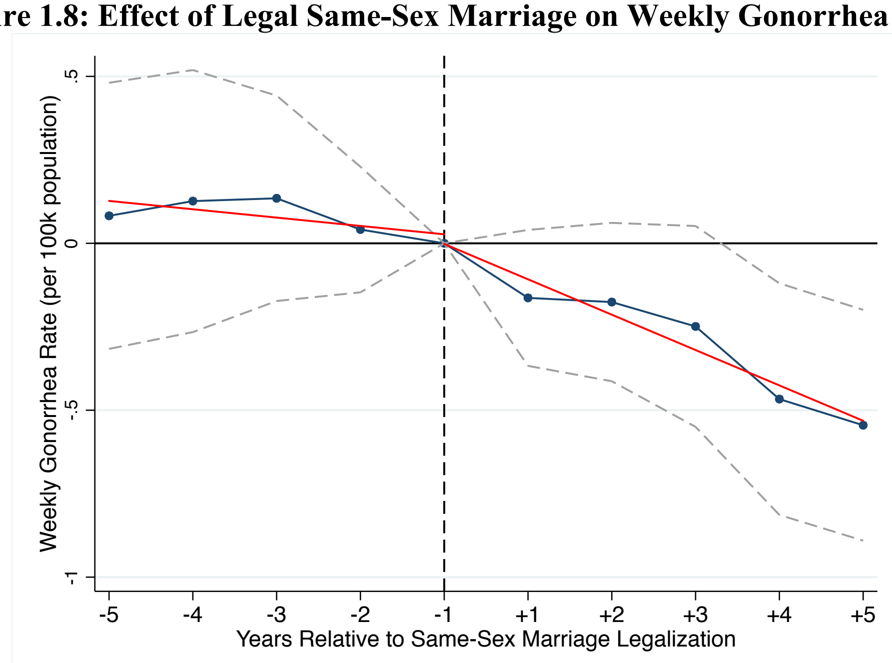
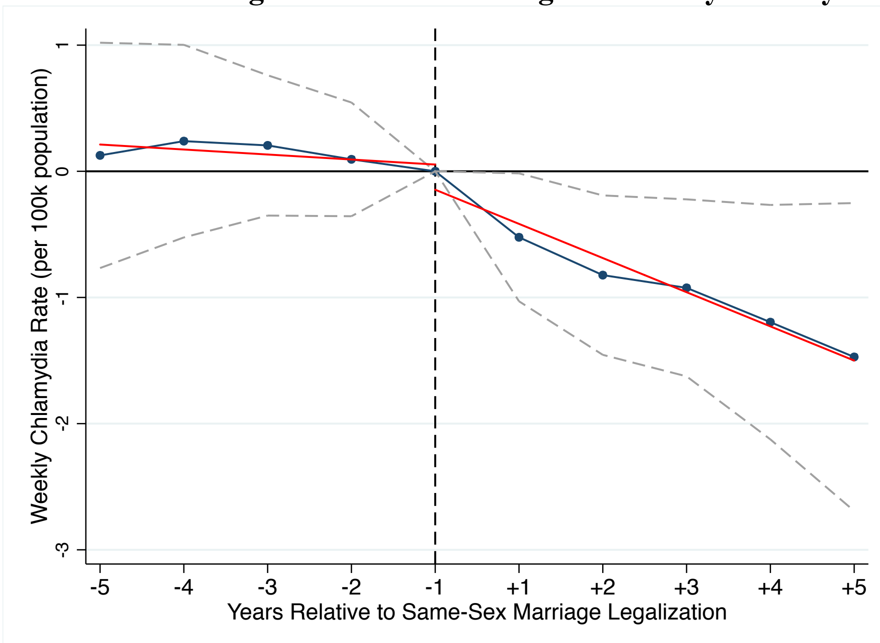


Figure 1.9: Effect of Legal Same-Sex Marriage on Weekly Chlamydia Rates



To further assess the robustness of the main difference-in-differences results for syphilis, they are put through a number of specification checks and sample restrictions in Tables 1.3 and 1.4. Unlike Figure 1.7, Figures 1.8 and 1.9 for gonorrhea and chlamydia unambiguously show no evidence of differential pre-trends. Figure 1.9 in particular provides clear evidence that the legalization of same-sex marriage was associated with significant decreases in chlamydia rates, strongly supporting the results in column 1 of the bottom panel in Table 1.1.

Table 1.2 explores potential sources of heterogeneity in the relationship between same-sex marriage and syphilis rates with respect to demographic and state subsamples.

Table 1.2: Heterogeneity in Relationship Between Legal Same-Sex Marriage and Syphilis
Coefficient on Legal Same-Sex Marriage
Outcome is Syphilis Rate (per 100,000 population)

	(1) Weekly Data	(2) Annual Data
Baselines		
Full Sample	-.023** (.009) {-24.5%}	-.685* (.379) {-9.9%}
By Sex (Annual Data, 1984-2016)		
Male	--	-.957* (.499) {-10.4%}
Female	--	-.449 (.336) {-9.6%}
By Age (Annual Data, 2000-2016)		
Age 25 to 34	--	-.709* (.377) {-21.9%}
Age 35 to 44	--	-1.240 (.813) {-12.9%}
Age 55+	--	-1.087 (.777) {-14.1%}
By State Characteristics		
States with rates below the median	.007 (.007) {18.5%}	.112 (.375) {3.9%}
States with rates above the median	-.034*** (.012) {-30.2%}	-.433 (.475) {-5.2%}
Only states that did not have CU/DP prior to SSM	-.020** (.009) {-20.9%}	-.347 (.431) {-4.7%}
Only states that adopted SSM via court order	-.025* (.013) {-26.5%}	-.272 (.543) {-3.9%}
Only states that adopted SSM via *Obergefell v. Hodges*	-.036 (.030) {-30.9%}	-5.318 (13.962) {-65.4%}

Notes: *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively. See notes to Table 1.1. CU/DP refers to civil union/domestic partnership policy. Observations for weekly data range from 1995 through September 2017. Observations for annual data range from 1984 through 2016 unless stated otherwise.

As with Table 1.1, column 1 presents results obtained using weekly data while column 2 presents results obtained using annual data. Effect sizes relative to the mean syphilis rate before the legalization of same-sex marriage are reported in curled brackets. The main difference-in-differences results from the top panel of Table 1.1 are again listed at the top of Table 1.2 as a baseline for reference followed by models looking at the state-week rate of new syphilis cases per 100,000 population broken down by: sex, age (age 25 to 34, age 35 to 44, and age 55 or greater), and whether states had above or below median syphilis rates. I also look for heterogeneity with respect to characteristics of the LGB policy environment in a state (states that did not offer civil unions or domestic partnerships to same-sex couples prior to legalizing same-sex marriage, states that legalized same-sex marriage via court order, and states that legalized same-sex marriage as a result of *Obergefell v. Hodges*). If it is the case that changes in the returns to monogamy brought about by same-sex marriage are driving reductions in syphilis rates, then the relationship between same-sex marriage and syphilis should be more pronounced in states that had larger gaps in the benefits afforded to same-sex and different-sex couples prior to the legalization of same-sex marriage. States that adopted civil union or domestic partnership policies prior to same-sex marriage closed that gap some by extending a subset of marriage benefits to same-sex couples. Therefore, finding a stronger relationship between legal same-sex marriage and syphilis rates in states that did not have civil union or domestic partnership policies would be consistent with the hypothesis that increasing returns to monogamy are contributing to reductions in syphilis.

Looking at the contents of column 1 in Table 1.2 for models using weekly data, reductions in syphilis rates appear to be driven by states with above median rates of syphilis.[15] The effect is roughly 16.7% smaller for states that did not have civil union or domestic partnership policies prior to legalizing same-sex marriage. This finding is inconsistent with the hypothesis that reductions in syphilis rates are being driven by increases in the returns to monogamy as such increases would have been greatest in states that lacked civil union and domestic partnership policies prior to legalizing same-sex marriage. The effect of legalizing same-sex marriage on syphilis rates within the subsample of states that did so via court order is similar to that of the full sample. This is likely because relatively few states legalized same-sex marriage through legislation or referendum.[16] Although the coefficient on legal same-sex marriage for the subsample of states that adopted same-sex marriage as a result of *Obergefell v. Hodges* is not statistically significant, it is larger in magnitude than the baseline results.

While the annual data in column 2 of Table 1.2 do not allow for the precise treatment coding of the weekly data, they do have the benefit of being available by demographic characteristics. When looking at syphilis rates by sex,

[15] States with above median syphilis rates are: Alabama, Arizona, Arkansas, California, Delaware, Florida, Georgia, Illinois, Indiana, Louisiana, Maryland, Massachusetts, Michigan, Mississippi, Missouri, Nevada, New Mexico, New York, North Carolina, Ohio, Oklahoma, South Carolina, Tennessee, Texas, and Virginia plus DC.

[16] Delaware, Hawaii, Illinois, Maryland, Minnesota, New Hampshire, New York, Rhode Island, Vermont, Washington, and DC legalized same-sex marriage via legislation. Maine was the only state to legalize same-sex marriage through a voter referendum.

the coefficient on legal same-sex marriage obtains statistical significance for the male sample and not the female sample. This is to be expected given that syphilis is much more common among men than among women, with rates of 16.9 and 2.3 cases per 100,000 population respectively (CDC, 2018). Moreover, the effect of legal same-sex marriage on syphilis rates for the male sample is about 40% larger than estimates for the general population, represented by the baseline estimates. Reductions are also more pronounced among young adults ages 25 to 34, with an effect size 12 percentage points greater than that of the full sample. Although the coefficients on legal same-sex marriage for the various state subsamples do not achieve statistical significance in the annual models, the magnitudes of the coefficients relative to each other are consistent with the findings of the weekly models, that reductions in syphilis are more pronounced in states with above median syphilis rates. Overall, the takeaways from Table 1.2 are that reductions in syphilis rates following the legalization of same-sex marriage: appear to be driven by reductions among men and young adults, are stronger in states with above median rates of syphilis, and do not appear consistent with the hypothesis that changes in the returns to monogamy are a primary driver in reducing syphilis rates.

Table 1.3 reports a series of analyses probing the robustness of the baseline specification used to obtain the main difference-in-differences results in the top panel of Table 1.1.

Table 1.3: Robustness of Association Between Same-Sex Marriage and Syphilis to Alternative Specifications
Coefficient on Legal Same-Sex Marriage
Outcome is Syphilis Rate (per 100,000 population) Unless Otherwise Stated

	(1) Baseline	(2) Without quadratic trends	(3) Without lagged dependent variable	(4) Without quadratic trends and lagged dependent variable	(5) With 2nd period lag of outcome	(6) Outcome is natural log of syphilis rates	(7) Year and week FE replaced with year-by-week FE
Weekly Data, 1995-2017							
SSM legal	-.023**	-.022*	-.024**	-.023*	-.023**	-.013**	-.024**
	(.009)	(.011)	(.010)	(.012)	(.009)	(.006)	(.010)
Pre-SSM mean	.094	.094	.094	.094	.094	.085	.094
R-squared	.367	.341	.367	.335	.368	.522	.389
N	60139	60139	60244	60244	60034	60139	60139
Annual Data, 1984-2016							
SSM legal	-.685*	-.452	-1.862*	-1.535	-0.631*	-.042	--
	(.379)	(.407)	(.945)	(1.481)	(0.357)	(.047)	
Pre-SSM mean	6.922	6.922	7.070	7.070	6.785	1.648	--
R-squared	.907	.900	.765	.705	0.939	.947	--
N	1632	1632	1683	1683	1581	1632	--

Notes: * and ** denote statistical significance at 10% and 5%, respectively. See notes to Table 1.1.

In Table 1.3, the top panel displays the coefficients on legal same-sex marriage from models using weekly data, and the bottom panel displays coefficients from annual models. For reference, column 1 reports results from the baseline specification. State-specific quadratic trends are excluded from the model in column 2. Column 3 instead excludes the lagged dependent variable. Column 4 excludes both the state-specific quadratic trends and lagged dependent variable. Column 5 instead adds a second period lag of the dependent variable. Column 6 returns to the baseline specification but changes the outcome to the natural log of the rate of new weekly syphilis cases. Finally, column 7 replaces the year and week fixed effects with year-by-week fixed effects. The top panel of Table 1.3 shows that the finding of an association between same-sex marriage and reductions in weekly syphilis rates is robust to the exclusion of quadratic trends, the exclusion of the lagged dependent variable, the simultaneous exclusion of both the quadratic trends and the lagged dependent variable, the addition of a second period lag of the dependent variable, taking the natural log of syphilis rates, and the inclusion of year-by-week fixed effects. Results from the annual models in the bottom panel are less robust. While the annual models are robust to the exclusion of the lagged dependent variable and the inclusion of a second period lag of the dependent variable, they are not robust to the exclusion of quadratic trends or taking the natural log of syphilis rates.[17]

[17] Column (7) in the bottom panel of Table 1.3 is blank as year-by-week fixed effects cannot be

Table 1.4 reports a series of analyses checking robustness to alternative treatment of California and the exclusion of the five most populous states one-by-one. As in Table 1.3, the top panel reports results from weekly models, and the bottom panel reports results from annual models. Again, column 1 of Table 1.4 reports the baseline results. Column 2 reports the output from a model where same-sex marriage is coded as 'off' in California until the final week of June 2013, as opposed to briefly turning 'on' in the third week of June 2008, when the California Supreme Court ruled that denying same-sex couples the right to marry contradicted the Constitution of Carolina, and then back 'off' again in the first week of November 2008 following the Proposition 8 referendum amending the state constitution to specify that the right to marry did not extend to same-sex couples. Previous work focusing on sexual minorities has been highly sensitive to the exclusion of California, the state with both the largest population and home to a disproportionally large share of LGB Americans (see, for example, Francis et al. 2012). In order to be confident that the findings in this paper are not being driven by any one state in particular, the five most populous states, California, Texas, Florida, New York, and Pennsylvania, are dropped one-by-one in columns 3 through 7.

---

applied to the annual models. I also estimated Poisson models on the raw count data for new syphilis cases. The coefficients on the legalization of same-sex marriage in such models were -.071 (.076) and -.133 (.112) for weekly and annual syphilis counts respectively.

Table 1.4: Additional Robustness of Association Between Same-Sex Marriage and Syphilis to Alternative Specifications
Coefficient on Legal Same-Sex Marriage
Outcome is Syphilis Rate (per 100,000 population)

	(1) Baseline	**(2) Recode California as off until 2013**	**(3) Drop California**	**(4) Drop Texas**	**(5) Drop Florida**	**(6) Drop New York**	**(7) Drop Pennsylvania**
Weekly Data, 1995-2017							
SSM legal	-.023**	-023**	-.019**	-.022**	-.026**	-.022*	-.025**
	(.009)	(009)	(.009)	(.010)	(.010)	(.012)	(.010)
Pre-SSM mean	.094	.094	.094	.093	.092	.093	.093
R-squared	.367	.367	.367	.340	.380	.360	.371
N	60139	60139	58959	58959	58959	58959	58959
Annual Data, 1984-2016							
SSM legal	-.685*	-.648*	-.524	-.784*	-.789*	-.512	-.805**
	(.379)	(.374)	(.374)	(.404)	(.411)	(.461)	(.370)
Pre-SSM mean	6.922	6.918	6.908	6.740	6.562	6.773	7.022
R-squared	.907	.907	.907	.904	.907	.904	.907
N	1632	1632	1600	1600	1600	1600	1600

Notes: * and ** denote statistical significance at 10% and 5%, respectively. See notes to Table 1.1.

The weekly models in the top panel of Table 1.4 are robust the alternative same-sex marriage coding for California and sequentially dropping the five most populous states. The annual models in the bottom panel of Table 1.4 are robust to the California recode and the exclusion of Texas, Florida, and Pennsylvania, but not to the exclusion of California or New York.[18]

1.4.2 Evidence on legal same-sex marriage and internet searches related to sexually transmitted infections

Table 1.5 explores the relationship between legal same-sex marriage and interest in internet searches related to syphilis in the top panel, gonorrhea in the middle panel, and chlamydia in the bottom panel. Each cell of Table 1.5 reports the coefficient on the LEGAL SAME-SEX MARRIAGE variable from a fully saturated model where the outcomes are state-month level interest in internet searches containing both the STI in the respective panel and the keywords listed in the column headings: man or men in column 1, cure or cures in column 2, treatment or treatments in column 3, test or testing in column 4, and symptom or symptoms in column 5.

[18] Weekly models are also robust to the sequential exclusion of the five states with the highest rates of syphilis over the study period, Louisiana, Mississippi, Georgia, Florida, and South Carolina. Annual models are robust to the exclusion of Louisiana, Georgia, Florida, and South Carolina, but slip under the threshold for statistical significance when excluding Mississippi. These results are available upon request.

Table 1.5: Effect of Same-Sex Marriage on STI Related Google Searches
2004-2017 Google Trends Data
Coefficient on Legal Same-Sex Marriage
Outcome is intensity of searches containing combinations of the STI listed in the first column (syphilis, gonorrhea, or chlamydia) and keywords listed at the top of following columns

	(1) man or men	(2) cure or cures	(3) treatment or treatments	(4) test or testing	(5) symptom or symptoms
Syphilis					
SSM legal	6.307**	4.530***	3.136*	-5.014	-1.905
	(2.265)	(.735)	(1.808)	(3.485)	(1.879)
R-squared	.306	.203	.250	.272	0.449
N	1670	1336	3674	3841	6012
Gonorrhea					
SSM legal	1.702	4.848***	2.608	-2.963	1.684
	(3.435)	(1.358)	(2.179)	(2.001)	(1.909)
R-squared	.497	.271	.381	.415	.622
N	4175	2171	4509	4008	6513
Chlamydia					
SSM legal	6.855***	6.766*	-.711	2.935	-1.593
	(1.319)	(3.738)	(2.170)	(2.255)	(2.533)
R-squared	.636	.375	.476	.535	.659
N	6346	4342	5678	6012	7348

Notes: *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively. Search intensities also include common misspellings of syphilis (syphillis) and chlamydia (clamidia). See notes to Table 1.1.

For example, the outcome for the model in column 1 of the top panel is interest in internet searches containing the words ‘syphilis’ and either ‘man’ or ‘men’. This measure would capture searches such as, “how common is syphilis in men” and “can a man get syphilis.” The output displayed in Table 1.5 suggests that the legalization of same-sex marriage was associated with statistically significant increases in internet searches relating to syphilis in men, cures for syphilis, and treatment for syphilis; increases in searches for cures for gonorrhea; and increases in searches for chlamydia in men and cures for chlamydia. This is consistent with same-sex marriage induced increases in STI awareness, particularly with respect to syphilis and chlamydia.

1.5 Discussion and Conclusion

In this paper I leveraged recent variation in access to same-sex marriage across states and time to test the competing hypotheses that legal same-sex marriage either increased or decreased the incidences of syphilis, gonorrhea, and chlamydia in a quasi-experimental framework. Results reveal a strong and significant association between the legalization of same-sex marriage and reductions in syphilis rates, an exceptionally strong proxy of risky sex between men. This finding appears to be driven largely by men and young adults. Results also suggest, to a lesser extent, an association between legal same-sex marriage and reductions in chlamydia rates, a more diluted signal of risky sex between men. I find little evidence of a relationship between same-sex marriage and rates of gonorrhea.

The null finding for gonorrhea, though consistent with Dee (2008) and Francis et al. (2012), raises the question: why would the legalization of same-sex marriage impact rates of syphilis and chlamydia but not gonorrhea, especially given that MSM almost certainly constitute a larger portion of gonorrhea cases than chlamydia cases (see Figure 1.3)? The answer may have something to do with the timing and severity of symptoms presented by each STI. While gonorrhea can be asymptomatic (Peterman et al., 2006), symptoms in men usually appear one to fourteen days after becoming infected (Harrison et al., 1979), and include discharges from the infected area, itching, and soreness. Symptoms of syphilis, on the other hand, can take anywhere from ten days to three months to manifest, and begin with the appearance of a single sore lasting three to six weeks (CDC, 2017c). These sores often go unnoticed as they are typically painless and can be located in areas that are difficult to see. Even more likely to go unnoticed is chlamydia, which is usually asymptomatic, manifesting symptoms in only an estimated 6 to 30 percent of infected women and 11 percent of infected men (Korenromp et al., 2002; Farley et al., 2003). Even in those cases where symptoms of chlamydia do manifest, it is not until several weeks after infection (CDC, 2017b). In other words, it is much more difficult for someone to be unaware that they have been infected with gonorrhea than it is to unknowingly be infected by syphilis or chlamydia, meaning syphilis and chlamydia are more likely to be discovered by precautionary screening than gonorrhea. The findings that same-sex marriage is associated with reductions in syphilis and chlamydia but not gonorrhea are consistent with the hypothesis that the legalization of same-sex marriage induced sexual minorities to care more about whether or not they are unknowingly carrying an STI, perhaps because sexual health is valuable in the new LGB marriage market or because sexual minorities are now more invested in the health of their partners. Either way, if individuals are more likely to seek STI screening after the legalization of same-sex marriage, then cases of infections with few or delayed symptoms, such as syphilis or chlamydia, are sure to be identified and treated earlier than they would have been otherwise, thereby shortening the window during which infected individuals are unknowingly spreading these infections.

