Suddenly Married: Joint Taxation and the Labor Supply of Same-Sex Married Couples After *U.S. v. Windsor*

Elliott Isaac¹

Tulane University

June 10, 2022

Abstract

Joint taxation can exacerbate the deadweight loss of taxation due to labor supply responses, but evidence is scarce. I estimate the labor supply effects and efficiency costs of joint taxation in the United States by leveraging tax variation created by federal same-sex marriage recognition following the 2013 *United States v. Windsor* Supreme Court ruling. I estimate significant compensated elasticities along the extensive, but not intensive, margin among both higher and lower earners. My findings suggest that joint taxation is less efficient and generates less tax revenue than individual taxation, and that lowering tax rates for secondary earners could improve efficiency.

JEL: J22, H24, H21, D10

Keywords: taxation, labor supply, same-sex marriage, sufficient statistics

¹Elliott Isaac is an assistant professor of economics at Tulane University and a research affiliate at IZA (eisaac@tulane.edu). I would like to thank Leora Friedberg, Amalia Miller, Jonathan Colmer, Gaurab Aryal, Patrick Button, Ben Leyden, Zach Sullivan, and Sarah Turner, and other seminar participants at the University of Virginia, Annual Congress of the International Institute of Public Finance, and Annual Conference of the National Tax Association for comments. I am also grateful to Emily Hart for excellent research assistance. This research was supported in part using high performance computing (HPC) resources and services provided by Technology Services at Tulane University, New Orleans, LA. The author has nothing else to disclose. An Online Appendix has been included. The data used in this article and additional replication materials are available online: https://doi.org/10.3886/E173183V1.
1 Introduction

Many developed countries institute progressive tax systems, forcing them to choose the unit of taxation: the individual or the family (Rosen 1977). Choosing the family as the unit of taxation (as in the United States) achieves greater tax equity across families, but can exacerbate efficiency costs through labor supply distortions. However, direct evidence of these efficiency costs is relatively rare due to a lack of natural experiments involving a large-scale switch in systems. Another complication is that such tax regime switches generally involve a change in marriage incentives as well, introducing another margin of distortion. In this paper, I provide direct evidence of the labor supply effects, efficiency costs, and tax revenue consequences of joint taxation in the United States by leveraging tax variation created by federal recognition of existing same-sex marriages following the 2013 United States v. Windsor Supreme Court ruling.

The United States’ 1996 Defense of Marriage Act prevented same-sex marriages from being recognized at the federal level, but states retained the authority to permit same-sex marriages at the state level. This legislative environment meant that same-sex married couples were still required to file federal taxes as two single individuals through 2012. In June 2013 over 71,000 marriages recognized by states since 2004 were suddenly recognized by the United States federal government as a result of the United States v. Windsor ruling (DeSilver 2013). The ruling required same-sex married couples to file federal taxes as married beginning in tax year 2013.1

This shift of already-married couples from the individual to the family as the unit of taxation is unusual because governments usually employ only one system at any particular time. I leverage separate variation in household after-tax income and marginal tax rates generated by federal same-sex marriage recognition to separately identify the income and substitution effects of taxation and, therefore, compensated (Hicksian) labor supply elasticities. I then apply my
estimates to an existing sufficient statistic framework to calculate the additional deadweight loss and tax revenue created by joint taxation relative to individual taxation implied by my estimates.

I use the 2012-2015 waves of the American Community Survey, which are the first of the U.S. Census Bureau surveys to explicitly identify same-sex married couples. I use a generalized difference-in-differences framework, which compares predicted changes in net-of-tax wage rates and household after-tax income for a treatment group of individuals in same-sex married couples to a control group of individuals in different-sex married couples. I restrict the sample to same- and different-sex married couples who had already married before the United States v. Windsor ruling in order to exploit the plausibly exogenous shock to federal marital recognition while abstracting away from actual marriage decisions. I also extend work by Baldwin, Allgrunn, and Ring (2011) and McClelland, Mok, and Pierce (2014) by separately analyzing the labor supplies of higher and lower earners, which increases the similarity between treatment and control groups because higher and lower earners exhibit different labor supply dynamics. Higher or lower earners may also respond to taxation differently due, for example, to differing work preferences or attachment to the labor force. Separately analyzing each partner yields separate estimates of the effects of taxation by earning status while abstracting away from sex.

Distinct variation in net-of-tax wage rates and household income is crucial in estimating compensated labor supply elasticities. Variation in an individual’s tax rates due to the Supreme Court ruling originates through two primary channels: differences in tax bracket definitions between the single and joint schedules, and the addition of both partners’ earnings to taxable income rather than only the individual’s earnings. On average, higher earners in same-sex couples in my sample faced a $0.01 increase (0.04%) and lower earners faced a $0.23 decrease
(1.0%) in their predicted net-of-average-tax wage rates as a result of the United States v. Windsor ruling.

To estimate the income effect, I measure changes in the couple’s predicted marriage subsidy (or penalty), which is a common measure of marriage incentives under joint taxation (Alm and Whittington 1999; Eissa and Hoynes 2003; Michelmore 2018; Isaac 2020; Friedberg and Isaac Forthcoming), but which is new to the literature estimating income effects of taxation because shifting from the single to joint tax schedule is endogenous in most circumstances. The marriage subsidy is defined as the difference between the sum of the individuals’ tax liabilities if they are single and the couple’s joint tax liability if they are married. The average same-sex married couple experienced a decrease of $448.51 in predicted household after-tax income due to a tax-based marriage penalty, with substantial variation across households depending on total household earnings and how those earnings are split between partners.

I estimate significant extensive margin Hicksian (wage) and income elasticities among both higher and lower earners, with larger elasticities among lower earners and women. I estimate relatively large extensive margin Hicksian elasticities of 0.527 and 0.791 for higher and lower earners, respectively, and income participation elasticities of -0.051 and -0.363 for higher and lower earners, respectively.iii These Hicksian elasticities are in line with estimates by Selin (2014), who studies the switch from joint to individual taxation in Sweden in 1971, but smaller than those suggested by LaLumia (2008), who studies the original switch to joint taxation in the United States in 1948, and larger than those reported by Chetty (2012), although he notes that there are “economically significant differences in structural extensive margin elasticities across groups.” In contrast, I do not estimate any significant responses to taxation along the intensive margin. Although I estimate relatively large extensive margin elasticities, my estimates suggest
that the introduction of joint taxation decreased labor force participation of lower earners by only 0.8 percentage points (1.1%) and increased labor force participation of higher earners by only 0.02 percentage points (0.02%), which is a smaller effect compared to past findings by LaLumia (2008) and Kališková (2014).

I also examine heterogeneity by sex and the presence of children. My estimates reveal that men’s labor supply is relatively less elastic than women’s labor supply, as other researchers have concluded. A back-of-the-envelope calculation suggests that sex explains at least 47% of the difference in elasticities between lower earning women and higher earning men, but earning status in the couple can still explain up to 53% of the difference. My estimates also suggest that children may create labor supply frictions among parents that reduce their labor supply responsiveness to taxes.

Finally, I use an existing sufficient statistic framework to calculate the changes in deadweight loss and tax revenue imposed by joint taxation relative to individual taxation as implied by my elasticities (Feldstein 1999; Immervoll et al. 2007; Chetty 2009; Kleven 2021). My estimates suggest that the shift to joint taxation among same-sex married couples induced $56.9 million in additional deadweight loss and cost $54.1 million in tax revenue, which are smaller effects than past estimated revenue effects of same-sex marriage legalization by Stevenson (2012) and Alm, Leguizamon, and Leguizamon (2014). Extending my elasticity estimates to all married couples in the United States suggests that joint taxation increases deadweight loss by $16.4 billion and generates $26.9 billion less in tax revenue relative to individual taxation, although I discuss caveats to this extrapolation in Section 6.2.

This paper is grounded in traditional labor supply and taxation questions, but also adds to the small, but growing, literature concerning same-sex couples and the labor supply effects of
same-sex marriage legalization (Hansen, Martell, and Roncolato 2019; Sansone 2019). This paper is the first, to the best of my knowledge, to leverage tax variation among same-sex married couples to identify the effects of taxation on the labor supply of married couples. My analysis provides direct evidence of the additional efficiency costs and reduced tax revenue of joint taxation relative to individual taxation, and suggests that lowering tax rates among secondary earners, so as to mitigate the efficiency costs along the extensive margin, could further improve efficiency of the United States tax system. However, whether increased efficiency is worth the lower associated tax equity across families remains an open question.

The remainder of the paper is organized as follows. Section 2 discusses policy background and prior research. Section 3 presents the empirical strategy, Section 4 discusses the data, and Section 5 presents the results and robustness checks. Finally, Section 6 presents the deadweight loss and tax revenue analysis, and Section 7 concludes.

2 Background

2.1 Same-Sex Marriage Legislation in the United States

The Defense of Marriage Act (DOMA) was established in 1996 and defined, for federal government purposes, “marriage” as the union between one man and one woman and “spouse” as a member of the opposite sex who is a husband or wife. Despite these definitions at the federal level, states retained the right to allow same-sex marriages, and Massachusetts became the first state to do so in 2004.\(^iv\)

DOMA’s definitions of marriage and spouse prevented same-sex married couples from obtaining any federal benefits of marriage available to different-sex married couples. In June 2013 the Supreme Court ruled that DOMA’s definitions of marriage and spouse were unconstitutional, thereby requiring the federal government to recognize same-sex marriages at
the federal level. This had immediate effects on the tax environment faced by same-sex couples who had already married by the time of the ruling, and affected all same-sex married couples regardless of where they resided or whether their state of residence allowed same-sex marriages. Where previously these couples filed federal tax returns as two single individuals, beginning in tax year 2013 they were required to file as either married, filing jointly or married, filing separately (U.S. Department of the Treasury, Internal Revenue Service 2013).

2.2 Literature on Labor Supply and Taxation

Keane (2011) notes that past researchers have generally found small, if any, effects of taxation on male labor supply, with compensated (Hicksian) elasticities ranging from 0.05-0.84 with an average of 0.31, and larger labor supply responses to taxation among women. Many of these studies exploit tax variation caused by small to moderate tax changes among a subset of tax filers. For example, Crossley and Jeon (2007), Saez, Slemrod, and Giertz (2012), and Saez (2017) use tax variation at the top of the earnings distribution to estimate labor supply and earnings responses among the highest earners. Others, such as Eissa and Liebman (1996), Ellwood (2000), Eissa and Hoynes (2004), Moulton, Graddy-Reed, and Lanahan (2016), and Bastian (2020) use variation in the Earned Income Tax Credit to estimate hours and participation responses to taxes among likely low earners. Fetter and Lockwood (2018) exploit the introduction of the Old Age Assistance Program to estimate labor force participation decisions among older workers.

