

# THE EFFECT OF TAXES ON TAXABLE EARNINGS: EVIDENCE FROM THE 2001 AND RELATED U.S. FEDERAL TAX ACTS

Perry Singleton

*This study measures the elasticity of taxable earnings to the marginal tax price, identified by a marriage penalty relief provision contained in the Economic Growth and Tax Relief Reconciliation Act of 2001. The change in joint taxable earnings in response to the provision implies a joint elasticity ranging from 0.217 to 0.304. The joint response is driven by the taxable wage earnings of males, and does not reflect a shift in income from tax-deferred to taxable compensation, but an increase in total wage earnings.*

*Keywords: taxation, tax elasticity, new tax responsiveness literature, Economic Growth and Tax Relief Reconciliation Act of 2001*

*JEL Codes: H21, H31, J22*

## I. INTRODUCTION

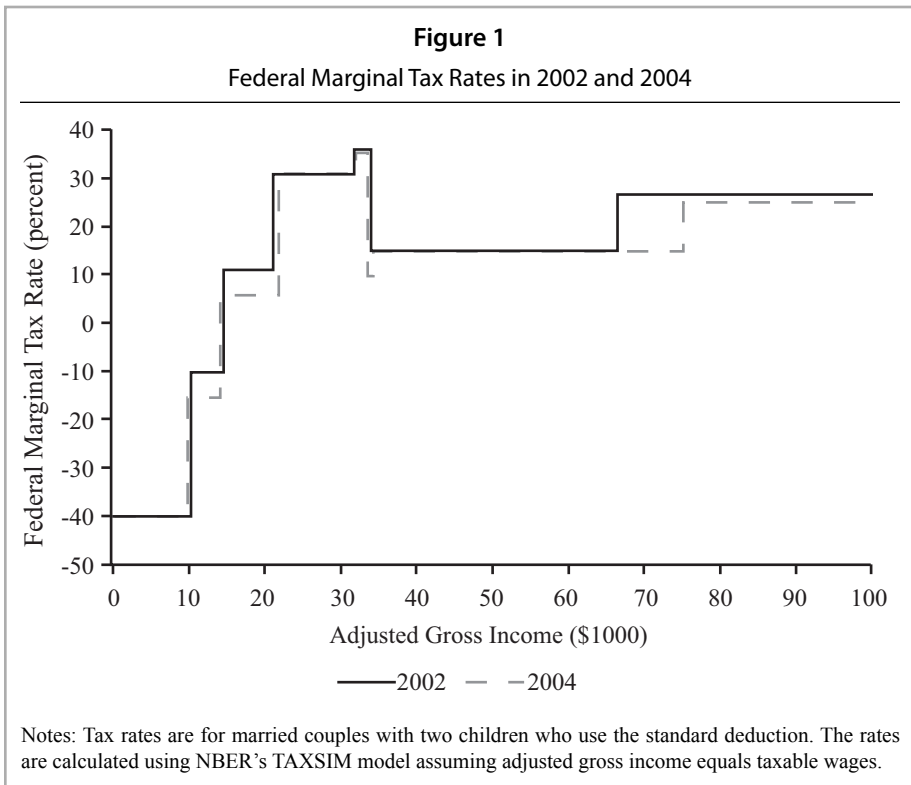
The elasticity of income to the marginal tax price is a critical parameter in the debate on the optimal level of taxation. A class of economic studies, termed the New Tax Responsiveness Literature (Goolsbee, 2000a), attempts to measure the elasticity of income using quasi-experimental methods. Nearly all studies rely on major tax reforms for identification, and nearly all studies use income data that are longitudinal.<sup>1</sup> In a review of the literature, Gruber and Saez (2002) note that most elasticity estimates range from zero to 0.8, reflecting differences across studies in the definition of income, the sample analyzed, and the source of identification.

---

<sup>1</sup> Reforms considered in the literature include the Economic Recovery Tax Act of 1981 (Lindsey, 1987; Navratil, 1995), the Tax Relief Act of 1986 (Feldstein, 1995; Auten and Carroll, 1999; Moffitt and Wilhelm, 2000; Kopczuk, 2005; Liebman and Saez, 2006), the Omnibus Budget Reconciliation Act of 1993 (Sammartino and Wiener, 1997; Carroll, 1998; Goolsbee, 2000b), the Economic Growth and Tax Relief Reconciliation Act (Auten, Carroll, and Gee, 2008; Heim, 2009), and various state and federal tax law changes during the 1980s (Gruber and Saez, 2002). All of the studies other than Lindsey use longitudinal data.

This study builds upon the New Tax Responsiveness Literature by exploiting variation in marginal tax rates due to the Economic Growth and Tax Relief Reconciliation Act (EGTRRA) of 2001. I focus in particular on a marriage penalty relief provision that was scheduled to phase in slowly over time, but was fully implemented by a subsequent tax reform in 2003. This provision lowered the marginal tax rate from 27 percent to 15 percent for couples with taxable income (TI) in 2002 between \$46,700 and \$55,334, but left tax rates below \$46,700 unchanged. The differential change in tax rates is illustrated in Figure 1, which plots marginal tax rates in 2002 and 2004 against adjusted gross income (AGI) in 2002. The AGI range associated with the greater decline in tax rates lies between \$65,000 and \$75,000 — the AGI range is greater than the TI range since TI is equal to AGI minus personal exemptions and deductions. Using a methodology similar to existing tax studies, I exploit this differential change in marginal tax rates to identify the effect of taxes on earnings.

An important contribution of this study is that the elasticity of earnings is identified for couples near the center of the income distribution. As Saez (2001) shows, optimal taxation depends not only on the elasticity of income to taxation, noted above, but also



on the density of the earnings distribution at a given level of income: the greater the density, the lower the marginal tax rate. According to administrative tax return data, the former and new tax-bracket thresholds span 8.6 percent of married couples.<sup>2</sup> In contrast, most existing tax studies focus on very high earners, because the largest tax changes in absolute value have occurred among the very wealthy (Saez, Slemrod, and Giertz, forthcoming).

Another important contribution of this study is that the response of earnings to taxation is examined for couples at the individual level. Studies that use administrative tax return data are limited in this regard because couples generally file and report income jointly. Although the elasticity of joint earnings is a sufficient statistic for estimating the deadweight loss of taxation (Saez, Slemrod, and Giertz, forthcoming), understanding the “anatomy” of the response has important policy implications (Slemrod, 1996). The policy implications follow because the response is not necessarily a behavioral parameter, but is affected by policy itself. For example, broadening the tax base limits tax avoidance behavior, thus limiting the overall response of taxable earnings to tax rates (Kopczuk, 2005). Tax policy can therefore reduce “revenue leakage” by constraining certain behavioral responses, but doing so requires identifying which margins account for the response at the aggregate level.

The main data for this study come from a unique match of administrative data to survey data. The administrative data, termed the Detailed Earnings Record (DER), contain longitudinal information on taxable wage earnings, taxable self-employment earnings and income deferred to 401(k) or similar type accounts. The DER is matched to the well-known March Supplement of the Current Population Survey (CPS). As noted, one advantage of these data over tax return data is that earnings are reported at the individual level. Another advantage is the availability of rich demographic information in the CPS, useful in determining whether treatment and comparison groups are comparable and, if not, in controlling for observable differences between them.<sup>3</sup> A disadvantage of the data, however, is the lack of tax information to define treatment and comparison groups precisely. To address this shortcoming, I use a two-sample strategy inspired by two-sample instrumental variable estimation.

At the joint level, the baseline estimate of the earnings elasticity ranges from 0.22 to 0.30, which is similar in magnitude to those found in previous studies. I then examine the response of earnings to taxation at the individual level. Because the objective is to understand the anatomy of the joint-level response, changes in individual-level earnings are measured to approximate the contribution of the individual-level change to the joint-level elasticity. The analysis reveals that the joint-level response is driven almost

---

<sup>2</sup> This figure, which is derived from the Statistics of Income of 2002, is the (weighted) percentage of couples whose joint taxable earnings fall between \$46,700 and \$54,194. The data are described in further detail below.

<sup>3</sup> Recent sources of tax return data contain more demographic information than earlier sources, including age, gender, and occupation. These data are used in Auten, Carroll, and Gee (2008), Giertz (2007), and Heim (2009).

entirely by changes in the taxable earnings of males. Moreover, the response of male earnings does not reflect self-employment earnings or income shifted from tax-deferred savings accounts, but an increase in overall wage earnings.

## II. BACKGROUND

### A. Source of Identification

The EGTRRA of 2001 was the largest tax cut in 20 years, costing an estimated \$1.35 trillion over the first ten years of its enactment.<sup>4</sup> Many provisions of EGTRRA were designed to phase in slowly over time and be repealed after 2010, the Act's sunset year. However, the subsequent Jobs Growth and Tax Relief Reconciliation Act of 2003 and the Working Family Tax Relief Act of 2004 fully implemented many EGTRRA provisions before their scheduled dates.