An analysis using Google Trends data returned evidence of an increase in internet searches relating to syphilis and chlamydia in men and treatment for syphilis after the legalization of same-sex marriage. This is also consistent with the hypothesis that legalizing same-sex marriage induced MSM to seek more frequent screening for syphilis and chlamydia. While increased screening has the effect of limiting the spread of an STI, it is also likely to catch some cases that would have gone unreported otherwise. Such cases would not have shown up in the NNDSS data prior to same-sex marriage, and would therefore bias the estimated effect of legal same-sex marriage on STIs towards zero. Nonetheless, a conservative back of the envelope calculation using the coefficient from the top panel of column 1 in Table 1.1 implies that legal same-sex marriage prevents 3,895 new cases of syphilis each year. Using the cost per syphilis case estimated by Owusu-Edusei et al. (2013), 3,895 fewer syphilis cases each year equates to annual savings of roughly $3.2 million (2018 dollars) in direct syphilis related medical expenditures. The coefficient in column 1 of panel 3 in Table 1.1 implies a reduction of 76,044 new cases of chlamydia each year. A similar calculation including reductions in chlamydia yields total annual savings in excess of $5.8 million (2018 dollars).[19] Overall the findings of this paper suggest that the legalization of same-sex marriage likely had the unintended consequence of improving public health with respect to sexually transmitted infections, improving the quality of life among sexual minorities and saving $5.8 million in direct costs every year. Relatedly, these findings suggest that efforts to repeal same-sex marriage, if successful, may negatively impact public health.

[19] Owusu-Edusei et al. (2013) estimates that the lifetime cost of a case of syphilis is $709 in 2010 dollars. They estimate that a case of chlamydia carries a lifetime cost of $30 for a man and $364 for a woman in 2010 dollars.

1.6 References

Alesina, A., & Fuchs-Schündeln, N. (2007). Good-bye Lenin (or not?): the effect of communism on people's preferences. *American Economic Review*, 97(4), 1507-1528.

Aksoy, C. G., Carpenter, C., De Hass, R., & Tran, K. (2018). Do Laws Shape Attitudes? Evidence From Same-Sex Relationship Recognition Policies in Europe. EBRD Working Paper, #219.

Badgett, M. V. L., & Herman, J. L. (2011) Patters of Relationship Recognition by Same-Sex Couples in the United States. *The Williams Institute.*

Barclay, S., Flores, A. R. (2017). Policy Backlash: Measuring the Effect of Policy Venues using Public Opinion. *Indiana Journal of Law and Social Equity*, 5(2), 391-408.

Buchmueller, T., & Carpenter, C. (2012). The Effect of Requiring Employers to Extend Health Benefit Eligibility to Same-Sex Partners of Employees: Evidence from California. *Journal of Policy Analysis and Management*, 31(2), 388-403.

Carpenter, C., Eppink, S. T., Gonzales, G., Jr., & McKay, T. (2018). Effects of Access to Legal Same-Sex Marriage on Marriage and Health: Evidence from BRFSS. NBER Working Paper, #24651.

Centers for Disease Control and Prevention. (2017). NCHHSTP AtlasPlus. Available at: https://www.cdc.gov/nchhstp/atlas/index.htm.

Centers for Disease Control and Prevention. (2017, November 30). Chlamydia – CDC Fact Sheet (Detailed). Available at: https://www.cdc.gov/std/chlamydia/stdfact-chlamydia-detailed.htm

Centers for Disease Control and Prevention. (2017, November 30). Syphilis – CDC Fact Sheet (Detailed). Available at: https://www.cdc.gov/std/syphilis/stdfact-syphilis-detailed.htm

Centers for Disease Control and Prevention. (2018). *Sexually Transmitted Disease Surveillance 2017*. Atlanta: U.S. Department of Health and Human Services.

Centers for Disease Control and Prevention. National Notifiable Diseases Surveillance System, Weekly Tables of Infectious Disease Data. Atlanta, GA. CDC Division of Health Informatics and Surveillance. Available at: https://www.cdc.gov/nndss/infectious-tables.html.

Centers for Disease Control and Prevention, National Center for HIV, STD and TB Prevention (NCHSTP), Division of STD/HIV Prevention, Sexually Transmitted Disease Morbidity 1984 - 2014 by Gender, CDC WONDER On-line Database.

Chesson, H., Harrison, P., & Kassler, W. J., (2000). Sex under the influence: the effect of alcohol policy on sexually transmitted disease rates in the United States. *Journal of Law and Economics*, 43(1), 215-238.

Dee, T. (2008). Forsaking all others? The effects of same-sex marriage partnership laws on risky sex. *Economic Journal*, 118(530), 1055-1078.

Dillender, M. (2014). The Death of Marriage? The Effects of New Forms of Legal Recognition on Marriage Rates in the United States. *Demography*, 51(2), 563-585.

Dillender, M. (2015). Health Insurance and labor Force Participation: What Legal Recognition Does for Same-Sex Couples. *Contemporary Economic Policy*, 33(2), 381-394.

Eskridge, W. N. (1996). The Case for Same Sex Marriage: From Sexual Liberty to Civilized Commitment, New York: The Free Press.

Farley, T. A., Cohen, D. A., & Elkins, W. (2003). Asymptomatic sexually transmitted diseases: the case for screening. *Preventive medicine*, 36(4), 502-509

Flood, S., King, M., Rodgers, R., Ruggles, S., & Warren J. R., (2018). Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]. Minneapolis, MN: IPUM

Fong, G. T., Hyland, A., Borland, R., et al. (2006). Reductions in tobacco smoke pollution and increases in support for smoke-free public places following the implementation of comprehensive smoke-free workplace legislation in the Republic of Ireland: findings from the ITC Ireland/UK survey. *Tobacco Control, 15*(Suppl. 3), 51-58.

Francis, A., & Mialon, H. (2010). Tolerance and HIV. *Journal of Health Economics*, 29(2), 250-267.

Francis, A., Mialon, H., & Peng, H. (2012). In Sickness and in Health: Same-Sex Marriage Laws and Sexually Transmitted Infections. *Social Science and Medicine*, 75(8), 1329-1341.

Gallus, S., Zuccaro, P., Colombo, P., Apolone, G., Pacifici, R., Garattini, S., et al. (2006). Effects of new smoking regulations in Italy. *Annals of Oncology*, 17(2), 346-347.

Gonzales, G. (2015). Association of the New York State Marriage Equality Act with Changes in Health Insurance Coverage. *JAMA Research Letter*, 314(7), 727-728.

Goodman-Bacon, A. (2018). Difference-in-Differences With Variation in Treatment Timing. NBER Working Paper, #25018

Harrison, W. O., Hooper, M. R., Wiesner, P. J., Campbell, A. F., Karney, W. W., Reynolds, G. H., Jones, O. G., Holmes, K. K. (1979). A trial of minocycline given after exposure to prevent gonorrhea. *N Engl J Med*, 300(19), 1074–1078.

Hatzenbuehler, M. L., O'Cleirigh, C., Grasso, C., Mayer, K., Safren, S., & Bradford, J. (2012). Effects of Same-Sex Marriage Laws on Health Care Use and Expenditures in Sexual Minority Men: A Quasi-Natural Experiment. *American Journal of Public Health*, 102(2), 285-291.

Hatzenbuehler, M. L., McLaughlin, K. A., Keyes, K. M., & Hasin, D. S. (2010). The Impact of Institutional Discrimination on Psychiatric Disorders in Lesbian, Gay, and Bisexual Populations: A Prospective Study. *American Journal of Public Health*, 100(3), 452-459.

Herdt, G., & Kertzner, R. (2006). I Do, but I Can't: The Impact of Marriage Denial on the Mental Health and Sexual Citizenship of Lesbians and Gay Men in the United States. *Sexuality Research and Social Policy Journal of NSRC*, 3(1), 33-49.

Judson, R. A., & Owen, A. L. (1999). Estimating dynamic panel data models: a guide for macroeconomists. *Economics Letters*, 65(1), 9-15.

Korenromp, E. L., Sudaryo, M. K., de Vlas, S. J., Gray, R. H., Sewankambo, N. K., Serwadda, D., Wawer, M. J., & Habbema, J. D. (2002). What proportion of episodes of gonorrhoea and chlamydia becomes symptomatic? *International journal of STD & AIDS*, 13(2), 91-101.

Kotsadam, A., & Jakobsson, N. (2011). Do laws affect attitudes? An assessment of the Norwegian prostitution law using longitudinal data. *International Review of Law and Economics*, 31(2), 103-115.

Müller, C. (2002). An economic analysis of same-sex marriage. *German Working Papers in Law and Economics*, No. 2002-1-1045.

Owusu-Edusei K., et al. (2012). The estimated direct medical cost of selected sexually transmitted infections in the United States. *Sex Transm Dis* 2013, 40(3), 197-201.

Peterman T., Tian, L., Metcalf, C., Satterwhite, C. L., Malotte, C. K., DeAugustine, N., Paul, S. M., Cross, H., Rietmeijer, C. A., and Douglas, J. M., Jr. (2006). High incidence of new sexually transmitted infections in the year following a sexually transmitted infection: a case for rescreening. *Ann Intern Med*, 145(8), 564–572.

Philipson, T. and Posner, R. A. (1993). Private Choices and Public Health: The AIDS Epidemic in an Economic Perspective, Cambridge: Harvard University Press.

Posner, R. A. (1992). Sex and Reason, Cambridge: Harvard University Press.

Rauch, J. (2004). Gay Marriage: Why it is Good for Gays, Good for Straights, and Good for America, New York: Times Books.

Riggle, E. D. B., Rostosky, S. S., & Horne, S. G. (2010). Psychological Distress, Well-Being, and Legal Recognition in Same-Sex Couple Relationships. *Journal of Family Psychology*, 24(1), 82–86.

Soss, J., & Schram, S. F. (2007). A public transformed? Welfare reform as policy feedback. *American Political Science Review*, 101(1), 111-127.

Sprigg, P., & Dailey, T. (Eds.). (2004). Is Homosexuality a Health Risk? In, *Getting it Straight: What the Research Shows about Homosexuality*. Washington DC: Family Research Council.

Svallfors, S. (2010). Policy feedback, generational replacement and attitudes to state intervention: eastern and western Germany, 1990-2006. *European Political Science Review*, 2(1), 119-135.

Tang, H., Cowling, D. W., Lloyd, J. C., Rogers, T., Koumjian, K. L., Stevens, C. M., et al. (2003). Changes of attitudes and patronage behaviors in response to a smoke- free bar law. *American Journal of Public Health*, 93(4), 611-617.

Wight, R., LeBlanc, A. J., & Badgett, M. V. L. (2013). Same-Sex Legal Marriage and Psychological Well-Being: Findings from the California Health Interview Survey. *American Journal of Public Health*, 103(2), 339-346.

Wolfers, J. (2006). Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results. *American Economic Review*, 96(5), 1802-1820

Appendix A

Table 1.6: Heterogeneity in Relationship Between Legal Same-Sex Marriage and Chlamydia

Coefficient on Legal Same-Sex Marriage
Outcome is Chlamydia Rate (per 100,000 population)

	(1) Weekly Data	(2) Annual Data
Baselines		
Full Sample	-.449* (.244) {-5.3%}	1.593 (6.773) {0.6%}
By Sex (Annual Data, 1984-2016)		
Male	--	-8.920 (6.752) {-7.1%}
Female	--	8.693 (9.869) {2.1%}
By Age (Annual Data, 2000-2016)		
Age 25 to 34	--	-.709* (.377) {-21.9%}
Age 35 to 44	--	-1.240 (.813) {-12.9%}
Age 55+	--	-1.087 (.777) {-14.1%}
By State Characteristics		
States with rates below the median	-.021 (.293) {-0.3%}	5.975 (6.554) {2.6%}
States with rates above the median	-.755** (.286) {-7.9%}	-2.078 (10.443) {-0.7%}
Only states that did not have CU/DP prior to SSM	-.425 (.328) {-4.8%}	-1.899 (8.198) {-0.7%}
Only states that adopted SSM via court order	-.405 (.349) {-4.7%}	.305 (9.833) {0.0%}
Only states that adopted SSM via *Obergefell v. Hodges*	1.263 (.852) {13.1%}	325.941*** (75.488) {104.7%}

Notes: *, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively. See notes to Table 1.1. CU/DP refers to civil union/domestic partnership policy. Observations for weekly data range from 1995 through September 2017. Observations for annual data range from 1984 through 2016 unless stated otherwise.

Table 1.7: Robustness of Association Between Same-Sex Marriage and Chlamydia to Alternative Specifications
Coefficient on Legal Same-Sex Marriage
Outcome is Chlamydia Rate (per 100,000 population) Unless Otherwise Stated

	(1) Baseline	**(2) Without quadratic trends**	**(3) Without lagged dependent variable**	**(4) Without quadratic trends and lagged dependent variable**	**(5) With 2nd period lag of outcome**	**(6) Outcome is natural log of chlamydia rates**	**(7) Year and week FE replaced with year-by-week FE**
Weekly Data, 1995-2017							
SSM legal	-0.449*	-0.467	-0.439*	-0.464	-0.462*	0.019	-0.483*
	(0.244)	(0.279)	(0.243)	(0.277)	(0.245)	(0.040)	(0.248)
Pre-SSM mean	8.579	8.579	8.571	8.571	8.578	2.134	8.579
R-squared	0.247	0.243	0.248	0.244	0.247	0.505	0.279
N	56983	56983	57049	57049	56930	56983	56983
Annual Data, 1984-2016							
SSM legal	1.593	5.619	16.409	20.876**	3.428	0.084	--
	(6.773)	(5.964)	(10.940)	(10.042)	(6.446)	(0.057)	
Pre-SSM mean	277.797	277.797	267.867	267.867	285.724	0.902	--
R-squared	0.963	0.960	0.942	0.932	0.962	5.363	--
N	1516	1516	1577	1577	1464	1516	--

Notes: * and ** denote statistical significance at 10% and 5%, respectively. See notes to Table 1.1.

Table 1.8: Additional Robustness of Association Between Same-Sex Marriage and Chlamydia to Alternative Specifications

Coefficient on Legal Same-Sex Marriage

Outcome is Chlamydia Rate (per 100,000 population)

	(1) Baseline	(2) Recode California as off until 2013	(3) Drop California	(4) Drop Texas	(5) Drop Florida	(6) Drop New York	(7) Drop Pennsylvania
Weekly Data, 1995-2017							
SSM legal	-0.449*	-.456*	-0.392	-0.455*	-0.389	-0.390	-0.440*
	(0.244)	(.242)	(0.263)	(0.250)	(0.255)	(0.310)	(0.242)
Pre-SSM mean	8.579	8.579	8.557	8.456	8.673	8.563	8.632
R-squared	0.247	.247	0.251	0.249	0.238	0.240	0.246
N	56983	56983	55853	55853	55853	56083	55853
Annual Data, 1984-2016							
SSM legal	1.593	2.622	-0.080	2.398	3.141	1.900	2.640
	(6.773)	(7.026)	(6.635)	(6.923)	(7.145)	(8.470)	(6.607)
Pre-SSM mean	277.797	278.427	280.353	274.201	276.694	278.858	280.256
R-squared	0.963	.963	0.962	0.961	0.962	0.967	0.962
N	1516	1516	1484	1486	1495	1484	1486

Notes: * denotes statistical significance at 10%. See notes to Table 1.1.

Chapter 2

Does It Get Better? Recent Estimates of Sexual Orientation and Earnings in the United States

Christopher S. Carpenter and Samuel T. Eppink*

Abstract

Using 2013-2015 National Health Interview Survey data, we reproduce a well-documented finding that self-identified lesbians earn significantly more than comparable heterosexual women. These data also show – for the first time in the literature – that self-identified gay men also earn significantly more than comparable heterosexual men, a difference on the order of 10 percent of annual earnings. We discuss several possible explanations for the new finding of a gay male earnings premium and suggest that reduced discrimination and changing patterns of household specialization are unlikely to be the primary mechanisms.

*Carpenter is Professor of Economics, Vanderbilt University. Eppink is a PhD Candidate in the Department of Economics, Vanderbilt University. All errors are our own.

2.1 Introduction

Improved attitudes toward the lesbian, gay, bisexual, and transgender (LGBT) communities have been some of the most striking and rapid social changes in the United States in the past several decades. These improved attitudes are perhaps most evident in the well-documented shift in public attitudes regarding same-sex marriage: the proportion of adults in the US who favored same-sex marriage increased from 35 percent to 55 percent from 2001 to 2016, the year after the US Supreme Court granted nationwide legal access to same-sex marriage in *Obergefell v. Hodges* in 2015 (Pew Research Center 2016). And historical data from the General Social Survey suggest that these shifts in attitudes began in the early 1990s: while in 1991 fully 72 percent of adults considered homosexual behavior 'always wrong', the associated share reporting this view in 2010 fell to 44 percent. The share of adults saying homosexual behavior was 'not wrong at all' increased over this same period from 14 percent to 41 percent (Smith 2011).

A natural question in the presence of these major changes in public attitudes toward sexual minorities is whether their socioeconomic position has improved at the same time.[20] For example, it is natural to ask if labor market discrimination against sexual minorities has fallen with the advancement of LGBT rights and public attitudes. A large body of research has addressed the question of whether gay men, lesbians, and bisexuals have different earnings than heterosexual individuals, controlling for observable characteristics such as education, experience, and other demographic and job characteristics. Most of these studies find that gay men in the United States have lower employment rates than comparable heterosexual men while lesbians have higher employment rates than comparable heterosexual women. Among full-time workers, most studies find that gay men earn less than otherwise similar heterosexual men, while lesbians earn more than otherwise similar heterosexual women (Badgett 1995, Klawitter and Flatt 1998, Allegretto and Arthur 2000, Black et al. 2003, Antecol et al. 2008, Jepsen and Jepsen 2017, Klawitter 2015, and others).

Few of these studies have provided information about whether these differentials have changed over time as attitudes have improved, with some notable exceptions. First, Clarke and Sevak (2013) use the National Health and Nutrition Examination Surveys (NHANES) from 1988 to 2007 and find that while men who reported same-sex sexual behavior had lower household incomes than

[20] Our paper's title references a viral media campaign started by LGBT rights activist Dan Savage whose 2010 'It Gets Better' project prompted a wide variety of individuals (including LGBT and non-LGBT celebrities, politicians, and others) to post YouTube videos encouraging young LGBT people that 'it gets better'. The project was started in response to the epidemic of LGBT youth suicide.

comparable men who reported different-sex sexual behavior at the start of their sample period, this penalty changed into a significant income *premium* by the end of the sample period. They suggest that improved attitudes toward sexual minorities may explain this pattern. Second, Klawitter (2015) performs a meta-analysis of all published studies on sexual orientation and income that use data through 2012. Her results suggest that lesbian premia and gay male penalties in older data have become smaller and closer to zero in studies using more recent data.

Third, patterns of evidence from well-designed correspondence studies point to the possibility that bias against LGBT individuals in the job search process has gone down. Tilcsik (2011) performed an audit study in which he sent over 1,700 resumes to job advertisements in 2005. He found that randomly assigned 'gay' resumes (as signaled through participation in an LGBT organization) received significantly fewer callbacks than otherwise similar resumes without such treatment; the magnitude of the difference was about as large as the black/white callback difference documented in Bertrand and Mullainathan (2004). More recently, however, two other audit studies that manipulated sexual orientation in different ways – one with participation in an LGBT student organization and another via a Facebook profile – found no significant differences in callback rates for gay candidates compared to other candidates in 2010 or 2013, respectively (Acquisti and Fong 2016; Bailey et al.

2013). Acquisti and Fong (2016) explicitly note that improved attitudes toward LGBT individuals might explain differences in the results of Tilcsik (2011) and the more recent audit studies.

In this paper we contribute to this growing literature by using data from the 2013-2015 National Health Interview Surveys (NHIS). These data are – to our knowledge – the only large, recent, nationally representative sample with information on both sexual orientation and earnings at the individual level. Prior studies have used similar data restricted to individual states such as California as opposed to national samples (e.g., Carpenter 2005), have relied on samples of same-sex couples as opposed to individuals (e.g., Antecol et al. 2008), or have examined household income as opposed to earnings (e.g., Clarke and Sevak 2013). Moreover, the vast majority of published work in the US – including all the work cited above – relies on data from the 1990s and 2000s. Although our data do not permit us direct tests of how earnings differences related to a minority sexual orientation have changed over a long period of time (since the NHIS did not ask about sexual orientation in earlier years), they do provide a new and very recent estimate from a high quality, national sample to this burgeoning literature.