The United States v. Windsor ruling, on the other hand, created large-scale variation that is more similar to that exploited by LaLumia (2008), who uses the introduction of joint taxation in the United States in the 1940s to estimate labor force participation responses among married couples, Selin (2014), who estimates the effects of a switch from joint to individual taxation in
Sweden in 1971, or Kališková (2014), who estimates the impact of moving to joint taxation using a voluntary switch to joint taxation in the Czech Republic in 2005. These are the only studies, to the best of my knowledge, offering direct evidence of the effects of joint taxation through natural experiments from changes in tax systems. The Supreme Court ruling effectively shifted same-sex married couples from an individual taxation system to one of joint taxation, and, as a result I am able to exploit a recent natural experiment to study changes in both tax rates and tax liabilities.

Most of the previously mentioned studies distinguish between male and female labor supply, often assuming that the wife is the secondary earner in the household. Blau and Kahn (2007) and Heim (2007), however, document shrinking labor supply elasticities among women over the last four decades, and suggest that women’s status as secondary earners has weakened over time. Baldwin, Allgrunn, and Ring (2011) and McClelland, Mok, and Pierce (2014) consider primary or secondary earner status regardless of sex and compare the results to those using the traditional male-female split. In this paper, I am able to break this link between sex and earning status by using a sample of same-sex couples and estimating heterogeneous effects by earning status and sex.

Finally, I use a sufficient statistic approach similar to Feldstein (1999) and Immervoll et al. (2007) and discussed more generally by Chetty (2009) and Kleven (2021) to estimate the additional deadweight loss and tax revenue created by joint taxation relative to individual taxation. I find that lower earners in same-sex couples face incentives to exit the labor force, thereby creating deadweight loss, which corroborates Kleven, Kreiner, and Saez’s (2009) conclusion that optimal rates can exhibit negative, rather than positive, jointness.

2.3 Economic Research on Same-Sex Couples and LGBT Individuals
Economic research concerning same-sex couples and LGBT individuals is scarce due to few available sources of data. Until recently, data editing procedures in prior waves of the census and American Community Surveys, such as changing the sex or marital status of same-sex partners, made it difficult to identify same-sex couples in the data.\(^{viii}\)

Oreffice (2011) and Antecol and Steinberger (2013) use the 2000 decennial census to study bargaining power in unmarried same-sex couples and labor supply differences relative to different-sex couples, respectively. Oreffice (2011) estimates greater bargaining power for richer and younger partners in same-sex couples, and Antecol and Steinberger (2013) find that children explain 52% of the hours gap between female secondary earners in different- and same-sex couples.

Stevenson (2012) and Alm, Leguizamon, and Leguizamon (2014) use the American Community Survey to predict labor supply and federal tax revenue consequences of same-sex marriage legalization. Stevenson (2012) predicts a $20-40 million increase in federal tax revenue, while Alm, Leguizamon, and Leguizamon (2014) conclude that the federal government could gain $5.7 million or lose up to $315.8 million. My calculations of the change in tax revenue as a result of *United States v. Windsor* are closer to estimates from Alm, Leguizamon, and Leguizamon (2014). Friedberg and Isaac (Forthcoming) leverage tax variation created by state same-sex marriage legalization, *United States v. Windsor*, and *Obergefell v. Hodges* to estimate the effect of taxes on marriage among same-sex couples. They find significant marriage responses to federal taxes, which are smaller than estimates from studies of different-sex couples.

More generally, Sansone (2019) and Hansen, Martell, and Roncolato (2019) estimate the overall effect of same-sex marriage legalization on labor supply, but do not consider the specific
role of taxes. These studies present mixed conclusions, with Sansone (2019) finding that same-
sex marriage legalization increased employment, and Hansen, Martell, and Roncolato (2019) finding decreased hours of work only among female couples. The overall effects implied by my estimates are more in-line with Hansen, Martell, and Roncolato (2019).ix

3. Empirical Strategy

I use a generalized difference-in-differences, or treatment intensity, framework to estimate hours and labor force participation responses to taxation among individuals in same-sex married couples compared to a control group of individuals in different-sex married couples. My empirical strategy leverages variation in the couple’s tax liability and in each partner’s marginal and average tax rates as a result of the switch from individual to joint taxation due to United States v. Windsor. I separately analyze the labor supplies of higher and lower earners, which increases the similarity between treatment and control groups because higher and lower earners exhibit different labor supply dynamics.

3.1 Specification

My specifications take the following general form:

\[ Y_{it} = \gamma_0 + \gamma_1(1 - \tau_{it})w_{it} + \gamma_2\Delta T_{ijt} + \gamma_3SSMC_i + \gamma_4X_{it} + \delta_t + \mu_s + \epsilon_{it} \]

Where \( Y_{it} \) is either annual hours of work or labor force participation, \( \Delta(1 - \tau_{it})w_{it} \) is the individual’s change in their net-of-tax wage rate in levels due to United States v. Windsor, \( \Delta T_{ijt} \) is the couple’s change in household income due to the marriage subsidy introduced by United States v. Windsor, \( SSMC_i \) is equal to one if the couple is a same-sex married couple, and \( X_{it} \) is a vector of additional covariates including age, age squared, race, education, sex, and the state unemployment rate. I also control for the individual’s own net-of-tax wage rate in levels (and its
square), their partner’s net-of-tax wage rate in levels (and its square), and the interaction between their own and their partner’s net-of-tax wage rate in levels (and its square) to ensure that identification of the effect of the change in net-of-tax wage rates is not driven by general earnings effects. \( \delta_t \) and \( \mu_x \) are year and state fixed effects, respectively. Year fixed effects will capture national-level time-varying shocks that may affect labor supply among same- and different-sex married couples. State fixed effects will capture, among other characteristics, overall state attitudes toward same-sex relationships and local labor market discrimination against LGBT individuals, which may be correlated with both individual labor supply and the state’s decision to recognize same-sex marriages (Gao and Zhang 2017). Of course, one limitation of these fixed effects is that they cannot control for all possible policy changes that may have differentially affected same-sex relative to different-sex married couples and, instead, can only control for differences affecting all couples in the sample.

\( \Delta(1 - \tau_{lt})w_{lt} \) measures the change in the net-of-tax wage rate (marginal or average) due to *United States v. Windsor* among individuals in same-sex married couples relative to individuals in different-sex married couples. I define this variable as the difference between the individual’s net-of-tax wage rate in year \( t \) and their counterfactual net-of-tax wage rate in year \( t \) if *United States v. Windsor* had not occurred:

\[
(2) \quad \Delta(1 - \tau_{lt})w_{lt} = (1 - \tau_{lt})w_{lt} - (1 - \tau_{l,\text{Pre-Windsor}})w_{lt}
\]

Where \( (1 - \tau_{lt})w_{lt} \) is the net-of-tax wage rate in year \( t \) and \( (1 - \tau_{l,\text{Pre-Windsor}})w_{lt} \) is the counterfactual net-of-tax wage rate the individual would face in year \( t \) if *United States v. Windsor* had not occurred. By construction, \( (1 - \tau_{lt})w_{lt} = (1 - \tau_{l,\text{Pre-Windsor}})w_{lt} \) and \( \Delta(1 - \tau_{lt})w_{lt} = 0 \) for the pre-policy treatment group (same-sex married couples before *United States v. Windsor*) because the ruling had not yet happened and for the control group (different-
sex married couples) because they were not affected by the ruling. For the post-policy treatment group (same-sex married couples after United States v. Windsor), the change in the net-of-tax wage rate reflects the wage effect of the introduction of joint taxation among same-sex married couples.

\[ \Delta T_{ijt} \] measures the couple’s change in household income due to the marriage subsidy introduced by United States v. Windsor:

(3) \[ \Delta T_{ijt} = T_{ij,t}^{\text{pre-Windsor}} - T_{ijt} \]

Where \( T_{ijt} \) is the couple’s total tax liability in year \( t \) and \( T_{ij,t}^{\text{pre-Windsor}} \) is the counterfactual total tax liability the couple would face in year \( t \) if United States v. Windsor had not occurred. By construction, \( T_{ij,t}^{\text{pre-Windsor}} = T_{ijt} \) and \( \Delta T_{ijt} = 0 \) for the pre-policy treatment group and for the control group.xi For the post-policy treatment group, the change in household income due to the marriage subsidy reflects the income effect of the introduction of joint taxation among same-sex married couples. A positive value indicates an increase in household income due to a lower joint tax liability after United States v. Windsor, whereas a negative value indicates a tax-based marriage penalty and a decrease in household income.xii

Of primary concern is that wage rates are unobserved for non-workers, meaning that net-of-tax wage rates in an extensive margin analysis will only vary among observed workers. I would, therefore, not be able to measure changes in net-of-tax wage rate among non-workers, some of whom may have been induced to leave the labor market due to large tax increases. Earnings are also endogenous to labor supply, meaning that higher tax rates will be correlated with higher wage rates and labor supply. I control for cross-sectional variation in wage rates, but it is still likely that wage rates and, therefore, the change in tax rates due to United States v. Windsor, suffer from endogeneity. To circumvent this endogeneity issue and obtain a potential
wage for non-workers, I predict individual earnings following the process outlined in Section 4, and use predicted earnings and the NBER TAXSIM simulator to estimate each individual’s tax rates and liabilities as well as the change in each variable due to *United States v. Windsor.* I bootstrap standard errors to address potential errors in the earnings prediction process.

The coefficients of interest are $\gamma_1$, which determines the total uncompensated (Marshallian) elasticity, and $\gamma_2$, which determines the income effect. Identification of $\gamma_1$ and $\gamma_2$ comes from individual- and couple-level tax rate and liability variation across individuals in same-sex married couples compared to individuals in different-sex married couples before and after the Supreme Court ruling conditional on other observable covariates. The identifying assumption I must make in this generalized difference-in-differences framework is that annual hours of work and labor force participation of same- and different-sex married couples would have evolved parallel to each other in the absence of the Supreme Court ruling, which I explore in Section 4. I also find that my estimates are robust to including state-by-year fixed effects, which allows me to control for more nuanced shocks that may vary at the state and year level while leaving identification of the effect of tax changes due to *United States v. Windsor* intact.