These tax laws resulted in numerous changes in federal marginal tax rates. To identify the effect of taxes on reported earnings, I focus on a marriage penalty relief provision that increased the TI threshold between the 15 percent marginal-tax-rate bracket and the 27 percent bracket for married filers.<sup>5</sup> In 2002, the TI threshold for married filers was less than double the threshold for single filers (the thresholds were \$27,950 and \$46,700 for single and joint filers, respectively), generating a marriage penalty for some couples.<sup>6</sup> The marriage provision eliminated this penalty by increasing the TI threshold for married filers to twice the threshold for single filers. The provision was scheduled to phase in over time, becoming fully effective in tax year 2009, but was fully implemented by the subsequent reform in 2003.

The implementation of the marriage penalty provision generated discrete variation in tax rates across the TI distribution. Absent any change in real TI, couples whose 2002 joint TI was roughly between \$46,700 and \$55,334 — the latter is the threshold of \$58,100 in 2004 expressed in 2002 dollars using the Consumer Price Index (CPI) — would have experienced a decline in the federal marginal tax rate from 27 percent to 15 percent by tax year 2004.<sup>7</sup> Couples whose joint TI was just below \$46,700 would have

---

<sup>4</sup> Relative to the size of the economy, the 1981 tax cut was about twice as large as the EGTRRA (Steuerle, 2004).

<sup>5</sup> A marriage penalty occurs when, holding individual incomes constant, a couple's joint tax liability increases upon marriage. Conversely, a marriage subsidy occurs when a couple's tax liability declines upon marriage. Marriage penalties in the U.S. tax code are detailed in Alm, Dickert-Conlin, and Wittington (1999).

<sup>6</sup> For example, if each partner earns \$30,000 in TI prior to marriage, then each partner's tax bill is \$4,746 (reflecting a 15 percent tax rate on TI up to \$27,950, and a 27 percent rate for the remainder); by comparison, upon marriage, the couple is liable for an additional \$1,104 in taxes (reflecting a 15 percent tax rate on TI up to \$46,700 and a 27 percent rate for the remainder, resulting in a total tax liability of \$10,596).

<sup>7</sup> The standard deduction increased in real terms from \$7,850 to \$9,700 (2002 dollars), so, for couples who use the standard deduction, the treatment range is \$46,700 to \$56,722. Tax return statistics from the Internal Revenue Service show that one-third of all taxpayers with adjusted gross income between \$50,000 and \$75,000 use the standard deduction (Parisi and Hollenbeck, 2002).

experienced no change in their marginal rate of 15 percent. I focus on this differential change in tax rates to measure the response of earnings to taxation.

## B. Theoretical Framework

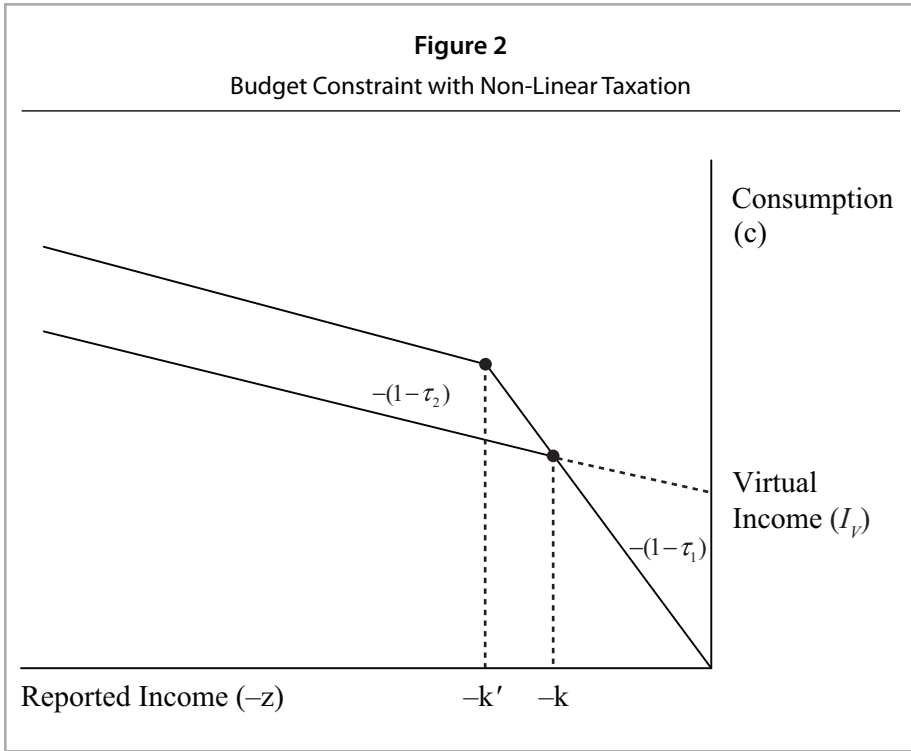
The model constructed by Feldstein (1999) provides a framework for examining the behavioral effects of taxation on reported income. In this model, individuals maximize a utility function that depends positively on consumption ( $c$ ) and negatively on reported income ( $z$ ). In Feldstein's model, reported income, which cannot be evaded, equals total income less exclusions and deductions. The negative association between utility and reported income reflects the costs of income-generating activities — for example, from forgone leisure. Individuals pay taxes on reported income and consume the remainder. If the tax schedule were linear, and if unearned income were untaxed, the budget constraint would be  $c = (1 - \tau)z + I$ , where  $\tau$  is the tax rate and  $I$  is unearned income. At the optimum, reported income can be represented by the function  $z(\tau, I)$ , and the response of reported earnings to a change in the marginal tax price  $(1 - \tau)$  can be characterized by the standard Slutsky equation

$$\frac{\partial z(\tau, I)}{\partial(1 - \tau)} = \left. \frac{\partial z(\tau, I)}{\partial(1 - \tau)} \right|_{u=\bar{u}} + z \frac{\partial z(\tau, I)}{\partial I}.$$

The first term on the right-hand side, the compensated substitution effect, is positive; the second term, the income effect, is negative, assuming  $z$  is an inferior good (e.g., leisure is a normal good). Thus, the net effect of taxes on reported income — the left-hand side of the equation — is theoretically ambiguous.

If the tax schedule is non-linear, however, then the standard Slutsky equation may not accurately reflect the response of reported earnings to taxation. The difference arises because, given a non-linear budget constraint, unearned income may no longer be represented by  $I$ , but rather by “virtual” income  $I_v$ . A simple example in Figure 2 highlights this distinction. As shown by the lower budget constraint, reported income  $z$  is taxed at rate  $\tau_1$  from zero to  $k$ , and rate  $\tau_2$  beyond  $k$ . Despite the resulting kink in the overall budget constraint, the constraint for individuals not at the kink can be represented locally by a linear equation. For example, individuals located on the second segment effectively disregard the first segment, since their reporting decisions do not change if the second segment were extended linearly. For these individuals, the locally linear budget constraint is  $c = (1 - \tau)z + I_v$ , where unearned income  $I_v$  is non-zero. But since unearned income in Figure 2 is in fact zero, income  $I_v$  is considered “virtual.” Unearned income  $I_v$  is now an implicit function of the marginal tax rate  $\tau_2$ , since an increase in the tax rate pivots the locally linear budget constraint counterclockwise at point  $k$ . Thus, optimal reported income becomes  $z(\tau, I(\tau))$ , and the Slutsky equation for the second segment is

$$\frac{\partial z(\tau_2, I)}{\partial(1 - \tau_2)} = \left. \frac{\partial z(\tau_2, I)}{\partial(1 - \tau_2)} \right|_{u=\bar{u}} + z \frac{\partial z(\tau_2, I)}{\partial I} + \frac{\partial z(\tau_2, I)}{\partial I} \frac{\partial I}{\partial(1 - \tau_2)},$$



where the final term is an additional income effect channeled through virtual income. The term  $\frac{\partial I}{\partial(1-\tau_2)}$  is equal to  $-k$  since  $I_v = k[(1-\tau_1) - (1-\tau_2)]$ , so the Slutsky equation reduces to

$$\frac{\partial z(\tau_2, I)}{\partial(1-\tau_2)} = \left. \frac{\partial z(\tau_2, I)}{\partial(1-\tau_2)} \right|_{u=\bar{u}} + (z-k) \frac{\partial z(\tau_2, I)}{\partial I}.$$

As shown by this modified Slutsky equation, the derivative of  $z$  with respect to  $I$  is now multiplied by  $(z-k)$ , which implies that the traditional income effect is counteracted by the virtual income effect. In particular, as income  $z$  approaches  $k$  from the right, the uncompensated response approaches the compensated response.

The potential response of reported earnings to the marriage provision — the source of identifying variation — can be evaluated within this theoretical framework. As described, the provision increased the taxable income threshold between the 15 percent marginal tax bracket and the 27 percent bracket. The effect of this provision on the budget constraint is shown graphically in Figure 2, where the kink is shifted from  $k$  to  $k'$ . Under standard utility assumptions, the tax provision (1) has no effect on reported earnings of individuals located just below the original threshold  $k$ ; (2) has an ambiguous

effect, dictated by the modified Slutsky equation, on earnings of individuals located between  $k$  and  $k'$ ; and (3) has an unambiguously negative effect on earnings of individuals located above the new threshold  $k'$ .<sup>8</sup> These predictions guide the empirical strategy outlined in the next section.

