The Labor Supply and Tax Revenue Consequences of Federal Same-Sex Marriage Legalization

Stevenson, Adam

University of Michigan

8 February 2012

The Labor Supply and Tax Revenue Consequences of Federal Same-Sex Marriage Legalization

Adam Stevenson*

January 4, 2012

Abstract

The issue of same-sex marriage legalization is increasingly part of the national political dialogue. This legalization would have a number of economic impacts, one of the most direct being a change in income tax payments, through the so-called marriage penalty. I estimate the effects of same-sex marriage legalization on federal income tax revenue. These estimates rely critically on the responsiveness of labor supply and marital choice to changes in the tax code. I present new evidence on both topics using changes in taxation generated from the 2003 Jobs and Growth Tax Relief Reconciliation Act. In addition, I propose a novel measure of the marriage penalty that incorporates the fact that agents will respond optimally to changes in marginal tax rates within the household.

JEL classifications: D10, H24, J12, J22
keywords: labor supply, marriage penalty, sexual orientation, DOMA

*Department of Economics, University of Michigan

1 Introduction

The debate over whether same-sex couples should be allowed the legal right to marry has become a hot-button issue in the United States in the last decade. The Defense of Marriage Act (DOMA) prohibits the recognition of same-sex marriage at the federal level, but a growing number of states and major cities have granted the right to either marry or enter into other officially-recognized relationships.[1] DOMA also allows states to disregard same-sex marriages granted in other states. While state- and local-level recognition can have profound social effects, the economic effects of federal recognition of same-sex marriage are potentially quite important. Changes in this arena are occurring at the federal level - President Obama recently ordered the Justice Department to cease defense of DOMA (Savage 2011). In this paper, I explore the effects of same-sex marriage legalization on federal income tax revenue.

Same-sex marriage legalization has the direct impact of allowing a change in a household's legal tax filing status, which can change the household's income tax burden, via the "marriage penalty" (which may be positive, a tax, or negative, a subsidy). This change in the tax schedule causes a change in labor supply, which also affects income tax revenue. Further, the desire to capture a marriage subsidy or avoid a marriage tax may cause marriage rates to vary along with the marriage penalty. To fully understand the effect of same-sex marriage legalization on income tax collection, we must understand the distribution of marriage penalties across same-sex households, the causal effect of these penalties on marriage likelihood, and how this change in taxation influences labor supply choices.

To capture both of these effects, I introduce a new measure of the marriage penalty, which I call the endogenous marriage penalty.[2] Changes in the tax schedule upon marriage will change incentives to work for both partners within a household. This in turn changes hours worked and further changes the tax rate. I estimate the marriage penalty

[1] These relationships are often called "civil unions", but different constituencies offer legal statuses that provide differing degrees of comparability to the state of marriage. Civil unions are similar to marriage, offering the same legal rights. Other legal statuses offering fewer and weaker legal rights include domestic partnerships or reciprocal/designated beneficiaries (Hawkins 2009, Badgett, Gates and Maisel 2008).

[2] Sjoquist and Walker (1995) have estimated marriage penalties at the aggregate level assuming that labor market choices differ between married and non-married cohabiting couples (finding little evidence of aggregate changes), but I'm not aware of any paper that estimates the behavioral response from microdata.

as a function of rational responses to anticipated changes in both the worker's own tax rate and their partner's, conditional on whether or not the couple is legally married. This endogenous marriage penalty is in contrast to the standard measure, which does not allow for labor market responses to tax changes and as such may be interpreted as an "instantaneous" measure. I provide a decomposition to illustrate the sources of the discrepancy between the instantaneous measure of the marriage penalty and the endogenous measure.

The question of the tax revenue consequences of same-sex marriage is not an idle one. Both the federal and a number of state governments have explicitly justified their same-sex marriage bans on the basis of the adverse tax revenue and expenditure effects (Hawkins 2009). In terms of measuring the tax revenue consequences of legalizing same-sex marriage, the paper that comes closest to this one is Alm, Badgett, and Whittington (2000), which is not based on micro-level income data but instead uses estimates of mean earnings from other papers and imposes varying assumptions about gay and lesbian household composition to tabulate a range of potential aggregate outcomes. The most systematic investigation of the net budgetary effects of same-sex marriage legalization is from the CBO (2004). It estimates an almost $1 billion annual budget windfall to legalization, $200-$400 million of which come from income tax revenues. The CBO's estimates of same-sex marriage penalties draw heavily on Alm, Badgett, and Whittington (2000).

I show that same-sex marriage legalization would increase federal income tax revenues somewhere between $100 and $175 million. Given the existing estimates of the increase in federal costs that would accompany legalization (CBO 2004), the policy at worst would be revenue-neutral, and, at the upper range of my estimates, may generate $140 million per year.

The Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) of 2003, an acceleration of the 2001 Economic Growth and Tax Relief Reconciliation Act (EGTRRA) (together frequently called the "Bush tax cuts") provides a useful natural experiment.[3] The Bush tax cuts implemented a number of reforms, but two primary changes were a general tax cut and a decrease in the average marriage penalty. The JGTRRA in particular served as an exogenous tax shock. The EGTRRA scheduled a slow phase-in of

[3]Keifer et al. (2002) provides a detailed description of the EGTRRA.

marriage penalty reforms and tax bracket adjustments, but the JGTRRA abruptly implemented all the planned income tax changes starting in tax year 2003. These cuts, the most substantial tax reforms in the U.S. of the last 15 years, have not been exploited as a source of variation in the question of household structure, and have only rarely been used in studies of labor supply choice (Heim 2009, Auten, Carroll, and Gee 2008).

The JGTRRA grants us the ability to make a number of methodological improvements to the previous literature. While a number of papers (described in section 4.2) have considered the effect of the marriage penalty on marriage likelihood, none to my knowledge have used a single, large, unanticipated cut to identify the effect. In combination with generalized propensity score methods (Hirano and Imbens 2005), I use the JGTRRA to construct novel estimates of the sensitivity of marriage to taxation. Similarly, while a small number of papers have studied the labor supply of same-sex couples, few have based their estimates on plausibly-exogenous shocks to net wages.[4]

2 Data and Descriptive Statistics

The main sources of data for this study are the U.S. Census 1% sample from the year 2000, and the American Community Survey from years 2001-2007, as prepared by the IPUMS project (Ruggles et al. 2010). I restrict the data to the time period 2000 to 2007 to include 4 years of data before and after the implementation of the JGTRRA while excluding the influence of other tax reforms (in particular the substantial Economic Stimulus Act of 2008). Following the literature on familial labor supply at prime working age, I drop all households where either member is older than 54, or where the head or partner is likely to be too young to participate in the labor market (Blau and Kahn 2007).[5] I use a 1% random sample of heterosexual households from the 2000 Census and a 2% sample from the ACS surveys so that my samples of heterosexual and homosexual couples are of comparable size in most years.

Since the Census surveys do not ask any questions about sexual behavior or identity, the only way to identify homosexuals in the data is through cohabitation and reported

[4] Antecol and Steinberger (2011) is the main exception, using an instrumental variable methodology very similar to my own, but considering only lesbian women, and only in a single cross-section.

[5] I definite potential experience as age - 6 - years of education, and drop households where one member has negative potential experience.

relationship status. There are well-documented problems with the Census relational data. Black, Gates, Sanders and Taylor (2007) point out that the sample of same-sex couples in the 2000 Census is substantially contaminated by mis-allocated individual gender or relationship status. The Census Bureau treats identification as a spouse or being in the state of marriage as logically inconsistent with being in a same-sex relationship. It re-codes either the gender or relationship status of a large number of individuals, creating a large number of "false positive" homosexual couples within the data. I define a gay or lesbian couple to be a pair of individuals of the same sex who live in the same household, and where the respondent classifies the other individual as an "unmarried parter." Following the standard response to this problem, I omit from the sample all individuals with allocated gender or marital status.

Standard theory says that labor supply choices are made in response to marginal wage rates. I calculate gross wage w by taking annual labor market earnings and dividing by annual hours worked, as calculated from typical weekly hours worked times weeks worked in the year. Following the literature, I drop individuals with "extreme" implied wage rates, below \$2 and above \$200, as measured in 2004 real dollars (Blau and Kahn 2007). A worker's marginal wage is $w * (1 - \mu)$, where μ represents the federal income tax rate, including the employee's share of the payroll (FICA) tax. Since primary earners overwhelmingly participate in the market, I drop all households where the primary earner has zero reported hours worked in the year.

Tax data comes from NBER's TAXSIM program (Feenberg and Couts 1993). These variables include total federal income tax and FICA contributions. Since my tax data is generated by TAXSIM, it is not true taxes paid, but predicted taxes based on observable data within the Census.[6] Potentially important variables not found in the Census include capital gains income and expenditures on child care, health care, mortgage interest and charitable donations.

Table 1 describes the characteristics of household members within the sample. When studying household income and specialization, it's useful to differentiate between primary and secondary earners (Antecol and Steinberger 2011). Most theoretical and empirical

[6]These tax-relevant data include the number of children (by age group), labor and business income, interest and dividend income, social security income, pension income, and payments made for rent or property tax. These data are all arguments of the tax function defined in section 3.

approaches treat the male as the primary earner in heterosexual households. I define the primary earner in homosexual households to be the member whose annual income is greater. In table 1 I separately describe male and female members of heterosexual households and primary and secondary earners in homosexual households.

The sample of gay and lesbian households is slightly younger than the married heterosexual couples, but older than the unmarried heterosexual households. Consistent with Black, Sanders and Taylor (2007), individuals in homosexual households have substantially fewer children and much more education than those in heterosexual households. Interestingly, in terms of education and wage rates, homosexual households are more specialized (along the dimension of market versus non-market human capital) than are heterosexual husbands and wives. Married men and women have identical average levels of education, and among those who work, married women earn about 71% of their spouse's wages, while gay secondary earners have 0.7 fewer years of education and median wages that are 52% of the primary earners'.[7] The household wage gap is smaller in lesbian households than in gay households, but greater than in married households.

3 Marriage and taxes

The first step towards estimating the effects of taxes on household status and labor supply is to measure the marriage penalty, which involves calculating actual and counterfactual tax burdens. These calculations allow me to estimate, first, marriage rates as a function of the tax treatment of marriage, and second, the change in the marginal tax rate due to marriage.

In principle, it is simple to define actual and counterfactual taxes, and therefore the marriage penalty. For any pair of individuals (following the notation of Berliant and Rothstein (2003)), define each individual's income tax liability when single as $T_t^s(y_e)$, where y_e indicates the vector of taxable income streams (and offsetting expenditures and characteristics) for earner $e \in \{1, 2\}$ and where $T_t^s(\cdot)$ is a function representing the federal tax schedule for unmarried persons in year t. If two individuals marry, the tax liability is

[7] Of course, since married women are less likely to work than secondary-earner homosexuals, those married women who are in the labor force are a relatively more select group than are homosexual secondary earners.

$T_t^m(y_1, y_2)$, where $T_t^m(\cdot, \cdot)$ is a function representing the federal tax schedule for married couples in year t. The marriage penalty P is the increase in taxes when a household of two single individuals instantaneously enter into the state of marriage,

$$P_t(y_1, y_2) \equiv T_t^m(y_1, y_2) - (T_t^s(y_1) + T_t^s(y_2)) \quad . \tag{1}$$

Clearly, when $P_t(y_1, y_2) < 0$, there is a negative marriage penalty, or a marriage subsidy. Since the y_e vectors are assumed fixed with respect to marital status, this calculation rules out the possibility of economies of household scale, for example in terms of rent or property taxes. Suppose that the first element of y_e is labor market income, denoted y_e^L. Typically, this value is also fixed in equation 1. If so, the measured penalty is best interpreted as an instantaneous change in tax burden among the partners, before any economic responses to the change in martial status can occur.