Our NHIS samples yield over 1,300 self-identified sexual minorities. These data also include detailed information on annual earnings, individual demographic characteristics, and job characteristics such as industry, occupation, job tenure, and firm size. We first replicate the literature's most common findings that gay men have lower employment rates and lesbians have higher employment rates than similarly situated heterosexual men and women, respectively. Next, we show that the 2013-2015 NHIS data reproduce the literature's other consistent finding that among full-time workers, lesbians have significantly higher earnings than similarly situated heterosexual women. Finally, we show that in these data self-identified gay men are also estimated to earn significantly *more* than similarly situated heterosexual men – a difference on the order of 10 percent of annual earnings. Our finding of a significant gay male earnings premium is – to our knowledge – the first such estimate in the literature, and we discuss and investigate several possible explanations for this finding. We argue that although there has likely been a reduction in the extent of labor market discrimination against gay men, this is unlikely to explain the overall patterns observed in the NHIS. We also discuss whether changes in household specialization or peculiarities of the NHIS data are likely to explain the gay male earnings premium.

This paper is organized as follows: Section II describes the NHIS data and the estimation framework. Section III presents the results, and Section IV offers a discussion and concludes.[21]

2.2 Data Description and Empirical Approach

Our data come from the 2013-2015 National Health Interview Surveys (NHIS). The NHIS is an annual survey of about 35,000 households in the United States. For our purposes, a key feature of these data is that the NHIS asked a sample adult in each household a direct question about sexual orientation. This improves on most prior work in the literature which has relied on less direct methods for identifying sexual minorities, such as same-sex sexual behavior (as in some public health surveys) or, more commonly, the presence of a cohabiting same-sex partner. Since people who do not engage in sexual relations can still identify as sexual minorities, and since non-partnered sexual minorities may have different outcomes than cohabiting partnered sexual minorities, our individual level data on self-reported sexual orientation offer a more comprehensive sample of the overall population of LGB individuals.

[21] We do not provide a detailed literature review, as several previous studies have described existing work in great detail (see, for example, Aksoy et al. 2016, forthcoming; Klawitter 2015; and others).

In the NHIS a sample adult in each household is asked: "Which of the following best represents how you think of yourself?" Response options for women include: 1: Lesbian or gay; 2: Straight, that is, not lesbian or gay; 3) Bisexual; 4) Something else; 5) I don't know the answer; and 7) Refused.[22] Approximately 2-3 percent of individuals 18 and older self-identified as gay, lesbian, or bisexual in each wave of the NHIS (Ward et al. 2013). This is similar to other large population-based surveys in the UK, US, and Canada (Joloza et al. 2010).

Individuals are also asked about their employment status, including whether they work full-time (defined as 35 hours or more per week). We also observe total earnings before taxes and deductions from all jobs and businesses in the prior calendar year which we define as annual earnings.[23] In addition to the critical questions on sexual orientation and earnings, the NHIS includes standard demographic characteristics such as sex, age, race/ethnicity, educational attainment, partnership/marital status, and the presence of children in the

---

[22] Response options for men were similar except they did not refer to 'lesbian'. Note that individuals who responded 'something else' or 'don't know' were further probed about the nature of those responses. These response are not included in the NHIS public use file, however, so we do not make use of them. The NHIS is a face-to-face survey with computer-aided personal interviewing (CAPI). Pilot testing by the National Center for Health Statistics showed no significant difference in sexual orientation responses by whether individuals were surveyed using CAPI or audio computer-assisted self-interviewing (ACASI). The sexual orientation question is asked in an 'Adult Selected Items' module that contains other questions deemed to be sensitive.

[23] Approximately 16 percent of individuals who are employed full time have missing data on earnings, which is fairly standard in surveys of this type. The NHIS imputes income for these individuals, but we restrict attention to individuals who gave a non-imputed response to the earnings question.

household. We restrict attention to individuals ages 25 to 64 to focus on individuals most likely to have completed their education.[24]

We first estimate the relationship between sexual orientation and employment by estimating linear probability models separately by sex.[25] These models take the form:

$$(2.1) \quad EMPLOYED_i = \alpha + \beta_1 X_i + \beta_2(GAY/LESBIAN)_i + \beta_3(BISEXUAL)_i + \varepsilon_i$$

where EMPLOYED is an indicator variable for having any employment or having full-time employment, depending on the model. X is a vector of demographic and job variables that (depending on the model) include: age and its square; education dummies (bachelor's degree or more, associate degree, some college, less than high school, don't know education, and refused education, with high school degree as the excluded category); race dummies (black only, American Indian or Alaskan Native only, Asian only, race group not releasable, and multiple race, with white as the excluded category);[26] a dummy variable for Hispanic ethnicity; relationship status dummies (widowed, divorced, separated, partnered, and marital

[24] In results not reported here but available upon request, we find that lowering our minimum age in the sample to 18 returns similar results.

[25] We drop a small share of observations that did not provide a valid employment status response.

[26] The race of NHIS respondents may be withheld due to respondent confidentiality or other reasons.

status missing, with never married as the excluded category)[27]; region dummies (Midwest, South, and Northeast, with West as the excluded category); and the presence of children in the household (indicators for the presence of children ages zero to five years old and children ages six to seventeen years old). We also include survey wave dummies and month of interview dummies in all models. Note that in this model the relevant excluded category for sexual orientation is composed of individuals who report a heterosexual orientation.[28] We estimate standard errors robust to heteroscedasticity.

To assess the relationship between sexual orientation and annual earnings we estimate earnings models separately for males and females among the sample of full-time workers, following the prior literature. These models take the form:

$$\text{(2.2)} \quad \text{LOG EARNINGS}_i = \alpha + \beta_1 X_i + \beta_2(\text{GAY/LESBIAN})_i + \beta_3(\text{BISEXUAL})_i + \varepsilon_i$$

[27] Partnership is based on a dummy variable indicating the person is in any type of partnership (married or living with a partner). This accounts for the fact that legal access to same-sex marriage for sexual minorities in our sample was not universal throughout the sample period under study. Of course, individuals can still describe themselves as 'married' even if they are not legally married, though we have no way of identifying these individuals, regardless of the sexual orientation of the respondent.

[28] In all models we separately include dummy variables for people who refused to provide a response to the sexual orientation question, or who reported 'something else' or 'I don't know'. Demographic characteristics for these respondents are reported in Appendix B Table 2.4 and reveal that both males and females across these groups tend to be less educated, are less likely to be partnered, and are less likely to have children compared to self-identified heterosexuals. The coefficients on these indicators in the earnings regressions are reported in Appendix B Table 2.5.

where all variables are as described above. In these models we also add to the X vector: the number of years of job tenure at the current firm (and its square); a series of 26 occupation and 24 industry dummy variables; firm size categories; and dummy variables for the sector of employment. The earnings models also include a dummy variable for whether the respondent's personal earnings or job tenure responses were topcoded.[29]

2.3 Results

Table 2.1 presents descriptive statistics for demographic and employment characteristics from the NHIS data broken down by self-reported sexual orientation and gender.[30] Self-identified lesbians are significantly more likely to have a bachelor's degree, less likely to have children in the household, more likely to be full time workers, and have higher average annual earnings than heterosexual women. Bisexual women are significantly younger, less likely to be partnered, less likely to live in the Northeast, and less likely to have children in the household than heterosexual women.

[29] The NHIS topcoded earnings at $150,000, $200,000, and $250,000 in 2013, 2014, and 2015, respectively. A model predicting the likelihood the individual's earnings response was topcoded showed that sexual orientation was not significantly related to the likelihood of being topcoded for women, though gay men were 3 percentage points more likely to have a topcoded earnings response. In all earnings models individuals with topcoded earnings were recoded to the median of the US earned income distribution above the topcode cutoff for their respective year. US earned income distributions for each year were constructed using IPUMS ACS data.

[30] We use the subsample of the NHIS respondents ages 25-64 for which we have earnings information.

Table 2.1: **Descriptive Statistics (among those with earnings information)**
2013-2015 NHIS, Adults ages 25-64

Variable	(1) Heterosexual women	(2) Bisexual women	(3) Lesbians	(4) Heterosexual men	(5) Bisexual men	(6) Gay men
Age	43.5 (11.1)	36.7[A] (10.6)	43.1 (10.5)	43.2 (11.1)	38.7[B] (13.1)	42.0 (12.1)
BA or more	.409 (.491)	.468 (.528)	.518[A] (.505)	.364 (.481)	.436 (.573)	.495[B] (.567)
Associate degree	.144 (.351)	.107 (.328)	.143 (.353)	.122 (.327)	.130 (.389)	.086[B] (.318)
Some college	.182 (.386)	.260 (.464)	.154 (.365)	.167 (.373)	.218 (.476)	.206 (.458)
High school degree	.191 (.393)	.114[A] (.336)	.161 (.371)	.241 (.428)	.138[B] (.399)	.177[B] (.433)
Less than high school degree	.059 (.236)	.036 (.197)	.017[A] (.130)	.084 (.277)	.028[B] (.190)	.031[B] (.197)
White	.784 (.411)	.829 (.398)	.812 (.395)	.812 (.390)	.837 (.427)	.855[B] (.399)
Partnered (living with a partner or married)	.668 (.471)	.462[A] (.528)	.661 (.479)	.744 (.436)	.406[B] (.567)	.474[B] (.566)
Any children in household	.471 (.499)	.339[A] (.501)	.275[A] (.451)	.443 (.496)	.302[B] (.530)	.078[B] (.304)
Northeast	.174 (.379)	.113[A] (.336)	.184 (.391)	.164 (.370)	.126 (.384)	.189 (.444)
Midwest	.230 (.421)	.235 (.448)	.202 (.406)	.237 (.425)	.247 (.498)	.152[B] (.407)
South	.369 (.482)	.399 (.518)	.352 (.483)	.362 (.480)	.310 (.534)	.367 (.547)
West	.227 (.419)	.253 (.460)	.262 (.445)	.237 (.425)	.317 (537)	.292[B] (.516)
Avg. Annual Earnings	39,902.80 (32,871.67)	38,802.90 (38,528.90)	47,026.12[A] (36,827.92)	57,032.58 (42,814.36)	49,766.34 (49,043.15)	59,618.16 (51,042.92)
Full-time worker	.732 (.443)	.737 (.465)	.812[A] (.394)	.871 (.334)	.731[B] (.512)	.843 (.412)
Sample Size	22,337	252	426	21,444	118	540

Weighted means (standard deviations). Not reported here (but included in the earnings models) are 721 females and 639 males, who when asked about sexual orientation, responded 'something else' or 'don't know', refused a response, or otherwise have missing data on sexual orientation. [A] The superscript letter A means statistically significant difference ($P < 0.05$) between the groups of lesbians and bisexual women in contrast to the heterosexual women. [B] The superscript letter B means statistically significant difference ($P < 0.05$) between the groups of gay men and bisexual men in contrast to the heterosexual men.

Self-identified gay men are significantly more likely to have a bachelor's degree, more likely to be white, less likely to be partnered, less likely to have any children in the household, more likely to live in the West, and less likely to live in the Midwest than heterosexual men.[31] Self-identified bisexual men are significantly younger, less likely to be partnered, less likely to have children in the household, and less likely to be full-time workers than heterosexual men.

In Table 2.2 we examine the relationship between sexual orientation, employment, and earnings. Columns 1 and 2 examine the likelihood of any employment, columns 3 and 4 examine the likelihood of full-time employment, and columns 5 and 6 examine log annual earnings among full-time workers. Odd numbered columns include only the controls for sexual orientation, month dummies, and survey wave dummies; even numbered columns add the demographic characteristics (age, education, race/ethnicity, relationship status, region, and the presence of children in the household), and in column 6 we also add job characteristics (job tenure, firm size, occupation, and industry controls). We estimate models separately for females in the top panel and for males in the bottom panel.

[31] These broad patterns of demographic characteristics replicate most of the patterns from credible population datasets (see, for example, Black et al. 2000). Note that Table 2.1 shows that a larger proportion of lesbians reports being partnered compared to gay men (66 percent of lesbians versus 47 percent of gay men). This pattern – that the lesbian partnership rate is similar to the partnership rate of heterosexual women and that the gay male partnership rate is substantially lower than the partnership rate of heterosexual men – has been replicated in several datasets (see, for example, Carpenter and Gates 2008 and Aksoy et al. 2016, forthcoming).

Table 2.2: **Sexual Orientation, Employment, and Earnings**
2013-2015 NHIS, Adults ages 25-64

	(1) Any employment	(2) Any employment	(3) Full time employment	(4) Full time employment	(5) Log annual earnings, among FT employed	(6) Log annual earnings, among FT employed
Females						
Lesbian	.091*** (.024)	.042* (.022)	.113*** (.027)	.059** (.025)	.140*** (.047)	.086** (.043)
Bisexual	.013 (.031)	-.024 (.030)	.001 (.035)	-.035 (.034)	-.081 (.109)	-.031 (.092)
R-squared	.00	.10	.00	.03	.10	.35
N	38,353	38,353	38,081	38,081	17,016	17,016
Males						
Gay	-.051** (.022)	-.047** (.022)	-.065*** (.024)	-.054** (.023)	.077* (.042)	.097** (.038)
Bisexual	-.104** (.051)	-.084* (.048)	-.150*** (.053)	-.119** (.050)	-.062 (.094)	-.021 (.068)
R-squared	.00	.13	.00	.14	.18	.35
N	32,247	32,247	31,975	31,975	18,981	18,981
Controls:						
Sexual orientation dummies	Yes	Yes	Yes	Yes	Yes	Yes
Month & year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Demographics		Yes		Yes		Yes
Job characteristics						Yes

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Models in columns 1, 3, and 5 control for: sexual orientation dummies (gay, bisexual, other sexual orientation, don't know sexual orientation, refused sexual orientation, and missing sexual orientation, with heterosexual as the excluded category); month of interview dummies; and survey wave dummies. Models in columns 2, 4, and 6 additionally control for: age and its square; race dummies (indicators for black only, American Indian or Alaskan Native only, Asian only, race group not releasable, and multiple race, with white as the excluded category); Hispanic ethnicity; education dummies (less than high school degree, some college, associate's degree, BA or more, don't know educational attainment, and refused to provide educational attainment, with high school degree as the excluded category); relationship/marital status (widowed, divorced, separated, partnered [married or living with a partner], and missing marital status, with never married as the excluded category); a dummy variable for any children 0-5 in the household; a dummy variable for any children 6-17 in the household; and region dummy variables (Northeast, Midwest, and South, with West as the excluded category). Additional controls in column 6 include: number of years of job tenure and its square; dummy variables for firm size (10-24, 25-49, 50-99, 100-249, 250-499, 500-999, 1000 or more, don't know firm size, refused to provide firm size, and missing firm size, with less than 10 workers as the excluded category); sector of employment (indicators for public sector, don't know sector, refused to provide sector, and missing sector, with private sector as the excluded category); 24 industry dummies; and 26 occupation dummies. All estimates are from OLS models with NHIS sample weights, and standard errors in parentheses are robust to heteroscedasticity.

In the top panel of columns 1 and 2 of Table 2.2 we find that lesbians are significantly more likely to have any employment than similarly situated heterosexual women, a difference on the order of 4.2 percentage points in the fully saturated model of the top panel of column 2. Estimates for bisexual women are not statistically significant. Turning to full time employment in the top panel of columns 3 and 4 of Table 2.2 we find qualitatively identical patterns to those for any employment in columns 1 and 2: lesbians are 5.9 percentage points more likely than otherwise similar heterosexual women to be in full-time work after controlling for detailed observable characteristics, while differences for bisexual women are smaller and not statistically significant. For women's annual earnings among full-time workers in columns 5 and 6 of the top panel of Table 2.2 we confirm another of the literature's most consistent findings: lesbians have significantly higher annual earnings than similarly situated heterosexual women, conditional on full-time work. In the fully saturated model we estimate the lesbian earnings difference to be about 9 percent. Bisexual women are estimated to have slightly lower annual earnings than comparable heterosexual women, though the estimate is not statistically significant. Overall, the results in the top panel of Table 2.2 concur with a large body of prior work showing that lesbians supply more labor than heterosexual women and have higher annual earnings conditional on full-time work (see, for example, Tebaldi and Elmslie 2006 and Antecol and Steinberger 2013).

For males in the bottom panel of Table 2.2 we find that both gay and bisexual men are significantly less likely to be in any employment (columns 1 and 2), and full-time employment (columns 3 and 4) than otherwise similar heterosexual men. Moreover, these differences are large: gay men are estimated to be 5.4 percentage points less likely than comparable heterosexual men to be in full-time work in the fully saturated model of the bottom panel of column 4 of Table 2.2, while bisexual men are 11.9 percentage points less likely to be in full-time work. Both estimates are statistically significant at conventional levels. Turning to earnings among men in full-time work in columns 5 and 6, we find that gay men are estimated to have significantly *higher* annual earnings than comparable heterosexual men, a difference on the order of 10 percent in the fully saturated model of column 6.[32] The associated estimate for bisexual men is smaller and statistically insignificant.[33] While some prior work has also found that sexual minority men have lower employment than heterosexual men, to our knowledge the finding in the bottom panel of column 6 of Table 2.2 is the first in the literature to find that gay men have significantly *higher* annual earnings than comparable heterosexual men.

[32] Appendix B Table 2.5 provides an expanded set of coefficient estimates from the fully saturated model (column 6 of Table 2.2). It shows that annual earnings are positively associated with age, education, white race, being partnered, residing in the Northeast, longer job tenure and working for larger firms. There is a wage penalty associated with having children in the household for women but not for men. These patterns are consistent with a large body of prior work.

[33] The estimated annual earnings premia for lesbians and gay men were qualitatively similar when we included 18-24 year olds or included individuals whose earnings had been imputed. These results are available upon request.

In Table 2.3 we investigate heterogeneity in the earnings differences shown in column 6 of Table 2.2 for lesbians and gay men compared to otherwise similar heterosexual women and men, respectively. Each entry in Table 2.3 is the coefficient on the GAY/LESBIAN indicator from a separate regression restricted to the sample described in each row (the models also include the other sexual orientation, demographic, and job controls as well, though they are not reported). The results in Table 2.3 for women indicate larger lesbian earnings premia in samples of white women, women in the Midwest, women working at smaller firms, and women in the private sector. For men, we estimate larger gay male earnings premia in samples of older men, whites, non-partnered men, men in the private sector, and men working at the largest firms. Few of the differences across groups are statistically significant, however, owing to small samples. We note that the bottom set of estimates in Table 2.3 shows that the lesbian earnings premium is largest in 2014, while the gay male earnings premium is largest in 2013, though these across-year differences again are not statistically different from each other. These patterns clarify that gay men (and lesbians) have not improved their economic position relative to heterosexuals in a year-by-year sense, but rather that they have had significantly higher earnings over the pooled 2013-2015 period.

Table 2.3: Heterogeneity in Log Annual Earnings Gaps
Coefficient on Gay/Lesbian for various Subsamples
2013-2015 NHIS, Fully Saturated Specification, Adults ages 25-64, Full Time Workers

	(1) Females (coefficient on Lesbian)	(2) Males (coefficient on Gay)
Baseline – Table 2.2, Col 6 ($N_{lesbian}$=340; N_{gay}=434)	.086** (.043)	.097** (.038)
25-44 year olds ($N_{lesbian}$=174; N_{gay}=254)	.088 (.054)	.038 (.054)
45+ year olds ($N_{lesbian}$=166; N_{gay}=180)	.076 (.066)	.182*** (.053)
Whites ($N_{lesbian}$=268; N_{gay}=358)	.116** (.046)	.111*** (.040)
Nonwhites ($N_{lesbian}$=72; N_{gay}=76)	-.047 (.100)	-.014 (.117)
At least a BA ($N_{lesbian}$=178; N_{gay}=236)	.062 (.058)	.101* (.053)
Less than a BA ($N_{lesbian}$=162 ; N_{gay}=198)	.096 (.062)	.106* (.061)
Partnered ($N_{lesbian}$=182; N_{gay}=146)	.086 (.055)	.053 (.049)
Not partnered ($N_{lesbian}$=158 ; N_{gay}=288)	.073 (.065)	.146*** (.056)
Northeast ($N_{lesbian}$=71; N_{gay}=82)	-.092 (.144)	.111 (.068)
Midwest ($N_{lesbian}$=62; N_{gay}=54)	.203*** (.058)	.090 (.090)
South ($N_{lesbian}$=112; N_{gay}=168)	.086 (.056)	.105* (.057)
West ($N_{lesbian}$=95; N_{gay}=130)	.079 (073)	.092 (.089)
Public sector ($N_{lesbian}$=79; N_{gay}=65)	.020 (.114)	.091* (.054)
Private sector ($N_{lesbian}$=261; N_{gay}=368)	.111*** (.042)	.103** (.043)
At least 500 workers at firm ($N_{lesbian}$=61; N_{gay}=105)	.021 (.056)	.154*** (.052)
Fewer than 500 workers at firm ($N_{lesbian}$=271; N_{gay}=322)	.104** (.051)	.077 (.049)
2013 ($N_{lesbian}$=116; N_{gay}=158)	.012 (.082)	.209*** (.056)
2014 ($N_{lesbian}$=110; N_{gay}=140)	.178*** (.059)	.064 (.065)
2015 ($N_{lesbian}$=114; N_{gay}=136)	.074 (.062)	-.000 (.081)

See notes to Table 2.2. Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.