### III. EMPIRICAL STRATEGY

#### A. Methodology

The empirical objective is to estimate the response of earnings to changes in marginal tax rates, identified by the implementation of the marriage penalty relief provision enacted in 2003. Following Feldstein (1995), Saez (2003), and others, the response of earnings to taxation is measured from two groups of tax filers: one group that faces large, predicted changes in tax rates, and another group that does not. Once defined, group assignment can be used as an instrument for changes in tax rates. If the only effect of group assignment on earnings is through differential changes in taxes — the exclusion restriction for instrumental variable estimation — then the relative change in earnings between the two groups can be attributed to the reform.

The marriage penalty relief provision considered in this study yields two obvious groups: couples located just above and below the original tax-bracket threshold. If joint taxable income had remained constant in real terms before and after the reform, couples just below the threshold would have experienced no change in the federal marginal tax rate, whereas couples just above the threshold would have experienced a decline from 27 to 15 percent. By defining  $T$  as an indicator of treatment, equaling one for couples in the treatment group and zero otherwise, the differential change in earnings among those affected by the reform can be estimated from the following reduced-form equation

$$(1) \quad \log\left(\frac{g_{2004}}{g_{2002}}\right) = \alpha_0 + \alpha_1 T + \varepsilon,$$

where  $g_{2002}$  and  $g_{2004}$  are joint earnings in 2002 and 2004, and  $\alpha_1$  represents the differential change in earnings between the two groups. If the exclusion restriction holds,  $\alpha_1$  measures the effect of the reform on reported earnings.

Technically, the exclusion restriction assumes that the treatment indicator and the error term are uncorrelated. Intuitively, the restriction implies that, absent the reform, the differential change in earnings between the two groups would have been zero. Violation of the exclusion restriction may occur, for example, from mean reverting trends in earnings at the lower end of the income distribution or from increasing income inequality at the high end of the income distribution. In the current study, however, treatment and comparison groups have similar incomes and are located at a dense segment of

<sup>8</sup> Similar to previous studies, bunching at kink points is not explicitly considered in the empirical analysis.

the income distribution, reducing the likelihood of differential trends between the two groups (Saez, 2003; Saez, Slemrod, and Giertz, forthcoming). Moreover, systematic differences between treatment and comparison groups — differences that may lead to violating the exclusion restriction — can be accounted for by including controls for observable characteristics in (1).

Welfare loss calculations depend on the compensated elasticity of reported income to the marginal tax price, so for  $\alpha_1$  to be a useful measure of response, it should be expressed as an elasticity. The elasticity can be derived by estimating the following first-stage equation

$$(2) \quad \log\left(\frac{1 - \tau_{2004}}{1 - \tau_{2002}}\right) = \beta_0 + \beta_1 T + v,$$

where  $\tau_{2002}$  and  $\tau_{2004}$  are the marginal tax rates in 2002 and 2004, and the coefficient  $\beta_1$  represents the differential change in marginal tax price due to the provision. The elasticity of earnings to the marginal tax price — equaling the percentage change in earnings relative to the percentage change in the marginal tax price — can be calculated by dividing  $\alpha_1$  by  $\beta_1$ .

The resulting elasticity estimate is the uncompensated response rather than the compensated response, since neither the reduced-form nor first-stage equations contain explicit controls for income. To estimate the compensated elasticity, some studies control for policy-induced changes in after-tax income in both the first-stage and reduced-form regressions (Gruber and Saez, 2002), but this strategy is limited to quasi-experiments involving more than one reform.<sup>9</sup> Given that a single reform is considered in this analysis, the compensated effect could be approximated by narrowly defining the treatment group just above the original kink. As shown in the theoretical framework section, the virtual income effect counteracts the traditional income effect near the threshold, so the response of couples just above the original threshold is closer to the compensated response than the standard Slutsky equation portrays. Thus, to approximate the compensated effect in the empirical analysis below, I define the treatment group to be just above the original kink.

## B. Data

The main data for the analysis come from cross-sectional survey data matched to administrative data on longitudinal earnings. The survey data come from the 2002 and 2003 March Supplements of the CPS, which are nationally representative of the non-institutionalized U.S. population and contain important demographic information such as age, education, race, and family size. The administrative data on longitudinal earnings

<sup>9</sup> The empirical issue is that, with only one reform, policy-induced changes in after-tax income are not independent from policy-induced changes in marginal tax rates, since both are defined by pre-policy income. As a result, the income effect cannot be separately identified from the compensated price effect.

come from the DER, compiled by the Internal Revenue Service (IRS) and maintained by the Social Security Administration. These data contain individual-level information on taxable earnings, deferred earnings and self-employment earnings necessary for computing Social Security benefits.<sup>10</sup> The match of administrative data to survey data does not depend on whether an individual has earnings, but on whether the CPS respondent consents to the match and provides a social security number. I refer to the match of CPS and DER data collectively as the CPS-DER.

The CPS-DER has important advantages over the administrative tax return data often used in tax studies. One advantage is that the CPS-DER contains important demographic information, whereas most administrative data sources do not. Demographic information can be used to control for observable differences between treatment and comparison groups. Another advantage is that the DER reports longitudinal earnings at the individual level, whereas tax return data report earnings of couples jointly. I can therefore examine whether the earnings response at the aggregate level reflects changes in male earnings, female earnings, or both.

Despite these advantages, a major disadvantage of the CPS-DER is that it does not contain the information on TI necessary to define treatment and comparison couples precisely. As an alternative, I define treatment and comparison groups based on the sum of taxable wage earnings and self-employment earnings, referred to as total taxable earnings. To define treatment and comparison groups using this alternative measure, I use another source of administrative data that contains this measure and precise information on TI — the Statistics of Income (SOI) of 2002, which is representative of tax returns filed for the 2002 tax year. Using the SOI, I identify one range of total taxable earnings that contains a significant number of treated couples and another range that does not. The choice of these two ranges implicitly determines the estimate of  $\beta_1$  from the first-stage (2).<sup>11</sup> To estimate the reduced-form (1), the earnings ranges from the SOI define the treatment and comparison groups in the CPS-DER. The effect of treatment-group status on longitudinal earnings is estimated from the DER, controlling for demographic information from the CPS. The coefficient from the reduced-form equation  $\alpha_1$ , estimated from the CPS-DER, relative to the first-stage equation  $\beta_1$ , estimated from the SOI, yields the elasticity of earnings to the marginal tax price.

The use of two samples to derive a single estimate is inspired by two-sample instrumental variable (IV) estimation.<sup>12</sup> The most notable assumption of this strategy — in addition to the exclusion restriction — is that the joint distribution of true treatment

---

<sup>10</sup> Wages and deferred earnings come from employer W-2 forms, and self-employment earnings come from Form 1040s filed with the IRS.

<sup>11</sup> The estimate of  $\beta_1$  measures the differential change in predicted tax rates, so the resulting elasticity estimate is technically reduced-form. In contrast, previous studies often instrument the actual change in tax rates with the predicted change. However, since the correlation between the predicted change and actual change is quite high (Saez, 2003; Gruber and Saez, 2002), the two strategies likely yield similar results.

<sup>12</sup> Two-sample IV estimation was pioneered by Angrist and Krueger (1992). A straightforward implementation of the estimation strategy is to replace the endogenous regressor in the structural equation with the predicted value from the first stage (Dee and Evans, 2003).

and total taxable earnings observed in the SOI is identical to the (unobserved) joint distribution in the CPS-DER. This assumption could be invalid if total taxable earnings are measured with error; however, measurement error is not a major concern in this case since total taxable earnings are calculated from administrative data in both instances. The identification assumption could also be problematic if the SOI and CPS-DER samples represent distinctively different populations. Therefore, the two samples should be comparable to the fullest extent possible.

### C. Data Restrictions

For the SOI and CPS-DER samples to be comparable, any restriction imposed on one sample should be imposed on the other. This imperative, combined with the population of interest, determines the analysis samples in a number of ways. First, the samples are restricted to married couples who, in addition to being the predominate target of the provision, would not have divorced because of the provision, since it increased the incentive for marriage. For the CPS-DER, the restriction is single-family households headed by a married couple, containing only the household heads and their own children. For the SOI, the restriction is joint 1040 returns, filed for the 2002 tax year, that do not claim a non-child exemption. Second, I exclude couples who earn farm income, determined by survey data in the CPS-DER and by Schedule F income in the SOI. Third, to match the geographic coverage of the CPS-DER, I exclude from the SOI tax returns from U.S. territories as well those without state of residence information.