Given the data in the Census, calculating $T_t^m(y_1, y_2)$ is simple, since the federal tax schedule pools non-labor income and children within married households. It is less simple to calculate counterfactual taxes for married couples (i.e., as if they were unmarried). Many assets in married households are pooled, making a clean distinction between each member's non-labor income difficult. Even if the asset holdings of each spouse were clearly delineated, the assets are unlikely to be split according to these divisions in divorce proceedings. Taxes if single are calculated, following the literature, as if a "divorce" and a split of assets occurred, under the assumption that households minimize total tax lability (Feenberg and Rosen 1994). This method dictates that unmarried couples (whether actual or counterfactually divorced) will equally split all non-work income between the two members.

Children are a major tax deduction, and we must determine which parent is assigned custody of each child when a couple is unmarried. The standard tax-minimization algorithm assigns all child deductions to the higher earner. This implies that the higher earner will take any non-biological children (i.e., the partner's child from a different relationship) into their household after divorce.[8] This is an unappealing assumption, especially among

[8]Another popular method for penalty calculations involves assigning children to the woman, as typically occurs in divorce settlements. This procedure provides no guidance, however, for homosexual households. Alm and Whittington (1996) explore the consequences of the opposing methods for allocating tax deductions. They show that while the different methods imply different marriage penalties, the

same-sex couples, where there simply cannot be shared biological children. I assign children to parents upon divorce in the following way. Parents always claim a child who is theirs and who is not the biological child of their partner. Shared biological children are assigned to the higher earner. Individuals with no children file as single while individuals with children file as head of household.

Figure 1 shows that there is a clear level shift in the average marriage penalty after 2003. The average size of this shift is essentially neutral with respect to martial status and sexual orientation. Figure 1 aggregates across the child groups described in table 2 and shows that there is no substantial time trend in marriage penalties before or after the JGTRRA. For every coresidency type, the fall in the average marriage penalty is on the order of $1000.[9]

Figure 2 illustrates that all four coresidency types became more likely to draw a marriage penalty of zero after the JGTRRA (shown in grey) than before (shown in black outline). For all groups, the modal outcome was to pay a small marriage penalty prior to the reforms, and to pay nothing afterwards.

Table 2 presents tax liability and marriage penalty statistics, by coresidency, prior to and after the implementation of the JGTRRA. I also separate the statistics by the number of children in the household. Homosexual couples are particularly likely to have no children, and we want to compare them to similar heterosexual couples in this regard. In terms of changes in actual tax lability, childless gay and lesbian couples experienced the largest average tax cuts of all groups, an average income tax cut of 8.5% for childless lesbian women, 6.1% for childless gay men. Similarly, the majority of heterosexual married couples with children paid a marriage penalty both before and after the reform, while the majority of homosexual couples (regardless of the presence of children) do not.

Treating the marriage penalty as a lump-sum payment as in equation 1 is useful in thinking about the economic incentive to marry. But in terms of labor supply choices, we typically want to know how marginal tax rates will change upon marriage. Here and throughout the paper, when I calculate individual marginal tax rates when married, I assume the primary earner is taxed first, starting at the lower marginal rates. Secondary

time trends and year-over-year changes across methods of calculation are virtually identical.

[9] I use the term 'coresidency type" to differentiate between individuals living together in heterosexual married, heterosexual unmarried, gay male and lesbian female households.

earners (by definition of being secondary) take the primary earner's income as given and so are taxed as if their first dollar is the primary earner's last dollar.[10] It always holds, then, that for marginal labor income tax rates, $\frac{\partial T_t^m}{\partial y_1^L} \leq \frac{\partial T_t^m}{\partial y_2^L}$. Among the unmarried, typically $\frac{\partial T_t^s}{\partial y_1^L} \geq \frac{\partial T_t^s}{\partial y_2^L}$ (since the prime earner typically earns more than the secondary earner, though this is not true in all heterosexual unmarried households). Holding constant all other elements of the y^e vectors, the more specialized a household is in terms of labor market earnings, the more the prime earner's tax rate will fall upon marriage and the more the secondary earner's will rise upon marriage (Eissa and Hoynes 2000). The effects of marriage on incentive to work will therefore differ across coresidency and earner type, depending on these factors.

As seen in table 3, each worker type experienced a marginal tax rate cut in 2003, ranging from 1.3 percentage points for secondary-earning gay men to 2.3 percentage points for unmarried heterosexual women. I define the "rate marriage penalty" to be the increase in marginal tax rates experienced by an individual when they enter into marriage, $\frac{\partial T_t^m}{\partial y_e^L} - \frac{\partial T_t^s}{\partial y_e^L}$. These penalties are calculated assuming no labor market response to any tax schedule changes that occur upon marriage.

For every coresidency type, the marginal tax rate for primary earners always falls upon marriage and the rate for secondary earners rises.More interesting are the differences across coresidency types. Since married couples are relatively specialized in terms of labor market effort, married women would pay relatively low tax rates if unmarried and high rates if married. As we'll see again in section 4.1.1, homosexual secondary earners are more similar to unmarried women than to married women in terms of labor market outcomes. Homosexual rate marriage penalties are significantly lower than married women's, but are statistically indistinguishable from heterosexual unmarried couples. The Bush tax reforms lowered the rate penalties for every coresidency type except unmarried couples, with married women and secondary-earning gay men experiencing the biggest decreases in rate marriage penalties.

[10]The labor supply estimation results for married men and women are insensitive to the use of this rate, rather than applying the pooled top marginal rate to both members of the household.

4 Empirical Results

4.1 Labor supply

There is a very large literature studying the effects of taxation on household labor supply choices. The methodology often takes the form of instrumental variables (IV) estimates (Mroz (1987) surveys the early literature in this vein).[11] Recent contributions to the literature using instruments similar to my own include Devereaux (2004), Blau and Kahn (2007), and Antecol and Steinberger (2011), with this last paper focusing in particular on lesbian households. Orrefice (2011), Leppel (2007) and Tebaldi and Elmslie (2006) all study differences in labor supply choices by sexual orientation, but do not use exogenous net wage variation as a source of identification.

In this section, I estimate the labor supply function (in the form of an hours equation) by coresidency type and primary/secondary worker status in section 4.1.1.[12] Using those estimates, in section 4.1.2 I introduce an alternative measure of the marriage penalty, the "endogenous" penalty, which allows individuals to make optimal labor market responses in response to the shift in tax schedule that occurs upon marriage or divorce.

4.1.1 IV estimates of own- and partner-wage effects

Suppose that the worker's labor supply function is given by the equation

$$h_{icet} = \alpha^1_{ce} \ln w^o_{icet} + \alpha^2_{ce} \ln w^p_{icet} + \alpha^3_{ce} I_{ict} + \beta_{ce} X_{icet} + \varepsilon_{icet} \quad , \tag{2}$$

where i indexes households across coresidency types c, earner types (primary or secondary) e , and time periods t. Letting w measure the net-of-taxes marginal wage, annual hours worked h is a function of (the log of) both the individual's own wage, w^o, and their partner's, w^p. I_{ict} measures the household's non-labor market income, and X_{icet} is a vector of individual and household characteristics.[13] I do not include indicators for educational attainment (education is one of my grouping instruments), and so the α^1_{ce} coefficient is

[11] Simple difference-in-difference estimates are also widely used in the literature, as in Eissa (1995), LaLumia (2009), and Crossley and Jeon (2007).

[12] I use the term 'coresidency type" to differentiate between individuals living together in heterosexual married, heterosexual unmarried, gay male and lesbian female households.

[13] In every specification reported, these covariates include indicators for race, Hispanic ethnicity, and time and geographical region indicator variables.

not interpretable as the compensated labor supply elasticity, but instead as the effect of a lifetime wage increase (Blundell and MaCurdy 1999). All coefficients are allowed to differ across the eight individual types *ce*.

I do not estimate the labor market participation choice. This is primarily for comparability with the previous literature. For the purposes of estimating tax revenue changes in section 5, this is equivalent to assuming that for the relevant range of policy options, no individual will opt into or out of the labor market due to changes in household taxation stemming from marital choices.

Standard economic theory makes clear that hours worked and wages are endogenously determined, as high-productivity workers will typically work more hours and draw higher wages, biasing OLS regression results. I use two sources of exogenous variation in net-of-taxation marginal wages. First, I estimate equation 2 using data only from 2003 and 2004, exploiting the expected break in tax rates, all else equal, caused by the JGTRRA.[14] Second, I use only changes in wages associated with exogenous characteristics. In particular, I use group-membership indicators as instruments for wages (Angrist 1991). The identifying assumption is that the demographic groups are defined such that they are unrelated and unchangeable with respect to the policy change, but the policy change has different effects on the different groups (Blundell, Duncan and Meghir (BDM) 1998). The intuition of the instruments can be seen in table 2, where the change in tax burden after the JGTRRA varies substantially across coresidency type and the number of children present.

Because of the progressivity of the tax schedule, the tax cuts will have a different effect on households of different income levels. I take each partners' birth year and level of terminal education to be shifters of earnings. Both are clearly exogenous with respect to current economic conditions. BDM (1998) suggest (but do not use) the presence of children as a grouping instrument, and the tax code explicitly changed with respect to the treatment of children. While it is unlikely that there was sufficient time to adjust family size in response to the tax cuts between 2003 and 2004, I take the precaution of grouping households by the presence of children aged 2-15 as another demographically exogenous characteristic. It is not possible to bear a child and rear it to two years of age in the

[14] In addition, a short time horizon reduces any potential impact of a violation of the "parallel-lines" assumption implicit in the class of instrumental variables estimators that I apply (BDM 1998).

two sample years in response to the policy shock, and since parents are generally legally responsible to rear a child until age 18, it is very difficult to remove a child under age 15 or 16 from the household in response to economic conditions.[15] I create three categories for each of these variables. Allowing group means to differ across time generates 54 (= 3 x 3 x 3 x 2) groups per *ce* type. Devereaux (2004) points out that in the presence of assortative mating, not only will own-wage be endogenous, but so will partner wage. Both he and Antecol and Steinberger (2011) instrument for both own- and parter-wages. I also will treat partner wages as endogenous, and include a set of 54 indicators for partner type in the vector of instruments, in addition to the own-type indicators.

Since individuals have reservation wages (due to the value of leisure, fixed costs of working, etc), those who draw low wage offers will opt out of the labor force. Low-skill workers will be particularly likely to draw very low offers, causing low-offer individuals and their parters to be omitted from the estimates of equation 2. This tends to bias α^1_{ce} downward for secondary workers.[16] Two commonly-applied corrections for this issue are wage imputation (Blau and Kahn 2007) and re-weighting schemes (Juhn 1992, Juhn and Murphy 1997). Both function to represent the entire wage offer distribution when we estimate equation 2. The implementation of these corrections is described in detail in appendix A.1. Since all prime earners work in my sample, these corrections are only made with respect to households whose secondary earners do not report a wage.

Table 4 presents estimates of the equation 2 labor supply model. Panel A gives the basic grouped-mean IV estimates while panels C and C add in the re-weighting and wage imputation censoring corrections, respectively.

By and large, primary earners are fairly similar across coresidency type. Own-wage coefficients tend to be small or statistically indistinguishable from zero at conventional levels. The insignificant heterosexual male coefficients suggest an uncompensated own-wage elasticity between 0.02 and 0.05. For gay men, the (also statistically insignificant) elasticity estimates range from 0.005 to -0.08. Children have little influence on labor supply decisions of prime earners. We can never reject the hypothesis that the sensitivity of hours worked to wages (either their own, or their partner's) between married men and

[15]This discussion clearly omits the small impact of adoption or child mortality.

[16]The bias in α^2_{ce} for secondary workers and in both log wage coefficients for primary workers is theoretically ambiguous, depending on the effects of productive substitution and complementarity, and assortative mating.

prime-earning homosexuals is equal. Correcting for selection of partners into the labor force reduces our estimates of the sensitivity of labor supply to own-wages and partner-wages among unmarried households, heterosexual and homosexual alike.