### 3.2 Elasticities

I use the estimated coefficients from Equation 1 to estimate uncompensated (Marshallian) elasticities, income elasticities, and compensated (Hicksian) elasticities along the intensive and extensive margins. Following Gruber and Saez (2002) and Keane (2011), the total effect of a wage rate change is given by the Slutsky equation below, which I have multiplied by $\frac{(1-\eta)w}{H}$ to convert into elasticities:
\[
\frac{\partial H}{\partial (1 - \tau)w} \frac{(1 - \tau)w}{H} = \frac{\partial H}{\partial (1 - \tau)w} \frac{(1 - \tau)w}{H} + \frac{\partial H}{\partial (1 - \tau)w} \frac{(1 - \tau)wH}{Y} \]

Where \( H \) is hours worked, \((1 - \tau)w\) is the net-of-tax wage rate, and \( Y \) is non-labor income. To estimate these elasticities, I interpret \( \gamma_1 \) as the effect of a change in the net-of-tax wage rate and \( \gamma_2 \) as the effect of a change in household non-labor income, leading to the intensive margin elasticities:

\[(5a) \quad \text{Marshallian:} \quad \varepsilon_M = \frac{\partial H}{\partial (1 - \tau)w} \frac{(1 - \tau)w}{H} = \frac{\partial H}{\partial \text{change in } (1 - \tau)w} \frac{(1 - \tau)w}{H} = \hat{\gamma}_1 \frac{(1 - \bar{\tau})\bar{w}}{H} \]

\[(5b) \quad \text{Income:} \quad \varepsilon_i = \frac{\partial H}{\partial Y} \frac{Y}{H} = \hat{\gamma}_2 \frac{\bar{Y}}{H} \]

\[(5c) \quad \text{Hicksian:} \quad \varepsilon_H = \frac{\partial H}{\partial (1 - \tau)w} \frac{(1 - \tau)w}{H} - \frac{\partial H}{\partial Y} (1 - \tau)w = \hat{\gamma}_1 \frac{(1 - \bar{\tau})\bar{w}}{H} - \hat{\gamma}_2 (1 - \bar{\tau})\bar{w} \]

where \( \bar{H} \) is the average hours worked in the sample, \( \bar{Y} \) is the average household non-labor income in the sample, and \((1 - \bar{\tau})\bar{w}\) is the average net wage rate in the sample.

Economic theory predicts that the Hicksian elasticity is positive and, if leisure is a normal good, that the income effect is negative. Therefore, I expect \( \hat{\gamma}_2 < 0 \) and \( \hat{\gamma}_1 \frac{(1 - \bar{\tau})\bar{w}}{H} - \hat{\gamma}_2 (1 - \bar{\tau})\bar{w} > 0 \), but cannot predict the sign of \( \hat{\gamma}_1 \) along the intensive margin a priori because the Marshallian elasticity can be positive or negative depending on the dominating effect.

The Hicksian and Marshallian elasticity concepts are the same along the extensive margin because the income effect term in Equation 4 is zero:

\[(6a) \quad \text{Hicksian:} \quad \eta_H = \frac{\partial \text{LFP}}{\partial (1 - \tau)w} \frac{(1 - \tau)w}{\text{LFP}} = \frac{\partial \text{LFP}}{\partial \text{change in } (1 - \tau)w} \frac{(1 - \tau)w}{\text{LFP}} = \hat{\gamma}_1 \frac{(1 - \bar{\tau})\bar{w}}{\text{LFP}} \]

\[(6b) \quad \text{Income:} \quad \eta_i = \frac{\partial \text{LFP}}{\partial Y} \frac{Y}{\text{LFP}} = \hat{\gamma}_2 \frac{\bar{Y}}{\text{LFP}} \]
where $\overline{LFP}$ is the average labor force participation rate in the sample. Along the extensive margin, economic theory predicts that $\hat{\gamma}_1 > 0$ and $\hat{\gamma}_2 < 0$.

4 Data

I use the 2012-2015 waves of the American Community Survey to estimate the effects of taxation on annual hours of work and labor force participation (Ruggles et al. 2019). My main sample includes individuals between 30 and 60 years old in same- and different-sex married couples who married in 2012 or earlier.\textsuperscript{xiv,xv}

By focusing on same-sex couples who had already married by the time of the Supreme Court ruling, I can abstract away from the marriage decision and use tax rate and household income changes associated with the change in federal marriage recognition. I include all individuals when examining labor force participation, and further restrict the sample to working individuals when examining annual hours of work, leaving me with 2,676,072 individuals in my extensive margin sample (10,496 in same-sex married couples and 2,665,576 in different-sex married couples), and 2,238,615 individuals in my intensive margin sample (8,966 in same-sex married couples and 2,229,649 in different-sex married couples).

4.1 Demographics

Online Appendix Table A.1 presents summary statistics of demographic variables among higher and lower earners in same- and different-sex married couples, for which I control in my main specifications.\textsuperscript{xvi} Notably, different-sex couples have been married for one year longer, on average, relative to comparable same-sex couples. Same-sex couples are also slightly older, earn more, have more years of education, and are slightly more likely to be Black, but are less likely
to be Hispanic or Asian and are less likely to have children, on average, than different-sex couples.

Table 1 presents summary statistics of observed and predicted labor supply variables. Observed labor force participation and annual hours of work conditional on working are both lower among lower earners in different-sex married couples compared to lower earners in same-sex married couples, suggesting a potential causal effect of taxes that I estimate below. Higher earners generally work full-time, full-year.

Figure 1 displays trends over time in the means of the outcome variables in Table 1 for higher and lower earners. Most outcome variables exhibit visually similar pre-period mean trends, suggesting that the samples remain relatively balanced. One exception is the slight relative decrease in labor force participation among higher earners in same-sex married couples in 2014, although labor force participation among higher earners is very high and this trend is a decrease from a 97% to a 96% participation rate.

Note that if a same-sex couple reports themselves to be married even though they reside in a state that does not recognize same-sex marriages, then I assume the couple married in a state that did recognize same-sex marriages. 48.3% of same-sex couples in my sample are observed living a state that legalized same-sex marriage before United States v. Windsor. Migration between states is infrequent: in my sample, at least one partner in 2.5% of same-sex couples moved between states. More telling, only 8% of those movers moved to a state that did not legalize same-sex marriage before United States v. Windsor from a state that did. Given this low migration rate, it seems more likely that couples observed living in states that did not legalize same-sex marriage before Windsor travelled to another state to marry. Although the bias that migration responses could introduce is theoretically ambiguous, I show in Online Appendix
D that my main results are robust to restricting the sample to couples observed in states that legalized same-sex marriage before *Windsor*, suggesting that selective migration is not biasing my main estimates.

4.2 Predicted Gross Wage Rates

In order to overcome the primary endogeneity concerns outlined above, I follow the general method of Gruber and Saez (2002), Dahl and Lochner (2012), and Friedberg and Isaac (Forthcoming) by predicting individual earnings based on predetermined characteristics, and use predicted earnings to obtain each individual’s predicted tax rates and liabilities. This process provides a predicted level of earnings for all individuals in the sample, including non-workers.xx

I use a machine learning least absolute shrinkage and selection operator (LASSO) to obtain the earnings predictions. The LASSO is a penalized regression, similar to a ridge regression, that selects a subset of the available variables that best fit the data (Tibshirani 2011).xxi This powerful machine learning approach allows me to provide a large number of variables and interactions while allowing the LASSO to select the most appropriate set of variables to predict earnings. Variables that I included, but which the LASSO may have ultimately ignored, include five-year age groups, four education groups, number of children, and fixed effects for race, sex, two-digit occupation, college major, and state, as well as pairwise interactions between all these variables. I limit the prediction sample to individuals observed in 2012, so that my predictions do not reflect labor supply responses following *United States v. Windsor*, estimate the LASSO separately for individuals in different- and same-sex couples, and follow a two-step procedure to predict individual earnings. xxii I first use a LASSO to predict whether each individual has positive earnings using a linear probability model. I convert these predicted probabilities into a binary variable by setting a threshold in the distribution of predicted
positive earnings such that the binary variable has the same observed sample mean of having positive predicted earnings. If the predicted probability is less than this threshold, I assign $0 in predicted earnings to that individual. For individuals who are predicted to have positive earnings, I use another LASSO regression to predict their earnings level, estimated on those in the sample with positive observed earnings.\textsuperscript{xxiii} I then divide predicted earnings by 2,080 (52 weeks multiplied by 40 hours per week) to obtain a predicted measure of each individual’s full-time gross wage rate: \( w_{it} = \frac{\text{Predicted earnings}}{2080} \).\textsuperscript{xxiv} Table 1 shows that these predicted earnings and gross wage rates understate the observed values for higher earners in the sample and overstate the observed values for lower earners in the sample, as is expected since not all observed lower earners are working full-time.\textsuperscript{xxv}

### 4.3 Net-of-Tax Wage Rates

The United States uses a progressive, family-based system of income taxation with four tax schedules: single, head of household (single with dependent children), married filing jointly, and married filing separately. Unmarried taxpayers are required to file as either single or head of household and married taxpayers are required to file as either married filing separately or married filing jointly. The vast majority of married taxpayers file jointly because the married filing separately schedule usually results in larger tax liabilities (Fisher, Gee, and Looney 2018).\textsuperscript{xxvi} Online Appendix Figure A.1 displays the four tax schedules for 2013.

I use the NBER TAXSIM simulator and predicted earnings to obtain each individual’s predicted federal marginal and average tax rates in year \( t \) and the counterfactual tax rates the individual would face in year \( t \) if \textit{United States v. Windsor} had not occurred (Feenberg and Coutts 1993). These tax rates are functions of only predicted earned income, number of children, and state of residence taking into account federal recognition of same-sex marriages. Same-sex
married couples observed in 2012 or 2013 (tax years 2011 or 2012) were required to file as single or head of household at the federal level and those observed in 2014 or 2015 (tax years 2013 or 2014) were required to file as married.