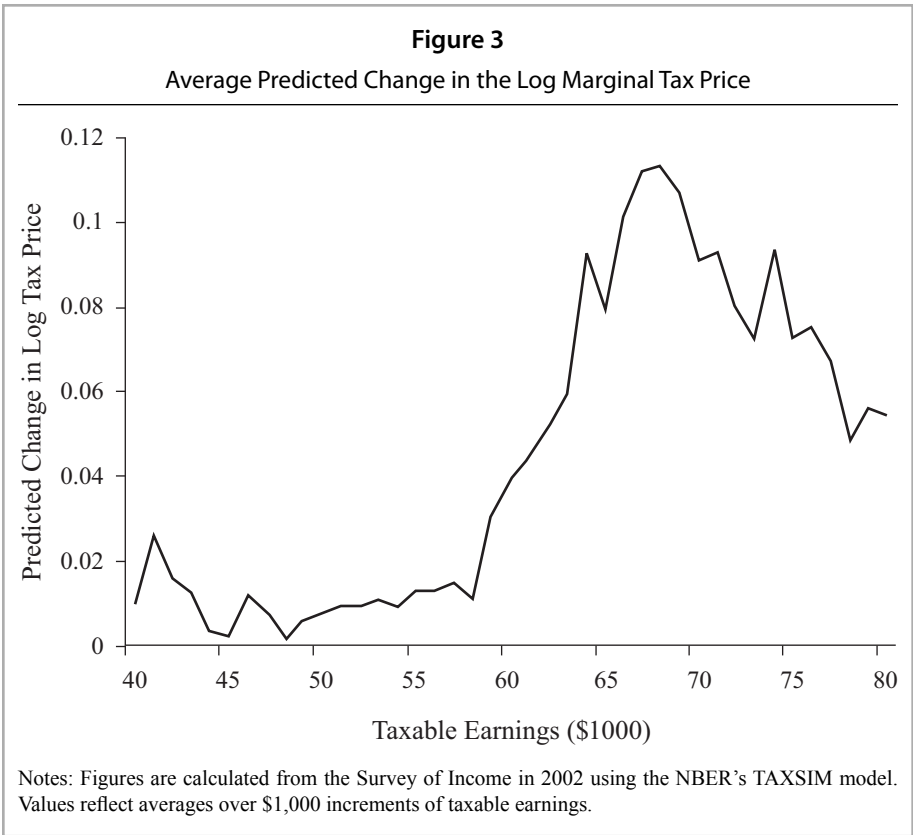
Two additional restrictions were made to one data source that could not be imposed on the other. First, the SOI is restricted to joint filers who either itemized their deductions or claimed the standard deduction — less than 1 percent of the weighted SOI sample report neither. And second, the CPS-DER is necessarily restricted to couples in which both spouses are matched to administrative data — the case for 71.8 percent of the remaining couples.<sup>13</sup>

### D. Analysis Samples

Treatment and comparison groups are defined using precise tax information from the SOI of 2002. To define the groups, I calculate federal marginal tax rates in 2002 and 2004 using the SOI and the TAXSIM model constructed by the National Bureau of Economic Research (NBER) (details are given in the Appendix). Figure 3 plots the predicted change in the marginal tax price — the outcome variable of the first-stage equation — by \$1,000 increments of total taxable earnings.<sup>14</sup> As shown, the predicted

<sup>13</sup> In most cases, both spouses are matched to the DER or both are not; only 6.2 percent of couples have only one spouse matched. Of couples with dependent children, 75.0 percent have both spouses matched to the DER, compared to 66.8 percent among couples without children. The match rate also increases with education.

<sup>14</sup> For example, the average change in the marginal tax price is calculated between \$50,000 and \$51,000 and is reported in Figure 3 at \$50,000.



change ranges from about zero to 0.01 from \$40,000 to \$60,000, but increases sharply thereafter. The increase in the marginal tax price largely reflects the effect of the marriage provision on federal marginal tax rates.

Based on Figure 3, I define the treatment group as couples with total taxable earnings between \$61,000 and \$67,000. As discussed, defining the treatment group just above the threshold yields an elasticity estimate closer to the compensated response than the uncompensated one. An additional empirical concern, and the reason for the narrow range of earnings, is that treatment is defined imprecisely by total taxable earnings. As a result, treatment and comparison groups may be contaminated by couples whose taxable income was previously above the new tax bracket, and who therefore face income effects with negligible substitution effects. In theory, these couples unambiguously reduce earnings, decreasing the overall change in earnings among the groups they contaminate. If the treatment group is more contaminated than the comparison group — a plausible concern given the groups' proximities in the earnings distribution — then the resulting earnings elasticity would be biased downward. The degree of relative contamination,

and therefore the bias, is mitigated by choosing a low upper bound for the range of earnings that defines the treatment group.

According to Figure 3, any range of total taxable earnings below \$60,000 serves as a candidate for the comparison group. I initially define the comparison group as having total taxable earnings between \$47,000 and \$53,000, though the resulting elasticity estimates are robust to different specifications.

## E. Summary Statistics

Summary statistics from SOI data for the two groups are presented in the left panel of Table 1. To present a rough measure of treatment across the two groups, I report the proportion of couples whose TI falls between \$46,700 and \$55,334 — the former and new thresholds in 2002 dollars.<sup>15</sup> As shown, 19.9 percent of treatment couples have taxable incomes that fall within this range, compared to 4.7 percent of comparison couples. The difference in treatment translates into an average change in the predicted marginal tax price of 0.072 and 0.007 for treatment and comparison couples, respectively, implying a first-stage estimate of  $\beta_1$  in (2) of 0.065.

I also calculate the percentage of couples with taxable income above the new threshold of \$55,334; again, these are couples who face income effects with negligible substitution effects. As reported in Table 1, the percentages of these couples in the comparison group and treatment group were 8.9 percent and 13.1 percent, respectively. Since these couples unambiguously reduced earnings, this difference is likely to bias downward the estimated earnings elasticity.

The definition of the treatment and comparison groups from the SOI then determines the groups in the CPS-DER. Summary statistics of the two groups from CPS-DER data are reported in the right panel of Table 1. As shown, the treatment and comparison groups are similar in terms of age and race, but differ by educational attainment. According to Table 1, individuals in the treatment group are approximately 9 percentage points more likely to have attended some form of higher education relative to their same-sex counterparts in the comparison group.

The comparability of the SOI and CPS-DER samples can be established from variables shared by both data sources. Variables shared by the SOI and DER include the proportion of couples who report any self-employment earnings and the average amount of self-employment earnings. As noted above, the self-employment income reported in the DER comes from tax returns filed with the IRS. As shown in Table 1, the percentage of comparison couples with any self-employment earnings differs significantly across samples (13.1 percent versus 16.3 percent for the SOI and CPS-DER, respectively). However, neither this difference for the treatment group, nor the difference in average self-employment earnings across samples for both groups is statistically significant, suggesting that the two samples are indeed comparable along this dimension.

Variables shared by both the SOI and CPS include income from wages, self-employment, social security, interest, dividends, rental property, and the number of dependents.

<sup>15</sup> As mentioned previously, the treatment range with respect to TI is slightly different for couples who take the standard deduction.

**Table 1**  
**Summary Statistics of Treatment and Comparison Groups from SOI and CPS-DER**

	SOI		CPS-DER	
	Comparison	Treatment	Comparison	Treatment
Joint taxable earnings	50,008	63,978	49,989	63,973
\$2002	(49.4)	(50.2)	(30.1)	(32.4)
Treated (percent)	4.71 (0.59)	19.9 (1.2)		
Above treated (percent)	8.9 (0.79)	13.1 (1.0)		
Federal Tax Rate				
Tax year 2002	14.7 (0.19)	19.5 (0.22)		
Tax year 2004	14.2 (0.19)	15.1 (0.14)		
Change in log tax price	0.007 (0.001)	0.072 (0.003)		
Any self employment earnings (percent)	19.6 (1.1)	13.1 (1.0)	18.9 (0.68)	16.3 (0.69)
Self employment earnings	2733 (253)	2364 (286)	2886 (158)	2575 (181)
Age, males			42.4 (0.18)	42.9 (0.18)
Age, females			40.1 (0.18)	40.8 (0.17)
Education, males (percent)			55.1 (0.86)	63.9 (0.90)
Education, females (percent)			56.7 (0.86)	66.2 (0.89)
White, males (percent)			88.3 (0.56)	89.3 (0.58)
White, females (percent)			87.0 (0.58)	88.5 (0.60)
Observations	1,302	1,138	3,311	2,841

Notes: The samples were constructed from the SOI and the CPS matched to DER. The comparison group is defined as joint taxable earnings between \$47,000 and \$53,000; the treatment group is defined as joint taxable earnings between \$61,000 and \$67,000. "Treated" is defined as having joint taxable earnings between \$46,700 and \$55,334. "Above Treated" is defined as having joint taxable earnings greater than \$55,334. "Education" is defined as having some form of higher education. Standard errors are in parentheses. Joint taxable earnings and self-employment earnings reported for the CPS-DER come from DER data. Figures from the SOI reflect sample weights. Standard errors are in parentheses.