The same broad-brush similarly does not hold among secondary earners. There are essentially two types of secondary earners: married women, and everyone else. Consistent with the literature (Mroz 1987, Blundell and MaCurdy 1999), in every specification, married women have upward-sloping supply curves. Panel C implies an uncompensated own-wage elasticity among working married women of 0.48. Secondary-earner homosexuals and unmarried heterosexual women all behave similarly in terms of labor supply. Own- and partner-wage coefficients among these groups are the opposite sign as married women's. Statistically, we always reject equivalence of wage effects between secondary-earning homosexuals and married women, and we can not reject equivalence to unmarried women's wage coefficients, except for among lesbian women in panel C. Across all three panels, homosexual secondary earners' hours are very sensitive to partner wages. The uncompensated partner-wage elasticities range from 0.38 among lesbian women in panel B, to 0.45 among lesbian in panel C. In contrast to prime earners, controlling for labor market selectivity increases our estimates of the sensitivity of labor supply to own-wages and partner-wages for secondary earners in all household types.

It is consistent with past literature to find negative labor supply elasticities among non-married secondary earners. Antecol and Steinberger (2011) find negative own-wage elasticities among lesbian women in the 2000 Census. Stevenson and Wolfers (2007) argue that childless couples are much more likely to form couples on the basis of assortative mating (rather than labor market specialization). Non-married secondary earners are more likely to work than married women (see table 1), and the positive partner-wage coefficients are consistent with assortative mating behavior. Black, Sanders and Taylor (2007) argue that since homosexual couples have fewer children, they tend to gravitate to higher-private-amenity, higher-cost areas. The rewards to living in a dual-earner households are presumably greater in these areas, and this could drive greater household sorting by ability and greater coordination of labor supply decisions.

4.1.2 Endogenous marriage penalty

Given estimates of hours worked as a function of net-of-taxes wages, we can return to the measurement of the marriage penalty. The standard marriage penalty calculation is instantaneous in the sense of assuming no behavioral response to the shift in tax schedule that occurs upon marriage. Table 3 shows that there are substantial changes in marginal tax rates upon marriage, and so this assumption is rather unappealing. Instead, table 4 tells us that workers will adjust hours in response to changes in the tax schedule. These hours adjustments will change the tax burden, and so will affect the marriage penalty. I describe here an alternate way of calculating the marriage penalty, allowing for endogenous hours response to marriage and divorce, and I calculate the average bias implicit in the standard instantaneous measure.

I use the coefficients from my estimates of equation 2 to predict labor supply given the change in tax schedule.[17] Letting μ be the marginal tax rate, a worker's expected labor income is $w^o_{icet}\widehat{h_{icet}} = w^o_{icet}(\widehat{\alpha^1}(w^o_{icet}(1-\mu^o_{icet})) + \widehat{\alpha^2}(w^p_{icet}(1-\mu^p_{icet})) + \widehat{\beta_{ce}}X_{icet})$, where non-labor income is included in the X_{icet} here.[18]

A worker's hours choice will change as a function of changes in both her own marginal tax rate and her partner's. The marginal tax rate μ^o_{icet} is in turn a function of the hours choice. I predict each worker's hours choice given married and single tax schedules, household wages, the X_{icet} variables, and the value of income tax deductions. Predicted hours imply a marginal tax rate, which in turn give am hours choice. I iterate over this process until marginal tax rates for both household members converge. This generates predicted tax labilities when single and when married.

Denote the predictions of hours under each tax system for each earner to be $\widehat{h^e_q}$, where $q \in (m, s)$ indicates marital status (and $-q$ implies the counterfactual marital status). This hours choice implies a gross income $\widehat{y^e_q}$ and average tax rate $\widehat{\tau^e_q}$. I show in appendix A.3 that we can state the bias (overstatement) of the instantaneous method relative to

[17] I use the Blau and Kahn adjustment to estimate potential wages throughout the paper. All the data used in this section if from the years 2004-2007, under the post-JGTRRA tax system.

[18] One way to justify the use of an instantaneous marriage penalty is to assume that hours are a function of gross wages rather than net-of-taxes wages. Since wages are typically assumed to be exogenous to marital status, h_{icet} does not change upon marriage if taxes are excluded from the hours equation.

the endogenous method to be

$$
\begin{aligned}
I - E = \sum_{e=1,2} w^e \Big[& h_q^e \left((\tau_m^e - \tau_s^e) - (\widehat{\tau_m^e} - \widehat{\tau_s^e}) \right) + (h_q^e - \widehat{h_q^e})(\widehat{\tau_m^e} - \widehat{\tau_s^e}) \\
& + \tau_{-q}^e \left((h_m^e - h_s^e) - (\widehat{h_m^e} - \widehat{h_s^e}) \right) + (\tau_{-q}^e - \widehat{\tau_{-q}^e})(\widehat{h_m^e} - \widehat{h_s^e}) \Big] \\
& + \; n_m^e(\tau_m^e - \widehat{\tau_m^e}) - n_s^e(\tau_s^e - \widehat{\tau_s^e})
\end{aligned} \tag{3}
$$

where variables without hats represent the observed values in the data, as used in the instantaneous method of calculation.

There are five types of bias described in equation 3. The first term, $w^e h_q^e \left((\tau_m^e - \tau_s^e) - (\widehat{\tau_m^e} - \widehat{\tau_s^e}) \right)$, describes the *tax rate bias* of the instantaneous method: the tax revenue impact of the fact that, because it does not allow workers to adjust labor market choices, the change in the tax rate upon marriage is overstated (evaluated at the observed level of hours worked). The second term, $w^e(h_q^e - \widehat{h_q^e})(\widehat{\tau_m^e} - \widehat{\tau_s^e})$, describes the *modeling bias*: the difference in collected tax revenue due to the difference in predicted labor market earnings in the observed marital state. Note that this does not primarily come from prediction error in the hours equation, which in expectation is mean-zero. Instead, it comes from the induced correlation between hours and wages that arises when we predict hours according to equation 2. The third term, $w^e \tau_{-q}^e \left((h_m^e - h_s^e) - (\widehat{h_m^e} - \widehat{h_s^e}) \right)$, describes the instantaneous *hours bias*: the value of the expected hours response to the changes in taxation upon marriage (since I hold deductions and non-work income constant). The fourth term, $w^e(\tau_{-q}^e - \widehat{\tau_{-q}^e})(\widehat{h_m^e} - \widehat{h_s^e})$ is the *counterfactual bias*, the value of the difference in tax rates calculated for the non-observed marital status. Finally, $n_m^e(\tau_m^e - \widehat{\tau_m^e}) - n_s^e(\tau_s^e - \widehat{\tau_s^e})$, the *non-work income bias* represents the difference in tax treatment of non-work income between the two methods. Table 5 decomposes the average difference in the two marriage penalty measures, by coresidency type, into these 5 sources of bias.

For every household type, allowing for an endogenous labor market response to changes in the tax schedule results in higher predicted marriage penalties, on the order of $900 per household. The role of the non-work income and counterfactual biases are very small. For secondary earners, the modeling bias is quite large. Since wages and hours are, by construction, much more highly correlated when we model hours as a function of wages, we tend to over-predict hours for high-wage workers and under-predict hours for low-wage

workers (illustrating one aspect of why the labor supply behavior of secondary earners is of such perennial research interest (Blundell and MaCurdy 1999)). This modeling bias is particularly large for heterosexual married couples and gay men, the two highest-earning household types.

The two biases of primary theoretical interest are the tax rate bias and the hours bias. The instantaneous method assumes no hours response to changes in taxation due to entrance or exit into marriage. In terms of tax revenue, for the average heterosexual couple this assumption is not unreasonable. The endogenous method predicts only small average changes in hours worked after changing marital status. The endogenous method predicts a $55 smaller marriage penalty as a result. Secondary earners in homosexual households are predicted to work substantially more when married $(\widehat{h_m^e} - \widehat{h_s^e}) > (h_m^e - h_s^e)$, generating $167 and $283 more in tax revenue from gay and lesbian households, respectively, than predicted by the instantaneous method. This is the largest component of (non-modeling) instantaneous bias for homosexual households. For heterosexual households, the tax rate bias is the largest (non-modeling) contributor to the instantaneous bias. Allowing for endogenous labor market responses, the tax rate when single will be lower, and the tax rate when married higher (since $(\widehat{\tau_m^e} - \widehat{\tau_s^e}) > (\tau_m^e - \tau_s^e)$) than the instantaneous method asserts, contributing $360 to the instantaneous under-prediction of marriage penalties.

Clearly, accounting for the responsiveness of labor supply to changes in the tax structure could have a large impact on our estimates of the tax revenue changes generated by same-sex marriage legalization.

4.2 Marriage choice

Very few studies have measured the consequences of the Bush tax cuts on the labor market. None, to my knowledge, have focused on its consequences in terms of the marriage market. In terms of labor market effects, Heim (2009) simulates the consequences of the JGTRRA based on estimates from the 2001 PSID. Auten, Carroll, and Gee (2008) estimate the sensitivity of taxable income to the tax change, but not the sensitivity of labor supply *per se*. A much larger literature exists with respect to the effects on consumption and savings with respect to the 2001 tax rebate (see for example Johnson, Parker and Souleles (2006) and its references). In addition to using plausibly exogenous and unexpected changes in

policy, my estimates are the first that I'm aware of to account for the fact that assignments to treatment groups (i.e., the size of the marriage penalty) is non-random with respect to observable characteristics.

Alm and Whittington (2003) undertake the most thorough longitudinal study of the effects of taxation on marriage (among heterosexuals). They show that while transitions from non-cohabitation to cohabitation (either married or not) are insensitive to taxation, the transition from unmarried cohabitation to marriage is significantly influenced by changes in the household's instantaneous marriage penalty. Badgett, Gates, and Maisel (2008) analyze the sensitivity of cohabitation status to economic factors among homosexual couples, and find little influence.

The marriage penalty is not distributed randomly with respect to marital status, and so simple OLS will provide a biased estimate of the effect of taxes on marriage.[19] Since marital choice is partly a function of the expected marriage penalty, we must estimate martial choice as a function of unexpected changes in the marriage penalty. Alm and Whittington (2003) used variation used from changes in tax law between 1983-1997 to reach their conclusions. They show in previous work (Alm and Whittington 1996), however, that no discrete changes in the marriage penalty occurred over that time period. The JGTRRA created the first major jump in the marriage penalty since 1977 (see figure 1 and Alm and Whittington (1996)), and was largely unexpected.

I apply the Hirano-Imbens (2005) estimator of the effects of a continuous treatment variable. The treatment is the "jump" between the expected marriage penalty as a function of the pre-JGTRRA tax schedule $\widehat{p}(y_1, y_2)$ and the actually-experienced marriage penalty after implementation of the JGTRRA, $P(y_1, y_2)$. In essence, the expectation $\widehat{p}(y_1, y_2)$ can be treated as a taxation propensity score, a function of observable characteristics. I identify responsiveness of marital choice to taxation by comparing households with identical $\widehat{p}(y_1, y_2)$ that are exposed to varying levels of the treatment, $P(y_1, y_2)$. To ensure the necessary condition of sample overlap is satisfied (Imbens 2004), I balance the data following Barsky, Bound, Charles, and Lupman (BBCL 2002). The weights constructed by the BBCL procedure impose the counterfactual assumption that the distribution of

[19]For example, in the standard Becker (1991) model of household formation, partners will choose to specialize in market or non-market production. All else equal, individuals who plan to marry will also a) tend to be more specialized, and b) will have higher-earning primary workers. Each of these facts affects the marriage penalty.

$P(y_1, y_2)$ among married household matches the distribution observed among unmarried households. Intuitively, this process places greater weight on the married households that are observably most similar to non-married households in terms of the treatment received.

Appendix A.2 describes the precise econometric procedure used to derive the estimates of this section.