2.4 Discussion and Conclusion

Our main objective in this paper was to provide a new and recent estimate of the association between sexual orientation and earnings using high quality, nationally representative data with individual level information on sexual orientation and earnings from the 2013-2015 National Health Interview Surveys. These data have not been used previously in the growing literature on sexual orientation and economic outcomes. We first documented that these data reproduce the literature's most consistent findings that lesbians have higher employment rates and higher earnings than comparable heterosexual women, while gay men have lower employment rates than comparable heterosexual men. The NHIS data also indicate, however, that gay men earn significantly higher wages than comparable heterosexual men, a difference on the order of 10 percent of annual earnings. To our knowledge, this is the first estimate in the literature that finds a significant gay male earnings premium using population representative data on self-reported sexual orientation and earnings.

What might explain the patterns observed in the NHIS, particularly with respect to the fact that these data are the first to uncover a robust gay male earnings premium? One possibility is simply that the NHIS data on sexual orientation and/or earnings are incorrect or otherwise idiosyncratic. This explanation is unlikely given that the data return estimates of the proportion of self-identified gay men, lesbians, and bisexuals – approximately 2-3 percent of the population – that are well in line with other credible population-based survey datasets that have been used extensively in the literature and that have returned the usual pattern of results (i.e., lesbian earnings premia and gay male earnings penalties) such as the UK Integrated Household Surveys (Aksoy et al. 2016 forthcoming), the Household, Income and Labour Dynamics in Australia (HILDA) study (Sabia et al. 2017, forthcoming), the Canadian Community Health Surveys (Carpenter 2008), and others. Moreover, the sexual orientation data in the NHIS has already been used for numerous publications in public health and medicine (see, for example, Ward et al. 2014, Jackson et al. 2016, and others) and has been the subject of several technical reports establishing the integrity of the sexual orientation data (see, for example, Dahlhamer, Galinsky, et al. 2014). Further evidence that the data are internally valid comes from Appendix B Table 2.2, which shows the expanded set of coefficient estimates from the fully saturated earnings model of column 6 of Table 2.2. Appendix B Table 2.2 shows that the NHIS data return reasonable and sensible estimates of the associations of education, age, race/ethnicity, and other characteristics with annual earnings; these estimates are in line with those from datasets that are more commonly used to estimate earnings models such as the Current Population Surveys or the American Community Surveys.

Somewhat related to the issue of data quality are the possibilities that our data and specification choices are related to the gay male earning premium. Recall that most prior work finding large gay male earnings penalties – in addition to using older data – also generally relied on individual information about same-sex sexual behavior or restricted attention to individuals in same-sex relationships due to data limitations. Studies using individual level reports of self-reported sexual orientation have generally found smaller earnings differences for gay men compared to heterosexual men (Carpenter 2005), and there is also evidence that partnership-based samples overstate the gay male earnings penalty found in prior work (Carpenter 2008). Our finding in Table 2.3 that the gay male earnings premium is larger in samples of non-partnered men supports this general pattern, though it remains true that no prior work has found evidence of a gay male earnings *premium*.

Regarding sample and specification choices, we focus on annual earnings among full-time workers as our main outcome. While the large majority of prior work also focuses on full-time workers, outcomes in the literature have ranged from annual personal income to hourly earnings, with studies focusing on labor market discrimination tending to focus on the latter when it is available. Our choice of annual earnings reflects the fact that hours information is missing for a large number of NHIS respondents. More research is needed from datasets with multiple earnings and income measures as well as information on labor effort to determine whether the literature's pattern of earnings estimates is systematically related to the type of outcome studied, though the remarkable stability of the lesbian earnings premium across different measures in the literature is suggestive that these measurement choices are unlikely to explain our finding of a gay male premium.

Another candidate explanation for the first estimate of a gay male earnings premium is the rapid improvement in attitudes toward the LGBT community over the past decade which has been coupled by major changes in public policies toward sexual minorities such as same-sex marriage legalization and increasing prevalence of non-discrimination policies in employment. In fact, Clarke and Sevak (2013) suggest these changing attitudes as a possible reason for their findings in the NHANES data of a statistically significant increasing trend in relative household incomes for men reporting same-sex sexual behavior compared to men not reporting same-sex sexual behavior from the 1990s into the 2000s.

But while this explanation of improving attitudes has intuitive appeal, there are several challenges with it as well. First, it is not clear why improving attitudes toward LGBT people would produce a gay male earnings *premium*. While we might have expected that the well-documented gay male earnings penalty would be reduced or even eliminated as compliance with non-discrimination policies increased and attitudes toward LGBT people improved, it is not clear how these factors would result in gay men earning significantly *more* than comparable heterosexual men. Second, and related to the first point, the NHIS data continue to indicate that gay men have significantly lower employment rates than comparable heterosexual men. To the extent that the lower employment partly reflects discrimination against gay men, it is hard to imagine earnings improving substantially but not employment. Third, the explanation of improving attitudes toward LGBT people is hard to square with the fact that our estimated lesbian earnings premium is right in line with prior estimates from different and older datasets in the United States (i.e., it is not substantially larger). That is, it seems unlikely that the LGBT civil rights movement would have substantially improved labor market outcomes for gay men but not for lesbians. Fourth, the pattern of progress for gay rights has not been universally positive for sexual minorities. As in many civil rights movements, there has been some backlash to the speed with which sexual minorities have achieved equality in the eyes of the law. For example, there were substantial increases in LGBT-related harassment reported to governments and police agencies in the wake of major policy rulings on same-sex marriage, and there is still pervasive anti-LGBT sentiment throughout the United States. Fifth, recent estimates of the association between sexual orientation and earnings using high quality data from other countries that have experienced similar improvements in attitudes toward the LGBT community do not show a similar pattern of a gay male premium. Aksoy et al. (2016, forthcoming), for example, find that self-identified gay men earn very similar wages to comparable heterosexual men in the United Kingdom from 2012-2014 (i.e., no earnings penalty and no earnings premium) despite that lesbians command a statistically significant 5.5 percent premium. These patterns are difficult to square with the simple hypothesis of reduced discrimination against sexual minorities.[34]

A related set of hypotheses pertains to selection into who identifies as a sexual minority to surveys. This is related to changing and improving attitudes toward the LGBT community, as it would increase the likelihood that some individuals would be willing to 'come out' to a survey interviewer about their sexuality. If these individuals are also increasingly likely to come out to family, friends, coworkers, and employers, this could also cause earnings patterns to vary, particularly if the unobserved characteristics associated with the changing nature of selection are systematically related to earnings potential.

It is, of course, nearly impossible to know the nature of selection into 'coming out' to a survey interviewer. Some researchers have argued that the only people who can afford to be out about their sexuality are people with high

[34] A variant of the 'improved attitudes' hypothesis is that the emergence of a gay male premium in these recent data could reflect across-cohort differences in experiences of discriminatory attitudes across the lifecycle. The fact that we estimate a larger gay male earnings premium in samples of older men compared to younger men (Table 2.3) is largely inconsistent with this hypothesis, however, since the older gay men would have experienced more discrimination in their early adulthood than the more recent cohorts.

education and earnings (and other unobservable characteristics positively associated with earnings potential). If so, the changing nature of selection would be inconsistent with relative improvements in gay male earnings compared to heterosexual male earnings. It is also possible, however, that high earning sexual minorities have the 'most to lose' by coming out about their sexual orientation to survey administrators and/or to employers. If so, changing selection based on improving attitudes could produce the patterns observed above for gay men. A remaining challenge with the 'changing nature of selection' hypothesis, however, is that – like the changing attitudes hypothesis more generally – it is generally inconsistent with the fact that our estimated lesbian earnings premium is right in line with prior estimates based on much older data. Finally, we note that there are key patterns that are strongly inconsistent with the selectivity hypothesis, most notably that the share of adults identifying as gay, lesbian, or bisexual in the 2013-2015 NHIS is not noticeably higher than in other, older surveys. Since increased willingness to identify as a sexual minority individual to a survey interviewer would have predicted a noticeable increase in the share self-identifying as LGB, this explanation seems unsatisfying.

It is clear that any hypothesis for the patterns we observe has to rationalize why the improvement in outcomes is observed much more strongly for gay men than for lesbians.[35] One such explanation that fits with some of these patterns is the changing nature of market-based specialization within households. Increasing access to same-sex marriage has given same-sex couples the same legal rights and responsibilities as different-sex couples. It is plausible that these legal changes differentially affected gay men compared to lesbians. Prior work shows that even in the absence of same-sex marriage, lesbians were more likely to be in same-sex partnerships than gay men, and among those in partnerships, lesbians were much more likely to formalize their partnership by registering with the government than gay men (Carpenter and Gates 2008). This could partly reflect the fact that lesbians had more to gain from official recognition because their households were much more likely to contain children (including children from prior heterosexual relationships) than gay men in same-sex relationships. Evidence on take-up of legal marriage among same-sex couples is very limited; Carpenter (2016) finds that when the Massachusetts Supreme Court legalized same-sex marriage in 2004 it induced large increases in marriage for lesbians and modest increases for gay

[35] We should note that even if discriminatory attitudes were reduced for both gay men and lesbians, this need not necessarily imply that the relative earnings position of those groups should have improved over time compared to their same-gender heterosexual counterparts. If, for example, part of the historically large gay male earnings penalty is due to a unique distaste for gay men (as opposed to a distaste for sexual minorities in general), then it could be that reductions in discrimination would be observed in a relative improvement in earnings for gay men compared to straight men but not for lesbians compared to straight women. Put differently, it could be that lesbians are somewhat immune from the large gender-based labor market penalty experienced by heterosexual women and that sexual orientation does not play a strong role in determining lesbian earnings differentials. Thus, the seemingly differential nature of the evolution of the gay male earnings differential compared to the lesbian earnings differential need not be entirely inconsistent with a role for changing attitudes and discrimination. Of course, why such dynamics would produce significant premia (as opposed to simply reducing penalties) remains unclear.

men. We are not aware of any published evidence on more recent legalizations of same-sex marriage in other states or from federal Supreme Court decisions in *Windsor* and *Obergefell*. But it is possible that increased legal access to marriage induced greater changes in gay coupling behavior than in lesbian coupling behavior, and moreover among couples it is plausible that recent legal changes induced more substantive changes to gay male couples' households than to lesbian households to the extent that lesbian households already functioned effectively as a 'married' unit. These Becker-based dynamics make it possible that changing legal access to gay marriage which occurred in our sample period would have induced larger changes in home versus market-based specialization within gay male households than the associated change in lesbian households.

The data produce some patterns consistent with this hypothesis. First, recall that gay men have significantly lower employment rates than otherwise similar heterosexual men. This would be expected if gay men are specializing more in the wake of legal access to same-sex marriage, as half of the gay male partners could be specializing relatively in home production. Second, the rate of gay male partnership in the NHIS (approximately 45 percent) is somewhat higher than in other datasets. This is consistent with increasing legal access to same-sex marriage increasing partnership among gay men.[36] Notably, the estimated rate of lesbian partnership in the NHIS (about 66 percent) is in line with prior published estimates from credible datasets. Of course, the one pattern that is entirely inconsistent with a hypothesis about the changing nature of household specialization is that the gay male premium is observed primarily in samples of non-partnered men (see Table 2.3). Since household specialization-based theories for the emergence of a gay male premium rely on the presence of a partner (or, in its most generous interpretation, on the expectation of a partner), however, this finding is broadly inconsistent with the household specialization hypothesis.

Our paper calls for more research and data collection on sexual orientation in high quality datasets that also include information on economic outcomes. As of the time of this writing, the NHIS is to our knowledge the only ongoing large federal survey available to researchers to include both a direct question about sexual orientation and information on earnings. Adding a sexual orientation question to large datasets such as the Current Population Surveys, the National Longitudinal Surveys of Youth, the Panel Study of Income Dynamics, the Health and Retirement Study, and others could yield fundamental insights into the basic structure of economic relationships. And understanding whether a gay male

[36] The public use NHIS data do not identify individual states, so we cannot directly test for how marriage equality affected partnership or marriage in these data, as the policy came into effect in different states at different times.

premium in adult earnings can be replicated across other recent datasets – and what might be causing it – should be important priorities for future work.

1.5 References

Acquisti, Alessandro and Christina Fong (2016). "An Experiment in Hiring Discrimination Via Online Social Networks," SSRN Working Paper #2031979.

Aksoy, Cevat G., Christopher S. Carpenter, and Jeff Frank (2016, forthcoming). "New Evidence on Sexual Orientation and Earnings: Evidence from the UK," Industrial and Labor Relations Review, forthcoming.

Allegretto, Sylvia and Michelle Arthur (2001). "An Empirical Analysis of Homosexual/ Heterosexual Male Earnings Differentials: Unmarried and Unequal?," Industrial and Labor Relations Review, 54(3): 631-646.

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

Antecol, Heather and Michael D. Steinberger (2013). "Labor Supply Differences Between Married Heterosexual Women and Partnered Lesbians: A Semi-Parametric Decomposition Approach," Economic Inquiry, 51(1): 783-805.

Badgett, M.V. Lee (1995). "The Wage Effects of Sexual-Orientation Discrimination," Industrial and Labor Relations Review, 48(4): 726-739.

Bailey, John, Michael Wallace, and Bradley Wright (2013). "Are Gay Men and Lesbians Discriminated Against When Applying for Jobs? A Four-City, Internet-Based Field Experiment," Journal of Homosexuality, 60(6): 873-894.

Black, Dan, Gary Gates, Seth Sanders, and Lowell Taylor (2000). "Demographics of the Gay and Lesbian Population in the United States: Evidence from Available Systematic Data Sources," Demography, 37(2): 139-154.

Black, Dan, Hoda Makar, Seth Sanders, and Lowell Taylor (2003). "The Earnings Effects of Sexual Orientation," Industrial and Labor Relations Review, 56(3): 449-469.

Carpenter, Christopher (2016). "The Direct Effects of Legal Same-Sex Marriage: Evidence from Massachusetts," working paper.

------ (2008). "Sexual Orientation, Work, and Income in Canada," Canadian Journal of Economics, 41(4): 1239-1261.

----- (2005). "Self-Reported Sexual Orientation and Earnings: Evidence from California," Industrial and Labor Relations Review, 58(2): 258-273.

Carpenter, Christopher, and Gary J. Gates (2008). "Gay and lesbian partnership: Evidence from California," Demography 45(3): 573-590.

Clarke, Geoffrey and Purvi Sevak (2013). "The disappearing gay income penalty," Economics Letters, 121: 542-545.

Dahlhamer, James M., Ward, Brian W., Adena M. Galinsky, and Sarah S. Joestl (2014). "Sexual Orientation Results from the National Health Interview Survey." Presentation to the Board of Scientific Counselors Meeting, October 29-30, 2014. Available at: https://www.cdc.gov/nchs/data/bsc/bscpres_dahlhamer_oct2014.pdf.

Dahlhamer, James M, Adena M. Galinsky, Sarah S. Joestl, and Brian W. Ward (2014). "Sexual orientation in the 2013 national health interview survey: a quality assessment," Vital and Health Statistics Series 2, 169: 1-32.

Jackson, Chandra L., Madina Agenor, Dayna A. Johnson, S. Bryn Austin, and Ichiro Kawachi (2016). "Sexual orientation identity disparities in health behaviors, outcomes, and services use among men and women in the United States: a cross-sectional study," BMC Public Health 16: 807.

Jepsen, Christopher and Lisa Jepsen (2017). "Self-employment, earnings, and sexual orientation," Review of Economics of the Household, 15(1): 287-305.

Joloza, Theodore, Joanne Evans, Rachel O'Brien, and Angela Potter-Collins (2010). "Measuring Sexual Identity: An Evaluation Report." UK Office for National Statistics.

Klawitter, Marieka (2015). "Meta-analysis of the Effects of Sexual Orientation on Earnings," Industrial Relations: A Journal of Economy and Society, 54(1): 4-32.

Klawitter, Marieka and Victor Flatt (1998). "The Effects of State and Local Antidiscrimination Policies on Earnings for Gays and Lesbians," Journal of Policy Analysis and Management, 17(4): 658-686.

Pew Research Center (2016). "Changing Attitudes on Gay Marriage: Public opinion on same-sex marriage," Pew Research Center Report, May 12, 2016. Available at: http://www.pewforum.org/2016/05/12/changing-attitudes-on-gay-marriage/. Last accessed December 21, 2016.

Sabia, Joseph, Mark Wooden, and Thanh Tam Nguyen (2017). "Sexual Identity, Same-Sex Relationships, and Labour Market Dynamics: New Evidence from Longitudinal Data in Australia," Southern Economic Journal, forthcoming.

Smith, Tom W. (2011). "Public Attitudes toward Homosexuality," Report of the National Opinion Research Council. Available at: http://www.norc.org/PDFs/2011%20GSS%20Reports/GSS_Public%20Attitudes%20Toward%20Homosexuality_Sept2011.pdf. Last accessed December 21, 2016.

Tebaldi, Edinaldo and Bruce Elmslie (2006). "Sexual orientation and labour supply," Applied Economics 38(5): 549-562.

Tilscik, Andras (2011). "Pride and Prejudice: Employment Discrimination Against Openly Gay Men in the United States," American Journal of Sociology, 117(2): 586-626.

Ward, Brian W., James M. Dahlhamer, Adena M. Galinsky, and Sarah S. Joestl (2014). "Sexual orientation and health among U.S. adults: National Health Interview Survey, 2013. National health statistics reports; no 77. Hyattsville, MD: National Center for Health Statistics.

Appendix B

Table 2.4: **Descriptive Statistics for the Other Sexual Orientation Response Categories (among those with earnings information)**

2013-2015 NHIS, Adults ages 25-64

How the sexual orientation question was answered ➔	**(1) Women Something else**	**(2) Women Don't know**	**(3) Women Refused**	**(4) Women No response**	**(5) Men Something else**	**(6) Men Don't know**	**(7) Men Refused**	**(8) Men No response**
Age	42.7 (12.5)	45.6 (14.4)	48.4[A] (11.3)	43.7 (11.2)	40.7 (12.3)	43.0 (13.8)	45.3 (12.7)	41.8 (10.3)
BA or more	.376 (.544)	.320 (.532)	.446 (.524)	.341[A] (.470)	.315 (.516)	.181[B] (.432)	.480 (.580)	.312 (.440)
Associate degree	.121 (.367)	.057[A] (.264)	.144 (.355)	.135 (.338)	.110 (.348)	.152 (.403)	.192 (.457)	.104 (.289)
Some college	.247 (.484)	.169 (.427)	.158 (.385)	.209 (.403)	.348 (.529)	.065[B] (.278)	.101 (.350)	.176 (.361)
High school degree	.151 (.402)	.225 (.477)	.208 (.428)	.239 (.423)	.161 (.408)	.153 (.404)	.124[B] (.383)	.319[B] (.442)
Less than high school degree	.103 (.342)	.216[A] (.469)	.034 (.191)	.048 (.212)	.045 (.231)	.387[B] (.547)	.081 (.317)	.073 (.247)
White	.750 (.487)	.711 (.517)	.832 (.394)	.729[A] (.441)	.828 (.420)	.764 (.477)	.732 (.514)	.786 (.390)
Partnered (living with a partner or married)	.424[A] (.55)	.444[A] (.567)	.373[A] (.510)	.570[A] (.491)	.543[B] (.554)	.474[B] (.561)	.493[B] (.580)	.719 (.426)
Any children in household	.123[A] (.369)	.343 (.542)	.212[A] (.431)	.410[A] (.488)	.147[B] (.393)	.439 (.558)	.310 (.537)	.422 (.469)
Northeast	.207 (.455)	.201 (.457)	.292 (.479)	.160 (.364)	.133 (.378)	.063[B] (.272)	.247 (.500)	.210 (.386)
Midwest	.185 (.436)	.170 (.428)	.146[A] (.372)	.236 (.421)	.284 (.501)	.192 (.443)	.238 (.495)	.261 (.417)
South	.333 (.530)	.397 (.558)	.314 (.490)	.409 (.487)	.177[B] (.424)	.333 (.530)	.245 (.499)	.393 (.464)
West	.275 (.502)	.233 (.482)	.248 (.455)	.195 (.393)	.406 (.546)	.412[B] (.553)	.269 (.515)	.137[B] (.326)
Avg. Annual Earnings	37,707.60 (27,591.28)	30,335.45[A] (24,922.17)	35,816.32 (29,525.85)	40,512.85 (33,244.90)	37,529.93[B] (27,681.97)	30,654.61[B] (34,248.06)	53,517.14 (39,152.12)	58,189.27 (40,155.29)
Full-time worker	.804 (.446)	.755 (.490)	.714 (.474)	.732 (.439)	.690[B] (.510)	.781 (.465)	.818 (.447)	.892 (.294)
Sample Size	52	88	88	493	49	88	80	422

Weighted means (standard deviations). [A] The superscript letter A means statistically significant difference ($P < 0.05$) between the group of women identified in the column header and the self-identified heterosexual women from Table 2.1. [B] The superscript letter B means statistically significant difference ($P < 0.05$) between the group of men identified in the column header and the self-identified heterosexual men from Table 2.1.