I denote the net-of-marginal-tax wage rate as \((1 - \tau_{it})w_{it}\) and the net-of-average-tax wage rate as \((1 - \bar{\tau}_{it})w_{it}\), where \(w_{it}\) is the predicted gross wage rate, \(\tau_{it}\) is the predicted marginal tax rate and \(\bar{\tau}_{it} = \frac{\text{Predicted tax liability}}{\text{Predicted earnings}}\) is the predicted average tax rate.\textsuperscript{xxvii} I use these predicted net-of-tax wage rates to construct \(\Delta(1 - \tau_{it})w_{it}\), defined in Equation 2, which is one of my main explanatory variables and reflects the wage effect of the introduction of joint taxation among same-sex married couples.

Variation in marginal and average tax rates due to the Supreme Court ruling comes from two sources: the shift from the single to joint tax schedule and the addition of both partners’ earnings into taxable income. Figure 2 displays variation in the observed net-of-tax wage rates to provide a sense of how net-of-tax wage rates changed in the sample following the introduction of joint taxation. In general, lower earners faced lower net-of-tax wage rates (that is, higher marginal and average tax rates) under joint taxation relative to individual taxation, while higher earners faced more muted changes, but generally higher net-of-tax wage rates. Table 1 shows that, on average, higher earners in same-sex married couples faced a $0.07 decrease (0.3%) in their predicted net-of-marginal-tax wage rate and a $0.01 increase (0.04%) in their predicted net-of-average-tax wage rate as a result of United States v. Windsor, while lower earners in same-sex married couples faced a $0.38 decrease (1.9%) in their predicted net-of-marginal-tax wage rate and a $0.23 decrease (1.0%) in their predicted net-of-average-tax wage rate as a result of United States v. Windsor with sizable variation among both higher and lower earners. As with the other earnings variables, the predicted net-of-tax wage rates understate the observed values for higher
earners and overstate the observed values for lower earners. The observed and predicted net-of-
average-tax wage rates are larger than the observed and predicted net-of-marginal-tax wage
rates, as is expected under a progressive tax system.

4.4 The Marriage Subsidy

The nature of the Supreme Court ruling allows me to use the marriage subsidy as a new source of
variation in household income to identify the income effect, and, by extension, the Hicksian
wage elasticity.\textsuperscript{xviii} The marriage subsidy is traditionally defined as the difference between the
sum of the individuals’ tax liabilities if they are single and the couple’s joint tax liability if they
are married. It is a common measure of tax incentives to marry or divorce in the family structure
literature (Alm and Whittington 1995; 1999; Ellwood 2000; Eissa and Hoynes 2003; Michelmore
2018; Isaac 2020; Friedberg and Isaac Forthcoming), but is new to the literature seeking to
estimate income effects of taxation because shifting from the single to joint tax schedule is
endogenous in most circumstances.\textsuperscript{xxix}

I use the NBER TAXSIM simulator and predicted earnings to obtain each individual’s
and couple’s predicted federal tax liability in year $t$ and the counterfactual tax liability the
individual and couple would face in year $t$ if \textit{United States v. Windsor} had not occurred. These
tax liabilities are functions of only predicted earned income, number of children, and state of
residence taking into account federal recognition of same-sex marriages. I use these predicted tax
liabilities to construct $\Delta T_{ijt}$, defined in Equation 3, which is one of my main explanatory
variables and reflects the income effect of the introduction of joint taxation among same-sex
married couples.

Variation in the marriage subsidy is driven not only by total household earnings, but also
by how those earnings are split between partners. Figure 3 displays variation in the marriage
subsidy for same-sex couples in my sample, separated by the couple’s earnings split and by the presence of children. Figure 3 shows that, in general, the marriage subsidy tends to increase as the earnings split becomes more uneven, as seen in panels A and B relative to panels E and F, and couples with children and more even earnings splits are more likely to face marriage penalties, as seen in panels D and F relative to panels C and E.\textsuperscript{xxx,xxxi}

Table 1 shows that, on average, same-sex married couples experienced a predicted decrease of $448.51 in household income due to a tax-based marriage penalty, with substantial variation across households, whereas the observed value is a $371.61 subsidy. Note, however, that the mean observed marriage subsidy reported in Table 1 will be affected by labor supply decisions following \textit{United States v. Windsor}. For example, if the Supreme Court ruling resulted in more unequal earnings splits within same-sex couples, then the observed marriage subsidy would tend to be higher, on average. The fact that the predicted value is a penalty is also likely due to the fact that the earnings prediction process produces relatively more equal earner couples, on average, which tends to decrease the marriage subsidy and can turn it into a penalty.

\section{Results}

Table 2 presents my baseline extensive margin estimates from Equation 1.\textsuperscript{xxxii} I estimate statistically significant effects of taxation on labor force participation among all groups. Using all individuals, I estimate that a $1 increase in the net-of-average-tax wage rate increases labor force participation by 5.6 percentage points (6.7\%) and a $1,000 increase in household income decreases labor force participation by 4.7 percentage points (5.6\%), implying a large Hicksian participation elasticity of 1.821 ($p < 0.01$) and an income elasticity of -0.241 ($p < 0.01$). Note, however, that the comparison between the treatment and control groups is less straightforward in the full sample because it combines higher and lower earners, who exhibit different labor supply
trends. Limiting the sample to higher earners, I estimate that a $1 increase in the net-of-average-tax wage rate increases labor force participation by 1.3 percentage points (1.3%) and a $1,000 increase in household income decreases labor force participation by 1.1 percentage points (1.1%), implying a Hicksian participation elasticity of 0.527 ($p < 0.01$) and an income elasticity of -0.051 ($p < 0.01$). Limiting the sample to lower earners, I estimate that a $1 increase in the net-of-average-tax wage rate increases labor force participation by 3.5 percentage points (5.0%) and a $1,000 increase in household income decreases labor force participation by 6.3 percentage points (9.0%), implying a Hicksian participation elasticity of 0.791 ($p < 0.01$) and an income elasticity of -0.363 ($p < 0.01$).

It is also possible to control for more nuanced state-by-year variation using state-by-year fixed effects because the *United States v. Windsor* ruling affected all same-sex married couples regardless of where they live or whether their state of residence recognized same-sex marriage, thereby leaving identification of the effect of the tax changes intact. State-by-year fixed effects will capture state-time varying shocks and unobservables that may affect labor supply of same- and different-sex married couples, such as changes in access to state welfare programs. However, as with separate state and year fixed effects, these state-by-year fixed effects remain limited to controlling for differences affecting all couples in the sample, not just same-sex couples. In other words, although state-by-year fixed effects can control for some more nuanced policy variation that may bias my baseline estimates, they cannot fully control for time-varying policy changes specific to same-sex marriage that could still be correlated with my treatment variables. With this caveat in mind, Tables 2-3, Columns 4-5, present these results. My elasticity estimates using state-by-year fixed effects are essentially unchanged, indicating that state-time varying shocks do not confound my original estimates.
My Hicksian participation elasticity estimates are comparable to estimates by Selin (2014), who finds participation elasticities between 0.5 and 1, but smaller than those implied by LaLumia (2008), who studies the original switch to joint taxation in the United States in 1948. My elasticity estimates are larger than, but remain in-line with, many others from samples pre-dating federal same-sex marriage recognition, such as those from Bastian (2020), who exploits the introduction of the Earned Income Tax Credit, and Keane (2011), who reports male Hicksian hours elasticities between 0.02 and 1.32, with an average of 0.31, and female Hicksian participation elasticities between 0.01 to 1.60. Chetty (2012), in his examination of micro and macro labor supply elasticities, notes that there are “economically significant differences in structural extensive margin elasticities across groups.”

My income participation elasticity estimates among lower earners are larger than those by Blau and Kahn (2007) and Heim (2007). Blau and Kahn (2007) estimate female income elasticities between -0.111 and -0.140 when accounting for taxes, although their outcome variable is annual hours of work including zeros, which captures a combination of the extensive and intensive margins. I estimated an alternative specification using annual hours of work (including zeros) as the outcome variable and estimated an income elasticity of -0.200 ($p < 0.01$) among all women, which is closer to Blau and Kahn’s (2007) estimate. My larger income participation elasticities seem reasonable, therefore, given that I estimate separate intensive and extensive margin responses in my main samples while estimating a combined effect that is similar to estimates from Blau and Kahn (2007).

Heim (2007) estimates a female income participation elasticity around -0.075 when accounting for taxes, which is also smaller than my estimates. One explanation for this difference may be the lower labor force participation rate among lower earners in my sample relative to
Heim’s (2007) sample of married women. Assuming an equivalent tax response, a lower labor force participation rate would tend to increase elasticity estimates overall.

At the mean values of the predicted tax change variables, my participation elasticities suggest that the introduction of joint taxation decreased labor force participation of lower earners by only 0.8 percentage points (1.1%) and increased labor force participation of higher earners by only 0.02 percentage points (0.02%). These average effects are smaller than past findings, such those by LaLumia (2008) and Kališková (2014) who estimate a 2 and 3 percentage point decline in female employment, respectively, due to the introduction of joint taxation.xxxiv

Table 3 presents my baseline intensive margin estimates. In contrast to the participation results, I do not estimate any significant effects on hours of either net-of-marginal-tax wage rates or changes in household income due to the marriage subsidy. My estimates translate into insignificant, imprecise, and sometimes opposite-signed Hicksian and income hours elasticities. However, the 95% confidence intervals of my elasticity estimates allow me to rule out Hicksian hours elasticities greater than 0.306 and 0.799 for higher and lower earners, respectively, and income hours elasticities larger than -0.017 and -0.050 for higher and lower earners, respectively, which are smaller than Blau and Kahn’s (2007) estimates.

The noisy and mixed hours elasticity estimates may be due to the fact that I use cross-sectional data and find significant participation responses, which necessarily erodes my ability to correctly estimate significant hours responses because it implies a selection issue on the intensive margin. The mixed hours elasticity estimates may also be due to measurement error in self-reported usual hours worked per week and weeks worked last year (in intervals).xxxv This means it may be more difficult to precisely estimate hours responses to taxes in this context. To the best
of my knowledge, labor force participation in the ACS is not as susceptible to this type of problem, leading to more precise estimates of the effect of taxes on labor force participation.