Figure 3 plots the likelihood of marriage as a function of the marriage penalty (the outcome of equation A.5, $E(I_m\widehat{(P(y_1, y_2))})$), for three types of heterosexual households: those with no children, those with one child, and those with more than one child. The plots are estimated via local polynomial regression. Under the assumptions of (weak) unconfoundedness and overlap (as summarized in Imbens (2004)) these functions describe the causal effect of marriage penalties on marital choice. The black curves are estimates based on the instantaneous penalty, while grey curves are estimates based on the endogenous penalty (with bootstrapped 95% confidence intervals in dashed grey lines).

There are three main conclusions from figure 3. First, the marriage penalty conveys a clear negative treatment effect among childless couples. Imposing the assumption of strict linearity in the treatment effect, the coefficient from regressing predicted marriage outcomes (as predicted via the results reported in table A.2) on the instantaneous marriage penalty is -0.000031 (s.e. = 0.00001). Evaluated at a typical marriage penalty cut of around -$800, this implies an increase in the probability of marriage of three-fifths of a percentage point for a childless couple (from 85.0% to 85.6%). This is a small effect, to be sure, but statistically significant and in line with the previous literature. This estimate is almost identical to that of Alm and Whittington (2003).

Second, the results are relatively insensitive to whether we use changes in the instantaneous or endogenous measure of the marriage penalty. Among childless couples, the linear treatment effect of the endogenous marriage penalty on marriage likelihood is -0.000045 (s.e. = 0.00001). This at once confirms that past studies of the relationship between taxation and marriage are robust to this new measure of the marriage penalty, while allowing us to apply more economically realistic measures of the tax revenue response to changes in marital status.

Third, among households with children, there is little evidence that unexpected changes in the marriage penalty cause changes in marriage likelihood. Both marriage penalty mea-

sures suggest an insignificant negative relationship between marriage and penalties among multi-child households. The instantaneous penalty implies a significantly positive relationship among single-child households (coefficient = 0.000033, s.e. = 0.00001), but the endogenous penalty implies a much smaller, negative, statistically insignificant effect. As the plot illustrates, the positive linear relationship of the instantaneous method comes almost entirely from the very imprecise estimates among low-penalty households.[20] The conservative conclusion is to reject the hypothesis of a consistent treatment effect, positive or negative, in this group.

5 Consequences of Same-Sex Marriage Legalization

A primary difficulty in estimating the effects of same-sex marriage legalization on federal income tax collection is knowing the marriage rate applicable to homosexual couples. Previous studies have applied a number of rates. The Williams Institute at UCLA assumes a 50% same-sex marriage rate (Badgett and Sears 2005, Badgett 2010).[21] Badgett (2010) also cites a number of studies that suggest that 80% of young gay individuals hope to marry. The Congressional Budget Office (CBO 2004) assumes that to a first approximation, all cohabiting same-sex couples would marry.

There are alternatives to these assumptions. We might assume that marriage likelihood is a function of demographic characteristics and is unrelated to sexual orientation. I impose this assumption by estimating the likelihood of marriage (among heterosexuals) via probit regression, as a function of age, terminal education, number of children in the household, and state of residence. Or, we could assume that marriage likelihood among same-sex couples is responsive to the marriage penalty in exactly the same way as it is among heterosexuals: the dose-response functions of figure 3 apply to same-sex couples. Finally (for the purposes of this paper) we could assume, as does the Williams Institute,

[20] Among one-child households, the lower the woman's earnings, the less likely she is to be married. The opposite is true among the childless and multiple child families. Since marriage penalties fall as the secondary earner's income fall (i.e., the tax code rewards specialization and penalizes dual-earners), there is a relatively large pool of high-subsidy unmarried one-child couples, which generates the inverted-U shape of the dose-response function using the instantaneous method. Notice the large confidence intervals around the average marriage likelihood in this group. This relates to the growth of out-of-wedlock births in poor urban areas (Stevenson and Wolfers 2007).

[21] The Williams Institute has released a series of state-level evaluations of the fiscal impact of state-level same-sex marriage legalization, which may be found at http://www2.law.ucla.edu/williamsinstitute//publications/Policy-Econ-index.html.

that the typical same-sex household has a 50% likelihood of marrying, but that marriage likelihood decreases in the marriage penalty, exactly as in figure 3. I take the figure 3 marital response functions and decrease them until they have a mean likelihood of 0.5.

Let p_i be the probability of marriage, and T_i^q be the tax lability of household i given marital status q. Given that all same-sex couples are single prior to legalization, the expected change in federal tax liability after legalization is $p_i(T_i^m - T_i^s)$, or p times the marriage penalty. I aggregate to the national level by multiplying this number times its household sampling weight and summing by year. My estimates are conservative for two reasons. The first is because of the under-count of same-sex couples in the Census due to the omission of individuals with allocated gender and marital status. Second, because we can only observe existing coresident homosexual couples in the Census and ACS, I can not account for the degree to which gay and lesbian will form households based on legalization of same-sex marriage.

Table 6 presents annual averages for the indicated time frames (2000-2002 and 2005-2007) for three measures of the marriage penalty: the instantaneous penalty, the endogenous penalty, and a"de-biased" penalty that removes the contribution of the modeling and counterfactual biases from the discrepancy between the instantaneous and endogenous methods, so that only the changes due to endogenous hours, tax rate, and nonwork income adjustments remain.

I present the income tax revenue estimates from 2000-2002 mainly to compare my estimates to previous work. The CBO (2004) estimated a pre-JGTRRA revenue boost between \$200-\$400 million, using a 100% marriage rate and an instantaneous penalty measure. My estimate in this case is a \$231 million revenue boost, within their range, but at the low end. Alm, Badgett, and Whittington (2000), who do not use individual-level tax data, estimate revenue boosts in the range of \$260 million to \$1.3 billion, with their preferred estimate being \$1 billion in additional income tax revenue. My instantaneous estimates are all below this range.

Between 2000 and 2002, the endogenous method of measuring the marriage penalty implies revenue boosts that are consistently more than double those of the instantaneous method. At the high end of the estimates, when we use the unadjusted endogenous penalty and assume a 100% marriage rate among observed same-sex couples, the federal

revenue gain is $491 million. At the low end, the bias-corrected endogenous penalty, assuming a 50% mean marriage rate and the empirical marginal response of marriage likelihood to the marriage penalty, the revenue gain to marriage legalization would have been $176 million. Notice that applying this "mean-50 conditional marriage expectation" almost always results in lower revenue estimates than the simple $p = 0.5$ estimates, since high-penalty households are the least likely to marry, all else equal.

My preferred results are the estimates based on the endogenous marriage penalty, and using the heterosexual conditional marriage expectation. This measure allows for optimizing agents to respond to the tax system in terms of both marriage and labor supply choice. The mean marriage probability among childless couples (the child-rearing category into which most same-sex couples fall) is around 85%. While this is much higher than the Williams Institute's preferred value, it is below that used by the CBO, and is in line with survey data on expressed willingness-to-marry among same-sex couples (Badgett 2010). The de-biased endogenous penalty corrects for some of the problems inherent in predicting hours choices (in particular, the induced wage-hour correlation), and in terms of revenue is always a more conservative estimate of revenue gains, so I tend to prefer those results in this context. Allowing for endogenous economic response, my preferred estimate implies an income tax revenue boost of $131 million (and no larger than $252 million, for any plausible set of assumptions). In the time period between the JGTRRA and the economic crisis of 2008, I estimate that the instantaneous method under-states the revenue gains from legalization, relative to the de-biased endogenous penalty, by $60-$120 million annually.

In the context of the federal budget, where income tax revenues were on the order of $1 trillion, these revenue gains are obviously very small. To put the same-sex numbers into context, in 2007, married heterosexual couples contributed the equivalent of $5.4 billion in (instantaneous) marriage penalty revenue to the federal government, or about $160 per married household. The expected gain in per-household revenue upon same-sex marriage legalization based on the instantaneous penalty and the heterosexual conditional marriage expectation is $118.

These estimates do not incorporate the changes in federal government expenditure that would arise from same-sex marriage legalization. These payments are projected to

be small, on the order of $100 million in total annually, split mainly between medicare and social security payments to spouses (CBO 2009) and health care and other partner benefits (CBO 2010). Interestingly, these cost increases almost exactly offset my preferred estimate of the revenue gains, suggesting that the federal legalization of same-sex marriage may be revenue-neutral.

6 Conclusion

Same-sex marriage legalization is a topic that inspires passionate responses from partisans on all sides. While the politics of the debate are likely to remain hot, it's necessary to have cool economic analysis at hand to understand the costs and benefits of the proposal. While many people have argued about the financial aspects of the law, I argue that credible economic and statistical identification is hard to find in the literature. I exploit a major exogenous break in the tax code to estimate the labor market response to changes in the tax rate that would accompany same-sex marriage, were it legalized. Using the same tax code shock, I present new evidence on the sensitivity of marriage to taxes among cohabiting couples.

I show that primary and secondary earners in homosexual households make very different labor market choices as a function of net wages. Further, the supply functions of homosexual couples are very different from heterosexual married couples. As such, it is reasonable to assume that homosexual and heterosexual couples' tax burdens might differ after marriage, all else equal. I present a new measure of the marriage penalty, the endogenous penalty, that measures the change in tax burden upon marriage as a function of the optimal labor market choices of both parters across households of different types.

I argue that the most reasonable estimates of the changes in tax revenue upon legalization are in the $100-$175 million range. Including reasonable measures of the additional costs that would come with same-sex marriage legalization, this implies the policy is at least, on net, revenue-neutral, and may generate a small boost to federal governmental coffers.

A Appendix

A.1 Labor Supply Estimation

A.1.1 Grouping instruments

I separate families with children aged 2-15 into three groups, those with 0 children, 1 child, or 2 or more children. I use three cohorts, defined by those born 1959 and earlier, those born 1960-1969, and those born 1970 or after. The educational groups are those with a high school diploma or less, those with more than high school but less than a bachelors degree (including college dropouts and those with associates degrees or similar credentials), and those with bachelor's and/or post-baccalaureate degrees. In every specification, I allow group effects to differ before and after the Bush tax (over two time periods t). Interacting the categories, I have 3 x 3 x 3 x 2 = 54 groups. I estimate equation 2 separately by coresidency group and earner type. Denote the vector of these group indicators g_{ce}. Due to the potential presence of assortative mating, in every labor supply regression, I instrument wages using both own- and partner-group indicators.

A.1.2 Selection: re-weighting

I follow the framework of Juhn and Murphy (1997) to implement the re-weighting correction for censoring at the bottom of the wage offer distribution. The idea is that those with low wage offers are likely to have low skill, and therefore would have both low wage offers and low hours worked, if they participated in the market. The people with reported wage data who are most similar to those without reported wages are those from the same group g_{ce} with low hours. I define "low hours" as 1 to 20 hours on the job in the typical week.

I define the number of individuals of each g_{ce} group who did not work as N_g^0 and the number who worked 1 to 20 hours as N_g^1. Let the weight adjustment $a_g = (1 + \frac{N_g^0}{N_g^1})$. This adjustment factor a_g allows households where the secondary earner works low hours (has low skill) to represent both working (non-censored) and non-working (censored) observations. Where ω_{icet} represents the sampling weights and u_{ict}^s is the usual weekly hours of work of the secondary earner in each observed household, define the adjusted weights by

$$\lambda_{icet} = \begin{cases} 0 & \text{if } u_{ict}^s = 0 \\ a_g * \omega_{icet} & \text{if } 1 \leq u_{ict}^s \leq 20 \\ \omega_{icet} & \text{if } u_{ict}^s > 20 \end{cases} \tag{A.1}$$

This adjustment is applied to both the primary and secondary earner to account for assortative mating or any other relationship between censored observations and observable household characteristics.

Juhn and Murphy (1997) argue that this method has the advantage of incorporating the distribution of all observable characteristics into the estimates, as opposed to only using information on mean wages, as done in the wage imputation method described below.