Table 2.5: Expanded set of Coefficient Estimates for Log Annual Earnings
Fully Saturated Model (i.e., Column 6 of Table 2.2), 2013-2015 NHIS, Adults ages 25-64, Full Time Workers

	(1) Females	(2) Males
Gay/Lesbian	.086** (.043)	.097** (.038)
Bisexual	-.031 (.092)	-.021 (.068)
S.O. Something Else	.149 (.105)	-.032 (.106)
S.O. Don't know	-.126** (.058)	-1.576 (1.345)
S.O. Refused	-.207 (.156)	.155* (.085)
S.O. Nonresponse	.088** (.039)	.002 (.099)
Age	.042*** (.006)	.046*** (.007)
Age-squared	-.000*** (.000)	-.001*** (.000)
Black/African American	-.044** (.020)	-.173*** (.021)
American Indian, Alaska Native	.020 (.058)	-.221*** (.076)
Asian	-.004 (.030)	-.138*** (.040)
Race group not releasable	.260*** (.098)	-127 (.127)
Multiple race groups	-.037 (.064)	-.072 (.057)
Hispanic	-.058*** (.021)	-.154*** (.023)
Less than high school degree	-.139*** (.034)	-.212*** (.038)
Some college	.057** (.023)	.022 (.026)
Associate degree	.088*** (.027)	.101*** (.023)
Bachelor's degree or more	.381*** (.027)	.254*** (.023)
Partnered	.055*** (.019)	.134*** (.023)
Widowed	.022 (.045)	.128*** (.044)
Divorced	.078*** (.022)	.129*** (.028)
Separated	-.005 (.035)	.022 (.045)
Marital Nonresponse	-.133 (.110)	-.676 (.426)
Presence of children ages 0 to 5	-.045* (.024)	.040** (.019)
Presence of children ages 6 to 17	-.050*** (.017)	-.009 (.016)
Northeast	.008 (.027)	.073*** (.025)
Midwest	-.097*** (.027)	.009 (.022)
South	-.090*** (.023)	-.038 (.027)
2014 survey wave	.015 (.016)	.060*** (.018)
2015 survey wave	.035** (.018)	.089*** (.018)
Job tenure	.045*** (.003)	.042*** (.003)
Job tenure squared	-.001*** (.000)	-.001*** (.000)
Firm size 10 to 24	.110*** (.030)	.197*** (.027)
Firm size 25 to 49	.185*** (.033)	.170*** (.041)

Firm size 50 to 99	.189*** (.033)	.262*** (.027)
Firm size 100 to 249	.225*** (.033)	.281*** (.029)
Firm size 250 to 499	.226*** (.036)	.270*** (.033)
Firm size 500 to 999	.265*** (.038)	.332*** (.031)
Firm size 1000 or more	.293*** (.032)	.324*** (.029)
Employed in public sector	.005 (.021)	.008 (.024)
R-squared	.35	.35
N	17,016	18,981

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. See notes to Table 2.2. The model also includes other control variables not listed here, including: occupation dummies, industry dummies, and others.

Chapter 3

Transgender Status, Gender Identity, and Socioeconomic Outcomes in the United States

Christopher S. Carpenter, Samuel T. Eppink, and Gilbert Gonzales*

ABSTRACT

We provide the first large-scale evidence on transgender status, gender identity, and socioeconomic outcomes in the United States using representative data from 31 states in the Behavioral Risk Factor Surveillance System (BRFSS) that asked identical questions about transgender status and gender identity in at least one year from 2014-2016. Over 1,500 adults ages 18-64 identified as transgender. Individuals who identify as transgender are significantly less likely to be college educated and less likely to identify as heterosexual than individuals who do not identify as transgender. Controlling for these and other observed characteristics, transgender individuals have significantly lower employment rates, lower household incomes, higher poverty rates, and worse self-rated health than otherwise similar men who are not transgender. Differences in household structure account for a substantial share of these differences.

*Carpenter is Professor of Economics at Vanderbilt University and Research Associate at the NBER. Eppink is a PhD Candidate in the Department of Economics, Vanderbilt University. Gonzales is Assistant Professor in the Department of Health Policy, Vanderbilt University. We thank Cevat Aksoy, Lee Badgett, Benjamin Cerf, Jeff Frank, Carrie Henning-Smith, Bill Jesdale, Marieka Klawitter, Josh Miller, Joe Sabia, and seminar participants at Vanderbilt University and the 2016 APPAM Fall Research Conference for very helpful comments. All errors are our own.

3.1 Introduction

Since Badgett's (1995) pioneering paper examining the wage effects of sexual orientation discrimination, a substantial literature has emerged on income differences between heterosexual individuals and gay men, lesbians, and bisexual individuals. A broad consensus has emerged from that literature that gay men have earnings and incomes that are lower than those of similarly situated heterosexual men, and lesbians have earnings and incomes that are higher than those of similarly situated heterosexual women.[37]

In contrast to the large and growing literature on sexual orientation and socioeconomic outcomes, there is far less research on transgender status, gender identity, and socioeconomic outcomes. While sexual orientation refers to a person's degree of different-sex versus same-sex romantic and sexual attraction, gender identity refers to one's innate sense of self as being male, female, both, or neither. A part of gender identity may involve gender expression, which refers to the external appearance (e.g., haircut, clothing, behavior) of one's gender identity (e.g., masculine, feminine, androgynous). Transgender individuals are people whose gender identity and/or gender expression or behavior differ from their sex assigned at birth or differ from gender-cultural norms attached to their sex

[37] We do not review that literature in detail here. See Klawitter (2015) for a meta-analysis. For a recent exception to this pattern, see Carpenter and Eppink (2017) who find income premia for both gay men and lesbians compared to otherwise similar heterosexuals in the 2013-2015 National Health Interview Surveys.

assigned at birth.[38] Cisgender individuals are individuals who are not transgender – i.e., people who identify with their sex assigned at birth. Importantly, gender identity and sexual orientation are distinct constructs: a transgender person need not have a minority sexual orientation (in fact, our data described below indicate that most individuals who identify as transgender also identify as heterosexual).

There are several possible channels through which transgender status could be related to socioeconomic outcomes. First, there is extensive anecdotal evidence that transgender individuals face pervasive discrimination in society, including in employment and labor markets (Center for American Progress and Movement Advancement Project 2015, James et al. 2016). This is compounded by the fact that there is relatively little antidiscrimination protection in private employment on the basis of transgender status and gender identity (Human Rights Campaign 2016).[39] These factors may make transgender individuals less likely to access employment and may harm their employment outcomes (e.g., wages, promotions, job satisfaction) conditional on being employed. Transgender

[38] Transgender and gender non-conforming individuals may include transsexuals, androgynous people, cross-dressers, genderqueers, and other gender non-conforming people who identify as transgender. Some, but not all, of these individuals may desire to undergo medical and/or legal sex changes. Transgender individuals whose gender identity does not match their sex assigned at birth and who desire to change from one sex to another are sometimes referred to as 'MTF' (for individuals who transition from male to female) or 'FTM' (for individuals who transition from female to male). There is wide variance in the use of these labels; for example, 'MTF' can be used by individuals who are male by birth and express a feminine identity but who have not taken steps to change their gender expression.

[39] At the time of this writing, only 20 states and the District of Columbia prohibited employment-based discrimination based on gender identity (see Figure 3.1) (Human Rights Campaign 2016).

individuals may also face discrimination in public accommodations and housing markets, which could also make it difficult to secure stable housing, employment, and income. Transgender individuals may also face harassment, bullying, and discrimination in educational environments (Kosciw et al. 2014), thus limiting their ability to accumulate the skills needed to succeed in the labor market.

Second, transgender individuals have unique health profiles that could independently affect their ability to work and/or their productivity at work. Specifically, transgender individuals are at increased risk of adverse health behaviors and outcomes due, in part, to the chronic stress associated with being a member of a marginalized population. For example, transgender individuals have a higher likelihood of activity limitations, mental health conditions (such as clinical depression and suicide attempts), and substance use disorders (Grant et al. 2011, James et al. 2016), and these could independently reduce labor market opportunities.

Third, willingness to identify as transgender in population-based studies may be related to unobservable characteristics that are correlated with economic outcomes. That is, among the set of all transgender individuals, it is likely that only a fraction choose to identify as such on a national survey, and this could be endogenously related to unobserved variables correlated with socioeconomic status.[40] There are several possible candidates for these unobservables, including: the presence of social supports from family and friends, local attitudes toward transgender individuals, and others. In this case we might observe that transgender individuals have systematically different socioeconomic outcomes compared with otherwise similar cisgender individuals, but these differences could be driven by unobserved factors correlated both with the decision to identify as transgender and socioeconomic outcomes. Finally, it is possible that socioeconomic outcomes may directly affect transgender status through the cultivation of identity (Akerlof and Kranton 2000). For example, greater economic resources may provide gender minorities more opportunities to cultivate a transgender identity.

This study provides the first multi-state, population-based evidence on transgender status, gender identity, and socioeconomic outcomes in the United States. To do so, we use data from the 2014, 2015, and 2016 Behavioral Risk Factor Surveillance System (BRFSS) which included sexual orientation and gender identity questions for 31 states (pooled across the three years). In these data we identify 1,500 transgender individuals aged 18 to 64 years. Of these individuals, approximately 50 percent identify as male-to-female (MTF), 30

[40] Coffman et al. (2017) provide direct evidence of this in the context of sexual orientation, but to our knowledge there is no good evidence on this phenomenon for gender minorities.

percent identify as female-to-male (FTM), and about 20 percent identify as gender non-conforming.

Our study makes important contributions relative to previous work. First, to our knowledge, we are one of the first studies to focus on the relationship between transgender status and socioeconomic outcomes using *population-based* survey data. As described below, some novel studies have examined income differences associated with the medical and/or legal processes of changing one's sex or gender. However, these studies necessarily omit individuals who would identify as transgender but who have not yet undertaken or will not undertake medical or legal actions to change their gender. Second, we provide the first *multi-state* study of differences in socioeconomic outcomes associated with gender minority status in the United States using representative data.[41] Our study includes data from 31 states and represents every region of the country, from New York and Louisiana to Wisconsin and Idaho. Relatedly, we use much larger samples than previous work: our data identify over 1,500 transgender individuals aged 18-64, which is at least 7 times larger than prior work using representative data.

Our primary empirical models compare employment, household income, poverty status, health insurance coverage, and self-rated health for cisgender men

[41] Throughout, we use the phrase 'gender minority' interchangeably with 'transgender'.

with cisgender women, transgender women (i.e., individuals who describe themselves as male-to-female transgender), transgender men (i.e., individuals who describe themselves as female-to-male transgender), and gender non-conforming individuals (i.e. individuals who do not describe their gender as only male or only female). These models return clear evidence that transgender individuals fare significantly worse than cisgender men with respect to employment, household income, poverty, and self-rated health. Notably, however, several of these associations are reduced when we control for the sex composition of other adults in the household, suggesting that household structure partly accounts for differences in socioeconomic outcomes experienced by transgender individuals. This is particularly true for employment, household income, and poverty status. Despite this, the most robust finding even after accounting for household sex composition is that individuals who identify as transgender, gender non-conforming have significantly lower employment rates and worse self-rated health than otherwise comparable cisgender men. This is a particularly interesting finding given that prior research relying on changes to gender identity in administrative and medical records is unlikely to capture these individuals (relative to individuals who identify as male-to-female or female-to-male).

The paper proceeds as follows: Section 2 provides a brief literature review of the handful of studies that have examined transgender individuals. Section 3 describes the data and outlines the empirical approach. Section 4 presents the results, and Section 5 offers a discussion and conclusion.

3.2 Literature Review

Very few representative studies have measured and compared any socioeconomic outcomes between transgender people and their cisgender (i.e. non-transgender) counterparts, largely due to the lack of credible data ascertaining transgender status or gender identity.[42] Using data from the 2007-2009 Massachusetts Behavioral Risk Factor Surveillance System with 131 self-identified transgender adults, Conron *et al.* (2012) found nonelderly (18-64 year old) transgender adults were more likely to be living in poverty (31% vs 9%) and unemployed (33% vs 12%) compared to their non-transgender peers.[43]

Two studies in economics have examined the association of earnings with transgender status in the US and the Netherlands. Schilt and Wiswall (2008) studied individuals who attended transgender conferences in the United States as well as individuals who participated in a transgender-focused internet site. They compared earnings trajectories for individuals who underwent medical procedures to change their gender expression. They found that individuals who transitioned

[42] To the best of our knowledge, no federal surveys in the United States ascertain gender identity or transgender status on a nationally representative scale.
[43] Other evidence from non-random samples of transgender populations collected through transgender-serving advocacy organizations supports these findings (Grant et al. 2011, Xavier et al. 2007).

from male to female experienced a large earnings decline (on the order of 30%), while individuals who transitioned from female to male experienced a small earnings increase. They interpret this as evidence of gender inequality in the workplace, whereby women report lower earnings compared to similarly situated men. Using administrative data from the Netherlands with larger samples of transgender individuals where transgender status is identified from a gender change in registry data, Geijtenbeek and Plug (2018, forthcoming) find a similar qualitative pattern as in Schilt and Wiswall (2008). Geijtenbeek and Plug (2018, forthcoming) also find that female-to-male transgender individuals in their sample earn more than similarly situated women but less than similarly situated men.[44]

Our study complements these prior studies. Because we lack panel data on individuals, we are unable to examine within-person differences in income associated with changing aspects of one's gender expression as in Schilt and Wiswall (2008) and Geijtenbeek and Plug (2018, forthcoming). Not all individuals who identify as transgender, however, take active steps (medical, legal, or otherwise) to change their gender identity, and we are able to capture these individuals with our broader measure. We also consider a wider range of socioeconomic outcomes besides wages than prior work, including: health

[44] In a related paper, Cerf Harris (2015) uses individuals who change their first names and sex coding with the Social Security Administration as a novel way to identify a population of individuals likely to be transgender. Cerf Harris (2015) does not estimate earnings models similar to Schilt and Wiswall (2008) or Geijtenbeek and Plug (2018, forthcoming), however.

insurance, poverty status, and self-rated health. Also, our multi-state sample allows us to explicitly examine whether the policy environment regarding transgender-specific protections – in particular the presence of trans-inclusive employment non-discrimination acts (ENDAs) – is associated with improvements in relative socioeconomic outcomes for transgender men and women.

3.3 Data Description and Empirical Approach

3.3.1 Data Description

We use data from the 2014 to 2016 Behavioral Risk Factor Surveillance System (BRFSS). The BRFSS is a large telephone survey fielded every year by state health departments and coordinated by the Centers for Disease Control and Prevention (CDC) who compiles them into an annual individual-level dataset that is designed to be representative at the state level. The main purpose of the BRFSS is to measure population health behaviors, access to care, and health outcomes, though the household screener and demographic portion of the survey also includes information about age, race/ethnicity, marital status, educational attainment, household income, household structure, and employment.

A key feature for our purposes is that since 2014 the CDC has released information on minority sexual orientation, transgender status, and gender identity on the public use BRFSS files. Specifically, several states administered an identical module about sexual orientation and gender identity (SOGI) in their statewide BRFSS survey.[45] Regarding gender identity, individuals are asked: "Do you consider yourself to be transgender?" We observe over 1,500 nonelderly individuals aged 18-64 years who self-identify as transgender from a total sample of over 390,000 respondents in states using the SOGI module. Individuals who identify as transgender are then asked whether they consider themselves to be male-to-female (MTF), female-to-male (FTM), or gender non-conforming.[46]

Several notes about the gender identity information in the BRFSS merit mention. Most of our empirical models compare men who do not identify as

[45] In 2014 the states were: Delaware, Hawaii, Idaho, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maryland, Minnesota, Montana, Nevada, New York, Ohio, Pennsylvania, Vermont, Virginia, Wisconsin, and Wyoming. In 2015 the states were: Colorado, Connecticut, Delaware, Georgia, Hawaii, Idaho, Illinois, Indiana, Iowa, Kansas, Maryland, Massachusetts, Minnesota, Missouri, Nevada, New York, Ohio, Pennsylvania, Texas, Virginia, West Virginia, and Wisconsin. In 2016 the states were: California, Connecticut, Delaware, Georgia, Hawaii, Idaho, Illinois, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Minnesota, Mississippi, Missouri, Nevada, New York, Ohio, Pennsylvania, Rhode Island, Texas, Vermont, Virginia, Washington, and Wisconsin. Note that some other states also included sexual orientation and gender identity questions on their state BRFSS surveys, but they did not use the CDC-provided module. As such, data from those states are not included, as the CDC only released data for states that adopted the unified CDC module. Other states may run their own state-wide public health survey with information on sexual orientation and gender identity (e.g., the California Health Interview Survey) but are not included here. Finally, a handful of states also administered the BRFSS SOGI module but did not give permission to the CDC to release their data in the public use file. Those states are also excluded here.

[46] If the survey respondent asked about the definition of transgender, the interviewer was instructed to read the following: "Some people describe themselves as transgender when they experience a different gender identity from their sex at birth. For example, a person born into a male body, but who feels female or lives as a woman would be transgender. Some transgender people change their physical appearance so that it matches their internal gender identity. Some transgender people take hormones and some have surgery. A transgender person may be of any sexual orientation – straight, gay, lesbian, or bisexual." Scholarship on transgender populations has suggested alternative ways to elicit gender identity in surveys (GenIUSS 2014).

transgender (i.e., cisgender men) as the excluded category against similarly situated women who do not identify as transgender (cisgender women), individuals who identify as transgender, male-to-female (transgender women), individuals who identify as transgender, female-to-male (transgender men), and individuals who identify as transgender, gender non-conforming (henceforth gender non-conforming). For transgender individuals, our main models use the gender identity asserted to the interviewer in the follow-up question after identifying as transgender. Notably, the BRFSS survey instrument does not directly ask cisgender respondents their sex or gender; instead, respondent sex is inferred by the interviewer as male or female based on the voice timbre of the respondent. For individuals who do not identify as transgender, we rely on the interviewer's assessment. This should be correct for the vast majority of cisgender respondents, but it is certainly possible that we miscode a small share of individuals for whom the interviewer's inference is incorrect.

In our context misreporting and miscoding are more concerning if these errors occur disproportionately for transgender-identified individuals. A small number of cisgender respondents inaccurately reporting their gender identity as transgender can inundate the relatively small number of transgender respondents with false positives. However, relying on the asserted gender as revealed in the follow-up response about being male-to-female or female-to-male gives us confidence that transgender respondents are actually transgender. These respondents are explicitly revealing to the interviewer both that they are transgender and the 'direction' of their gender transition. For these individuals, there should be few false positives.[47]

We consider a range of socioeconomic outcomes available in the BRFSS, including indicator variables for being employed or self-employed, being unemployed, or being unable to work.[48] We also study annual household income which is reported in categorical ranges in the BRFSS.[49] We use information on

---

[47] A limitation of using asserted gender for transgender-identified individuals is that a small but nontrivial share of transgender-identified individuals (less than 20 percent) describe themselves as 'gender non-conforming', and for these individuals we do not have further information on their current gender expression. We experimented with models in which we used the concordance or discordance between inferred gender and asserted gender for transgender-identified respondents as a possible measure of 'passing', but sample sizes were too small to produce meaningful results.

[48] The employment variable is coded one for individuals that indicated they were employed for wages or self-employed and coded zero for individuals who indicated being out of work, a homemaker, a student, retired, or unable to work. The unemployment variable is coded one for individuals indicating they were out of work or unable to work and coded zero for individuals who indicated they were employed for wages, self-employed, a homemaker, a student, or retired. Respondents who either refused to answer the employment question or for whom no response is recorded are excluded from both analyses.