5.1 Heterogeneity by Sex and Children

The different Hicksian elasticities for higher and lower earners in Tables 2-3 may also reflect different labor supply preferences between men and women or between couples with and without children. The top panel of Table 4 presents heterogeneous participation elasticity estimates for men and women. I estimate larger and significant Hicksian participation elasticities for women and continue to estimate larger elasticities for lower earners than for higher earners. I estimate insignificant Hicksian participation elasticities among both higher and lower earning men, but find that higher earning men, who are often assumed to be primary earners in prior studies, have the smallest point estimates. My income participation elasticities are similar to my baseline results among both men and women. For hours of work in the bottom panel of Table 4, I continue to find insignificant or opposite-signed Hicksian hours elasticities and income elasticities.

Comparing the differences in participation elasticities between men and women of different earning statuses in Table 4 can separate the effect of sex from that of earning status in the couple. A back of the envelope calculation suggests that sex explains at least 47% of the difference in elasticities between lower earning women and higher earning men, but earning status in the couple can still explain up to 53% of the difference.\textsuperscript{xxxvi} This finding differs from McClelland, Mok, and Pierce (2014), who instead conclude that being a secondary earner is a more important determinant of labor force participation than sex.

Table 5 presents heterogeneous elasticity estimates for men and women by both earning status and presence of children. Note that, although these estimates are useful in exploring
heterogeneous effects by the presence of children, they should be interpreted cautiously because having children is a choice that is likely related to labor supply. I estimate significant participation elasticities among childless women, which mirror the heterogeneous elasticity estimates by sex in Table 4. In contrast, I estimate smaller and statistically insignificant participation elasticities among all men and among women with children. Along the intensive margin, I continue to find generally insignificant or opposite-signed Hicksian hours elasticities.

These estimates suggest that the presence of children may create labor supply frictions among parents and reveal a monotonic increase in participation elasticities by sex and earning status among childless couples. I find that higher earning childless men have the smallest participation elasticity, followed by lower earning childless men, higher earning childless women, and, finally, lower earning childless women, who have the largest participation elasticity.

5.2 Robustness Checks and Alternative Specifications

I estimate several robustness specifications meant to alleviate concerns about measurement error in the ACS, inconsistent state and federal tax policies, and sample selection with regard to married couples. I also consider an alternative process of calculating wage rates. In all cases I continue to estimate insignificant or opposite-signed Hicksian hours elasticities, and I therefore limit my discussion below to the participation elasticity estimates. Overall, I find qualitatively similar participation results, indicating that my main findings are robust to a number of potential concerns.

There are two main measurement error concerns related to same-sex couples that need to be explored. First, Black et al. (2007) and Gates and Steinberger (2009) document substantial measurement error in identifying same-sex couples in the 2000 decennial census and the 2005-
2007 ACS, respectively, due to different-sex couples mis-marking the sex of one partner so that the couple appears to be a same-sex couple in the data. To address this concern, I restrict the sample to couples who responded using a computer assisted telephone or personal interview process (CATI/CAPI) rather than the traditional mail-in form, which has been shown to exhibit much lower gender mis-marking (Gates and Steinberger 2009; Kreider and Lofquist 2015; Lofquist 2015). Online Appendix Table A.2, Columns 1-2, presents these elasticity estimates, which are smaller than, but qualitatively similar to, my main results.\textsuperscript{xxxvii}

The second measurement error concern is that the year of marriage, which I use to select the sample of couples who married before 2013, may be allocated and result in a sample that includes some same-sex married couples who married after 2013. To address this concern, Online Appendix Table A.2, Columns 3-4, display estimates from a specification with additional controls for the length of marriage and an indicator variable for whether both partners’ year of marriage was allocated. Columns 5-6 display estimates using a sample that excludes couples where both partners’ year of marriage are allocated. In both cases, I obtain estimates that are essentially unchanged relative to my original estimates, suggesting that my main results are robust to these potential data quality issues.

Another concern from the same-sex couple literature is that compliance barriers may have limited take-up of joint filing among same-sex married couples after the United States v. Windsor (Fisher, Gee, and Looney 2016; 2018). Ten states in 2014 required married couples to file jointly on their state tax returns if they filed jointly on their federal tax returns, while simultaneously not allowing same-sex married couples to file jointly on state tax returns, which can create compliance barriers to filing jointly in some states and may bias my coefficient estimates toward zero.\textsuperscript{xxxviii} To address this concern, I excluded same-sex couples residing in
states with inconsistent tax policies. Online Appendix Table A.2, Columns 7-8, show that the elasticity estimates using this sample restriction are similar to my baseline estimates, indicating little bias due to inconsistent tax policies.

There are also two primary sample selection concerns that may bias my estimates. First, my main sample includes same- and different-sex married couples who married in 2012 or earlier, meaning that couples observed in 2015 will all have been married for at least three years whereas couples observed in 2012 may be newly married. The concern is that the results may be driven by a selected sample comparing more stable couples who have been married for longer in the post-Windsor period to less stable couples who are recently married in the pre-Windsor period. To address this concern, I restrict the sample to married couples who have been married for at least three years to impose the same marriage duration constraints across all observations. Note that missing variables in the 2012 wave of the ACS allow me to use only observations from 2013 or later in this specification. Online Appendix Table A.2, Columns 9-10, show that the elasticity estimates using this sample restriction remain stable, indicating little bias due to greater marriage duration constraints on couples observed in later waves.

Second, the original different-sex couple control group shares a similar legal relationship status and a similar post-ruling tax environment with same-sex married couples, but same-sex cohabiting couples have similar pre-ruling tax environments, leaving unclear which group serves as the best control group. Therefore, I re-estimate my model using same-sex cohabiting couples as an alternative control group. Note that I am unable to observe the duration of a cohabiting couple’s relationship, limiting my ability to restrict this sample in the same way as married couples. Online Appendix Table A.2, Columns 11-12, present the results. I estimate slightly smaller participation elasticities of 0.419 for higher earners and 0.701 among lower earners.
Although somewhat smaller, it does not appear that the elasticity estimates are significantly different at conventional levels compared to my main findings.

Finally, my main estimates use a predicted wage measure calculated as predicted earnings divided by 2,080, but this process does not differentiate part- and full-time workers who both have positive earnings, leading to potentially inaccurate predicted wage rates, especially for lower earners. I recalculated predicted gross wages, predicted net wages, and the change in predicted net wages by instead dividing predicted earnings by mean hours observed in the sample conditional on earning status and being in a same- or different-sex couple. Online Appendix Table A.2, Columns 13-14, display estimates using these alternative wage measures, which remain qualitatively similar to my original specification.

6 Deadweight Loss and Tax Revenue Implications

In this section, I apply my Hicksian elasticities and the relative tax changes created by United States v. Windsor to an existing sufficient statistic framework explored in more detail by Feldstein (1999), Immervoll et al. (2007), Chetty (2009), and Kleven (2021). My calibrations below provide further insight into the welfare and tax revenue consequences of joint taxation relative to individual taxation implied by my Hicksian elasticity estimates.

6.1 Deriving the Sufficient Statistic Formula

The total change in deadweight loss due to a small tax change is \( dDWL = -(dW + dTR) \), where \( dW \) is the change in consumer welfare due to the tax change and \( dTR \) is the change in tax revenue. A higher tax rate increases tax revenue through a mechanical revenue effect, \( dM \), but decreases tax revenue through behavioral responses, \( dB \), so that the total change in tax revenue is \( dTR = dM - dB \). Appealing to the envelope theorem, the change in consumer welfare due to
taxation, \(dW\), is equal to the mechanical revenue effect, \(-dM\). Substituting these expressions into \(dDWL\), we obtain \(dDWL = -(dM + dM - dB) = dB\), which describes how deadweight loss changes following a change in tax rates. In order to calculate \(dB\), I consider Chetty’s (2009) simple model to derive the sufficient statistic formula and adjust the calculation for a large reform following Kleven (2021).

Assume there are two types of individuals in a married couple, primary earners \((i = 1)\) and secondary earners \((i = 2)\). Each partner chooses hours of work, \(h^i\), to maximize her utility, \(u^i = C^i - \psi(h^i)\), which is a function of consumption, \(C^i\), and a disutility of work, \(\psi(h^i)\), subject to a constant marginal tax rate, \(\tau_i\), an hourly wage, \(w_i\), and non-labor income, \(y_i\). The price of consumption is normalized to 1. Finally, assume that the government redistributes the tax revenue through a lump sum transfer, \(T_i = \tau_i w_i h^i\), to each individual. The individual budget constraint is, therefore, \((1 - \tau_i)w_i h^i + y_i + T_i \geq C^i\).

I consider a utilitarian welfare function for a continuum of couples indexed by \(c\), in which consumer welfare is equal to the sum of the partners’ utilities subject to their budget constraints. Total welfare, \(W\), is the sum of consumer welfare and the tax revenue raised by the government integrated over all couples:

\[
W(\tau) = \int \sum_{i=1}^{2} \left\{ \left[ (1 - \tau_i)w_i h^i + y_i + T_i - \psi(h^i) \right] + \frac{\tau_i w_i h^i}{\text{Tax Revenue}} \right\} dc
\]

The term in braces above is consumer welfare. Following Kleven’s (2021) notation, consider tax policy and, therefore, marginal tax rates as a function of a treatment parameter, \(\theta\). Appealing to the envelope theorem, the change in the welfare cost of taxation is equal to:
\( (8) \quad dW = -dM = \int_c^d \frac{\sum_{i=1}^2 (1 - \tau_i) w_i h^i + y_i + T_i - \psi(h^i)}{d\theta} dc = \int_c^d \sum_{i=1}^2 -w_i h^i \frac{d\tau_i}{d\theta} dc \)

The change in tax revenue is due to the mechanical change in tax revenue, \( dM \), and the behavioral change, \(-dB\):

\( (9) \quad dTR = \int_c^d \frac{\sum_{i=1}^2 \tau_i w_i h^i}{d\theta} dc = \int_c^d \left[ \sum_{i=1}^2 w_i h^i \frac{d\tau_i}{dM} + \tau_i w_i \frac{dh^i}{d\theta} \left( \frac{-w_i}{d\theta} \right) \right] dc \)

From equations 8 and 9, we obtain:

\( (10) \quad dDWL = -dW(\tau) = -(dW + dTR) = \int_c^d \sum_{i=1}^2 \left[ \frac{\tau_i}{1 - \tau_i} - w_i h^i \epsilon^i_h \frac{d\tau_i}{d\theta} \right] dc \)