A.1.3 Selection: wage imputation

This procedure follows the discussion in Blau and Kahn (2007), and is widely applied in the literature. Again start with the set of grouping instruments g_{ce}.

Again assume that those who do not work are "low skilled" and most comparable to workers (of their g_{ce} type) who work low hours ($1 \leq u^s_{ict} \leq 20$). Among secondary workers, I estimate the conditional mean of net-of-tax log wages using the equation

$$\ln w_{icet} = \psi_{gct} g_{cet} + \gamma Z_{icet} + \upsilon_{icet} \tag{A.2}$$

where ψ_{gct} is the vector of coefficients associated with group indicators, Z_{icet} is a vector of variables (whose effects are common across g_{ce} types) that shift wages. I include race and geographic region indicators in Z_{icet}, and I allow for the existence of coresidency-group-specific time trends in wages.

I use the parameter estimates from equation A.2 to impute $\ln w_{icet}$ to secondary earners without wage data due to zero reported work hours. These conditional mean wages $\widehat{\ln w_{icet}}$ are then included (and the households no longer dropped) when I run the labor supply regression, equation 2. These wages are used both as $\ln(w^o_{icet})$ in the secondary-earner regressions and as $\ln(w^p_{icet})$ in primary earner regressions.

A.2 Marriage Choice Estimation

To estimate marriage choice as a function of the marriage penalty, I apply three corrections to ensure the exogeneity of the marriage penalty with respect to marital status. Following the Hirano and Imbens (2005) model for continuous treatment variables, I estimate the conditional marriage function

$$E(I_m(\pi)|\ p) = \alpha_0 + \alpha_1 \pi + \alpha_2 \pi^2 + \alpha_3 p(y_1, y_2) + \alpha_4 p(y_1, y_2)^2 + \alpha_5 \pi \cdot p(y_1, y_2) \tag{A.3}$$

where I_m is an indicator variable taking the value of one when the individual is married, π represents the treatment variable, the marriage penalty, and $p = p(y_1, y_2)$ is the expected marriage penalty as a function of taxable assets and offsets of each household member, as in equation 1. $p(y_1, y_2)$ is also referred to as the generalized propensity score. My first source of exogeneity is to only use changes in the marriage penalty around the expected level to identify the effects of marriage penalty on marriage.

Since "true" marriage penalty π is actually a deterministic function of y_1 and y_2, it might appear that any differences between π and the our estimate of this generalized propensity score $\widehat{p}(y_1, y_2)$ must come via functional-form misspecifications of $p(y_1, y_2)$. This would be true if the tax schedule were fixed across time, but my second source of exogeneity is the JGTRRA tax reforms of 2003. I estimate the coefficients of the propensity score function based on data from 2002 and 2003 (the time between the EGTRRA and JGTRRA reforms) and then predict $\widehat{p}(y_1, y_2)$ based on y_1 and y_2 data observed in 2003 and 2004 (immediately before and after the JGTRRA). Thus, the primary source of variation in π conditional on $\widehat{p}(y_1, y_2)$ in equation A.3 is the unexpected shock to the tax schedule in tax year 2003. I estimate the propensity score via OLS using quadratic terms in the income of the higher-earner and the earnings gap between household members (see Eissa and Hoynes 2000), linear terms in non-work income and expenditure variables, and indicator variables for every possible number of children. The propensity score regression

results are given in table A.1. Adding higher-order polynomial terms in any variable does not affect explanatory power, R^2. The results clearly do not capture all the non-linearities of the tax schedule (else, the R^2 would equal 1), but it does explain the great bulk of the determinants of the marriage penalty.

While equation A.3 describes the response of marriage to unexpected changes in the marriage penalty, it does not imply anything about how common each level of treatment is among those who are married and those who are unmarried. In particular, since the Census is purely cross-sectional, individuals who are married tend to be much older, higher-income, and more specialized in the labor market. Each of these factors implies that the married are over-represented among low-penalty households and under-represented among high-penalty households. In order to "balance" the sample in order to create of sample overlap (a necessary condition for identification) and to place appropriate weight on those married households most observably similar to unmarried households, my third correction is to re-weight the data following the method of BBCL (2002).

Intuitively, we ask: given π, what is the relative likelihood that the individual is unmarried, as a proportion of the population representation? In practice, I partition the unmarried in the 2003 Census into 20 marriage penalty quantile bins defined by π_b. Since the tax schedule changed in 2003, people with the same observable characteristics will have different treatment levels π in 2003 and 2004. Holding the tax schedule constant (which involves generating counterfactual taxes via TAXSIM under the assumption that there was no JGTRRA) I take every heterosexual in 2004 and every married household in 2003 and place them into the appropriate 2003 unmarried penalty bin π_b. Then for each bin I construct the weight adjustment factor

$$\phi(\pi) = \frac{Pr(I_m = 0|\ \pi)/Pr(I_m = 0)}{Pr(I_m = 1|\ \pi)/Pr(I_m = 1)} \quad \text{if } \pi_{b-1} < \pi \leq \pi_b. \tag{A.4}$$

Given sample weights ω_{icet}, the weights applied in the estimation of equation A.3 are $\omega_{icet} \times \phi(\pi_{ict})$. This process balances the sample with respect to the determinants of π, and the tax schedule changes after 2003. BBCL stress that this weight imposes the assumption that the distribution of the treatment (the marriage penalty) among the married matches the distribution among the unmarried. I set $\phi(\pi)$=0 for households whose marriage penalty is either greater than the maximum observed among the unmarried in 2003, or is lower than the minimum, to enforce a common support of the treatment variable.

The entire process described above allows me to recover the average potential outcome (or "dose-response" function) by predicting the expected marriage outcome over the treatment levels of interest. This dose-response function is given by

$$\begin{aligned} \widehat{E(I_m(\pi))} =& E\left[\widehat{\alpha}_0 + \widehat{\alpha}_1\pi_{ict} + \widehat{\alpha}_2\pi_{ict}^2 \right. \\ & \left. + \widehat{\alpha}_3\widehat{p}(y_{1ict}, y_{2ict}) + \widehat{\alpha}_4\widehat{p}(y_{1ict}, y_{2ict})^2 + \widehat{\alpha}_5\pi_{ict} \cdot \widehat{p}(y_{1ict}, y_{2ict})\right] \end{aligned} \tag{A.5}$$

where the term inside the brackets on the righthand side is evaluated for each individual, and the expectation over these values is evaluated by local polynomial regression over 50 centiles of the marriage penalty π. Since $\widehat{E(I_m(\pi))}$ is estimated using generated regressors and parameters, I recover standard errors by bootstrapping the entire process described above 1000 times over the sample of heterosexual households observed between 2002 and 2004. Equation A.5 estimates the causal effects of marriage penalties on marital choice.

I do this by penalty type (both instantaneous and endogenous) and by the number of children present in the household (none, one, or greater than one).

The regression output from equation A.3 is given in table A.2. These results are used to predict the righthand side of equation A.5. The estimates of the average potential outcome $\widehat{E(I_m(\pi))}$ are plotted in figure 3.

A.3 Bias of Instantaneous Marriage Penalty

To state the bias of the instantaneous method, I make two assumptions. First, workers do not opt into or out of the labor market due to the marital tax schedule changes. Second, I assume that an individual's (household's) income tax deduction is independent of their level of income, at least on the margin of changes being considered. Wages are also treated as fixed here. Let $e \in \{1, 2\}$ denote the primary and secondary earner, respectively, and $q \in \{m, s\}$ denote martial status (married, single). The instantaneous method assumes that within each household, each individual's hours are fixed at the observed level $h^e_q = h^e$. Taking wages w^e and non-labor income n^e as given, write instantaneous gross income in martial state q as $y^e_q = w^e h^e + n^e$. Given the average tax rate τ^e_q, we can calculate the instantaneous marriage penalty

$$I = [\tau^1_m y^1_m + \tau^2_m y^2_m] - [\tau^1_s y^1_s + \tau^2_s y^2_s] \quad . \tag{A.6}$$

When labor supply choice is allowed to be a function of the marginal tax rate, I write expected hours and income given taxes as $\widehat{h^e_q}$ and $\widehat{y^e_q} = w^e \widehat{h^e_q} + n^e$. Acknowledging that the average tax rate is in turn dependent on this endogenous hours choice, $\widehat{\tau^e_q}$, we can write the endogenous marriage penalty as

$$E = [\widehat{\tau^1_m}\widehat{y^1_m} + \widehat{\tau^2_m}\widehat{y^2_m}] - [\widehat{\tau^1_s}\widehat{y^1_s} + \widehat{\tau^2_s}\widehat{y^2_s}] \quad . \tag{A.7}$$

The bias of the instantaneous method of measurement is then

$$I - E = \left[(\tau^1_m y^1_m - \widehat{\tau^1_m}\widehat{y^1_m}) - (\tau^1_s y^1_s - \widehat{\tau^1_s}\widehat{y^1_s})\right] + \left[(\tau^2_m y^2_m - \widehat{\tau^2_m}\widehat{y^2_m}) - (\tau^2_s y^2_s - \widehat{\tau^2_s}\widehat{y^2_s})\right] \tag{A.8}$$

One particularly useful way of stating this bias arises by first pulling non-work income out of y^e_q, and adding and subtracting, for an individual of observed martial status q (and $-q$ representing the other marital status) the terms $w^e h^e_q \tau^e_{-q}$, $w^e \widehat{h^e_q}\widehat{\tau^e_{-q}}$, $w^e y^e_{-q}(\widehat{\tau^e_m} - \widehat{\tau^e_s})$, and $w^e \tau^e_q(\widehat{h^e_m} - \widehat{h^e_s})$. The result is that we can state the instantaneous bias as

$$\begin{aligned} I - E = \sum_{e=1,2} w^e \Big[& h^e_q \left((\tau^e_m - \tau^e_s) - (\widehat{\tau^e_m} - \widehat{\tau^e_s})\right) + (h^e_q - \widehat{h^e_q})(\widehat{\tau^e_m} - \hat{\tau^e_s}) \\ & + \tau^e_{-q}\left((h^e_m - h^e_s) - (\widehat{h^e_m} - \widehat{h^e_s})\right) + (\tau^e_{-q} - \widehat{\tau^e_{-q}})(\widehat{h^e_m} - \widehat{h^e_s})\Big] \\ & + \; n^e_m(\tau^e_m - \widehat{\tau^e_m}) - n^e_s(\tau^e_s - \widehat{\tau^e_s}) \end{aligned} \tag{A.9}$$

which is equation 3 given in the text.

References

[1] Alm, James, M. V. Lee Badgett, and Leslie A. Whittington, 2000. "Wedding Bell Blues: The Income Tax Consequences of Legalizing Same-Sex Marriage," *National Tax Journal* 53(2), pp. 201-214.

[2] Alm, James, and Leslie A. Whittington, 1996. "The Rise and Fall and Rise... of the Marriage Tax," *National Tax Journal* 49(4), pp. 571-589.

[3] Alm, James, and Leslie A. Whittington, 2003. "Shacking Up or Shelling Out: Income Taxes, Marriage, and Cohabitation," *Review of Economics of the Household* 1(3), pp. 169-186.

[4] Angrist, Joshua D. 1991. "Grouped-data estimation and testing in simple labor-supply models," *Journal of Econometrics* 47, pp. 243-266.

[5] Antecol, Heather and Michael Steinberger, *2011.* "Labor Supply Differences between Married Heterosexual Women and Partnered Lesbians: a Semi-Parametric Decomposition Approach," *Economic Inquiry.*

[6] Auten, Gerald, Robert Carroll, and Geoffrey Gee, 2008. "The 2001 and 2003 Tax Rate Reductions: An Overview and Estimate of the Taxable Income Response," *National Tax Journal* 61(3), pp. 345-364.

[7] Badgett, M. V. Lee, 2001. *Money, Myths and Change: The Economic Lives of Lesbians and Gay Men*, The University of Chicago Press, Chicago, IL.