Income values may not correspond directly between the SOI and CPS since the former are administrative data and the latter are survey-based data and thus subject to non-classical measurement error.<sup>16</sup> I therefore only estimate and compare the percentage of couples with a given type of income, which may be more reliable than the income values themselves. These figures are reported in Table 2. Although some estimates between the data sources are noticeably different, none are egregiously so. The number of dependent children is also greater in the CPS-DER compared to the SOI, but since dependents are imputed from the CPS data, this difference may reflect imputation bias rather than real differences in the number of dependents across samples.<sup>17</sup>

**Table 2**  
Sources of Income (percent) and Number of Dependents by  
Sample and Treatment Status

	SOI		CPS-DER	
	Comparison	Treatment	Comparison	Treatment
Wages	98.2 (0.37)	99.0 (0.30)	98.3 (0.22)	99.0 (0.19)
Self-employment	19.6 (1.1)	13.1 (1.0)	14.0 (0.60)	11.6 (0.60)
Social Security	8.07 (0.76)	4.00 (0.58)	5.77 (0.41)	3.48 (0.34)
Interest	68.4 (1.29)	69.4 (1.4)	66.0 (0.82)	74.2 (0.82)
Dividend	28.5 (1.25)	31.3 (1.4)	22.3 (0.72)	27.1 (0.83)
Rental	9.95 (0.83)	7.10 (0.76)	7.58 (0.46)	8.31 (0.52)
Number of dependents	1.18 (0.034)	1.14 (0.035)	1.38 (0.021)	1.40 (0.021)

Notes: The samples were constructed from the SOI and the CPS matched to the DER. The comparison group is defined as joint taxable earnings between \$47,000 and \$53,000; the treatment group is defined as joint taxable earnings between \$61,000 and \$67,000. Joint taxable earnings and self-employment earnings reported for the CPS-DER come from CPS data. Standard errors are in parentheses. Figures from the SOI reflect sample weights.

<sup>16</sup> Validation studies on measurement error in survey data include Bound et al. (1990), Bound and Krueger (1991), and Bollinger (1998).

<sup>17</sup> Child dependents in the CPS-DER were imputed based on the characteristics of own children residing in the home. Children are considered dependents if they are less than 19 years old or less than 24 years old and a full time student. This imputation does not reflect dependent children who are nonresidents which, according to the SOI, are rare.

#### IV. BASELINE RESULTS

##### A. Description of Earnings

The effect of the marriage penalty relief provision on earnings is measured from the DER. The top panel of Table 3 contains statistics of joint earnings by treatment status. (Earnings in 2004 are adjusted to 2002 using the CPI.) The first two rows report the

	CPS-DER	
	Comparison	Treatment
<b>Joint Total Taxable Earnings</b>		
Zero earnings 2002 (percent)	0.0	0.0
Zero earnings 2004 (percent)	1.2	0.53
Average earnings 2002	49,989	63,973
Average earnings 2004	50,066	64,723
Change in log of average	0.002	0.012
Average of change in log	-0.0670 (0.0086)	-0.0471 (0.0082)
<b>Male Total Taxable Earnings</b>		
Zero earnings 2002 (percent)	3.0	1.9
Zero earnings 2004 (percent)	7.1	4.2
Average earnings 2002	34,080	42,703
Average earnings 2004	33,519	43,309
Change in log of average	-0.011	0.009
Average of change in log	-0.0660 (0.0077)	-0.0311 (0.0064)
<b>Female Total Taxable Earnings</b>		
Zero earnings 2002 (percent)	18.0	14.3
Zero earnings 2004 (percent)	20.2	16.1
Average earnings 2002	15,909	21,269
Average earnings 2004	16,547	21,414
Change in log of average	0.013	0.002
Average of change in log	-0.0086 (0.0047)	-0.0132 (0.0041)
<b>Total Observations</b>	<b>3,311</b>	<b>2,841</b>

Notes: The sample is constructed from the CPS matched to the DER. The comparison group is defined as joint taxable earnings between \$47,000 and \$53,000; the treatment group is defined as joint taxable earnings between \$61,000 and \$67,000. Standard errors are in parentheses.

percentage of couples with zero joint earnings in 2002 and 2004. By construction, no couple reports zero earnings in 2002; in 2004, 1.2 percent of comparison couples and 0.53 percent of treatment couples report zero joint earnings. The next two rows contain average joint earnings (including zeros). As shown, average joint earnings increased by \$751 among the treatment group and by \$77 among the comparison group — a relative increase of \$674. Thus, in levels, couples in the treatment group, who face relative declines in marginal tax rates, exhibited relative increases in joint earnings.

The relative increase in joint earnings can be quantified alternatively by two approximations to percentage change. The first measure is calculated by taking the log of average earnings in 2002 and 2004 and then differencing these two values. This measure, labeled in the table as *Change in Log of Average*, is similar to the measure reported by Feldstein (1995). The second measure is calculated by taking the log of earnings in 2002 and 2004 for each couple, differencing these two values, and then averaging this difference across couples.<sup>18</sup> This measure, labeled in the table as *Average of Change in Log*, corresponds with the reduced-form (1). As shown, both measures are greater for the treatment group relative to the comparison group (0.012 versus 0.002 and  $-0.047$  versus  $-0.067$  for the first and second measures, respectively). Thus, in both levels and logs, earnings of the treatment group increased more than earnings of the comparison group.

An advantage of the DER is that earnings are reported at the individual level. I can therefore examine whether the relative increase in treatment group earnings reflects those of males or females. Statistics of earnings by sex, similar to those reported for joint earnings, are reported in the second and third panels of Table 3.

I first consider how the marriage provision affects the likelihood of reporting zero earnings at the individual level. On one hand, the marriage provision encourages non-earning spouses to work — from zero to positive earnings — by lowering marginal tax rates. In principle, the increased incentive to work applies to both treatment and comparison groups; for the latter group, the incentive depends on the extent to which additional earnings increase joint earnings into the region of lower marginal tax rates. On the other hand, the marriage provision may discourage work — from positive earnings to zero — through an income effect, which only applies to couples in the treatment group.

According to the results presented in Table 3, approximately 21 percent of comparison couples and 16 percent of treatment couples have a non-earning spouse, most of whom are female. And from 2002–2004, the percentage of zero earnings increased among males and females in both groups. However, the increase in zero earnings was smaller among the treatment group, and the differential increase between treatment and comparison groups was greater among males (1.8 and 0.4 percentage points for males and females, respectively). If these differences are interpreted as causal effects of the provision, then the data suggest that lower marginal tax rates increased employment among treatment males.

I next examine changes in male and female earnings in levels. According to the second panel of Table 3, earnings of treatment males increased by \$606, whereas

<sup>18</sup> This measure is not defined for couples who report zero earnings in 2004, so these couples are dropped from the calculation.

earnings of comparison males decreased by \$561. This amounts to a relative increase of \$1,167 among treatment males. Treatment females, on the other hand, exhibited relative declines in earnings. As reported in the third panel of Table 3, earnings of treatment group females increased by \$145, whereas earnings of comparison group females increased by \$638 — a relative decline of \$493. Thus, the relative increase in joint earnings among the treatment group, reported in the top panel of Table 3, appears to be driven by male earnings.

The two alternative measures of earnings changes — the *Change in Log of Average* and the *Average of Change in Log* — also reveal that the change in joint earnings is driven by male earnings. The first measure, the change in the log of average earnings, is calculated as  $\log((\bar{g}_{2002} + \Delta\bar{g}_l)/\bar{g}_{2002})$ , where  $\bar{g}_{2002}$  is average joint earnings in 2002 and  $\Delta\bar{g}_l$  is the change in average earnings at spousal earnings level  $l$ . The second measure, the average of the change in log earnings, is calculated as the average of  $\log((g_{2002} + \Delta g_l)/g_{2002})$ , where  $g_{2002}$  is joint taxable earnings and  $\Delta g_l$  is the change in earnings at the spousal level. Computing the individual level changes in this manner effectively decomposes the aggregate measure into its individual components, since the sum of the individual measures approximates the joint level measure — they do not sum precisely since the logarithm approximates the percentage change and since the measure is undefined in some cases.<sup>19</sup> Note that the decomposition measure used here, which measures changes in individual earnings relative to baseline joint earnings, differs from the individual earnings elasticity, which measures changes in individual earnings relative to baseline individual earnings. As such, a large component of income will have a larger decomposition measure than a small component of income, even if the two components have the same elasticity. Note also that the decomposition measure is defined even in cases when individual-level income is zero in both years — an elasticity measure, in contrast, is not. As shown in Table 3, both measures reveal a relative increase in earnings among treatment males and a relative decrease in earnings among treatment females.

## B. Regression Results

An important advantage of the CPS-DER is the availability of demographic information useful in determining whether the relative changes in earnings reported in Table 3 are robust to observable differences between treatment and comparison groups. To control for observable differences, I consider regressions with and without demographic controls at each level of earnings: joint, male, and female. The controls include age (added linearly), educational attainment (some form of higher education), race (white), and the number of child exemptions (added linearly).<sup>20</sup>

Regression estimates are presented in Table 4. The three sets of columns — columns one and two, columns three and four, and columns five and six — correspond to

<sup>19</sup> More formally,  $\log\left(\frac{g_{2004}}{g_{2002}}\right) \approx \frac{\Delta g_i}{g_{2002}} + \frac{\Delta g_j}{g_{2002}} \approx \log\left(\frac{g_{2002} + \Delta g_i}{g_{2002}}\right) + \log\left(\frac{g_{2002} + \Delta g_j}{g_{2002}}\right)$ .

<sup>20</sup> State fixed effects are considered auxiliary variables here since they are not controlled for in Gruber and Saez (2002) and Saez (2003). State fixed effects are included in the regression specification in the robustness section below.

