THE EFFECT OF INCOME TAXES ON HOUSEHOLD INCOME

by

Gerald Auten and Robert Carroll
U.S. Department of the Treasury

OTA Working Paper 75 March 1998

Office of Tax Analysis
U.S. Treasury Department
Washington, DC 20220

We thank Jim Cilke and Gordon Wilson for their help in the development of the data and tax calculators, and Pete Sailer for his help in the development of the occupation codes. We are grateful to Len Burman, Martin Feldstein, Jonathan Gruber, Douglas Holtz-Eakin, David Joulfaian, C. Erik Larson, Karl Scholz, Joel Slemrod, Paul Smith, Jerry Tempalski, Eric Toder, and participants of the 1994 NBER Summer Institute, the 1995 Tax Policy Research Symposium sponsored by the University of Michigan's Office of Tax Policy Research, and the Harvard-MIT Public Finance Seminar for helpful comments.

ABSTRACT

Virtually all research on income and wealth distributions during the 1980s has identified a trend towards increasing inequality. Some of this research indicates that the increasing inequality during the 1980s is the continuation and acceleration of trends spanning several decades. This paper explores to what extent behavioral responses to the tax changes during the 1980s may also explain the rising inequality. The 1986 Tax Reform Act is used as a natural experiment to explore the roles played by both taxes and a variety of non-tax factors, including changing returns to education. Our principal finding is that both tax rates and non-tax factors appear to have had substantial effects on relative income growth during the late 1980s.

THE EFFECT OF INCOME TAXES ON HOUSEHOLD INCOME

I. Introduction

Although income inequality increased dramatically during the 1980s, the causes of the rise in income inequality during this period are unclear. Two views have emerged as possible explanations of this rise in income inequality. One view is that it reflects the continuation and acceleration of trends spanning several decades. According to this view, a variety of economic factors having little to do with taxes explain the trend towards higher income inequality. For example, Murphy and Welch (1992) and Katz and Murphy (1992) found that the returns to human capital or education increased as a result of increased demand for labor with more education and skill. Other explanations include technological change (Bound and Johnson, 1992), declining union membership (Freeman, 1993), increasing import competition, lessening of government pressure to increase minority employment (Bound and Johnson, 1992), increasing immigration (Topel, 1994), changing supplies of college educated workers (Katz and Murphy, 1992), and continuation of the unwinding of wage compression during World War II (Goldin and Margo 1992). In general, this research has not considered the role played by taxes.

Another view suggests that the tax rate reductions during the 1980s played an integral role in the rise in income inequality during the 1980s.[1] According to this view, increases in reported incomes among high income households were, in large part, the result of tax-induced behavioral responses, rather than a more fundamental change in the shape of the income distribution (see, for example, Lindsey (1987) and Feldstein (1995a)). These behavioral responses may have taken many different forms including increased labor supply and

---

[1]In the Economic Recovery Tax Act of 1981 (ERTA), the top individual income tax rate was reduced from 70 percent to 50 percent. In the Tax Reform Act of 1986 (TRA), Congress broadened the tax base and further reduced the top statutory tax rate from 50 percent to 28 percent.

participation, increased savings, altering the timing of income, changes in the form of compensation, rearranging of portfolios, reduced tax evasion and avoidance, and changes in taxpayer decisions about deductions (e.g., how much to donate to charity and borrow for housing).[2] This literature, however, has generally incorporated only limited controls for nontax factors that may have affected income changes.

Evaluating the effects of the tax changes of the 1980s and any appropriate policy response to increased inequality will depend on the extent to which these views can explain the increase in income inequality during the 1980s. If taxpayers reduce their reported incomes in response to higher tax rates, policies that increase tax rates to increase the progressivity of the tax system and reduce income inequality may be less effective. In addition, to the extent households report higher incomes due to lower tax rates, the U.S. Treasury would receive more revenue than otherwise. Moreover, if taxpayers are very responsive to changes in tax rates and go to great lengths to rearrange the way they receive and earn their income, the welfare cost of higher tax rates could be large compared to tax revenues collected (Feldstein, 1995b). Indeed, many of these behavioral responses might be viewed as favorable outcomes of tax rate reductions to the extent that they might also be associated with a smaller welfare loss.