[8] Badgett, M. V. Lee, 2010. "The Economic Value of Marriage for Same-Sex Couples," *Drake Law Review* 58, pp. 1081-1116.

[9] Badgett, M. V. Lee, Gary J. Gates, and Natalya C. Maisel, 2008. "Registered Domestic Partnerships Among Gay Men and Lesbians: the Role of Economic Factors," *Review of Economics of the Household* 6, pp. 327-346.

[10] Badgett, M. V. Lee, and R. Bradley Sears, 2005. "Putting a Price on Equality? The Impact of Same-Sex Marriage on California's Budget," *Stanford Law and Policy Review* 16(1).

[11] Barsky, Robert, John Bound, Kerwin Kofi Charles, and Joseph P. Lupton, 2002. "Accounting for the Black-White Wealth Gap: A Nonparametric Approach," *Journal of the American Statistical Association* 97, pp. 663-673.

[12] Becker, Gary. 1991. *A Treatise on the Family. Enlarged edition.* Cambridge: Harvard University Press.

[13] Berliant, Marcus, and Paul Rothstein, 2003 "Possibility, Impossibility, and History in the Origins of the Marriage Tax," *National Tax Journal* 56(2), pp. 303-317.

[14] Black, Dan A., Gary Gates, Seth Sanders, and Lowell Taylor, 2007. "The Measurement of Same-Sex Unmarried Partner Couples in the 2000 U.S. Census," *California Center for Population Research On-Line Working Paper Series* CCPR-02307, University of California - Los Angeles.

[15] Black, Dan A., Seth Sanders, and Lowell Taylor, 2007. "Economics of Lesbian and Gay Families," *Journal of Economic Perspectives* 21(2), pp. 53-70.

[16] Blandford, John M., 2003. "Sexual Orientation and Gender in the Determination of Earnings," *Industrial and Labor Relations Review* 56(4), 2003, pp. 622-642.

[17] Blau, Francine D. and Lawrence M. Kahn, 2007. "Changes in the Labor Supply Behavior of Married Women: 1980-2000," *Journal of Labor Economics* 25(3), pp. 393-437.

[18] Blundell, Richard, Alan Duncan, and Costas Meghir, 1998. "Estimating Labor Supply Responses using Tax Reforms," *Econometrica* 66(4), pp. 827-861.

[19] Blundell, Richard, and Thomas MaCurdy, 1999. "Labor supply: A review of alternative approaches," in *Handbook of Labor Economics, vol. 3A*, Orley Ashenfelter and David Card, eds., Amsterdam, Elsevier.

[20] Carpenter, Christopher S., 2008. "Sexual Orientation, Work, and Income in Canada," *Canadian Journal of Economics* 41(4), pp. 1239-1261.

[21] Crossley, Thomas and Sung-Hee Jeon, 2007. "Joint Taxation and the Labour Supply of Married Women: Evidence from the Canadian Tax Reform of 1988," *Fiscal Studies* 28(3), pp. 343-365.

[22] Congressional Budget Office, 2004. *The Potential Budgetary Impact of Recognizing Same-Sex Marriages*, Washington D.C., http://www.cbo.gov/doc.cfm?index=5559.

[23] Devereux, Paul J., 2004. "Changes in Relative Wages and Family Labor Supply," *Journal of Human Resources* 39, pp. 696-722.

[24] Eissa, Nada, 1995. "Taxation and Labor Supply of Married Women: the Tax Reform Act of 1986 as a Natural Experiment," *NBER working paper* 5023, Cambridge, MA.

[25] Eissa, Nadia, and Hilary Williamson Hoynes, 2000. "Explaining the Fall and Rise in the Tax Cost of Marriage: the Effects of Tax Laws and Demographic Trends, 1984-97," *National Tax Journal* 53(3), pp. 683-711.

[26] Feenberg, Daniel Richard, and Elizabeth Coutts, 1993. "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management* 12(1), pp. 189-194. http://www.nber.org/taxsim/.

[27] Feenberg, Daniel R., and Harvey S. Rosen, 1994. "Recent Developments in the Marriage Tax," *NBER working paper* 4705, Cambridge, MA.

[28] Hawkins, Jennifer N., 2009. "Tenth Annual Review of Gender and Sexuality Law, Family Law Chapter: Legal Recognition of Same-Sex Relationships," *The Georgetown Journal of Gender and the Law* 10, pp. 751-800.

[29] Heim, Bradley T., 2009. "Structural Estimation of Family Labor Supply with Taxes," *Journal of Human Resources* 44(2), pp. 350-385.

[30] Hirano, Keisuke and Guido W. Imbens, 2005. "The Propensity Score with Continuous Treatments," chapter 7 of *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, A. Gelman and X.-L. Meng, eds., John Wiley and Sons, Chichester, UK.

[31] Imbens, Guido W., 2004. "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," *The Review of Economics and Statistics* 86(1), pp. 4-29.

[32] Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles, 2006. "Household Expenditure and the Income Tax Rebates of 2001," *The American Economic Review* 96(5), pp. 1589-1610.

[33] Juhn, Chinhui, 1992. "Decline of Male Labor Market Participation: The Role of Declining Market Opportunities," *Quarterly Journal of Economics* 107, pp. 79-121.

[34] Juhn, Chinhui, and Kevin M. Murphy, 1997. "Wage Inequality and Family Labor Supply," *Journal of Labor Economics* 15(1), pp. 72-97.

[35] Keifer, Donald, Robert Carroll, Janet Holtzblatt, Allen Lerman, Janet McCubbin, David Richardson, and Jerry Tempalski, 2002. "The Economic Growth and Tax Relief Act of 2001: Overview and Assessment of Effects on Taxpayers," *National Tax Journal* 55(1), pp. 89-117.

[36] LaLumia, Sara, 2008. "The Effects of Joint Taxation of Married Couples on Labor Supply and Non-wage Income," *Journal of Public Economics* 92, 1698-1719.

[37] Leppel, Karen, 2008. "Labour Force Status and Sexual Orientation," *Economica* 76, pp. 197-207.

[38] Mroz, Thomas, 1987. "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica* 55(4), pp. 765-799.

[39] Orrefice, Sonia, 2011. "Sexual Orientation and Household Decision Making. Same-sex Couples' Balance of Power and Labor Supply Choices," *Labour Economics* 18, pp. 145-158.

[40] Ruggles, Steven J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek, 2010. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota.

[41] Savage, Charlie, 2011. "Obama Orders End to Defense of Gay Marriage Law," The New York Times, Feb 23, 2011. http://www.nytimes.com/2011/02/24/us/24marriage.html.

[42] Sjoquist, David L., and Mary Beth Walker, 1995. "The Marriage Tax and the Rate and Timing of Marriage," *National Tax Journal* 48(4), pp. 547-558.

[43] Stevenson, Betsey and Justin Wolfers, 2007. "Marriage and Divorce: Changes and their Driving Forces," *Journal of Economic Perspectives* 21(2), pp. 27-52.

[44] Tebaldi, Edinaldo and Bruce Elmslie, 2006. "Sexual Orientation and Labour Supply," *Applied Economics* 38(5), pp. 549-562.

Table 1: Census 2000-2007 summary statistics for individuals in co-resident households, by sexual orientation

	Heterosexual				Homosexual			
	Married		Unmarried		Men		Women	
	Men	Women	Men	Women	Prime earner	Second earner	Prime earner	Second earner
Panel a: mean statistics								
Black	0.078	0.07	0.143	0.11	0.045	0.054	0.066	0.064
	(0.002)	(0.002)	(0.010)	(0.009)	(0.004)	(0.004)	(0.004)	(0.004)
Age	40.3	38.5	34.3	32.6	39.1	37.9	38.3	37.6
	(0.057)	(0.056)	(0.207)	(0.200)	(0.127)	(0.128)	(0.129)	(0.132)
Years of education	13.7	13.7	12.6	12.8	15.0	14.3	15.0	14.5
	(0.021)	(0.019)	(0.056)	(0.055)	(0.048)	(0.043)	(0.042)	(0.042)
Number of own biological children present	1.49	1.51	0.32	0.53	0.10	0.06	0.22	0.17
	(0.009)	(0.008)	(0.018)	(0.025)	(0.010)	(0.005)	(0.011)	(0.009)
Particpate in labor force (at survey date)	0.963	0.717	0.953	0.793	0.973	0.845	0.966	0.84
	(0.002)	(0.003)	(0.007)	(0.010)	(0.003)	(0.006)	(0.003)	(0.006)
Annual hours worked, if >0	2,233	1,698	2,062	1,859	2,229	1,995	2,151	1,928
	(4.5)	(5.5)	(16.7)	(15.9)	(9.6)	(10.6)	(8.5)	(9.2)
HH has non-labor income	0.366		0.337		0.453		0.456	
	(0.003)		(0.011)		(0.007)		(0.007)	
Panel b: median statistics								
Gross wage if hours > 0	19.87	14.24	13.89	12.29	24.21	16.34	20.54	14.92
	(0.16)	(0.13)	(0.39)	(0.29)	(0.44)	(0.35)	(0.34)	(0.24)
Gross labor income if hours > 0	43,883	24,306	28,046	23,232	53,428	32,692	44,000	28,743
	(348)	(213)	(600)	(534)	(917)	(497)	(698)	(423)
HH labor income	66,792		50,126		83,847		70,319	
	(412)		(910)		(1247)		(978)	
HH non-labor income, if any	2,194		2,804		3,000		3,076	
	(84)		(274)		(264)		(147)	
N	37649	37649	3263	3263	7082	7082	7058	7058

Note: Individual statistics are weighted by Census person weights, household characteristics by household weights. Standard errors are in parenthesis. All financial variables are stated in terms of 2004 dollars.

Table 2: Household tax liability and marriage penalty

	No children				One child				More than one child			
	Hetero married	Hetero partner	Gay men	Lesbian women	Hetero married	Hetero partner	Gay men	Lesbian women	Hetero married	Hetero partner	Gay men	Lesbian women
actual household federal income taxes												
2000-2003	20,624	14,342	28,748	21,781	18,751	7,779	13,896	12,476	16,325	5,978	8,056	11,925
	(424)	(684)	(710)	(644)	(476)	(680)	(3447)	(882)	(324)	(654)	(2267)	(1717)
2004-2007	18,638	14,701	26,738	19,545	16,858	6,701	18,011	15,632	15,420	4,417	15,199	13,455
	(370)	(641)	(489)	(370)	(316)	(535)	(2260)	(944)	(252)	(435)	(2136)	(1003)
change after JGTRRA	**-1,986**	359	**-2,010**	**-2,236**	**-1,894**	-1,077	4,116	**3,157**	**-905**	-1,561	**7,143**	1,531
	(610)	(1064)	(967)	(842)	(623)	(1001)	(4523)	(1520)	(449)	(926)	(3549)	(2223)
average marriage penalty payment												
2000-2003	183	469	1098	817	1692	735	1317	1393	1230	623	871	1604
	(35)	(46)	(62)	(56)	(54)	(145)	(434)	(127)	(39)	(197)	(300)	(233)
2004-2007	-613	-220	156	-176	757	234	621	439	259	-467	731	302
	(32)	(45)	(45)	(32)	(40)	(95)	(231)	(124)	(31)	(161)	(274)	(195)
change after JGTRRA	**-796**	**-689**	**-942**	**-993**	**-935**	**-501**	-696	**-954**	**-971**	**-1,090**	-140	**-1,301**
	(48)	(64)	(76)	(65)	(54)	(158)	(482)	(173)	(43)	(247)	(505)	(302)
N	9855	1773	6541	5388	9272	747	293	956	18524	737	234	690

Note : all tax data is calculated through TAXSIM on Census data, using the variables described in footnote 5 of the text. Bold numbers indicate that the change in the variable after the 2003 JGTRRA (whose effects first appear in the 2004 ACS) is statistically significant at the 5% level. Financial variables are stated in terms of 2004 dollars. Standard errors are in parenthesis.