**Table 4**  
Regression Estimates of Total Taxable Earnings

	Joint Earnings		Male Earnings		Female Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.0199*	0.0142	0.0349***	0.0309***	-0.0046	-0.0046
	(0.012)	(0.012)	(0.010)	(0.010)	(0.006)	(0.006)
Age, males		-0.0038***		-0.0038***		-0.0003
		(0.001)		(0.001)		(0.001)
Age, females		-0.0026*		-0.0012		-0.0013*
		(0.001)		(0.001)		(0.001)
Education, males		0.046***		0.0514***		-0.0032
		(0.013)		(0.011)		(0.007)
Education, females		0.0558***		0.0259**		0.01
		(0.014)		(0.011)		(0.007)
White, males		-0.0061		0.0061		0.0056
		(0.025)		(0.021)		(0.013)
White, females		-0.0131		-0.03		0.0057
		(0.025)		(0.019)		(0.013)
Number of child Exemptions		0.0187***		0.0015		0.0137***
		(0.005)		(0.005)		(0.003)
Tax-adjusted Estimates	0.304*	0.217	0.535***	0.473***	-0.071	-0.070
	(0.182)	(0.184)	(0.153)	(0.159)	(0.092)	(0.095)
Observations	6,097	6,097	6,114	6,114	6,148	6,148
R-Squared	0.0005	0.0332	0.0019	0.0248	0.0001	0.0128

Notes: The sample is constructed from the CPS matched to the DER. The comparison group is defined as joint taxable earnings between \$47,000 and \$53,000; the treatment group is defined as joint taxable earnings between \$61,000 and \$67,000. Robust standard errors are in parentheses. Asterisks denote statistical significance at the 10 percent (\*), 5 percent (\*\*), and 1 percent (\*\*\*) levels.

regressions with joint, male and female earnings as the outcome variable, respectively. Within each column set, the first column excludes the demographic controls; the second includes them. The relative change in earnings among the treatment group is measured by the variable labeled *Treatment*.

At the joint level, the relative change in earnings without controls is 0.0199. The estimate, which is identical to the difference in the *Average of Change in Log* reported in Table 3, is statistically significant at the 10 percent level. The inclusion of control variables, however, reduces the magnitude of the estimate and renders it statistically insignificant. The estimate, presented in column two, is 0.0142. As shown, age is associated with lower earnings growth; education and the number of exemptions are associated

with greater earnings growth. When the controls are included in the regression equation singly, education appears to decrease the elasticity estimate, and age appears to increase it.<sup>21</sup> Controls for race and the number of child exemptions have no notable effects.

At the individual level, the outcome variable is  $\log((g_{2002} + \Delta g_i)/g_{2002})$ , so the results correspond with the *Average of Change in Log* reported in Table 3. Without control variables, the relative increase among treatment males is 0.0349 and statistically significant at the 1 percent level. The inclusion of control variables decreases the magnitude of the estimate to 0.0309, but the estimate remains statistically significant. As shown, the education level of each spouse independently affects the growth of male earnings.

The fifth and sixth columns of estimates contain regression results of female earnings. Both estimates suggest a relative decline in earnings among treatment females, but both are statistically insignificant. Moreover, the inclusion of control variables has no notable effect.

To compare the magnitudes of these estimates with previous studies, I divide the estimates by the differential (predicted) change in the marginal tax price reported in Table 1. The resulting figures are reported in a separate row in Table 4 labeled *Tax-adjusted Estimates*. At the joint level, the resulting tax-adjusted figure is the joint-level elasticity; at the individual levels, the resulting figure is the approximate contribution of each earnings level to the joint elasticity. As shown, the elasticity of joint earnings with and without demographic controls is 0.304 and 0.217, which fall within the range of those from previous studies.<sup>22</sup> As expected, the tax-adjusted estimates are larger for males.

## V. ADDITIONAL CONSIDERATIONS

### A. Robustness

The baseline results suggest that the aggregate earnings response is driven mainly by the earnings of males. In this section, I examine the robustness of the baseline estimates for males to different empirical specifications. (The substantive results at the joint and female levels remain unchanged.) Sensitivity tests include redefining the earnings range of the sample, including additional control variables, dropping Social Security beneficiaries, and allowing the error terms of the earnings equations to be correlated across spouses. Estimates from these sensitivity tests are reported in Table 5.

The top panel of the table reports robustness results from various specifications of the earnings ranges. The first row contains the baseline, tax-adjusted estimate of 0.535.

<sup>21</sup> When only education controls are included in (2), the coefficient decreases from 0.0199 to 0.0097 (robust standard error: 0.012). When only age controls are included in (2), the coefficient increases to 0.024 (robust standard error: 0.012).

<sup>22</sup> Most other studies, particularly those using administrative data, measure the elasticity of taxable income to taxation, not taxable earnings. Since taxable income reflects more margins of behavioral response than taxable earnings — and has been shown to be more sensitive to taxation than broader measures of income (Gruber and Saez, 2002) — the elasticity presented here likely underestimates the compensated elasticity of taxable income.

**Table 5**  
Robustness Analysis of Male Earnings Elasticity

Earnings Range	Control Variables	Errors	Observations	Tax-adjusted Estimate
<i>A. Change Earnings Ranges</i>				
Baseline range	None	Robust	6,114	0.535*** (0.153)
Comparison: \$47,000–\$58,000	None	Robust	8,776	0.419*** (0.136)
Comparison: \$42,000–\$53,000	None	Robust	8,720	0.528*** (0.136)
Comparison: \$42,000–\$58,000	None	Robust	11,382	0.450*** (0.128)
Treatment: \$61,000–\$71,000	None	Robust	7,903	0.353*** (0.118)
<i>B. Include Additional Controls</i>				
Baseline range	Baseline Controls	Robust	6,114	0.473*** (0.159)
Baseline range	Expanded Controls	Robust	6,114	0.341** (0.159)
<i>C. Drop Social Security Beneficiaries</i>				
Baseline range	Baseline Controls	Robust	5,789	0.348** (0.147)
<i>D. Use SUR Model</i>				
Baseline range	Baseline Controls	SUR	5,645	0.493*** (0.155)

Notes: Baseline controls include age (linearly), education (some college or more) and race (white) for both spouses, and the number of child exemptions. Expanded controls include a retirement age indicator (age 55 or older) and major-industry fixed effects for both spouses, an indicator of homeownership, and state-level fixed effects. SUR stands for seemingly unrelated regression. Asterisks denote statistical significance at the 10 percent (\*), 5 percent (\*\*), and 1 percent (\*\*\*) levels.

In the next three rows, I change the upper and lower bounds of the comparison-group range by \$5,000 increments. As indicated, the tax-adjusted estimate is robust to these comparison-group specifications. In the final row of the panel, I increase the upper bound of the treatment-group range by \$5,000, which, as shown, decreases the estimate. As discussed earlier, the decline in the estimate may reflect that the traditional income effect dominates the virtual income effect among treatment couples further from the original

threshold, and that couples previously above new threshold, who face income effects with little substitution effects, may contaminate the treatment-group range. Indeed, the percentage of couples previously above the new threshold increases to 23.1 percent at \$71,000 in joint taxable earnings, and to 41.0 percent at \$75,000. Since income effects cannot be controlled separately, adding these couples may be biasing downward the tax-adjusted price estimate.

In the second panel, I use a richer set of control variables than those used in the baseline regression. The additional controls include a retirement age indicator for both the male and female, equaling one for individuals aged 55 and older and zero otherwise, an indicator of homeownership, state fixed effects, and major-industry fixed effects for both the male and female.<sup>23</sup> The age indicators control for the effect of retirement on earnings, and the state and industry effects control for region- and industry-specific trends in earnings. Adding these additional controls decreases the baseline estimate of 0.473 to 0.341, but the estimate remains statistically significant. Much of the change in the estimate is accounted for by the retirement age indicator and the indicator of homeownership.<sup>24</sup>

In the third panel, the CPS-DER and the SOI samples are restricted to couples who do not receive income from Social Security. As shown, the estimate is 0.348, significant at the 5 percent confidence level.

In the final panel, I allow the errors of the earnings equations to be correlated across spouses using seemingly unrelated regression (SUR). Because SUR requires the regression equations across spouses to differ by at least one variable, I model each spouse's earnings equation containing only the age, education level, and race of oneself. As shown, the estimate for males is robust.

## B. Mechanism

The finding that male earnings account for much of the change in total earnings may seem counterintuitive. Indeed, a review of the labor supply literature suggests that secondary earners, who are predominately female, are more sensitive to taxes than primary earners, who are predominately male. However, the current analysis differs from previous studies in two important ways. First, the individual-level response is measured by its contribution to the joint-level elasticity, not by the individual-level elasticities. As noted

---

<sup>23</sup> Housing status is defined by four categories: (1) not in the universe, (2) owned or being bought, (3) rent, and (4) no cash rent. Homeownership refers to the second category. To code the major industry of the respondent's employment, the CPS uses 23 industry categories plus a "not in the universe" category.