Where \( \epsilon^i_h = \frac{\frac{dh^i}{dw_i(1 - \tau_i)} w_i(h(1 - \tau_i))}{h^i} \) is the Hicksian hours elasticity for individual \( i \). Equation 10 shows that the change in deadweight loss along the intensive margin due to a small tax change can be expressed as a function of the marginal tax rate, earnings, and the compensated Hicksian hours elasticity. Employing Kleven’s (2021) notation for large tax changes, where \( \tau_i \) is rewritten as \( \tau_i + \theta \Delta \tau_i \), the trapezoidal approximation to the integral, and iso-elastic utility, it is possible to rewrite Equations 9 and 10 for a large reform, such as the switch to joint taxation I study in this paper:

\( (11) \quad dTR = \sum_{i=1}^2 \left\{ w_i h^i + \frac{1}{2} \Delta \left[ w_i h^i \right] - \epsilon^i_h \left[ \frac{\tau_i}{1 - \tau_i} w_i h^i + \frac{1}{2} \Delta \left[ w_i h^i \frac{\tau_i}{1 - \tau_i} \right] \right] \right\} \Delta \tau_i \)

\( (12) \quad dDWL = \sum_{i=1}^2 \epsilon^i_h \left[ \frac{\tau_i}{1 - \tau_i} w_i h^i + \frac{1}{2} \Delta \left[ w_i h^i \frac{\tau_i}{1 - \tau_i} \right] \right] \Delta \tau_i \)

Where \( \Delta \left[ w_i h^i \right] = w_i^1 h^1_i - w_i^0 h^0_i \) is the change in \( w_i h^i \) due to the reform, and
\( \Delta \left[ w_i h^i \frac{\tau_i}{1 - \tau_i} \right] = w_i^1 h^1_i \frac{\tau_i + \Delta \tau_i}{1 - \tau_i} - w_i^0 h^0_i \frac{\tau_i}{1 - \tau_i} \) is the change in \( w_i h^i \frac{\tau_i}{1 - \tau_i} \) due to the reform.
A symmetric equation holds for extensive margin responses with the extensive compensated elasticity, \( \eta_H \), substituted for \( \varepsilon_H \), and the average tax rate, \( \alpha \), substituted for \( \tau_i \), which is similar to the formulation derived by Immervoll et al. (2007).

### 6.2 Empirical Implementation and Findings

I calibrate Equations 11 and 12 using population-weighted observations of married couples observed in 2014 (corresponding to tax year 2013), my Hicksian elasticity estimates, and the sample means of observed \( w_i h_i \), \( \tau \), and \( \alpha \). I also use the sample mean of the simulated \( d\tau \) and \( d\alpha \) obtained from the NBER TAXSIM simulator and reflecting the tax changes created by *United States v. Windsor*. Using the simulated \( d\tau \) and \( d\alpha \) in this way means that my calibration of Equations 11 and 12 reflect the tax revenue and deadweight loss consequences of the switch from individual to joint taxation.

I consider heterogeneous changes in deadweight loss and tax revenue according to the household earnings split between partners. Figure 4A displays my deadweight loss calculations, where a positive value indicates more deadweight loss under joint taxation relative to individual taxation and a negative value indicates less deadweight loss, and Figure 4B plots my tax revenue calculations, where a positive value represents an increase in tax revenue and a negative value indicates a decrease.

Figure 4A shows that, in general, joint taxation creates a similar amount of deadweight loss, compared to individual taxation, for relatively equal earning couples, and more deadweight loss than individual taxation for more unequal earner couples, which increases monotonically until single-earner couples. I estimate that *United States v. Windsor* increased deadweight loss by $56.9 million among same-sex married couples. Figure 4B shows that joint taxation increases tax revenue relative to individual taxation for relatively equal-earning couples due to the
marriage penalty, but decreases tax revenue for other couples. I estimate that *United States v. Windsor* cost the federal government $54.1 million among same-sex married couples.

It is also possible to extrapolate my deadweight loss and tax revenue calculations to different-sex married couples to obtain a rough understanding the full efficiency and tax revenue costs of joint taxation in the United States. To do this, I must assume that my estimates are applicable to different-sex married couples, but it is, of course, not possible to establish external validity conclusively. My regression specification is meant to distinguish the effects of being in a same-sex married couple rather than a different-sex married couple from the tax-related incentives, yet different-sex married couples who have always faced the prospect of joint federal taxation post-marriage may not react in the same way to changes in tax incentives as my estimates suggest. One indication of such a difference is the larger probability of having children among different-sex couples, which I show in Table 5 tends to reduce labor supply responsiveness to taxes. With these caveats in mind, my calculations suggest that joint taxation increases deadweight loss by $16.4 billion and generates $26.9 billion less in tax revenue among married couples overall. These results corroborate Kleven, Kreiner, and Saez’s (2009) conclusion that optimal tax rates can exhibit negative, rather than positive, jointness.

Finally, the tax revenue effects above are due only to behavioral responses to changes in tax rates, and do not include any measure of the marriage subsidy associated with the *United States v. Windsor* ruling. My calculations suggest that federal tax revenue would have declined by $35.9 million due to the marriage subsidy alone following the Supreme Court ruling, which is smaller than other authors’ simulations of the federal tax revenue consequences of same-sex marriage legalization, such as those by Stevenson (2012) and Alm, Leguizamon, and Leguizamon (2014), who estimate decreases of $38 million and $95-237 million in federal tax
revenue, respectively, using measures similar to this instantaneous measure. Note, however, these estimates assume no labor supply responses to taxation and only reflects the immediate tax revenue consequences among already married couples; it does not include possible changes in tax revenue from other couples who may have married following *United States v. Windsor*.

7 Conclusion

The June 2013 Supreme Court decision in *United States v. Windsor* has been heralded as a landmark civil rights case. While many commentators have focused on the legal and social effects of defining who can marry, the Supreme Court ruling also had immediate consequences for the tax environment faced by same-sex couples who were already married by the time of the ruling. I leverage improved data quality of same-sex married couples and new tax variation among this understudied population to examine classic economic questions concerning the labor supply effects, efficiency costs, and tax revenue consequences of joint taxation.

I use the 2012-2015 waves of the American Community Survey, which are the first of the Census Bureau surveys to explicitly identify same-sex married couples, in order to compare a treatment group of individuals in same-sex married couples to a control group of individuals in different-sex married couples. I also predict individual earnings using a machine learning LASSO approach to construct tax change measures that address the endogeneity of earnings. I quantify variation in federal tax rates due to the sudden movement from the single tax schedule to the joint filing tax schedule, and variation in household after-tax income due to the marriage subsidy. These sources of variation allow me to separate the income and substitution effects of taxation.

I estimate a significant compensated (Hicksian) participation elasticities of 0.791 and 0.527 for lower and higher earners, respectively, and significant income participation elasticities
of -0.363 and -0.051 for lower and higher earners, respectively. Along the intensive margin, I estimate insignificant and oftentimes opposite-signed Hicksian hours elasticities among all samples. My estimates are comparable to, for example, estimates by Selin (2014), who studies the switch from joint to individual taxation in Sweden in 1971, but smaller than those suggested by LaLumia (2008), who studies the original switch to joint taxation in the United States in 1948. My estimates are also larger than, but remain in-line with, others from the literature studying different-sex couples and using different sources of variation, suggesting that my estimates are applicable not just for same-sex married couples, but for different-sex married couples as well. My results are robust to a number of potential confounding factors and alternative specifications. My estimates suggest that the introduction of joint taxation had small, but significant, effects on labor force participation, decreasing participation of lower earners by 0.8 percentage points (1.1%) and increasing participation of higher earners by 0.02 percentage points (0.02%).

Finally, I apply my Hicksian elasticity estimates to an existing sufficient statistic framework to calculate the additional deadweight loss and tax revenue created by joint taxation relative to individual taxation, which is equal across marital status rather than family income. My estimates suggest that United States v. Windsor induced $56.9 million in additional deadweight loss and cost $54.1 million in tax revenue from same-sex married couples relative to individual taxation. Extending my elasticity estimates to all married couples, I find that joint taxation increases deadweight loss by $16.4 billion and reduces tax revenue by $26.9 billion relative to individual taxation.

My findings suggest that there may be efficiency gains to lowering tax rates, especially among secondary earners, so as to mitigate the efficiency costs along the extensive margin. Improved data quality combined with important legal victories for same-sex couples in the
United States offer the possibility of further understanding not only the particular economic changes and challenges faced by same-sex couples, but also an opportunity to learn more about married couples’ responses to legislation more generally.
References


<table>
<thead>
<tr>
<th>Higher earners</th>
<th>Lower earners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Same-sex couples</td>
</tr>
<tr>
<td>Positive earnings</td>
<td>0.964</td>
</tr>
<tr>
<td></td>
<td>(0.187)</td>
</tr>
<tr>
<td>Conditional annual hours worked</td>
<td>2,204.013</td>
</tr>
<tr>
<td></td>
<td>(599.597)</td>
</tr>
<tr>
<td>Observed annual earnings</td>
<td>94,744.451</td>
</tr>
<tr>
<td></td>
<td>(100,779.845)</td>
</tr>
<tr>
<td>Predicted annual earnings</td>
<td>68,293.403</td>
</tr>
<tr>
<td>Observed gross wage</td>
<td>45.887</td>
</tr>
<tr>
<td></td>
<td>(105.573)</td>
</tr>
<tr>
<td>Predicted gross wage</td>
<td>32.833</td>
</tr>
<tr>
<td></td>
<td>(15.504)</td>
</tr>
<tr>
<td>Observed net wage</td>
<td>33.960</td>
</tr>
<tr>
<td></td>
<td>(73.630)</td>
</tr>
<tr>
<td></td>
<td>(10.603)</td>
</tr>
<tr>
<td>Observed average net wage</td>
<td>37.990</td>
</tr>
<tr>
<td></td>
<td>(81.863)</td>
</tr>
<tr>
<td>Predicted Δ net wage a</td>
<td>-0.067</td>
</tr>
<tr>
<td></td>
<td>(1.607)</td>
</tr>
<tr>
<td>Predicted Δ average net wage a</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.914)</td>
</tr>
<tr>
<td>Observed federal marriage subsidy a</td>
<td>371.614</td>
</tr>
<tr>
<td></td>
<td>(6,534.080)</td>
</tr>
<tr>
<td>Predicted federal marriage subsidy a</td>
<td>-448.507</td>
</tr>
<tr>
<td></td>
<td>(1,860.132)</td>
</tr>
<tr>
<td>Observations</td>
<td>5,248</td>
</tr>
<tr>
<td>Worker observations</td>
<td>5,057</td>
</tr>
</tbody>
</table>