Table 3: Marginal income tax rates, marriage, and the Bush tax reforms

	Heterosexual				Homosexual			
	Married		Unmarried		Men		Women	
	Men	Women	Men	Women	Prime earner	Second earner	Prime earner	Second earner
Individual marginal tax rate (Federal income tax+ individual portion of FICA)								
2000-2003	0.255	0.283	0.275	0.216	0.312	0.246	0.304	0.239
	(0.001)	(0.001)	(0.003)	(0.006)	(0.002)	(0.004)	(0.002)	(0.004)
2004-2007	0.23	0.257	0.259	0.199	0.297	0.235	0.283	0.22
	(0.001)	(0.001)	(0.003)	(0.004)	(0.001)	(0.002)	(0.002)	(0.002)
change after JGTRRA	**-0.025**	**-0.026**	**-0.016**	**-0.017**	**-0.015**	**-0.011**	**-0.021**	**-0.019**
	(0.001)	(0.001)	(0.005)	(0.009)	(0.002)	(0.003)	(0.003)	(0.005)
"Rate marriage penalty"								
2000-2003	-0.036	0.099	-0.056	0.043	-0.046	0.064	-0.052	0.058
	(0.001)	(0.001)	(0.004)	(0.005)	(0.002)	(0.003)	(0.002)	(0.003)
2004-2007	-0.046	0.084	-0.059	0.034	-0.053	0.045	-0.058	0.047
	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)
change after JGTRRA	**-0.010**	**-0.015**	-0.003	-0.009	**-0.007**	**-0.019**	**-0.006**	**-0.011**
	(0.001)	(0.002)	(0.005)	(0.009)	(0.002)	(0.004)	(0.002)	(0.004)
N	37651	37651	3264	3264	7082	7082	7058	7058

Note : all tax data is calculated through TAXSIM on Census data, using the variables described in footnote 5 of the text. Bold numbers indicate that the change in the variable after the 2003 JGTRRA (whose effects first appear in the 2004 ACS) is statistically significant at the 5% level. Financial variables are stated in terms of 2004 dollars. The "rate marriage penalty" is the marginal tax rate if married minus the marginal rate if single. Standard errors are in parenthesis.

Table 4: Labor supply function by coresidency type and primary/secondary earner status

	Primary earners				Secondary earners			
	Hetero married men	Hetero parter men	Gay men	Lesbian women	Hetero married women	Hetero partner women	Gay men	Lesbian women
Panel A: grouping instrument results								
Log own marginal wage	53.28	-339.98	-191.61	-130.79	224.47	-231.04	-167.50	-461.74
	(69.16)	(132.06)	(114.84)	(103.44)	(68.80)	(106.93)	(163.14)	(114.01)
Log partner marginal wage	120.59	395.33	410.26	278.14	-189.74	316.07	515.90	532.14
	(66.30)	(122.02)	(141.41)	(114.50)	(64.15)	(114.35)	(131.74)	(106.47)
HH non-work income (in thousands)	1.11	0.14	-1.21	-5.09	3.88	-3.13	-2.97	-12.04
	(2.05)	(4.29)	(1.64)	(2.33)	(3.87)	(3.85)	(1.72)	(2.48)
I(no children)	-11.27	-199.85	-98.34	14.46	139.68	-45.00	-94.82	135.86
	(39.36)	(132.85)	(214.84)	(94.61)	(44.07)	(126.41)	(207.98)	(95.34)
Number of children in HH	13.28	-95.64	-59.90	-22.95	-46.87	-9.94	-6.79	85.68
	(15.36)	(59.29)	(74.83)	(54.01)	(19.48)	(66.27)	(110.15)	(51.67)
Panel B: grouping + re-weighting results								
Log own marginal wage	114.547	-287.12	-29.209	-70.653	469.421	-537.56	-352.31	-591.5
	(62.93)	(125.81)	(115.13)	(104.12)	(100.81)	(112.15)	(173.11)	(110.43)
Log partner marginal wage	102.067	321.441	224.476	255.602	-276.2	419.375	769.024	742.103
	(63.58)	(111.44)	(128.56)	(109.85)	(86.97)	(117.65)	(139.79)	(119.25)
HH non-work income (in thousands)	0.3	0.861	-1.996	-3.826	4.644	-3.711	-3.573	-12.38
	(1.97)	(3.95)	(1.97)	(2.56)	(3.80)	(4.58)	(2.01)	(2.82)
I(no children)	-13.05	-205.33	-77.192	-18.193	193.424	32.571	-189.02	273.379
	(42.61)	(136.58)	(236.30)	(132.92)	(62.16)	(140.02)	(226.01)	(161.64)
Number of children in HH	16.968	-100.81	-43.713	-69.804	-54.02	-16.455	-7.972	148.287
	(16.85)	(59.24)	(82.38)	(89.69)	(25.58)	(73.41)	(106.07)	(86.36)
Panel C: grouping + wage imputation results								
Log own marginal wage	108.599	-186.05	12.75	116.34	586.17	-420.55	-403.82	-837.14
	(52.04)	(109.60)	(101.98)	(76.90)	(86.83)	(147.75)	(175.62)	(91.37)
Log partner marginal wage	85.882	251.885	104.882	27.328	-329	536.734	841.083	884.275
	(61.12)	(114.54)	(134.56)	(82.76)	(71.61)	(168.34)	(126.53)	(98.39)
HH non-work income (in thousands)	-0.139	-5.244	-2.992	-5.953	-0.799	-22.411	-11.278	-17.137
	(1.60)	(4.97)	(2.19)	(2.37)	(4.41)	(10.24)	(3.00)	(3.31)
I(no children)	-17.145	-196.89	473.948	31.822	228.235	353.31	140.174	232.037
	(35.47)	(126.03)	(223.13)	(98.23)	(54.24)	(179.16)	(237.01)	(158.31)
Number of children in HH	14.573	-139.52	214.306	-40.82	-95.331	11.833	-82.215	89.32
	(13.05)	(53.65)	(104.73)	(61.29)	(22.01)	(96.66)	(115.75)	(84.32)

Note : Dependent variable is annual hours of work. All regressions use own- and parter-group indicators as instruments for own- and partmer-wages. Each regression also includes indicator variables controlling for race, hispanic ethnicity, year, and census region, using data from year 2003 and 2004. The definition of grouping instuments, re-weighting procedure (for panel b) and wage imputation methodology (for panel c) are all described in Appendix A.1. Robust standard errors are given in parenthesis.

Table 5: Bias in instantaneous marriage penalty calculation

	Heterosexual						Homosexual					
	Married			Unmarried			Men			Women		
	Men	Women	HH Total	Men	Women	HH Total	Prime earner	Second earner	HH Total	Prime earner	Second earner	HH Total
(1) Tax rate bias:	-60	-286	-346	-14	-506	-520	6	-95	-89	-71	-179	-250
$w * h_q * ((\tau_m - \tau_s) - (\hat{\tau}_m - \hat{\tau}_s))$	(11)	(14)	(17)	(40)	(66)	(71)	(12)	(23)	(26)	(16)	(22)	(27)
(2) Model bias:	101	-655	-554	-2	-269	-271	26	-585	-559	31	-249	-218
$w * (h_q - \hat{h}_q) * (\hat{\tau}_m - \hat{\tau}_s)$	(8)	(28)	(29)	(31)	(94)	(96)	(18)	(69)	(72)	(18)	(51)	(57)
(3) Hours bias	4	69	73	46	-58	-12	10	-306	-296	-35	-345	-380
$w * \hat{\tau}_{-q} * ((h_m - h_s) - (\hat{h}_m - \hat{h}_s))$	(1)	(2)	(2)	(4)	(6)	(3)	(1)	(9)	(7)	(1)	(9)	(9)
(4) Counterfactual bias	2	8	10	-1	3	2	1	-14	-13	0	-10	-10
$w * (\tau_{-q} - \hat{\tau}_{-q}) * (\hat{h}_m - \hat{h}_s)$	(0)	(1)	(1)	(2)	(4)	(3)	(0)	(3)	(3)	(0)	(3)	(3)
(5) Non-work income bias	-5	-10	-15	13	-39	-26	-1	-67	-68	1	-39	-38
$(\tau_m - \hat{\tau}_m) * n_m - (\tau_s - \hat{\tau}_s) * n_s$	(2)	(3)	(4)	(8)	(18)	(15)	(4)	(14)	(13)	(4)	(8)	(8)
Total bias: (1) + (2) + (3) + (4) + (5)	42	-874	-832	42	-869	-827	42	-1067	-1025	-74	-822	-896
	(12)	(26)	(28)	(50)	(82)	(85)	(19)	(72)	(72)	(21)	(58)	(58)
N		22,198			1,962			4,506			4,546	

Note : Decomposition of instantaneous marriage penalty minus endogenous marriage penalty in post-JGTRRA years (2004-2007), as derived in appendix A.3. Subscript q indicates the actual marital status of the household, -q indicates the other marital status.

Table 6: changes in federal income tax revenue after same-sex marriage legalization

	Instantaneous penalty	Endogenous penalty	Endogenous, de-biased
	Pre-JGTRRA (2000 - 2002)		
p = 0.5	116 (7)	237 (13)	214 (11)
p = 1	231 (10)	474 (17)	428 (15)
p = f(demograhpics)	205 (10)	416 (17)	372 (14)
p = heterosexual conditional marriage expectation	179 (9)	326 (15)	310 (13)
p = mean-50 conditional marriage expectation	104 (7)	202 (12)	192 (10)
	Post-JGTRRA (2005 - 2007)		
p = 0.5	16 (5)	126 (7)	78 (7)
p = 1	31 (7)	252 (11)	156 (10)
p = f(demograhpics)	36 (7)	223 (10)	138 (9)
p = heterosexual conditional marriage expectation	13 (7)	180 (9)	131 (9)
p = mean-50 conditional marriage expectation	7 (5)	109 (7)	78 (7)

Note : units are millions of 2004 dollars. Stadard errors are given in parenthesis, generated by taking random draws from U[0,1] to determine who marries, summing the change in annual tax payment, and bootstrapping 1000 repetitions over this measure. The de-biased measure is equal to the endogenous penalty, less the effects of the modeling and counterfactual biases described in table 5.