<sup>24</sup> The inclusion of the retirement age indicators and home ownership indicator, but not industry fixed effects, reduces the estimated elasticity to 0.384. The estimates declines further to 0.341 by adding industry fixed effects. As stated in the text, the additional decline may reflect differential trends in earnings across industries, but may also reflect differential treatment effects across industries (Auten and Carroll, 1999). It is possible to estimate heterogeneous effects from the CPS-DER sample, but interpreting them would require estimating heterogeneous first-stage estimates from the SOI. Unfortunately, this is not possible since the SOI does not contain industry information.

above, a larger individual elasticity does not necessarily imply a greater contribution to the joint-level elasticity. Second, and more importantly, the outcome variable of interest is not labor supply, but earnings. As Feldstein (1995) and others note, labor supply affects wage earnings, but so do work intensity, types of compensation, and tax evasion and avoidance behavior. Thus, an inelastic labor-supply response does not necessarily imply an inelastic earnings response.

In regard to the second point, one possible explanation for the different results by sex is that males can respond to taxation on more margins than females. Indeed, compared to females, males are more likely to have self-employment earnings, more likely to defer income into 401(k) plans or similar accounts, and more likely to provide health insurance coverage to the family.<sup>25</sup> This explanation is consistent with the finding that the response of income to taxation increases as the definition of income narrows, reflecting that narrower definitions of income imply more margins for behavioral responses (Saez, 2003; Gruber and Saez, 2002).

To explore this possible explanation for the results, I use CPS-DER data to examine which margins the response of male earnings reflects. I first examine which of the two components of total taxable earnings — taxable wage earnings or taxable self-employment earnings — account for the earnings response.<sup>26</sup> To do so, I estimate a set of regressions with the outcome variable defined by both spouse and income type. More precisely, the outcome variable is defined as  $\log((g_{2002} + \Delta g_{ld})/g_{2002})$ , where  $d$  indicates earnings type and  $l$  indicates spousal earnings level when applicable.

Regression estimates, similar to those for total taxable earnings in Table 4, are presented for wage earnings and self-employment earnings in Table 6 and Table 7, respectively. As shown, the only notable estimates are those for wage earnings among males.<sup>27</sup> With control variables, the estimate is 0.0394 and statistically significant at the 10 percent level of confidence. Relative to the differential change in tax rates, the estimate implies a tax-adjusted estimate of 0.604. All other estimates are relatively small and statistically insignificant.

I then examine what accounts for the increase in taxable earnings. Taxable earnings are defined as total wage earnings and self-employment earnings minus income deferred to 401(k) plans or similar accounts. Thus, the relative increase in earnings may reflect either an increase in overall compensation in taxable wage form, or a shift of income from a tax-deferred to a taxable wage form. To differentiate these two mechanisms, I

<sup>25</sup> Among treatment males, the average (including zeros) self-employment and deferred income was \$1,960 and \$1,512, respectively. Among treatment females, the respective averages were \$616 and \$656. Additionally, 74.7 percent of treatment males are covered by private health insurance, compared to 43.2 percent of treatment females.

<sup>26</sup> LaLumia (2009) measures the response of self-employment earnings to the Earned Income Tax Credit.

<sup>27</sup> Self-employment earnings represent a small fraction of total taxable earnings, so they would have to be extremely responsive to taxation to fully account for the aggregate response. To address this concern, I report in the Appendix Table the percent of observations with zero self-employment earnings, as well as the average self-employment earnings, in 2002 and 2004. As shown, neither measure changed significantly in levels among the treatment group relative to the comparison group, supporting the interpretation that male wages, not self-employment earnings, are responsible for the overall earnings response.

	Joint Earnings		Male Earnings		Female Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.0063 (0.013)	0.0001 (0.014)	0.0408* (0.022)	0.0394* (0.022)	-0.0006 (0.009)	0.0008 (0.009)
Age, males		-0.0028** (0.001)		-0.0042** (0.002)		0.0002 (0.001)
Age, females		-0.0026* (0.001)		-0.0037* (0.002)		-0.0017** (0.001)
Education, males		0.0531*** (0.014)		0.0017 (0.026)		-0.0103 (0.008)
Education, females		0.0511*** (0.015)		0.0646** (0.029)		0.003 (0.008)
White, males		-0.0059 (0.025)		0.0248 (0.032)		0.0055 (0.014)
White, females		-0.0213 (0.024)		-0.0409* (0.025)		0.002 (0.014)
Number of child Exemptions		0.0188*** (0.006)		0.0068 (0.009)		0.0147*** (0.003)
Tax-adjusted Estimates	0.097 (0.200)	0.002 (0.209)	0.625* (0.336)	0.604* (0.333)	-0.010 (0.133)	0.0121 (0.143)
Observations	6,096	6,096	6,123	6,123	6,149	6,149
R-squared	0	0.0236	0.0006	0.0109	0	0.0058

Notes: The sample is constructed from the CPS matched to the DER. The comparison group is defined as joint taxable earnings between \$47,000 and \$53,000; the treatment group is defined as joint taxable earnings between \$61,000 and \$67,000. Education is defined as some form of higher education. Robust standard errors are in parentheses. Asterisks denote statistical significance at the 10 percent (\*), 5 percent (\*\*), and 1 percent (\*\*\*) levels.

estimate a set of regressions with the outcome variable as total earnings: the sum of taxable wage earnings, taxable self-employment earnings, and deferred income. If the increase in taxable earnings among males reflects income shifted from a tax-deferred to taxable form, then there should be no differential increase in total earnings between treatment and comparison groups; if instead the increase reflects an increase in overall compensation, then total earnings should increase among the treatment group relative to the comparison group.

For males, regression estimates for total earnings (not shown) are similar in magnitude to those in Table 4, suggesting that the response of taxable earnings does not reflect a shift in income, but an increase in overall compensation. (For females, the response of

**Table 7**  
**Regression Estimates of Taxable Self-Employment Earnings**

	Joint Earnings		Male Earnings		Female Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.0065 (0.005)	0.0065 (0.005)	0.0032 (0.005)	0.0032 (0.005)	0.0019 (0.002)	0.0017 (0.002)
Age, males		-0.0005 (0.001)		-0.0001 (0.001)		-0.0003** (0.0001)
Age, females		-0.0001 (0.001)		-0.0005 (0.001)		0.0003 (0.0002)
Education, males		-0.0001 (0.004)		0.0042 (0.004)		0.0021 (0.002)
Education, females		0.0026 (0.004)		0.0001 (0.004)		-0.0002 (0.002)
White, males		-0.0156 (0.012)		-0.0073 (0.009)		-0.0004 (0.003)
White, females		0.0138 (0.013)		0.0094 (0.010)		0.0021 (0.004)
Number of child Exemptions		-0.0008 (0.002)		-0.0008 (0.002)		-0.0006 (0.001)
Tax-adjusted Estimates	0.100 (0.069)	0.100 (0.070)	0.049 (0.072)	0.049 (0.071)	0.029 (0.027)	0.026 (0.028)
Observations	6,140	6,140	6,146	6,146	6,152	6,152
R-squared	0.0003	0.0018	0.0001	0.0016	0.0002	0.0012

Notes: The sample is constructed from the CPS matched to the DER. The comparison group is defined as joint taxable earnings between \$47,000 and \$53,000; the treatment group is defined as joint taxable earnings between \$61,000 and \$67,000. Education is defined as some form of higher education. Robust standard errors are in parentheses. Asterisks denote statistical significance at the 10 percent (\*), 5 percent (\*\*), and 1 percent (\*\*\*) levels.

total earnings is negligible and statistically insignificant.) An alternative interpretation is that deferred income is small relative to total taxable earnings, so deferred income would have to be extremely sensitive to taxation to fully account for the joint-level response. To rule out this interpretation, I report in appendix Table A1 the percent of observations with zero deferred income, as well as average deferred earnings, in 2002 and 2004. As indicated, neither measure changed significantly in levels among the treatment group relative to the comparison group. Thus, the response of male total taxable earnings does not reflect a shift in income from a tax-deferred to taxable form.

## VI. DISCUSSION AND CONCLUSION

Similar to previous studies, this paper identifies the response of earnings to taxation using policy-induced changes in marginal tax rates. The use of survey data matched to administrative data on longitudinal earnings is a unique aspect of the paper. In contrast to data used in previous studies, these data allow the analysis to examine the response of earnings for males and females separately and to control for demographic differences between treatment and comparison groups.

At the joint level, the estimated elasticity of earnings to taxation is comparable to those found in previous studies. With baseline control variables, the elasticity estimate is 0.217. As noted earlier, most joint elasticity estimates range from zero to 0.8. Previous research also suggests that itemizers are more responsive than non-itemizers (Gruber and Saez, 2002). Unfortunately, the CPS-DER does not report itemization status, so it is difficult to estimate the elasticity separately along this dimension, but for the sake of comparison, 74.2 percent of treatment couples in the sample whose taxable income falls between \$46,700 and \$55,334 itemize their deductions.

This study draws two main conclusions. First, the joint-level response appears to be driven by the wage earnings of males, suggesting that the change in earnings among couples at the joint level appears to obscure larger behavioral responses at the individual level. The response among males may reflect that they have more behavioral margins on which to respond, but more research is clearly warranted. And second, the analysis sheds light on whether the response of earnings to taxation reflects a change in real earnings behavior or simply a shift in earnings from a non-taxable to taxable form. According to the results, the rise in taxable earnings among males appears driven by a real change in wage earnings.