This paper examines the responsiveness of pre-tax income (net of capital gains) to changes in tax rates and a variety of non-tax factors, including proxies for a taxpayer's human capital or education, during the late 1980s. By including both taxes and non-tax factors, we are able to incorporate both of these views and evaluate the extent to which changes in tax rates can

---

[2]Feldstein (1995a) and Slemrod (1994) discuss the different types of tax-induced changes in income reported for tax purposes.

explain the changes in income during the 1980s.

We use a panel of taxpayers constructed from the Statistics of Income (SOI) Individual Tax Files before and after the Tax Reform Act of 1986 (TRA). The use of panel data allows us to control for unobserved heterogeneity by modeling individual-specific and time-specific effects. The use of tax return data allows us to measure tax rates accurately. Detailed occupation information allows us to develop a proxy of a taxpayer's human capital based on the educational content of a taxpayer's job. Information on each taxpayer's age from a match with Social Security records allows us to account for life cycle effects. Finally, unlike most previous studies of income changes, our sample of tax returns contains many high income taxpayers.

In many respects, TRA provides a natural experiment to test whether tax rates or non-tax factors, or both, explain changes in income during this period. TRA lowered tax rates for most taxpayers, but high income taxpayers had the largest percentage reduction in tax rates. Prior to TRA, there were 14 tax brackets with rates ranging from 11 to 50 percent. TRA reduced the number of tax brackets to two, a 15 percent rate and a 28 percent rate bracket. A 33 percent tax rate (the so-called "bubble") was created by the phase-out of the lower 15 percent rate and personal exemptions for higher income taxpayers. Because tax rates changed differently for different taxpayers, the change in the tax rate will not only be a function of income. Moreover, the exogenous variation in tax rates due to TRA can be captured by calculating the change in tax rates holding income constant. This change in tax rates, which, by construction, is independent of the level of income, is the primary source of exogenous variation needed for identification of the tax rate response.[3]

Several researchers have used recent changes in tax policy as a natural experiment to predict how taxpayers respond to changes in tax rates. CBO (1986), Lindsey (1987) and Navratil (1995) find evidence of tax-induced behavioral responses for ERTA, while Feldstein (1995a), Feenberg and Poterba (1993), and Eissa (1995) find evidence of tax-induced responses for TRA. These studies, however, must be viewed with some caution. Lindsey (1987), which found very large responses, did not use panel data, but instead created a synthetic panel by grouping similarly situated taxpayers by income and estimating elasticities based on differences across these groups. Studies that assume the level of income that would have been achieved absent the tax changes are suspect because all unexpected income changes are assumed to be due to tax changes and the counter factual assumption about income growth influences the estimated elasticities (CBO, 1986; Lindsey, 1987; and Gravelle, 1993). Two recent studies that use panel data avoid some of these problems, but rely on samples of taxpayers that have relatively few high-income returns, and, similar to Lindsey (1987), may not adequately control for non-tax factors (Feldstein, 1995a and Navratil, 1995). Using successive cross-sections, Feenberg and Poterba (1993) investigate the rising share of income reported on very high-income tax returns. They find sharp increases in this group's share of reported income in 1987 and 1988 and conclude that at least part of the increase was due to the reduction in tax rates in TRA. Their

---

[3]Hausman and Poterba (1987) estimate that 40 percent of taxpayers faced either the same or higher tax rates under TRA, and 11 percent of taxpayers had their marginal tax rates lowered by 10 percentage points or more.

Other sources of independent variation that help identify the tax rate effect include state tax rates and the uneven effects of base broadening under TRA including the repeal of the sales tax deduction, higher floors under medical and miscellaneous deductions, and restrictions on deductions for IRA and Keogh contributions.

conclusion is subject to qualification, however, because they do not adjust the definition of income for statutory changes. Using Current Population Survey data, Eissa (1995) found evidence that the labor supply of high-income married women increased due to TRA.