Figure 1: Mean marriage penalty across time by coresidency type

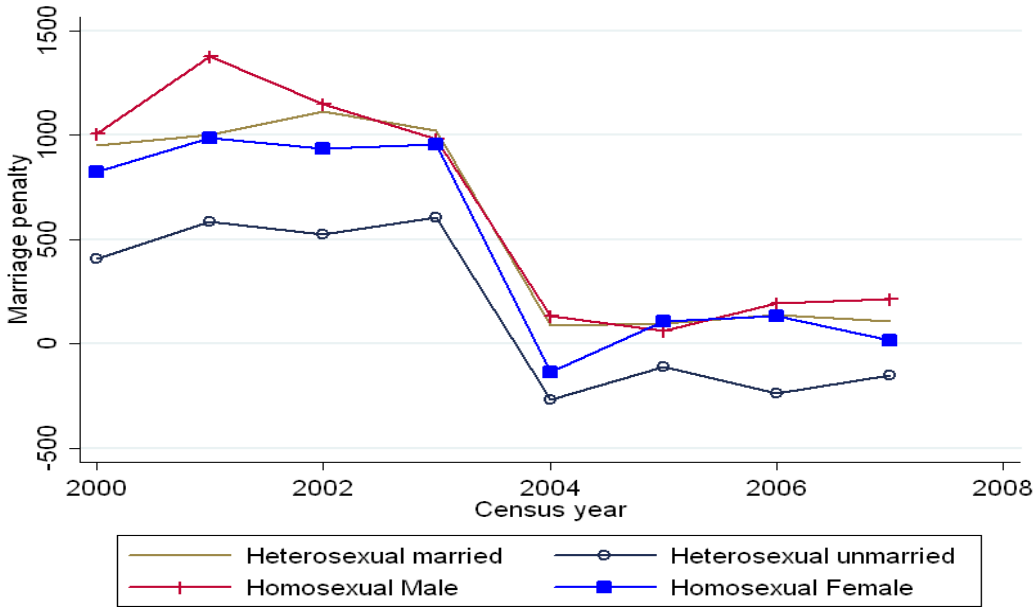
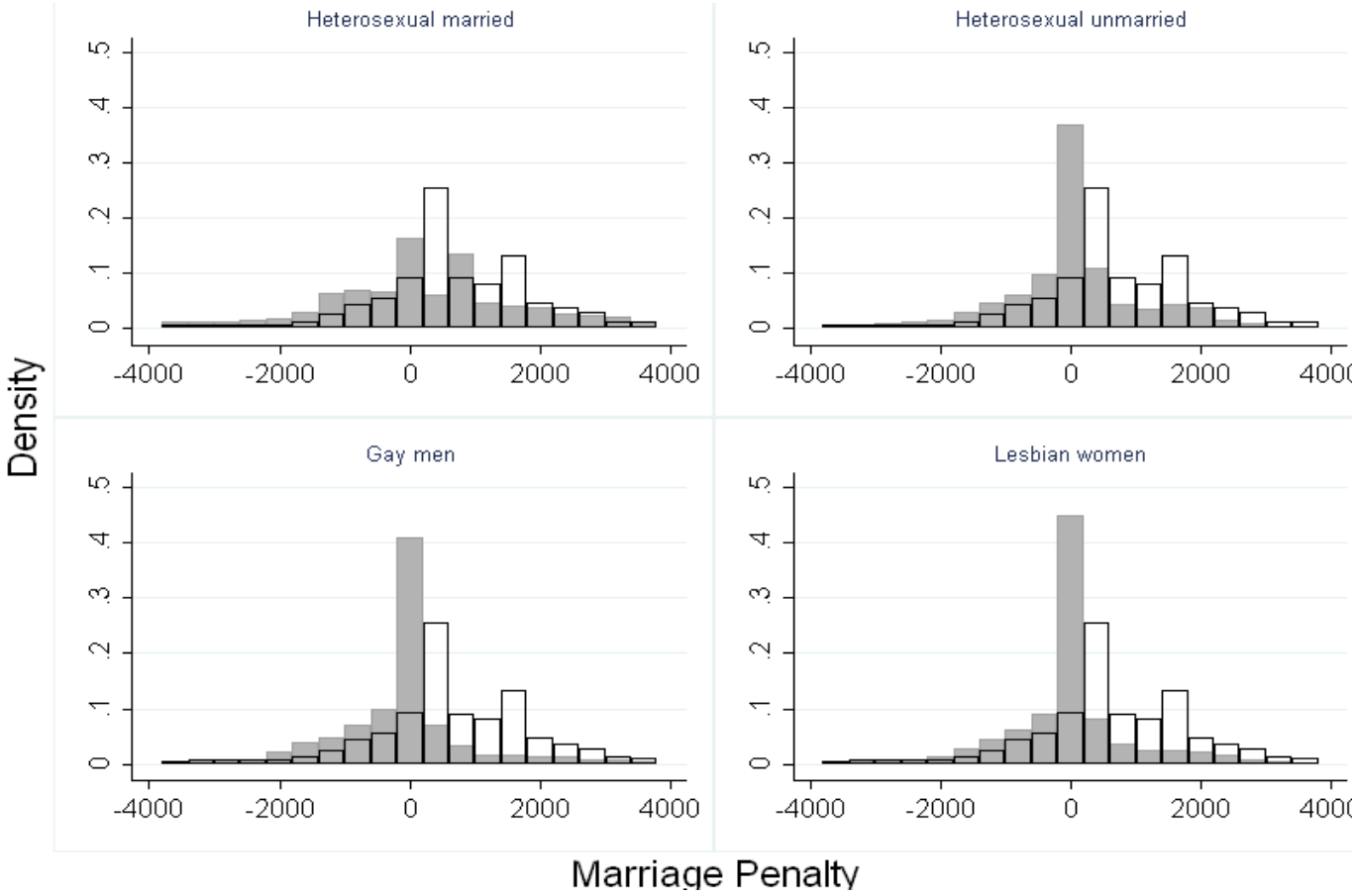
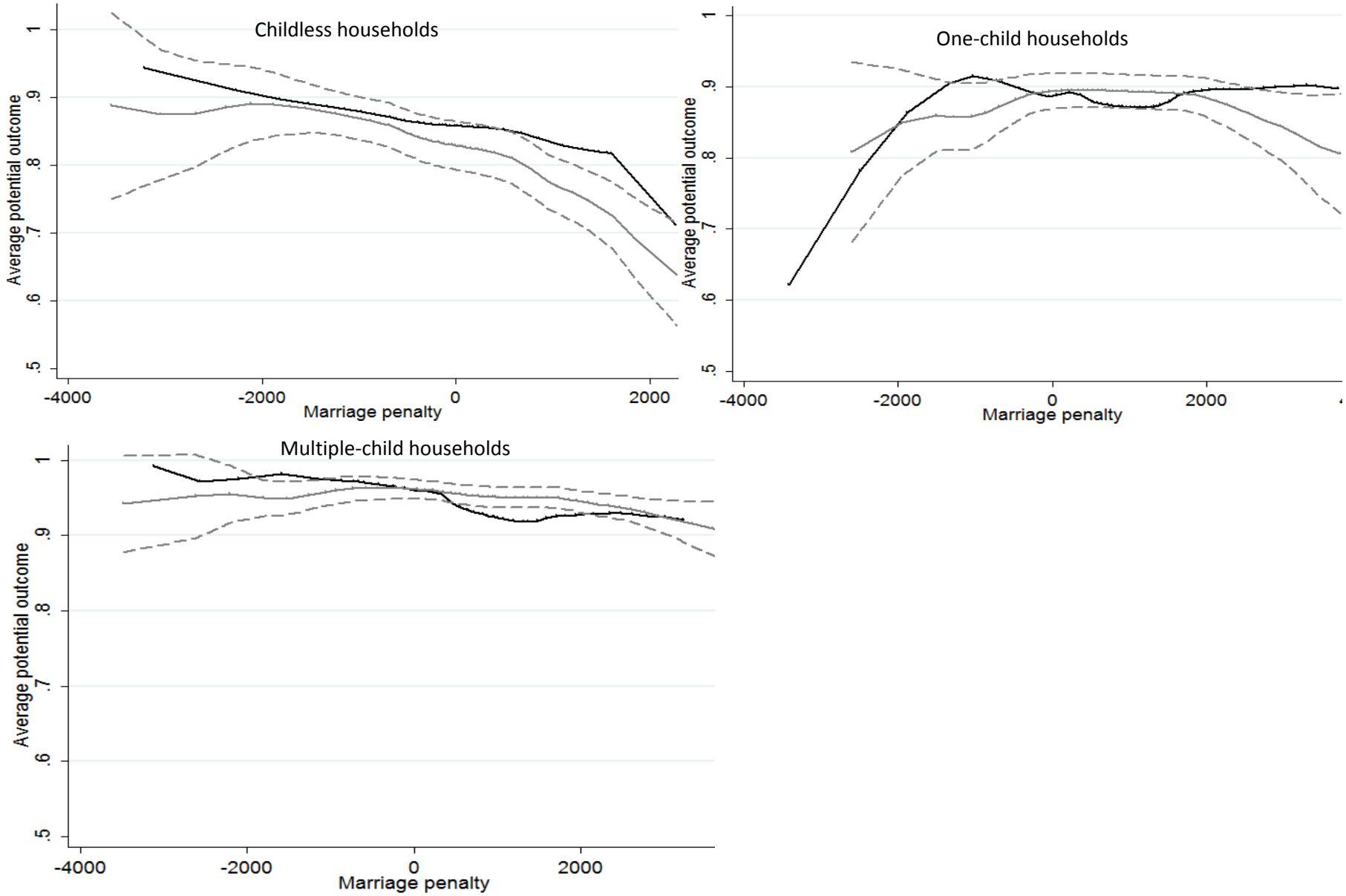


Figure 2: Distribution of marriage penalties, before and after tax reform



Note : Black outlined bars indicate the distribution before implementation of the JGTRRA, grey bars the distribution after. The distributions in this figure have been truncated for illustrative purposes - penalty data ranges from -$10,569 to $25,898.

Figure 3: Average potential marriage outcome, given marriage penalty



Note : Plots are of average potential marriage outcome as a function of marriage penalty, for heterosexual households with zero, one, or multiple children. The black solid line indicates the treatment effect as measured by the instantaneous marriage penalty, and the grey solid line measures the treatment effect as measured by the endogenous marriage penalty. The dashed grey lines represent a 95% confidence interval (CI) on the estimates using the endogenous penalty (I have suppressed the CI for the instantaneous method for graphical clarity). The methodology for generating average treatment outcomes follows Hirano and Imbens (2004), and is described in detail in appendix A.2.

Table A.1: pre-JGTRRA determinants of expected marriage penalty

	Instantaneous penalty	Endogenous penalty
Labor income of high earner /1,000	29.850	10.138
	(2.234)	(2.702)
(Labor income of high earner) ^{2} /1,000,000	0.391	-0.129
	(0.018)	(0.019)
(High labor income - low labor income) /1,000	-65.974	-21.792
	(1.536)	(2.018)
(High labor income - low labor income) ^{2} /1,000,000	-0.332	0.157
	(0.016)	(0.018)
Household investment income	0.048	0.033
	(0.006)	(0.003)
HH retirement income	0.042	0.051
	(0.006)	(0.006)
HH other income	0.039	0.026
	(0.005)	(0.006)
HH rent payment	-0.172	0.279
	(0.037)	(0.056)
HH roperty tax payment	-0.054	0.024
	(0.007)	(0.011)
Constant	-344.360	184.276
	(47.630)	(61.511)
N	8738	8738
R^2	0.7557	0.1656

Note : Dependent variable is marriage penalty. Estimate is based on heterosexual households in 2002 and 2003, over the common supposert of penalties between married and unmarried households. Estimate also includes indicator variables for every observed number of children. Standard errors are in parenthesis.

Table A.2: conditional marriage probability given treatment and expected treatment

	Instantaneous penalty			Endogenous penalty		
	No children	One child	More than one child	No children	One child	More than one child
Marriage penalty/1000	-0.135	-0.077	-0.007	-0.132	0.071	-0.094
	(0.089)	(0.103)	(0.055)	(0.034)	(0.073)	(0.072)
(Marriage penalty) ^{2}/1,000,000	0.034	-0.341	-0.237	-0.057	-0.067	-0.036
	(0.065)	(0.041)	(0.042)	(0.017)	(0.017)	(0.014)
Expected penalty/1000	0.030	-0.398	-0.517	-0.311	-0.372	-0.832
	(0.079)	(0.198)	(0.105)	(0.117)	(0.236)	(0.569)
(Expected penalty) ^{2}/1,000,000	-0.013	-0.263	-0.168	0.035	-0.083	-0.147
	(0.036)	(0.070)	(0.036)	(0.063)	(0.102)	(0.226)
Actual penalty * expected penalty	-0.029	0.779	0.519	0.128	0.039	0.127
	(0.098)	(0.116)	(0.085)	(0.089)	(0.073)	(0.060)
Constant	1.114	1.652	2.168	1.054	1.748	2.754
	(0.047)	(0.155)	(0.102)	(0.053)	(0.169)	(0.350)
N	3,036	2,218	4,408	3,036	2,218	4,408

Note : probit estimate of marriagie likelihood, according to appendix equation A.5. Estimate is based on heterosexual households in 2003 and 2004 over the common supposert of penalties between married and unmarried households. Expected marriage penalty is calculated based on the results of table A.1. Unadjusted standard errors are in parenthesis.