[49] The ranges are: less than \$10,000; \$10,000 to less than \$15,000; \$15,000 to less than \$20,000; \$20,000 to less than \$25,000; \$25,000 to less than \$35,000; \$35,000 to less than \$50,000; \$50,000 to less than \$75,000; and \$75,000 or more. Two key limitations regarding household income are worth noting. First, it is measured at the household level instead of at the individual level. It is plausible that transgender status is correlated with the number of adults in the household (e.g., if transgender individuals are differentially likely to have spouses or partners), though Tables 3.1a and 3.1b do not indicate significant differences in this respect. To address this limitation, we control for the total number of adults in the household in all specifications. Notably, to test differences in household structure by transgender status, we estimated models where we predicted the number of adults in the household as a function of observed demographic characteristics (e.g. age, race/ethnicity, educational attainment) and found no significant difference associated with gender minority status. The second key limitation of the income measure is that the question asks about income from all sources and does not separately identify labor earnings, government transfers, investment income, or other sources. Despite these limitations, household income is a useful summary measure of financial and economic wellbeing and resources available to individuals. Because there are – to our knowledge – no other large national surveys with information on labor market earnings for gender minorities, the BRFSS represents the first – and

household size in addition to household income to examine whether the individual's household is less than or equal to 100% poverty, following Conron *et al*. (2012). We also examine an indicator variable for having health insurance as well as indicator variables for 'excellent' or 'very good' self-rated health and 'fair' or 'poor' self-rated health.

We acknowledge some concerns about the quality of the economic outcomes in the BRFSS – which is primarily a health survey and not a labor survey – relative to datasets that are more commonly used by labor economists such as the March Current Population Surveys. In Appendix C Table 3.6 we provide some comparisons for employment and average household income from the BRFSS (using only individuals in the states and years that released the SOGI module) and the March CPS (using a random adult from each household for those same states and years). The CPS returns significantly higher average employment rates and household incomes (using midpoints of household income ranges in the BRFSS) – differences on the order of 5.7 and 8.5 percent relative to the BRFSS means, respectively. Moreover, differences in the distributions of household income are even larger. For example, the share of individuals with household incomes below $25,000 is 26.2 percent in the BRFSS but only 20 percent in the CPS, a much larger proportional difference. When we compared the average

as of this writing, the only – opportunity to systematically examine gender minorities and any type of income in a large multi-state representative sample from the United States.

employment rates at the state/year level between BRFSS and CPS, the correlation was .64. Moreover, in results not reported but available upon request, we estimated models predicting employment in the two datasets and found qualitatively similar relationships between employment and age, race/ethnicity, education, and marital status (i.e., the variables defined similarly in both data).

Several points are worth making in light of these patterns. First, our results on economic outcomes should be interpreted with caution, particularly for the household income and poverty variables given the differences with the more commonly used and widely accepted March CPS as observed in Appendix C Table 3.6. Second, we refer readers to several economics publications that have examined these same economic variables in the BRFSS either as outcomes (e.g., Evans and Garthwaite 2014) or as regressors (Ruhm 2005). Not surprisingly, an even larger number of economics papers have used the BRFSS health insurance and self-rated health variables (see, for example, Garthwaite et al. 2014, Bitler et al. 2005, and others). Finally, we reiterate that – at the time of this writing – the BRFSS is to our knowledge the only large-scale representative dataset that includes information on both transgender status and any economically relevant outcomes in the US. Thus, while we acknowledge the legitimate concerns about the quality of the economic data in the BRFSS – particularly with respect to income – we still see value in the descriptive evidence it affords us on this interesting population.

3.3.2 Empirical Approach

To understand the relationship between transgender status and socioeconomic outcomes, we first explore determinants of transgender status, allowing for the possibility of endogenous selection. We then examine how transgender status is related to the socioeconomic outcomes described above. Since we have no credible identification strategy for accounting for the endogeneity of transgender identification, we view these analyses as descriptive and exploratory attempts to provide the first documentation of transgender status and a range of socioeconomic and health outcomes from representative data.

To estimate the association between individual characteristics and transgender identification, we estimate a simple linear probability model of the form:

$$(3.1) \qquad TRANSGENDER_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

where $TRANSGENDER_i$ is an indicator for identifying as transgender. X_i is a vector of individual characteristics available in the BRFSS, including: age, age squared, sexual orientation (gay or lesbian, bisexual, and other sexual orientation, with heterosexual or straight as the excluded category), race/ethnicity (black non-Hispanic, Asian non-Hispanic, Hispanic, other/multiple non-Hispanic, with white non-Hispanic as the excluded category), education (less than a high school degree, some college, college degree and above, with high school degree as the excluded category), marital status (married/partnered, divorced, widowed, separated, with never married as the excluded category), number of adults in the household, and Census region (Northeast, Midwest, South, with West as the excluded category). X_i also includes a control for survey year and an indicator equal to one if the respondent was contacted on a cellphone (versus a landline phone).[50] We estimate this model with and without an indicator for the respondent's gender being female. In all models (including those described below) we estimate White standard errors that are robust to heteroscedasticity, and we use BRFSS sample weights throughout all analyses.

To estimate the association between gender minority status and socioeconomic outcomes, we estimate specifications similar to the cross-sectional

[50] Below, we report results from a robustness exercise that also controls for the sex composition of adults in the household. When we do so, we are forced to drop the cellphone sample because those individuals were not administered the portion of the household screener that asks individuals to state the number of adult men and the number of adult women in the household. In our full sample, 29.8 percent of respondents participated by cellphone.

models of Geijtenbeek and Plug (2018, forthcoming). These models take the form:[51]

$$(3.2) \qquad Y_i = \beta_0 + \beta_1 X_i + \beta_2(\text{CISGENDER WOMAN})_i + \beta_3(\text{TRANSGENDER})_i + \varepsilon_i$$

where Y_i are employment, household income, poverty, and health outcomes for individual *i* and X_i is as defined above. CISGENDER WOMAN_i is an indicator variable equal to one for women who do not identify as transgender. TRANSGENDER is an indicator variable equal to one for individuals who report being transgender. In some models we replace the single TRANSGENDER indicator with three separate indicators for TRANSGENDER WOMAN, TRANSGENDER MAN, and TRANSGENDER GENDER NON-CONFORMING.[52] In these models, TRANSGENDER WOMAN_i is an indicator variable equal to one for individuals who report being transgender, male-to-female. TRANSGENDER MAN_i is an indicator variable equal to one for individuals who report being transgender, female-to-male. TRANSGENDER

[51] For the dichotomous socioeconomic outcomes we estimate linear probability models. For household income we estimate interval regressions on the categorical responses. We drop a very small share of individuals with missing data on the demographic characteristics. Note that about 13.7 percent of the sample did not give a usable response to the household income question, which is common in such surveys. We estimated a model predicting nonresponse (i.e., refused, 'don't know', or missing) to the household income question and found that transgender status was not significantly related to the likelihood of a valid income response.

[52] In all models we include separate indicators for individuals who report that they 'don't know', refused to provide, or were missing a response to the transgender status question. We do the same for similar responses to the sexual orientation question.

GENDER NON-CONFORMING$_i$ is an indicator variable equal to one for individuals who report being transgender, gender non-conforming. The excluded category throughout is composed of cisgender men.

To investigate heterogeneity in the associations between minority gender identity and socioeconomic outcomes we interact individual characteristics (e.g., education, race, and minority sexual orientation) with all of the key indicator variables in equation (3.1): CISGENDER WOMAN, TRANSGENDER WOMAN, TRANSGENDER MAN, and TRANSGENDER GENDER NON-CONFORMING. And to examine if trans-inclusive Employment Non-Discrimination Acts (ENDAs) are associated with relative improvements in outcomes for transgender individuals, we estimate the following models:

(3.3) $Y_{is} = \beta_0 + \beta_1 X_{is} + \beta_2(\text{CISGENDER WOMAN})_i + \beta_3(\text{TRANSGENDER WOMAN})_i + \beta_4(\text{TRANSGENDER MAN})_i + \beta_5(\text{TRANSGENDER GENDER NON-CONFORMING})_i + \beta_6(\text{TRANS-INCLUSIVE ENDA})_s + \beta_7(\text{CISGENDER WOMAN * TRANS-INCLUSIVE ENDA})_{is} + \beta_8(\text{TRANSGENDER WOMAN * TRANS-INCLUSIVE ENDA})_{is} + \beta_9(\text{TRANSGENDER MAN * TRANS-INCLUSIVE ENDA})_{is} + \beta_{10}(\text{TRANSGENDER GENDER NON-CONFORMING * TRANS-INCLUSIVE ENDA})_{is} + \varepsilon_{is}$

where all variables are as described above and where *s* indicates state. TRANS-INCLUSIVE ENDA is an indicator variable for individuals living in one of the 20 states plus Washington DC that had a trans-inclusive Employment Non-Discrimination Act (ENDA) at the time of the survey (Human Rights Campaign 2016).[53] Here the primary coefficients of interest are β_8 through β_{10} on the interactions between the TRANS-INCLUSIVE ENDA indicator and the various TRANSGENDER dummies. A positive coefficient suggests that trans-inclusive ENDAs are protective for economic outcomes (e.g., employment).

3.4 Results

3.4.1 Descriptive Statistics

We begin by presenting descriptive statistics for the key demographic characteristics and socioeconomic outcomes in the 2014 to 2016 BRFSS data separately by gender identity in Table 3.1a for men and Table 3.1b for women.

[53] The 20 states with trans-inclusive ENDAs in public and private employment sectors as of 2014 were: California, Colorado, Connecticut, Delaware, Hawaii, Illinois, Iowa, Maine, Maryland, Massachusetts, Minnesota, Nevada, New Jersey, New Mexico, New York, Oregon, Rhode Island, Utah, Vermont, Washington, and Washington DC (Human Rights Campaign 2016).

Table 3.1a: Descriptive Statistics
2014-2016 BRFSS, Adults ages 18-64

Variable	**(1) Men who do not identify as transgender (i.e., cisgender)**	**(2) Men who identify as transgender, female-to-male**	**(3) Individuals who identify as transgender, gender non-conforming**
Age	40.454† (13.707)	38.592 (14.177)	35.520† (14.724)
White, non-Hispanic	.629*† (.483)	.475* (.522)	.511† (.497)
Gay or lesbian	.023*† (.149)	.145* (.360)	.136† (.332)
Bisexual	.015*† (.123)	.124* (.336)	.258† (.424)
Partnered (married or a member of an unmarried couple)	.546*† (.498)	.420* (.513)	.420† (.494)
High school degree or less	.446* (.497)	.676* (.487)	.419 (.493)
Some college	.291 (.454)	.232 (.439)	.384 (.486)
College degree or more	.263*† (.440)	.092* (.301)	.197† (.397)
Number of adults in HH	2.336 (1.185)	2.590 (1.530)	2.598 (1.505)
Any children in HH	.404*† (.491)	.536* (.517)	.314† (.465)
Employed or self-employed	.755*† (.430)	.597* (.516)	.512† (.498)
Average household income	68,098.37*† (44,144.16)	41,795.24* (38,895.91)	57,816.24† (43,799.88)
At or below 100% poverty	.151*† (.358)	.403* (.515)	.241† (.413)
Has health insurance	.845* (.361)	.735* (.458)	.852 (.352)
Very good or excellent health	.539*† (.498)	.422* (.512)	.389† (.487)
Fair or poor health	.145*† (.352)	.229* (.436)	.283† (.450)
Lives in a state with:			
Trans-inclusive ENDA	.361 (.480)	.351 (.495)	.401 (.489)
Marriage equality before *Obergefell*	.699† (.459)	.644 (.497)	.811† (.391)
Republican governor	.529† (.499)	.464 (.518)	.409† (.491)
Sample size	175,924	473	322

Weighted means (standard deviations). Note average household income and poverty status are determined using the midpoint of each household income range or the 80th percentile of annual household income for those who reported the highest income category; percent of poverty is calculated by dividing household income by household size specific U.S. Census Bureau poverty thresholds, following Conron et al. (2012). * indicates the means are significantly different between columns 1 and 2 at p<.05. † indicates the means are significantly different between columns 1 and 3 at p<.05.

Table 3.1b: Descriptive Statistics
2014-2016 BRFSS, Adults ages 18-64

Variable	(1) Women who do not identify as transgender (i.e., cisgender)	(2) Women who identify as transgender, male-to-female	(3) Individuals who identify as transgender, gender non-conforming
Age	41.062† (13.653)	41.045 (13.078)	35.520† (13.359)
White, non-Hispanic	.620† (.486)	.576 (.446)	.511† (.451)
Gay or lesbian	.014*† (.115)	.040* (.172)	.136† (.301)
Bisexual	.031*† (.173)	.129* (.295)	.258† (.384)
Partnered (married or a member of an unmarried couple)	.563† (.496)	.499 (.454)	.420† (.448)
High school degree or less	.381* (.486)	.612* (.442)	.419 (.447)
Some college	.325 (.469)	.273 (.404)	.384 (.441)
College degree or more	.294*† (.456)	.115* (.289)	.197† (.360)
Number of adults in HH	2.302 (1.120)	2.502 (1.176)	2.598 (1.365)
Any children in HH	.481*† (.500)	.352* (.434)	.314† (.422)
Employed or self-employed	.612† (.488)	.633 (.437)	.512† (.451)
Average household income	62,961.78* (48,869.43)	48,869.43* (37,953.53)	57,816.24 (39,695.82)
At or below 100% poverty	.208* (.406)	.327* (.426)	.241 (.375)
Has health insurance	.876* (.330)	.801* (.362)	.858 (.320)
Very good or excellent health	.531*† (.499)	.363* (.436)	.389† (.442)
Fair or poor health	.159† (.365)	.165 (.337)	.283† (.408)
Lives in a state with:			
Trans-inclusive ENDA	.359* (.480)	.294* (.413)	.401 (.444)
Marriage equality before *Obergefell*	.697*† (.460)	.605* (.444)	.811† (.355)
Republican governor	.529*† (.499)	.624* (.439)	.409† (.445)
Sample size	215,806	765	322

Weighted means (standard deviations). Note average household income and poverty status are determined using the midpoint of each household income range or the 80th percentile of annual household income for those who reported the highest income category; percent of poverty is calculated by dividing household income by household size specific U.S. Census Bureau poverty thresholds, following Conron et al. (2012). * indicates the means are significantly different between columns 1 and 2 at p<.05. † indicates the means are significantly different between columns 1 and 3 at p<.05.

The format of Tables 3.1a and 3.1b is as follows: column 1 reports weighted means for cisgender individuals, column 2 reports weighted means for transgender individuals who describe themselves as female-to-male (Table 3.1a) or male-to-female (Table 3.1b), and column 3 reports weighted means for transgender individuals who describe themselves as gender nonconforming.[54]

To our knowledge, the descriptive statistics in Tables 3.1a and 3.1b represent one of the first large multi-state comparisons of demographic characteristics between transgender and cisgender individuals in the United States. Columns 1 and 2 of Table 3.1a show that transgender men are significantly less likely to be white, more likely to identify as gay or bisexual, less likely to be partnered, less likely to have a college degree, more likely to have children in the household, less likely to be employed, more likely to be in poverty,

[54] Appendix C Table 3.7 shows that the characteristics of individuals living in states that did (31) and did not (19) contribute SOGI data to the public use BRFSS at some point from 2014-2016. Although there are several statistically significant differences in demographic characteristics, most are small in magnitude relative to the means. Moreover, the direction of the differences does not uniformly indicate positive or negative selection (e.g., individuals in states that did not provide SOGI data are significantly more likely to be white but have significantly lower average household incomes and rates of college degrees than individuals in states that did provide SOGI data). There are, however, very large political differences across the two samples, though again the pattern is not a simple one. For example, individuals in states that provided SOGI data at some point from 2014-2016 are significantly less likely to be from states with Republican governors and are significantly more likely to be from states with trans-inclusive ENDAs than individuals in states that did not provide SOGI data from 2014-2016. They are also, however, significantly less likely to live in states that adopted legal same-sex marriage prior to the 2015 United States Supreme Court decision in *Obergefell vs. Hodges*. It is important to keep these differences in mind when interpreting the findings we document below, as we cannot claim that our results are necessarily representative of the United States as a whole. It could be that transgender individuals are more likely to live in (or identify themselves as transgender in) states that are more socially liberal, in which case our results are likely to capture a disproportionate share of transgender adults (given the patterns in the bottom of Appendix C Table 3.7).

less likely to be insured, less likely to have excellent or very good self-rated health, and more likely to have fair or poor self-rated health than men who are not transgender (i.e., cisgender). Transgender men also have significantly lower average household incomes than cisgender men. Columns 1 and 3 of Table 3.1a show that individuals who identify as transgender, gender non-conforming are significantly younger, less likely to be white, more likely to have a minority sexual orientation, less likely to be partnered, less likely to have a college degree or more, less likely to have children in the household, less likely to be employed, more likely to be in poverty, less likely to have excellent or very good self-rated health, and more likely to have fair or poor self-rated health than men who are not transgender. Individuals who identify as transgender, gender non-conforming also have significantly lower household incomes than men who are not transgender.

Columns 1 and 2 of Table 3.1b show that transgender women are significantly more likely to report a minority sexual orientation, less likely to have a college degree, less likely to have children in the household, more likely to be in poverty, less likely to be insured, and less likely to have excellent or very good self-rated health compared with cisgender women. Transgender women also have significantly lower household incomes than women who are not transgender. Columns 1 and 3 of Table 3.1b show that individuals who identify as transgender, gender non-conforming are significantly younger, less likely to be white, more likely to have a minority sexual orientation, less likely to be partnered, less likely to have a college degree or more, less likely to have any children in the household, less likely to be employed, less likely to have excellent or very good self-rated health, and more likely to have fair or poor self-rated health compared with cisgender women.

3.4.2 Correlates of the Likelihood of Identifying as Transgender

In Table 3.2 we present results of regressions where we predict the likelihood an individual identifies as transgender in the BRFSS (from equation (3.1) above). We present these results in two columns: one where we control for a WOMAN indicator and one where we exclude it.

Table 3.2: Predictors of Transgender Status
2014-2016 BRFSS, Adults ages 18-64

Variable	**(1) Transgender**	**(2) Transgender**
Woman	---	-.002*** (.000)
Gay/Lesbian	.022*** (004)	.022*** (.004)
Bisexual	.030*** (.004)	.030*** (.004)
Other sexual orientation	.060*** (.013)	.060*** (.013)
Age	-.000 (.000)	-.000 (.000)
Age squared	.000 (.000)	.000 (.000)
Black	.001 (.001)	.001 (.001)
Asian	.001 (.001)	.001 (.001)
Other race	.003** (.001)	.003** (.001)

Hispanic ethnicity	.000 (.001)	.000 (.001)
Less than HS degree	.002 (.001)	.002 (.001)
Some college	-.001** (.001)	-.001** (.001)
College degree or more	-.003*** (.000)	-.002*** (.000)
Partnered	-.000 (.001)	-.000 (.001)
Divorced	.000 (.001)	.001 (.001)
Widowed	.001 (.001)	.001 (.001)
Separated	.000 (.001)	.001 (.001)
Northeast Census Region	.000 (.001)	.000 (.001)
Midwest Census Region	.001 (.001)	.001 (.001)
South Census Region	.001 (.001)	.001 (.001)
Very good or excellent health	-.001*** (.001)	-.001*** (.001)
Fair or poor health	-.001 (.001)	-.001 (.001)
Trans-inclusive ENDA	-.000 (.000)	-.000 (.000)
# adults in the household	.001** (.000)	.001** (.000)
2015 Survey Wave	.000 (.001)	.000 (.001)
2016 Survey Wave	-.001 (.001)	-.001 (.001)
In the cellphone-only sample	-.000 (.001)	-.000 (.001)
R-squared	.01	.01
N	390,029	390,029

** and *** denote statistical significance at 5% and 1%, respectively. Estimates are from linear probability models. All models control for 2015 and 2016 survey wave indicators, sexual orientation (indicators for gay/lesbian, bisexual, other sexual orientation, don't know sexual orientation, refusal to provide sexual orientation, and missing sexual orientation), age, age squared, race/ethnicity (indicators for black non-Hispanic, Asian, Hispanic, other race, and refusal to provide race/ethnicity), educational attainment (indicators for less than high school, some college, college or more, and refusal to provide educational attainment), marital status (indicators for married/partnered, divorced, widowed, separated, and refusal to provide marital status), Census region, whether the respondent was contacted on a cellphone or landline, and the number of adults in each household. Standard errors are robust to heteroscedasticity.