The next section briefly outlines the empirical model. The third section presents the data and describes the construction of the variables used in the empirical model. The fourth section presents our estimation results, robustness checks, and an estimate of the revenue maximizing tax rate. The fifth section concludes the paper. Both taxes and non-tax factors are found to explain income changes during the late 1980s.

II. The Model

Taxpayers' incomes are influenced by many factors, some of which are within their control. First, taxpayers can decide how much labor to supply. Second, taxpayers can change the form of their compensation to affect the amount of income subject to tax. High tax rates may induce some taxpayers to shift more compensation to tax-favored sources, such as fringe benefits or stock options. Third, taxpayers can alter how much they save and in what types of assets they invest. If taxes are high on income from assets that yield ordinary income, taxpayers can purchase assets that yield capital gains. Fourth, taxpayers can reduce income by making purchases that receive favorable tax treatment. Taxpayers can choose to contribute more generously to charity or to increase home mortgage debt. Fifth, taxpayers can simply reduce compliance with the tax law. Taxpayers' incomes are also influenced by factors beyond their control, such as business cycles, changes in interest rates, demographic changes, and the growth or decline in different industries, occupations, and regions.

Our empirical model assumes that a taxpayer's income, Y_{it}, is explained by four factors that attempt to capture these influences: (1) an individual specific effect, ι_i, (2) a time specific effect, γ_t, (3) individual characteristics that do not change over time, but whose relationship to income may change over time, X_i, and (4) the marginal tax rate, τ_{it}, which varies across both individuals and across time. Income for the ith individual in time period t is given by,

$$Y_{it} = \iota_i + \gamma_t + \alpha_t X_i + \beta\tau_{it} \tag{1}$$

As is widely known, the individual effect reflects individual specific characteristics such as tastes, physical characteristics, geographic location and skills that do not vary over time and whose relationship to income do not change over time (Hsiao, 1986). The individual effect is eliminated from the model by first differencing. The estimating model becomes,

$$\Delta Y_i = \Delta\gamma + \Delta\alpha X_i + \beta\Delta\tau_i \tag{2}$$

where Δ denotes the change in a variable between period t and period t-1. The time effect controls for factors that affect all individuals in the same fashion at each point in time. Although the time effect remains in the model, it is embedded in the constant term. A taxpayer's marginal tax rate, τ_{it}, plus the individual characteristics that remain constant over time, but whose relationship to income may have changed, X_i, comprise our set of independent variables.

Although taxpayer characteristics such as wealth, region or skill level may have remained constant between 1985 and 1989, the relationship of these variables to income may have changed. Earnings potential in many cases may be linked to regional labor markets (Topel, 1994). If a region or industry were in decline (e.g., the Texas oil market during the 1980s), earnings potential for those employed or with businesses within this region or industry would likely fall. In addition, earnings of highly skilled workers would likely grow more rapidly than the earnings of low skilled workers if their skills become relatively more valuable, even though their education level remained constant. Understanding how these factors influence the estimated elasticity for the tax variable or "tax price" will not only allow us to evaluate previous research, but also allow us to evaluate the role non-tax factors, such as changes in the demand for highly educated workers, played during the 1980s.

III. The Data

We use a panel of individual Federal income tax returns that consists of the same taxpayers for years both before and after the Tax Reform Act of 1986. In addition to detailed tax return data, the panel over samples high income taxpayers and includes the age and occupation of each taxpayer.[4] The panel of tax returns was constructed by matching taxpayers present on both the 1985 and 1989 Statistics of Income (SOI) Individual Tax Files.[5] Tax returns from 1985 should precede any tax-induced behavior in response to anticipated tax rate changes. Tax return data for 1989 were used because 1989 is the first year in which the rate changes and most other features of TRA were fully phased in.[6] Most temporary tax-induced responses to the tax rate

---

[4]The tax return data are supplemented with the age of the taxpayer using an exact match with data provided by the Social Security Administration. The occupation classifications are based on the taxpayer-provided occupation description on the signature line of Form 1040 and the industry of a taxpayer's employer from W-2 Forms.