## ACKNOWLEDGEMENTS

I would like to thank the members of my dissertation committee, Mark Duggan, William Evans, Jonah Gelbach, and Melissa Kearney. I am also grateful for helpful comments and support from Daniel Feenberg, Eric French, William Gentry, Susan Grad, Judy Hellerstein, Jeffrey Kubik, Michael Lovenheim, Joyce Manchester, Giovanni Mastrobuoni, David Pattison, Jae Song, Till von Wachter, several seminar and conference participants, and two anonymous referees. I thank Pat Cole and Joel Packman for data assistance. All errors are my own.

## REFERENCES

- Alm, James, Stacy Dickert-Conlin, and Leslie A. Whittington, 1999. "Policy Watch: The Marriage Penalty." *Journal of Economic Perspectives* 13 (3), 193–204.
- Angrist, Joshua D., and Alan B. Krueger, 1992. "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples." *Journal of the American Statistical Association* 87 (418), 328–336.

- Auten, Gerald, and Robert Carroll, 1999. "The Effect of Income Taxes on Household Income." *The Review of Economics and Statistics* 81 (4), 681–693.
- Auten, Gerald, Robert Carroll, and Geoffrey Gee, 2008. "The 2001 and 2003 Tax Rate Reductions: An Overview and Estimate of the Taxable Income Increase." *National Tax Journal* 61 (3), 345–364.
- Bollinger, Christopher R., 1998. "Measurement Error in the Current Population Survey: A Non-parametric Look." *Journal of Labor Economics* 16 (3), 576–594.
- Bound, John, Charles Brown, Greg J. Duncan, and Willard L. Rodgers, 1990. "Measurement Error in Cross-Sectional and Longitudinal Labor Market Surveys: Validation Study Evidence." In Hartog, Joop, Geert Ridder, and Jules Theeuwes (eds.), *Panel Data and Labor Market Studies*, 1–19. Elsevier, New York, NY.
- Bound, John, and Alan B. Krueger, 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9 (1), 1–24.
- Carroll, Robert, 1998. "Tax Rates, Taxpayer Behavior, and the 1993 Tax Act." Unpublished manuscript. U.S. Department of the Treasury, Office of Tax Analysis, Washington, DC.
- Dee, Thomas S., and William N. Evans, 2003. "Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates." *Journal of Labor Economics* 21 (1), 178–209.
- Feenberg, Daniel, and Elizabeth Coutts, 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12 (1), 189–194.
- Feldstein, Martin, 1995. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act." *Journal of Political Economy* 103 (3), 551–572.
- Feldstein, Martin, 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics* 81 (4), 674–680.
- Giertz, Seth H., 2007. "The Elasticity of Taxable Income over the 1980s and 1990s." *National Tax Journal* 60 (4), 743–768.
- Goolsbee, Austan, 2000a. "It's Not about the Money: Why Natural Experiments Don't Work on the Rich." In Slemrod, Joel (ed.), *Does Atlas Shrug? The Economic Consequence of Taxing the Rich*, 141–164. Cambridge University Press, Cambridge, MA.
- Goolsbee, Austan, 2000b. "What Happens when You Tax the Rich? Evidence from Executive Compensation." *Journal of Political Economy* 108 (2), 352–378.
- Gruber, Jon, and Emmanuel Saez, 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics* 84 (1), 1–32.
- Heim, Bradley T., 2009. "The Effect of Recent Tax Changes on Taxable Income: Evidence from a New Panel of Tax Returns." *Journal of Policy Analysis and Management* 28 (1), 147–163.

- Kopczuk, Wojciech, 2005. "Tax Bases, Tax Rates and the Elasticity of Reported Income." *Journal of Public Economics* 89 (11–12), 2093–2119.
- LaLumia, Sara, 2009. "The Earned Income Tax Credit and Reported Self-Employment Income." *National Tax Journal* 62 (2), 191–217.
- Liebman, Jeffrey, and Emmanuel Saez, 2006. "Earning Responses to Increases in Payroll Taxes." NBER Working Paper. National Bureau of Economic Research, Cambridge, MA.
- Lindsey, Lawrence, 1987. "Individual Taxpayer Response to Tax Cuts: 1982–1984, with Implications for the Revenue Maximizing Tax Rate." *Journal of Public Economics* 33 (2), 173–206.
- Moffitt, Robert A., and Mark O. Wilhelm, 2000. "Taxation and the Labor Supply Decisions of the Affluent." In Slemrod, Joel (ed.), *Does Atlas Shrug? The Economic Consequence of Taxing the Rich*, 193–239. Cambridge University Press, Cambridge, MA.
- Navratil, John, 1995. "The Economic Recovery Tax Act of 1981: Evidence on Individual Taxpayer Behavior from Panel Tax-Return Data." Unpublished manuscript. Harvard University, Cambridge, MA.
- Parisi, Michael, and Scott Hollenbeck, 2002. "Individual Income Tax Returns, 2002." *Bulletin Article: Statistics of Income Tax Stats*. Internal Revenue Service.
- Saez, Emmanuel, 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68 (1), 205–229.
- Saez, Emmanuel, 2003. "The Effect of Marginal Tax Rates on Income: A Panel Study of 'Bracket Creep.'" *Journal of Public Economics* 87 (5–6), 1231–1258.
- Saez, Emmanuel, Joel Slemrod, and Seth Giertz, forthcoming. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature*.
- Sammartino, Frank, and David Wiener, 1997. "Recent Evidence on Taxpayers' Response to the Rate Increase in the 1990s." *National Tax Journal* 50 (3), 683–705.
- Slemrod, Joel, 1996. "High Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response." In Feldstein, Martin, and James Poterba (eds.), *Empirical Foundations of Household Taxation*, 169–192. University of Chicago, Chicago, IL.
- Steuerle, C. Eugene, 2004. *Contemporary U.S. Tax Policy*. Urban Institute Press, Washington, DC.

**APPENDIX: SOI DATA AND THE NBER TAXSIM MODEL**

I use SOI data from 2002 to measure the relationship between treatment and total taxable earnings, defined as taxable wage earnings and taxable self-employment earnings. The relationship is then used to define treatment and comparison groups in the CPS-DER.

The Social Security Administration uses the DER to compute benefits. Therefore, the DER only reports 92.35 percent of positive self-employment earnings, the relevant portion of self-employment earnings subject to payroll taxation. To derive a comparable measure from the SOI, I factor self-employment income reported on Schedule C by 0.9235, and define total taxable earnings as the maximum between this value and zero.

I then use NBER TAXSIM model to compute marginal tax rates in 2002 and 2004 for the SOI. TAXSIM calculates marginal federal, state, and payroll tax rates based on 21 input parameters — Feenberg and Coutts (1993) provide a detailed description of the TAXSIM model. To calculate tax rates in 2004 using 2002 tax-return data, I inflate all monetary input variables using the CPI. Due to lack of information in the SOI, six of the 21 TAXSIM parameters were not used: number of filers aged 65 and older, secondary wages separately, non-taxable transfer income, rent paid, property taxes paid, and other property taxes paid.

**Table A1**  
Self-Employment and Deferred Earnings from the CPS-DER

	Self-Employment Earnings		Deferred Earnings	
	Comparison	Treatment	Comparison	Treatment
<b>Joint Earnings</b>				
Zero earnings 2002 (percent)	81.1 (0.68)	83.7 (0.69)	45.7 (0.87)	34.6 (0.89)
Zero earnings 2004 (percent)	82.0 (0.67)	84.1 (0.69)	47.0 (0.87)	36.6 (0.90)
Average earnings 2002	2,885.6 (158)	2,575.3 (181)	1,301.5 (35)	2,168.0 (52)
Average earnings 2004	2,683.0 (170)	2,566.9 (200)	1,366.2 (41)	2,269.2 (61)
<b>Male Earnings</b>				
Zero earnings 2002 (percent)	87.5 (0.57)	89.4 (0.58)	61.3 (0.85)	51.7 (0.94)
Zero earnings 2004 (percent)	87.8 (0.57)	89.9 (0.57)	62.1 (0.84)	53.5 (0.94)
Average earnings 2002	2,302.0 (145)	1,959.5 (162)	910.4 (30)	1,511.6 (44)
Average earnings 2004	2,077.6 (153)	1,797.6 (174)	941.7 (33)	1,558.9 (51)
<b>Female Earnings</b>				
Zero earnings 2002 (percent)	92.1 (0.47)	93.1 (0.48)	76.6 (0.74)	67.7 (0.88)
Zero earnings 2004 (percent)	92.9 (0.45)	92.7 (0.49)	76.8 (0.73)	67.8 (0.88)
Average earnings 2002	583.6 (57)	615.8 (69)	391.1 (21)	656.3 (29)
Average earnings 2004	605.5 (71)	769.3 (83)	424.5 (25)	710.3 (32)

Notes: The sample is constructed from the CPS matched to the DER. The comparison group is defined as joint taxable earnings between \$47,000 and \$53,000; the treatment group is defined as joint taxable earnings between \$61,000 and \$67,000. Standard errors are in parentheses.