The results from this exercise reveal that several demographic characteristics are significantly related to the likelihood of identifying as transgender. Sexual minorities are significantly more likely to identify as transgender than heterosexual adults. Perhaps surprisingly, age is not significantly related to the likelihood of identifying as transgender. Race/ethnicity is similarly insignificantly related to the probability of identifying as transgender, with the exception that individuals identifying as 'other race' are significantly more likely to identify as transgender. Education is strongly related to transgender identification: individuals with some college or college degrees are significantly less likely to identify as transgender than individuals with a high school degree. Marital status is not significantly related to the likelihood of identifying as transgender. Individuals in the Midwest and South (of the sample of states represented in the BRFSS who released their SOGI data) are significantly more likely to identify as transgender compared to individuals in the West. Individuals who report very good or excellent health are significantly less likely to identify as transgender than individuals reporting good health. Finally, individuals with more adults in the household are significantly more likely to identify as transgender than individuals with fewer adults in the household.

To our knowledge, the findings in Table 3.2 represent the first evidence on how observable characteristics are related to the likelihood of identifying as transgender. Moreover, the patterns in Table 3.2 provide some evidence on the nature of transgender identification relative to the hypotheses discussed in the introduction. For example, the fact that more education is negatively associated with transgender identification is generally inconsistent with the idea that higher income (which is strongly correlated with higher education) allows for the cultivation of transgender identity. It could be that higher educated transgender individuals have the most to lose from identifying as such, which would be consistent with the patterns in Table 3.2. Another notable pattern from Table 3.2 is the fact that minority sexual orientation is the demographic characteristic most strongly related to the likelihood of transgender identification. It could be that individuals who have already faced the societal pressures associated with coming out as sexual minorities are more likely to be able to navigate the potentially related challenges of coming out as transgender.

3.4.3 Main Regression Results

Table 3.3 presents the regression results from equation (3.2). This table essentially asks whether transgender individuals have different employment, household income, insurance status, and self-rated health profiles than cisgender men even after accounting for the fact that they have significantly different observable characteristics (e.g., lower levels of education and higher likelihood of having a minority sexual orientation).

Table 3.3: Transgender Status and Socioeconomic outcomes
2014-2016 BRFSS, Adults ages 18-64

	(1) Employed or self-employed	**(2) Un-employed**	**(3) Unable to work**	**(4) Log (household income)**	**(5) Poverty**	**(6) Insured**	**(7) Excellent or very good health**	**(8) Fair or poor health**
Mean of outcome:	.684	.137	.070	10.759	.180	.860	.533	.153
Model 1:								
Cisgender woman	-.150***	.014***	.015***	-.186***	.068***	.021***	-.014***	.017***
	(.003)	(.002)	(.001)	(.006)	(.003)	(.002)	(.003)	(.002)
Transgender	-.095***	.032* (.018)	.021* (.013)	-.226***	.100***	-.005	-.075***	.014
	(.024)			(.045)	(.024)	(.019)	(.021)	(.019)
R-squared	.12	.08	.09	.11	.19	.14	.10	.09
N	387,197	387,197	387,197	337,789	337,045	389,239	390,029	390,029
Model 2:								
Cisgender woman	-.150***	.014***	.015***	-.186***	.068***	.021***	-.015***	.017***
	(.003)	(.002)	(.001)	(.006)	(.003)	(.002)	(.003)	(.002)
Transgender, MTF	-.063*	.031	.018	-.252***	.116***	.003	-.095***	-.034
	(.035)	(.026)	(.018)	(.059)	(.035)	(.028)	(.030)	(.024)
Transgender, FTM	-.094** (.041)	.013	.004	-.290***	.143***	0.040	-.014	.025
		(.033)	(.018)	(.084)	(.046)	(.035)	(.040)	(.033)
Transgender, GNC	-.172** (.045)	.060* (.035)	.051	.080	.024	.026	-.115***	.111**
			(.031)	(.109)	(.045)	(.029)	(.043)	(.043)
R-squared	.12	.08	.09	.11	.19	.14	.10	.09
N	387,197	387,197	387,197	337,789	337,045	389,239	390,029	390,029

*, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively. Estimates in columns 1-3 and 5-7 are from linear probability models. Estimates in column 4 are from interval regression models on log annual household income. All models control for 2015 and 2016 survey wave indicators, sexual orientation (indicators for gay/lesbian, bisexual, other sexual orientation, don't know sexual orientation, refusal to provide sexual orientation, and missing sexual orientation), age, age squared, race/ethnicity (indicators for black non-Hispanic, Asian, Hispanic, other race, and refusal to provide race/ethnicity), educational attainment (indicators for less than high school, some college, college or more, and refusal to provide educational attainment), marital status (indicators for married/partnered, divorced, widowed, separated, and refusal to provide marital status), Census region, whether the respondent was contacted on a cellphone or landline, and the number of adults in each household. McFadden's Adjusted R-squared values are reported for log income models. Standard errors are robust to heteroscedasticity.

Each column of Table 3.3 is from a similarly specified model with a different outcome: employed or self-employed in column 1, unemployed in column 2, unable to work in column 3, log annual household income in column 4, under the poverty threshold in column 5, has health insurance in column 6, excellent or very good self-rated health in column 7, and fair or poor health in column 8. The top panel presents results from the model where we include a single dummy variable for anyone who identifies as transgender; the bottom panel presents results from the model where we separately control for the three subcategories: transgender, male-to-female; transgender, female-to-male; and transgender, gender non-conforming.

The results in the top panel of Table 3.3 indicate that cisgender women and transgender individuals are significantly less likely to be employed, more likely to be unemployed, more likely to be unable to work, more likely to live in poverty, and less likely to have excellent or very good self-rated health compared to otherwise comparable cisgender men. Cisgender women and transgender individuals also have significantly lower household incomes than comparable cisgender men. A notable null finding is that transgender individuals are not significantly more or less likely to be insured compared with otherwise similar cisgender men; this result is somewhat surprising given the large employment differential in column 1 and the fact that most adults in the United States obtain insurance through their employer.[55] Moving to the bottom panel of Table 3.3, we see that the significant transgender employment difference persists for individuals who identify as transgender, male-to-female; transgender, female-to-male; and transgender, gender non-conforming. The significant differences related to log household income and poverty status also emerges for individuals who identify as transgender, male-to-female and transgender, female-to-male in the bottom panel of Table 3.3. In contrast, the differences related to excellent or very good self-rated health obtain for individuals who identify as transgender, male-to-female and transgender, gender non-conforming in the bottom panel of Table 3.3.[56]

In Table 3.4 we further investigate the differences associated with transgender status documented in Table 3.3 to probe the sources of the differences in socioeconomic outcomes. In particular, we focus on the role of household structure and the sex composition of other adults in the household.

---

[55] In results not reported but available upon request, we also estimated coarsened exact matching models (Blackwell et al. 2010) as an alternative approach to account for the differing patterns of observable characteristics between transgender and non-transgender individuals documented in Tables 3.1a and 3.1b. Those models returned patterns that were qualitatively identical to the baseline models reported in the top panel of Table 3.3 with one exception: the matching models returned evidence that transgender status was associated with a significantly higher probability of reporting fair or poor health, unlike in the baseline model in Table 3.3 (which returns no meaningful difference for that outcome associated with transgender status).

[56] Appendix C Table 3.8 provides an expanded set of coefficient estimates on the baseline model for employment. Coefficient estimates on the control variables are as expected (e.g., education and age are positively associated with employment).

Table 3.4: Exploring the Role of Household Sex Composition
2014-2016 BRFSS, Adults ages 18-64

	(1) Employed or self-employed	(2) Log of annual household income	(3) Poverty	(4) Excellent or very good health
1: Baseline model				
Cisgender woman	-.150*** (.003)	-.186*** (.006)	.068*** (.003)	-.015*** (.003)
Transgender, male-to-female	-.063* (.035)	-.252*** (.059)	.106*** (.035)	-.095*** (.030)
Transgender, female-to-male	-.094** (.041)	-.290*** (.084)	.143*** (.046)	-.014 (.040)
Transgender, gender non-conforming	-.172*** (.045)	-.080 (.109)	.024 (.045)	-.115*** (.043)
2: 1, but only the landline sample (where we know the sex composition of adults in the household)				
Cisgender woman	-.149*** (.003)	-.176*** (.008)	.058*** (.003)	-.009** (.004)
Transgender, male-to-female	-.006 (.046)	-.201** (.084)	.096* (.049)	-.095** (.038)
Transgender, female-to-male	-.082* (.049)	-.363*** (.112)	.135*** (.048)	.003 (.049)
Transgender, gender non-conforming	-.180*** (.054)	-.184 (.137)	.037 (.057)	-.152*** (.053)
3: 2 + control separately for number of adult men and adult women in the household				
Cisgender woman	-.137*** (.005)	-.134*** (.010)	.036*** (.004)	.011** (.005)
Transgender, male-to-female	-.034 (.049)	-.200* (.104)	.057 (.044)	-.065 (.045)
Transgender, female-to-male	.100 (.066)	-.143 (.125)	.058 (.052)	.015 (.069)
Transgender, gender non-conforming	-.105* (.061)	-.142 (.157)	-.083** (.039)	-.174** (.072)

*, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively. See notes to Table 3.3. The full sample consists of 387,197 respondents, 272,402 of which are in the landline sample.

One reason for doing so is the interesting pattern in the bottom panel of Table 3.2 that individuals who identify as transgender, gender non-conforming have significantly lower employment rates but do not have significantly different household incomes or poverty rates than otherwise comparable cisgender men. One possibility is that the presence of other employed individuals in the household is buffering the employment gap (although we have controlled for the total number of adults in the household in the baseline models of Table 3.3, the sex composition of the other adults in the household may partially explain these differences).

A challenge in accounting for the sex composition of adults in the household is that this information was only obtained during the 'household screener' portion of the BRFSS, and the household screener was not administered to individuals who participated by cellphone. In the top panel of Table 3.4 we reprint the baseline estimates from the bottom panel of Table 3.3 for the four outcomes with the most significant differences associated with transgender status: employment (column 1 of Table 3.4), log household income (column 2 of Table 3.4), poverty status (column 3 of Table 3.4), and excellent or very good self-rated health (column 4 of Table 3.4). In the middle panel of Table 3.4 we show the results from the same specifications but limited to the individuals who participated in the BRFSS via the landline sample (for whom we observe the sex composition of the adults in the household). Notably, most of the patterns are similar in sign, magnitude, and statistical significance, with one notable exception: the employment difference for individuals who are transgender, male-to-female does not remain in the landline sample. Overall, however, we conclude that the landline sample returns qualitatively similar patterns about the relationship between the various categories of transgender individuals and their relationship with the socioeconomic outcomes under study.

In the bottom panel of Table 3.4 we show the results from models for that same landline sample as in the middle panel of Table 3.4, but instead of controlling for the total number of adults in the household, we separately control for the number of adult men and the number of adult women in the household. The effects of doing so relative to the results in the middle panel of Table 3.4 are striking: the association between transgender status and the socioeconomic outcomes under study is substantially reduced in the vast majority of cases. For employment, for example, we estimate that the 8.2 percentage point lower likelihood of employment for individuals who are transgender, female-to-male becomes positive and statistically insignificant. The associated coefficient for individuals who are transgender, gender non-conforming (an 18 percentage point lower likelihood of employment compared to cisgender men) falls by over 40 percent in magnitude, though notably it remains negative and marginally significant at the ten percent level. For log household income, the large and significant difference for individuals who are transgender, male-to-female falls by over half and is no longer statistically significant. For poverty, whereas individuals who identify as transgender, male-to-female and individuals who identify as transgender, female-to-male were both significantly more likely to be in poverty in the middle panel of Table 3.4, when we account separately for the household sex composition in the bottom panel of Table 3.4, both coefficients fall substantially in magnitude and neither is statistically significant. Moreover, in the bottom panel of Table 3.4 we actually estimate that individuals who identify as transgender, gender non-conforming are 8.3 percentage points *less* likely to be in poverty than otherwise similar cisgender men after we control for the sex composition of the household.[57] Finally, column 4 of the bottom panel of Table 3.4 shows that the significantly lower likelihood of individuals who are transgender, male-to-female reporting excellent or very good self-rated health compared to cisgender men is not robust to controlling for the sex composition of adults in the household. Taken together, the results in Table 3.4 indicate that household sex composition accounts for a very large share of the observed differences in socioeconomic outcomes experienced by transgender individuals, particularly for the economic outcomes. Another general finding from the bottom panel of Table 3.4 is that the largest differences associated with transgender status in socioeconomic outcomes are observed for individuals who identify as

[57] Because poverty was calculated using thresholds already explicitly accounting for the number of adults in a household, the results reported in column 3 of panel 3 are from a model controlling for the share of adults in a household that are men rather than the number of adult men and adult women in the household.

transgender, gender non-conforming (significantly lower employment rates and significantly worse self-rated health compared to otherwise similar cisgender men) in comparison to the associated differentials for individuals who identify as transgender, male-to-female or individuals who identify as transgender, female-to-male.

In Table 3.5 we turn to investigating heterogeneity in the transgender employment differential documented in Tables 3.3 and 3.4. We focus on employment because it is an individual level outcome that is not confounded by the household-level measure required for the BRFSS income variable. Specifically we ask whether the employment differential for transgender individuals compared to cisgender men varies systematically by education, race/ethnicity, minority sexual orientation, and state policy environment (defined as the presence of a statewide trans-inclusive Employment Non-Discrimination Act, or ENDA). Columns 1-3 of Table 3.5 report estimates from models where we interact the CISGENDER WOMAN, TRANSGENDER WOMAN, TRANSGENDER MAN, and TRANSGENDER GENDER NON-CONFORMING indicators with indicators for: having less than a high school education in column 1; being nonwhite in column 2, and identifying as gay, lesbian, bisexual, or other non-heterosexual sexual orientation (i.e., any minority sexual orientation) in column 3.

Table 3.5: Heterogeneity Analyses

Outcome is Employed or Self-Employed; 2014-2016 BRFSS, Adults ages 18-64 year olds

	(1) Education	(2) Race	(3) Minority sexual orientation	(4) Trans-inclusive ENDA
Cisgender woman	-.130*** (.003)	-.138*** (.003)	-.153*** (.003)	-.151*** (.004)
Transgender, MTF	-.078** (.032)	-.057* (.035)	-.043 (.038)	-.058 (.045)
Transgender, FTM	-.082* (.044)	-.116** (.050)	-.112** (.045)	-.107* (.055)
Transgender, GNC	-.148*** (.049)	-.164*** (.053)	-.107* (.058)	-.156*** (.055)
Less than high school education	-.083*** (.008)	--	--	--
Cisgender woman * Less than HS degree	-.155*** (.010)	--	--	--
Transgender, MTF * Less than HS degree	.027 (.100)	--	--	--
Transgender, FTM * Less than HS degree	-.084 (.108)	--	--	--
Transgender, GNC * Less than HS degree	-.182 (.114)	--	--	--
Nonwhite	--	-.014*** (.005)	--	--
Cisgender woman * Nonwhite	--	-.034*** (.006)	--	--
Transgender, MTF * Nonwhite	--	-.023 (.078)	--	--
Transgender, FTM * Nonwhite	--	.042 (.082)	--	--
Transgender, GNC * Nonwhite	--	-.020 (.091)	--	--
Any minority sexual orientation	--	--	-.074*** (.011)	--
Cisgender woman * Trans-inclusive ENDA	--	--	.071*** (.015)	--
Transgender, MTF * Any minority S.O.	--	--	-.066 (.089)	--
Transgender, FTM * Any minority S.O.	--	--	.102 (.103)	--
Transgender, GNC * Any minority S.O.	--	--	-.108 (.089)	--
Trans-inclusive ENDA	--	--	--	.013*** (.004)
Cisgender woman * Trans-inclusive ENDA	--	--	--	.003 (.005)
Transgender, MTF * Trans-inclusive ENDA	--	--	--	-.018 (.67)
Transgender, FTM * Trans-inclusive ENDA	--	--	--	.038 (.081)
Transgender, GNC * Trans-inclusive ENDA	--	--	--	-.040 (.094)
R-squared	.13	.12	.12	.12
N	387,197	387,197	387,197	387,197

*, **, and *** denote statistical significance at 10%, 5%, and 1%, respectively. See notes to Table 3.3. The outcome in all columns is an indicator for being employed or self-employed.

These individual characteristics are independently associated with lower employment; thus, a negative and significant interaction between the transgender indicator and these other at-risk groups would be consistent with 'double disadvantage' theories.

The results in columns 1-3 of Table 3.5 do not return evidence of statistically significant 'double disadvantages' in employment, though some of the coefficients are very large in magnitude (e.g., the interaction of less than a high school education with the indicator for identifying as transgender, gender non-conforming). In all of the cases of columns 1-3 of Table 3.5 we do find that the main effect of having less than a high school education, being a racial minority, and identifying as a sexual minority, respectively, are individually statistically significant.[58]

Column 4 of Table 3.5 investigates the role of state-level employment non-discrimination acts (ENDAs) that explicitly protect transgender individuals in public and private employment sectors. ENDAs may potentially offset adverse labor market experiences and employment-based discrimination commonly reported by transgender workers. To test for the association between state-level ENDAs and employment, we present estimates from equation (3.3) in which we

[58] Note that in the model of column 3 of Table 3.5 individuals who refused or reported they don't know their sexual orientation have been dummied out as a separate group and included in the model, though we do not report the coefficients. They are available upon request.

interact the various TRANSGENDER dummy variables with an indicator for living in a state with trans-inclusive non-discrimination protections covering private sector employment.[59] These results are presented in column 4 of Table 3.5 and provide no evidence that trans-inclusive non-discrimination protections in private employment are associated with significantly higher employment rates for transgender women or transgender men compared to cisgender men in those states.[60] Notably, the coefficient on the main effect of residing in a state with a trans-inclusive ENDA is positive and significant, suggesting that there are important unobserved characteristics about the types of states that adopt such protections and/or the types of jobs or people that live in those places.[61]

[59] We do not observe the sector of employment in the BRFSS, but: 1) the vast majority of employed individuals are private sector workers; and 2) all states with trans-inclusive employment protections covering the private sector also cover employees in the public sector.

[60] We also tested alternative definitions of state policy environments, including having any ENDA regardless of whether the ENDA was trans-inclusive or limited to public sector employment, and results were similar.

[61] States adopting trans-inclusive ENDAs may have different educational attainment patterns, different industries, or different demographics associated with higher employment rates compared to states not adopting trans-inclusive ENDAs. Meanwhile, transgender individuals may systematically choose to live in or move to states with trans-inclusive employment non-discrimination protections. Transgender individuals may also be more likely to come out to an interviewer as transgender in the presence of such protections. We do not find strong evidence of this in the data, however. Specifically, a model predicting the likelihood that an individual identifies as transgender as a function of demographic characteristics and the presence of a trans-inclusive ENDA did not uncover meaningful relationships between the trans-inclusive policy and the likelihood of identifying as transgender (see Table 3.2).

3.5 Discussion and Conclusion

We used newly available data on self-identified transgender individuals from 31 states in the 2014-2016 BRFSS to provide the first large-scale evidence using population-based and representative data on how transgender status is related to socioeconomic outcomes such as employment, household income, poverty, health insurance coverage, and self-rated health in the US. We also provide some of the first descriptive information for self-identified transgender individuals, showing that they have much lower education levels than individuals who do not identify as transgender. We first model the decision to identify as transgender and find that sexual minorities are more likely to identify as transgender while highly educated individuals are less likely to identify as transgender. Our regression models for socioeconomic outcomes that account for observable demographic characteristics return evidence that – compared with cisgender men – transgender individuals report significantly lower employment rates, lower household incomes, higher rates of poverty, and lower rates of having excellent or very good self-rated health. Accounting for the differential sex composition of adults in the household can explain a substantial portion of these differences, particularly for the economic outcomes.