Notes: The data come from the 2012-2015 waves of the American Community Survey. The samples include individuals between 30 and 60 years old in same- and different-sex married couples who married in 2012 or earlier. Annual hours worked is the product of “usual hours worked per week” and “weeks worked last year.” “Worker observations” is the number of observations of individuals with positive annual hours of work.

a: These summary statistics are for post-period observations in 2014-2015.
Table 2  
Generalized Difference-in-Differences Effects of the  
*United States v. Windsor* Ruling on Labor Force Participation

<table>
<thead>
<tr>
<th></th>
<th>Separate State and Year FEs</th>
<th>State-by-Year FEs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Higher earners</td>
</tr>
<tr>
<td><strong>Outcome: Individual’s labor force participation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted Δ net wage</td>
<td>0.056***</td>
<td>0.013***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Change in HH income due to marriage subsidy ($1,000s)</td>
<td>-0.047***</td>
<td>-0.011***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>State FEs</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Year FEs</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>State-by-Year FEs</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Hicksian elasticity</td>
<td>1.821***</td>
<td>0.527***</td>
</tr>
<tr>
<td>(substitution effect)</td>
<td>(0.218)</td>
<td>(0.169)</td>
</tr>
<tr>
<td>Income elasticity</td>
<td>-0.241***</td>
<td>-0.051***</td>
</tr>
<tr>
<td>(income effect)</td>
<td>(0.018)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Mean of dep. var.</td>
<td>0.837</td>
<td>0.974</td>
</tr>
</tbody>
</table>

*Notes:*, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Bootstrapped standard errors are in parentheses (1,000 repetitions). The data come from the 2012-2015 waves of the American Community Survey. The sample includes individuals between 30 and 60 years old in a same- or different-sex married couple who married in 2012 or earlier. Labor force participation is equal to 1 if the individual has positive annual hours of work, where annual hours of work is the product of “usual hours worked per week” and “weeks worked last year.”
# Table 3
Generalized Difference-in-Differences Effects of the United States v. Windsor Ruling on Annual Hours of Work

<table>
<thead>
<tr>
<th>Outcome: Individual’s annual hours of work</th>
<th>Separate State and Year FEs</th>
<th>State-by-Year FEs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Higher earners</td>
</tr>
<tr>
<td>Predicted Δ net wage</td>
<td>11.318</td>
<td>-9.023</td>
</tr>
<tr>
<td></td>
<td>(7.304)</td>
<td>(7.203)</td>
</tr>
<tr>
<td>Change in HH income due to marriage subsidy ($1,000s)</td>
<td>6.848</td>
<td>3.764</td>
</tr>
<tr>
<td>State FEs</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Year FEs</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>State-by-Year FEs</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Marshallian elasticity (income + substitution effects)</td>
<td>0.160</td>
<td>-0.144</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.115)</td>
</tr>
<tr>
<td>Hicksian elasticity (substitution effect)</td>
<td>-0.036</td>
<td>-0.277</td>
</tr>
<tr>
<td></td>
<td>(0.249)</td>
<td>(0.298)</td>
</tr>
<tr>
<td>Income elasticity (income effect)</td>
<td>0.011</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Mean of dep. var.</td>
<td>1,999,065</td>
<td>2,211,771</td>
</tr>
<tr>
<td>Observations</td>
<td>2,238,615</td>
<td>1,303,691</td>
</tr>
</tbody>
</table>

Notes: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Bootstrapped standard errors are in parentheses (1,000 repetitions). The data come from the 2012-2015 waves of the American Community Survey. The sample includes individuals between 30 and 60 years old with positive annual hours of work in a same- or different-sex married couple who married in 2012 or earlier, where annual hours of work is the product of “usual hours worked per week” and “weeks worked last year.”
Table 4
Heterogeneous Effects of the *United States v. Windsor* Ruling by Sex

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Higher earners</td>
<td>Lower earners</td>
</tr>
<tr>
<td><strong>Outcome: Individual’s labor force participation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hicksian elasticity</td>
<td>0.368</td>
<td>0.409</td>
</tr>
<tr>
<td>(substitution effect)</td>
<td>(0.243)</td>
<td>(0.381)</td>
</tr>
<tr>
<td>Income elasticity</td>
<td>-0.042***</td>
<td>-0.300***</td>
</tr>
<tr>
<td>(income effect)</td>
<td>(0.016)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Mean of dep. var.</td>
<td>0.979</td>
<td>0.764</td>
</tr>
<tr>
<td>Observations</td>
<td>971,973</td>
<td>365,899</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Higher earners</td>
<td>Lower earners</td>
</tr>
<tr>
<td><strong>Outcome: Individual’s annual hours of work</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marshallian elasticity</td>
<td>-0.077</td>
<td>0.036</td>
</tr>
<tr>
<td>(income + substitution effects)</td>
<td>(0.179)</td>
<td>(0.235)</td>
</tr>
<tr>
<td>Hicksian elasticity</td>
<td>-0.220</td>
<td>0.217</td>
</tr>
<tr>
<td>(substitution effect)</td>
<td>(0.459)</td>
<td>(0.526)</td>
</tr>
<tr>
<td>Income elasticity</td>
<td>0.007</td>
<td>-0.015</td>
</tr>
<tr>
<td>(income effect)</td>
<td>(0.018)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>Mean of dep. var.</td>
<td>2,273.124</td>
<td>1,915.631</td>
</tr>
<tr>
<td>Observations</td>
<td>951,453</td>
<td>279,385</td>
</tr>
</tbody>
</table>

Notes: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Bootstrapped standard errors are in parentheses (1,000 repetitions). The data come from the 2012-2015 waves of the American Community Survey. The sample includes individuals between 30 and 60 years old in a same- or different-sex married couple who married in 2012 or earlier. The hours of work sample is further restricted to individuals with positive annual hours of work. Labor force participation is equal to 1 if the individual has positive annual hours of work, and annual hours of work is the product of “usual hours worked per week” and “weeks worked last year.” All specifications include year and state fixed effects.
### Table 5
Heterogeneous Effects of the *United States v. Windsor* Ruling by the Presence of Children

<table>
<thead>
<tr>
<th></th>
<th>Childless couples</th>
<th>Couples with children</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td></td>
<td>Higher earners</td>
<td>Lower earners</td>
</tr>
<tr>
<td><strong>Outcome: Individual’s labor force participation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hicksian elasticity</td>
<td>0.301</td>
<td>0.707</td>
</tr>
<tr>
<td>(substitution effect)</td>
<td>(0.278)</td>
<td>(0.435)</td>
</tr>
<tr>
<td>Income elasticity</td>
<td>-0.057***</td>
<td>-0.389***</td>
</tr>
<tr>
<td>(income effect)</td>
<td>(0.022)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Mean of dep. var.</td>
<td>0.975</td>
<td>0.745</td>
</tr>
<tr>
<td>Observations</td>
<td>769,347</td>
<td>303,700</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Childless couples</th>
<th>Couples with children</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td></td>
<td>Higher earners</td>
<td>Lower earners</td>
</tr>
<tr>
<td><strong>Outcome: Individual’s annual hours of work</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marshallian elasticity</td>
<td>-0.209</td>
<td>-0.065</td>
</tr>
<tr>
<td>(income + substitution effects)</td>
<td>(0.202)</td>
<td>(0.257)</td>
</tr>
<tr>
<td>Hicksian elasticity</td>
<td>-0.547</td>
<td>0.179</td>
</tr>
<tr>
<td>(substitution effect)</td>
<td>(0.587)</td>
<td>(0.586)</td>
</tr>
<tr>
<td>Income elasticity</td>
<td>0.018</td>
<td>-0.018</td>
</tr>
<tr>
<td>(income effect)</td>
<td>(0.024)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Mean of dep. var.</td>
<td>2,268.479</td>
<td>1,905.485</td>
</tr>
<tr>
<td>Observations</td>
<td>750,489</td>
<td>226,401</td>
</tr>
</tbody>
</table>

*Notes:*, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. Bootstrapped standard errors are in parentheses (1,000 repetitions). The data come from the 2012-2015 waves of the American Community Survey. The sample includes individuals between 30 and 60 years old in a same- or different-sex married couple who married in 2012 or earlier. The hours of work sample is further restricted to individuals with positive annual hours of work. Labor force participation is equal to 1 if the individual has positive annual hours of work, and annual hours of work is the product of “usual hours worked per week” and “weeks worked last year.” All specifications include year and state fixed effects.
Notes: The data come from the 2012-2015 waves of the American Community Survey. Each data point presents the mean value of the demographic variable in that year. The samples include individuals between 30 and 60 years old who are in same- or different-sex married couples who married in 2012 or earlier.
Figure 2

Variation in Net-of-Tax Wage Rates Among Higher and Lower Earners

Notes: The data come from the 2014-2015 waves of the American Community Survey (the post-Windsor years) and include individuals between 30 and 60 years old in same-sex married couples who married in 2012 or earlier. Each data point is an individual and the black line displays the mean value within a $5 net wage rate bin. Individuals with net wage rates greater than $100 or changes in net wage rates larger than ±$15 have been excluded from the figure for illustration purposes only.
Figure 3
Marriage Penalty Variation Among Same-Sex Couples

Notes: The data come from the 2014-2015 waves of the American Community Survey (the post-Windsor years) and include same-sex couples with both partners between 30 and 60 who married in 2012 or earlier. Each data point is a same-sex couple, and the black line displays the mean value within a $10,000 earnings bin. Couples with outlier values of the marriage subsidy or with household income greater than $350,000 have been excluded from the figure for illustration purposes only.
Figure 4
Additional Deadweight Loss and Tax Revenue Under Joint Taxation Relative to Individual Taxation

Notes: Each data point represents the average net additional deadweight loss or net additional tax revenue created by joint taxation relative to individual taxation per family for households in the same earnings split bins, taking into account both intensive and extensive margin responses. Households with earnings splits between 50-50 and 59-41 are included in the 50-50 split bin, households with earnings splits between 60-40 and 69-31 are included in the 60-40 split bin, and so on.
Endnotes

1 My analysis ends in 2015 and is therefore not affected by the Obergefell v. Hodges ruling, which required all states to permit same-sex marriages. The United States v. Windsor ruling only required the federal government to recognize at the federal level same-sex marriages that were permitted by states.

ii The net-of-tax rate is one minus the tax rate.

iii In Section 5.2, I examine several potential influences that might confound my estimates and find that my original results are robust to a number of alternative specifications and samples.

iv Online Appendix figure A.2 presents a timeline of same-sex marriage legalization in the U.S. leading up to federal recognition in 2013.

v The federal tax code recognizes marriages according to where they occurred (the place-of-celebration rule) not according to the couple’s current state of residence.

vi Martinez, Saez, and Siegenthaler (Forthcoming) and Sigurdsson (2019) also use large-scale tax variation created by income tax holidays in Switzerland and Iceland, respectively, but estimate intertemporal labor supply (Frisch) elasticities. Estimating Frisch elasticities depends upon temporary wage variation, whereas the variation I use in this paper is more akin to a permanent switch in tax regimes.

vii Bick and Fuchs-Schündeln (2017) and Borrella, Nardi, and Yang (2019) conclude that jointness in the U.S. tax system and Social Security, respectively, suppresses married women’s labor supply. Both of these studies corroborate LaLumia’s (2008), Selin’s (2014), and Kališková’s (2014) conclusions, as well as my own results.

viii In the 1990 census, if a couple appeared to be a same-sex married couple then one partner’s sex was changed so that the couple appeared to be a different-sex married couple. In the 2000 and 2010 censuses and in pre-2012 waves of the American Community Survey, if a couple appeared to be a same-sex married couple then their marital status was changed to “unmarried partner,” sometimes without an accompanying data quality flag (U.S. Census Bureau 2009).

ix Other research on LGBT individuals has focused on workplace discrimination (Badgett 1995; Carpenter 2007; Plug, Webbink, and Martin 2014), health outcomes (Buchmueller and Carpenter 2010; Gonzalez and Blewett 2014; Carpenter et al. Forthcoming), or differences in labor market behavior between same- and different-sex couples (Tebaldi and Elmslie 2006).

x I use marginal tax rate measures when examining annual hours of work and average tax change measures when examining labor force participation.

xi For same-sex married couples before United States v. Windsor, both measures of the couple’s predicted total tax liability are the sums of each partner’s individual tax liability because they are taxed individually. For different-sex married couples, both measures of the couple’s total tax liability are joint tax liabilities because they were taxed jointly before and after the ruling. For same-sex married couples after United States v. Windsor, $T_{ij,u}$ is the couple’s joint tax liability and $T_{ij,pre-Windsor}$ is the sum of each partner’s individual tax liabilities because they were taxed individually before the ruling and taxed jointly afterward.

xii Note that the value of the variable $\Delta T_{ij,t}$ will be the same for both partners in a same-sex married couple.

xiii I have also estimated models that use the observed changes in net-of-tax rates based on reported earnings instead of predicted earnings. However, these estimates are likely to be biased due to unobservable factors that influence specialization within the household. I obtain larger intensive margin Hicksian elasticities and smaller extensive margin Hicksian elasticities, both of which suggest that unobservable factors that increase specialization in the household are biasing the estimates. These results are available upon request.

xiv The age restriction is to focus on prime working years while ignoring individuals in age groups that may still be in college or graduate school or eligible for Social Security.

xv For same-sex couples observed in 2012, I rely on a data quality flag indicating that the couple was a same-sex married couple whose marital status was adjusted to “unnamed partnership.” For same-sex married couples observed in 2013-2015 and for different-sex married couples, I rely on the year of marriage to condition on being married before 2013. In Section 5.2, I find that my results are robust to controlling for years married and to excluding couples where both partners’ years of marriage are allocated. A Komorov-Smirnov test fails to reject the null hypothesis that the distribution of marriage years in each year of the ACS are the same among same-sex married couples observed in 2013-2015 (Komorov 1933; Smirnov 1933).

xvi If neither spouse has positive earnings or if both spouses’ earnings are exactly equal, then I assign the self-reported household head as the higher earner.

xvii Online Appendix Figures A.3 and A.4 display trends over time in the demographic variables.
Online Appendix C presents estimates from an event study specification to test for significant differences in the outcome variables over time between individuals in same- and different-sex couples. These coefficients are not statistically different from zero in the pre-period, lending some support in favor of the parallel trends assumption in this context.

Online Appendix D contains more detail about potential migration responses and presents the spatial distribution and migration statistics of same-sex couples in my sample.

The prediction procedure allows me to assign policy variables to non-workers, but it raises the Heckman concern that non-workers may have systematically lower wages than workers. This concern is not a substantial problem in my context because identification of the labor supply responses I estimate come from variation in tax rates, rather than wage rates.

The LASSO uses an L1 norm constraint rather than the L2 norm constraint used by the ridge regression. Online Appendix B contains more detailed technical information about the LASSO prediction process.

Estimating the LASSO separately allows the coefficient of every covariate and interaction to differ between individuals in different- and same-sex couples, which is equivalent to interacting an indicator for being in a same-sex relationship with every covariate and interaction that enters the LASSO.

The $R^2$ of the positive earnings regression is 0.460. The $R^2$ of the earnings regression conditional on positive earnings is 0.268.

In Section 5.2, I find that my results are robust to an alternative wage definition in which I divide predicted earnings by mean observed hours conditional on earning status and being in a same- or different-sex couple.

Online Appendix Figure A.5 demonstrates that the LASSO prediction process understates reported earnings, with greater concentrations of household earnings and earnings split around $70,000 split relatively evenly between the two partners.

I assume that married couples in my sample file jointly when possible, and that the partner with higher predicted earnings claims any dependent children for tax purposes in this case and files as “head of household.” Fisher, Gee, and Looney (2018) estimate that 96% of married couples in 2014 filed jointly. In reality, the biological parent would claim the children, but I cannot observe biological relationship status in the ACS.

If predicted earnings are equal to $0, then I define the predicted average tax rate as 0%.

The marriage subsidy may, instead, be a penalty for some couples, and is often referred to as the “marriage penalty.” My empirical specification, however, uses the subsidy measure, and I will refer to it throughout the paper as the marriage subsidy.

Same-sex married couples would not have faced state tax changes due to the ruling in United States v. Windsor because the ruling affected only federal marriage recognition. The 2015 Supreme Court ruling in Obergefell v. Hodges, however, would have altered the state legislative landscapes faced by same-sex married couples, and it is possible that same-sex couples faced marriage incentives from their state tax code. I do not consider state incentives in my empirical strategy, but they would, in general, be substantially smaller than those created by the federal tax code. Indeed, Light and Omori (2008) and Friedberg and Isaac (Forthcoming) find that marriage penalties created by state taxes do not have significant effects on family structure decisions.

40% of same-sex married couples and 62% of different-sex married couples have children, which will enter into my calculations of the tax rate and tax liability changes.

Another visible trend in Figure 3 is that low-earning equal earners with children face larger marriage penalties than low-earning equal earners without children. This has to do with tax credits targeted at low-earning households that are more generous for families with children, such as the Child Tax Credit and the Earned Income Tax Credit. Bastian (2017) and Michelmore (2018) more thoroughly explore marriage incentives in the Earned Income Tax Credit.

Online Appendix Tables A.4-A.11 display the full set of coefficient estimates for the tables below.

These results are available upon request.

Using, instead, the mean values of the observed tax change variables (not reported in Table 1), my elasticities imply that joint taxation decreased labor force participation of lower earners by 2.5 percentage points (3.5%) and increased labor force participation of higher earners by 0.7 percentage points (0.8%). These estimates are more comparable in magnitude to those by LaLumia (2008) and Kališková (2014), but are likely biased upward because I cannot observe net wage rate changes for individuals who do not work.

Baum-Snow and Neal (2009) document reporting errors in self-reported usual hours worked per week, making measurement of annual hours relatively noisier. Note that my predicted net-of-tax wage rates do not rely on observed annual hours of work because I divide predicted earnings by 2,080 (52 weeks multiplied by 40 hours per week) to avoid this endogeneity issue.

51
The combined effect of earning status and sex is $0.913 - 0.368 = 0.545$. The effect of being female versus male is $0.623 - 0.368 = 0.255$ when comparing higher earners and $0.913 - 0.409 = 0.504$ when comparing lower earners. Therefore, the fraction of the total effect that is due to sex is between $\frac{0.255}{0.545} = 0.47$ and $\frac{0.504}{0.545} = 0.92$. The combined effect of earning status and sex is $0.913 - 0.368 = 0.545$. The effect of being a lower versus higher earner is $0.409 - 0.368 = 0.041$ when comparing men and $0.913 - 0.623 = 0.290$ when comparing women. Therefore, the fraction of the total effect that is due to earning status is between $\frac{0.041}{0.545} = 0.08$ and $\frac{0.290}{0.545} = 0.53$.

There are 3,293 out of 5,248 same-sex couples (62.7%) in the CATI/CAPI sub-sample. The reduction in statistical significance evident in Online Appendix Table A.2 may therefore be due to the reduction in sample size of the treatment group.

These states are Alabama, Georgia, Kansas, Kentucky, Louisiana, Michigan, Missouri, Nebraska, North Dakota, and Ohio.

2012 ACS data editing procedures specify that the “year of marriage” variable is not defined for same-sex married couples in this wave of the American Community Survey. The absence of the year of marriage does not allow me to determine how long the couple has been married, and so I exclude 2012 observations of all couples in this specification.

Online Appendix Table A.3 presents summary statistics for these new wage measures alongside the original statistics. Note that I use the mean hours displayed in row 4 of Online Appendix Table A.3, which is unconditional on labor force participation because I use predicted wage rates for both workers and non-workers in this specification to better reflect mean hours in the sample.

Online Appendix Figure A.6 displays how deadweight loss and tax revenue change following a small tax increase.

This utility function ignores income effects in order to make the derivation tractable, and is also money metric so that social welfare, below, is measured in dollars. This simplification is common in the sufficient statistic literature studying income taxation, and means that the Marshallian and Hicksian elasticity concepts are the same along the intensive margin.

Online Appendix Figure A.7 displays the distribution of earnings splits among same- and different-sex married couples in my sample.

These calculations do not tell us the total efficiency cost of taxation, only the additional cost of joint taxation relative to individual taxation. Feldstein (1999), for example, finds that the personal income tax creates deadweight loss equal to 32.2% of tax revenue.