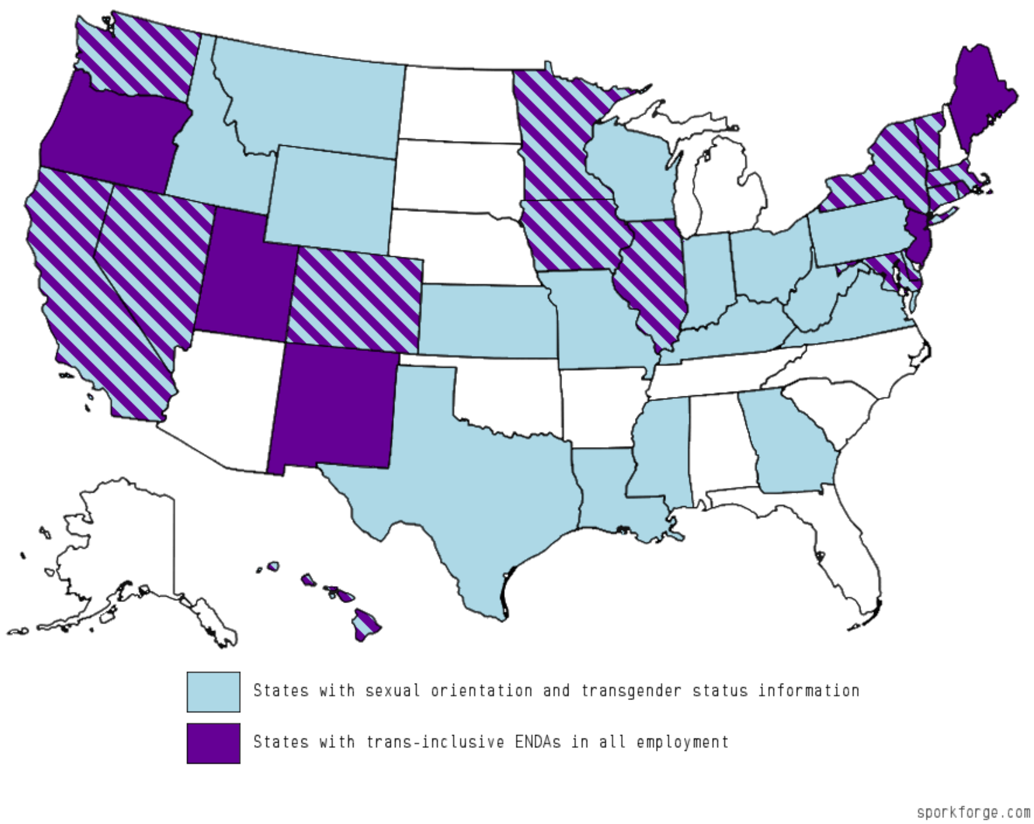
It is interesting to compare our findings with those from the small existing literature on differences in socioeconomic outcomes associated with a minority gender identity. Both Schilt and Wiswall (2008) and Geijtenbeek and Plug (2018, forthcoming) examine within-person variation in gender expression as measured by the timing of when individuals seek physical and/or legal changes to one's sex, and both find individuals transitioning from male-to-female experience significant reductions in earnings, while individuals transitioning from female-to-male experience no reduction in earnings and perhaps a small increase. Our results complement those of Schilt and Wiswall (2008) and Geijtenbeek and Plug (2018, forthcoming) by examining a related but distinct set of socioeconomic outcomes. Unlike their studies, we cannot identify before/after changes within the same individuals. Also, we do not know how long, for example, transgender women experienced the advantages in human capital accumulation and labor market treatment from presenting as male. Despite this limitation, our finding that the largest and most significant differentials in socioeconomic outcomes accrue to individuals who identify as transgender, gender non-conforming (bottom panel of Table 3.4) is particularly interesting because these individuals are not likely to be included in the 'before/after gender transition' approaches for identifying transgender individuals in previous studies.

Our study is subject to some notable limitations, many of which pertain to the data. First, there is some debate amongst scholars in this area about the most appropriate way to ask gender identity and transgender status in large surveys. We are limited to the specific questions asked in the BRFSS. Second, and closely related, we are limited in using a self-reported transgender status measure. As described above, there may be systematic selection associated with disclosing to a survey administrator about being transgender (e.g., it is plausible that transgender individuals with financial and social resources are the only ones who can 'afford' to come out as transgender). We do note that the two questions about gender identity never require individuals to say the word 'transgender' or 'male-to-female' (i.e., they can indicate 'yes' or 'no' to the transgender status question and can identify asserted gender by referring to the numerical choice among multiple stated options). A related issue is measurement error associated with transgender-identified individuals who are 'false positives' (i.e., not transgender but marked as such by the survey). This problem is particularly severe given the very low prevalence of transgender identification in the survey: a few tenths of one percent of the sample identifies as transgender. Even very small rates of measurement error could substantially contaminate the transgender sample, and if these errors are correlated with socioeconomic status (for example because low educated individuals may not know what transgender is despite incorrectly identifying as such), then transgender status would be mechanically negatively correlated with economic outcomes. While transgender individuals are significantly less likely to be white non-Hispanic in Table 3.1 than non-transgender individuals (possibly indicating language-related measurement errors in identifying as transgender), we do note that there are other patterns that suggest the signal in the transgender identification measure is valid. For example, transgender-identified individuals are significantly more likely to identify as non-heterosexual, a pattern corroborated in other independently drawn data.

Third, our sample of transgender adults only includes non-institutionalized adults randomly selected among landline and cell phone users in US households. Missing from our analysis were homeless adults and adults residing in institutionalized medical facilities, incarceration facilities, and homeless shelters. Data from non-representative samples of transgender individuals suggest that these exclusions may disproportionately affect transgender individuals, since transgender individuals report high rates of homelessness and incarceration compared to the general population (Grant et al. 2011, Burwick et al. 2014, James et al. 2016). Finally, our results may not be generalizable to the entire transgender population, as our study only includes data from 31 states. Our sample has reasonable coverage of the Northeast and Midwest but disproportionately excludes the Southern United States (Figure 3.1).

Despite these limitations, our paper makes an important contribution to understanding how a minority gender identity is independently related to a range of socioeconomic outcomes. By showing that transgender individuals do have significantly different socioeconomic outcomes compared with cisgender men, our findings should spur additional research on this important population.

Figure 3.1: Map of BRFSS States with Sexual Orientation and Transgender Status Information and with Trans-Inclusive ENDAs, 2014-2016



Specifically, future research should further explore the determinants and consequences of the transgender employment, income, poverty, and health differentials we document here, including whether and how public policies besides trans-inclusive ENDAs might improve relative socioeconomic outcomes for transgender individuals. Alternative policies, for instance, may include investments in support services; protections from discrimination for transgender people in healthcare, education, and/or housing; or workforce development and training.[62]

[62] For example, the state of California recently supported an employment program with incentives targeted at hiring transgender individuals (Duran 2016).

3.6 References

Akerlof, George and Rachel Kranton (2000). "Economics and Identity," The Quarterly Journal of Economics, 115(3): 715-753.

Badgett, M.V. Lee (1995). "The Wage Effects of Sexual-Orientation Discrimination," Industrial and Labor Relations Review, 48(4): 726-739.

Balingit, Moriah (2016). "Another 10 States Sue Obama Administration Over Bathroom Guidance for Transgender Students," The Washington Post, July 8, 2016.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes (2005). "Welfare Reform and Health," Journal of Human Resources, 40(2): 309-334.

Blackwell, Matthew, Stefano Iacus, Gary King, and Giuseppe Porro (2010). "cem: Coarsened Exact Matching in Stata," February 22, 2010 working paper available at: http://gking.harvard.edu/files/gking/files/cem-stata.pdf.

Blandford, John (2003). "The Nexus of Sexual Orientation and Gender in the Determination of Earnings," Industrial and Labor Relations Review 56(4): 622-642.

Burwick, Andrew, Gary Gates, Scott Baumgartner, and Daniel Friend (2014). Human Services for Low-Income and At-Risk LGBT Populations: An Assessment of the Knowledge Base and Research Needs. OPRE Report Number 2014-79. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Bye Larry, Elizabeth Gruskin, Gregory Greenwood, Victoria Albright and Karol Krotki (2005). "California Lesbians, Gays, Bisexuals, and Transgender (LGBT) Tobacco Use Survey—2004." Sacramento, CA: California Department of Health Services.

Carpenter, Christopher (2008). "Sexual Orientation, Work, and Income in Canada," Canadian Journal of Economics, 41(4): 1239-1261.

----- (2007). “Revisiting the Earnings Penalty for Behaviorally Gay Men: Evidence from NHANES III,” Labour Economics, 14(1): 25-34.

----- (2005). “Self-Reported Sexual Orientation and Earnings: Evidence from California,” Industrial and Labor Relations Review, 58(2): 258-273.

----- (2004). “New Evidence on Gay and Lesbian Household Incomes,” Contemporary Economic Policy, 22: 78-94.

Carpenter, Christopher and Samuel T. Eppink (2017). “Does It Get Better? Recent Estimates of Sexual Orientation and Earnings in the United States,” Southern Economic Journal, 84(2): 426-441.

Carpenter, Christopher, and Gary J. Gates. "Gay and lesbian partnership: Evidence from California." Demography 45.3 (2008): 573-590.

Center for American Progress and Movement Advancement Project (2015). “Paying an Unfair Price: The Financial Penalty for Being Transgender in America.” Available at: http://www.lgbtmap.org/file/paying-an-unfair-price-transgender.pdf.

Cerf Harris, Benjamin (2015). “Likely Transgender Individuals in U.S. Federal Administrative Records and the 2010 Census,” working paper.

Clain, Suzanne and Karen Leppel (2001). “An investigation into sexual orientation discrimination as an explanation for wage differences,” Applied Economics, 33(1): 37-47.

Coffman, Katherine B., Lucas C. Coffman, and Keith M. Marzilli Ericson (2017). “The Size of the LGBT Population and the Magnitude of Antigay Sentiment Are Substantially Underestimated,” Management Science, 63(10): 3168-3186.

Conron, Kerith, Gunner Scott, Grace Sterling Stowell, and Stewart J. Landers (2012). “Transgender Health in Massachusetts: Results from a Household Probability Sample of Adults,” American Journal of Public Health, 102(1): 118-122.

Dahlhamer, James M., Adena M. Galinsky, Sarah S. Joestl, and Brian W. Ward (2014). “Sexual Orientation in the 2013 National Health Interview Survey: A Quality Assessment,” Vital Health Stat 2(169).

Duran, Leo (2016). "California Restaurants Launch Nation's First Transgender Jobs Program," Radio broadcast episode in 'Food for Thought' segment of Morning Edition, National Public Radio, October 19, 2016. Also available at: http://www.npr.org/sections/thesalt/2016/10/19/494102719/california-restaurants-launch-nations-first-transgender-jobs-program. Last accessed November 8, 2016.

Drydakis, Nick. "Sexual orientation discrimination in the labour market." Labour Economics 16.4 (2009): 364-372.

Equal Employment Opportunity Commission (2012). Macy v. Holder. Appeal No. 0120120821 (April 20, 2012). Available at: https://www.eeoc.gov/decisions/0120120821%20Macy%20v%20DOJ%20ATF.txt

Evans, William N. and Craig L. Garthwaite (2014). "Giving Mom a Break: The Impact of Higher EITC Payments on Maternal Health," American Economic Journal: Economic Policy, 6(2): 258-90.

Flores, Andrew R., Jody L. Herman, Gary J. Gates, and Taylor N. T. Brown (2016). "How Many Adults Identify as Transgender in the United States," Williams Institute Report. Available at: http://williamsinstitute.law.ucla.edu/research/how-many-adults-identify-as-transgender-in-the-united-states/

Garthwaite, Craig, Tal Gross, and Matthew J. Notowidigdo (2014). "Public Health Insurance, Labor Supply, and Employment Lock," Quarterly Journal of Economics, 129(2): 653-696.

Grant, Jaime M., Lisa A. Mottet, Justin Tanis, Jack Harrison, Jody L. Herman, and Mara Keisling (2011). "Injustice at Every Turn: A Report of the National Transgender Discrimination Survey. Washington, DC: National Center for Transgender Equality and National Gay and Lesbian Task Force.

Geijtenbeek, Lydia and Erik Plug (2018, forthcoming). "Is there a penalty for registered women? Is there a premium for registered men? Evidence from a sample of transsexual workers," European Economic Review, forthcoming.

The GenIUSS Group (2014). Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys. J. L. Herman (Ed.). Los Angeles, CA: The Williams Institute. Available at: http://williamsinstitute.law.ucla.edu/wp-content/uploads/geniuss-report-sep-2014.pdf.

Herman, Jody L. (2016). "LGB within the T: Sexual Orientation in the National Transgender Discrimination Survey and Implications for Public Policy." In Trans Studies: The Challenge to Hetero/Homo Normatives. Yolanda Martinez-San Miguel and Sarah Tobias, Eds. Rutgers University Press.

Human Rights Campaign (2016). "2015 State Equality Index: A Review of State Legislation Affecting the Lesbian, Gay, Bisexual, and Transgender Community and a Look Ahead in 2016." Report available at: http://www.hrc.org/sei2015.

James, Sandy E., Jody L. Herman, Susan Rankin, Mara Keisling, Lisa Mottet, and Ma'ayan Anaf (2016). "The Report of the 2015 U.S. Transgender Survey." Washington, DC: National Center for Transgender Equality.

Joloza, Theodore, Joanne Evans, Rachel O'Brien, and Angela Potter-Collins (2010). "Measuring Sexual Identity: An Evaluation Report." UK Office for National Statistics.

Klawitter, Marieka (2015). "Meta-analysis of the Effects of Sexual Orientation on Earnings," Industrial Relations: A Journal of Economy and Society, 54(1): 4-32.

Klawitter, Marieka and Victor Flatt (1998). "The Effects of State and Local Antidiscrimination Policies on Earnings for Gays and Lesbians," Journal of Policy Analysis and Management, 17(4): 658-686.

Kosciw, Joseph, Emily Greytak, Neal Palmer, and Madelyn Boesen (2014). "The 2013 National School Climate Survey." Report available at: http://www.glsen.org/article/2013-national-school-climate-survey

Kosciw, Joseph G., Emily A. Greytak, Noreen M. Giga, Christian Villenas, and David J. Danischewski (2016). "The 2015 National School Climate Survey." Report available at: https://www.glsen.org/sites/default/files/GLSEN%202015%20National%2

0School%20Climate%20Survey%20%28NSCS%29%20-%20Full%20Report.pdf

Kowalski, Robin M. and Susan P. Limber (2013). "Psychological, Physical, and Academic Correlates of Cyberbullying and Traditional Bullying," Journal of Adolescent Health, 53: S13-S20.

Liptak, Adam (2016). "Supreme Court to Rule in Transgender Access Case," The New York Times, October 28, 2016. Also available at: http://www.nytimes.com/2016/10/29/us/politics/supreme-court-to-rule-in-transgender-access-case.html. Last accessed on November 8, 2016.

Meyer, Ilan (1995). "Minority Stress and Mental Health in Gay Men," Journal of Health and Social Behavior, 36(1): 38-56.

Olsen, Emily O., Laura Kann, Alana Vivolo-Kantor, Steve Kinchen, and Tim McManus (2014). "School Violence and Bullying Among Sexual Minority High School Students, 2009-2011," Journal of Adolescent Health, 55(3): 432-438.

Reisner Sari, Tonia Poteat, JoAnne Keatley, Mauro Cabral, Tampose Mothopeng, Emilia Dunham, Claire Holland, Ryan Max, Stefan Baral (2016). "Global Health Burden and Needs of Transgender Populations: A Review." The Lancet, published online.

Ruhm, Christopher J. (2005). "Healthy Living in Hard Times," Journal of Health Economics, 24: 341-363.

Russell, Stephen T., Caitlin Ryan, Russell B. Toomey, Rafael M. Diaz, and Jorge Sanchez (2011). "Lesbian, Gay, Bisexual, and Transgender Adolescent School Victimization: Implications for Young Adult Health and Adjustment," Journal of School Health, 81(5): 223-230.

Schilt, Kristen and Matthew Wiswall (2008). "Before and After: Gender Transitions, Human Capital, and Workplace Experiences," The B.E. Journal of Economic Analysis & Policy, 8(1): Article 39.

United States Department of Education Office for Civil Rights (2014). "Questions and Answers on Title IX and Sexual Violence," Report available at: http://www2.ed.gov/about/offices/list/ocr/docs/qa-201404-title-ix.pdf

Weichselbaumer, Doris (2003). "Sexual Orientation Discrimination in Hiring," Labour Economics, 10: 629-642.

Xavier, Jessica, Julie A. Honnold and Judith Bradford (2007). "The Health, Health-Related Needs, and Lifecourse Experiences of Transgender Virginians: Virginia Transgender Health Initiative Study Statewide Survey Report," January 2007. Available online at: http://www.vdh.virginia.gov/epidemiology/diseaseprevention/documents/pdf/THISFINALREPORTVol1.pdf.

Appendix C

Table 3.6: Comparisons of Economic Outcomes in BRFSS and March CPS, 2014-2016

	(1) BRFSS	(2) CPS
Employed	.682* (.510)	.721* (.272)
Average Household Income	65,474.84* (49,199.43)	71,024.01* (27,016.96)
Share in HH income ranges		
0 – 9,999	.059* (.260)	.073* (.158)
10,000 – 14,999	.047* (.235)	.040* (.119)
15,000 – 19,999	.073* (.288)	.042* (.123)
20,000 – 24,999	.084* (.308)	.046* (.127)
25,000 – 34,999	.095* (.325)	.088* (.173)
35,000 – 49,999	.130* (.374)	.125* (.202)
50,000 – 74,999	.153* (.400)	.179* (234)
75,000 or more	.359* (.532)	.406* (.299)
Share below HH income thresholds		
Below 10,000	.059* (.260)	.073* (.158)
Below 15,000	.106* (.341)	.112* (.193)
Below 20,000	.178* (.425)	.155* (.221)
Below 25,000	.262* (.488)	.200* (.244)
Below 35,000	.357* (.532)	.289* (.276)
Below 50,000	.488* (.554)	.414* (.300)
Below 75,000	.641* (.532)	.593* (.300)
Sample Size	393,304	37,596

Weighted means (standard deviations). * indicates the means are significantly different between columns 1 and 2 at p<.05.

Table 3.7: Comparing States that Did and Did Not Contribute SOGI Data to the Public Use BRFSS

2014-2016 BRFSS, Adults ages 18-64

Variable	(1) Individuals in the 19 states that did not contribute SOGI data to the 2014, 2015, or 2016 public use BRFSS	(2) Individuals in the 31 states that contributed SOGI data to the 2014, 2015, and/or 2016 public use BRFSS
Age	41.000* (15.569)	40.672* (12.824)
Male	.496 (.567)	.499 (.469)
White, non-Hispanic	.638* (.545)	.595* (.461)
Gay or lesbian	--	.018 (.134)
Bisexual	--	.024 (.153)
Partnered	.551* (.564)	.555* (.466)
High school degree or less	.417 (.559)	.414 (.462)
Some college	.328* (.532)	.310* (.434)
College degree or more	.256* (.494)	.276* (.420)
Number of adults in HH	2.285* (1.259)	2.352* (1.091)
Any children in HH	.433* (.561)	.449* (.467)
Employed or self-employed	.659* (.537)	.675* (.440)
Average household income	60,983.23* (49,562.87)	64,692.24* (41,775.78)
At or below 100% poverty	.193 (.456)	.194 (.369)
Has health insurance	.836* (.420)	.852* (333)
Very good or excellent health	.528* (.566)	.533* (.468)
Fair or poor health	.168* (.424)	.155* (.340)
Respondent's state characteristics:		
Trans-inclusive ENDA	.199* (.453)	.454* (.467)
Marriage equality before *Obergefell*	.760* (.484)	.704* (.429)
Republican governor	.914* (.318)	.455* (.467)
Sample size	336,277	552,482

Weighted means (standard deviations). Note average household income and poverty status are determined using the midpoint of each household income range or the 80th percentile of annual household income for those who reported the highest income category; percent of poverty is calculated by dividing household income by household size specific U.S. Census Bureau poverty thresholds, following Conron et al. (2012). * indicates the means are significantly different between columns 1 and 2 at p<.05.

Table 3.8: Expanded set of coefficient estimates for Employed or Self-Employed
2014-2016 BRFSS, Adults ages 18-64

	(1) **Baseline from Table 3.3, Column 1, Bottom Panel**
Cisgender woman	-.150*** (.003)
Transgender, male-to-female	-.063* (.035)
Transgender, female-to-male	-.094** (.041)
Transgender, gender non-conforming	-.172*** (.045)
Gay/Lesbian	-.026** (.011)
Bisexual	-.040*** (.011)
Other sexual orientation	-.078*** (.029)
Age	.048*** (.001)
Age squared	-.001*** (.000)
Black	-.033*** (.005)
Asian	-.069*** (.007)
Other race	-.058*** (.008)
Hispanic ethnicity	-.000 (.005)
Less than HS degree	-.149*** (.006)
Some college	.025*** (.004)
College degree or more	.131*** (.003)
Partnered	.053*** (.004)
Divorced	.017*** (.005)
Widowed	-.044*** (.010)
Separated	-.016* (.010)
Northeast Census Region	-.004 (.004)
Midwest Census Region	.015*** (.004)
South Census Region	-.011** (.005)
# adults in the household	-.004*** (.012)
2015 Survey Wave	.000 (.003)
2016 Survey Wave	.018*** (.004)
In the cellphone-only sample	.027*** (.004)
R-squared	.12
N	237,732

** and *** denote statistical significance at 5% and 1%, respectively. See notes to Table 3.3.