[5]The SOI Individual Income Tax Files are annual stratified random samples of over 100,000 individual tax returns filed in each year. See IRS (1989) for a description of the 1989 sample procedure.

[6]Not all provisions, however, were fully phased-in by 1989. For example, the repeal of the deduction of consumer interest for itemizers was not fully phased-in until 1990.

changes, such as shifting of income and of portfolios, likely occurred prior to 1989.[7] Consequently, a comparison of 1985 and 1989 should exclude most of the transitory effects of TRA on taxpayer incomes.[8]

We exclude some tax returns from our panel because of data limitations and to abstract from taxpayer characteristics that are not likely to be related to tax-induced behavioral responses, but affect a taxpayer's reported income.[9] For example, taxpayers other than single or joint filers with the same marital status in both years are excluded to avoid fluctuations in income related to changes in household composition due to marriage, divorce, or death of spouse.[10] To abstract from the effect of TRA on retirement decisions, taxpayers who are over age 55 in 1985 are excluded.[11] We also exclude all taxpayers who are under age 25 in 1985, since income changes for many of these taxpayers reflect the completion of schooling. Tax returns with low taxable incomes are excluded because many low income taxpayers drop out of the sample over time and certain types of income (e.g., Social Security benefits) may not be reported. We exclude all

[7]Most of the transitory shifting of income around TRA occurred in 1986, 1987 and 1988. In 1986, taxpayers had an incentive to accelerate capital gains and deductions, and defer ordinary income. In 1987, ordinary income was increased by 1986 deferrals, but decreased by taxpayers desiring the still lower rates in 1988. Income in 1988 was increased by deferrals from 1987.

[8]We may not capture effects that are fully long-run because the effects of the tax rate changes on capital formation and other slowly changing factors are not likely to be fully realized by 1989.

[9]These sample exclusions are similar, but less restrictive than those imposed by Feldstein (1995a) and Navratil (1995).

[10]We also exclude tax returns with a change in their secondary Social Security number because the household is no longer composed of the same individuals, which may indicate the taxpayer divorced and remarried between 1985 and 1989.

[11]We also exclude taxpayers who are identified as deceased or retired based on information from each taxpayer's self-described 1989 occupation. The construction of the occupation codes is discussed in greater detail below.

taxpayers with incomes below the income threshold for the 22 percent marginal tax rate in 1985 (i.e., $21,020 for joint filers and $15,610 for single filers). Taxpayers who are on the alternative minimum tax (AMT) in 1985 are also excluded. Although these taxpayers have relatively low marginal tax rates (in 1985 the AMT rate was 20 percent), their average tax rates are high relative to other taxpayers. Deleting AMT taxpayers avoids the difficulty of distinguishing between differences in the effect of marginal tax rate changes and average tax rate changes. Finally, we exclude taxpayers who report S corporation income after TRA, but not before. Many business owners elected S corporation status for their businesses after 1986. For taxpayers who converted businesses from C corporations to S corporations, we can observe their post-TRA S corporation income, but not their pre-TRA C corporation income. After these adjustments, 15,348 observations remain in our sample.[12] We examine the sensitivity of our results to these sample criteria below.[13]

As a result of the stratified sampling procedure used for SOI Individual Tax Files, which over samples tax returns with high incomes and/or tax returns accompanied by particular tax schedules, our panel includes a large number of high income tax returns and allows us to

[12]In most cases, the sample restrictions are based on 1985 characteristics because the 1989 characteristics can be affected by how a taxpayer responded to TRA and, therefore, can be endogenous. For example, whether a taxpayer is subject to the AMT in 1989 will, in part, depend on how the taxpayer responded to TRA. Nevertheless, in robustness tests presented below we relax many of these sample restrictions or apply them symmetrically based on both 1985 and 1989 characteristics to test the sensitivity of our results.

[13]The sample exclusions reduce the size of the panel from 56,003 tax returns to 15,348. The most significant sample exclusions are: (1) low income taxpayers (19,191 taxpayers), (2) age 55 or over (13,493 taxpayers), (3) age 25 or less (5,808 taxpayers), (4) taxpayers subject to the AMT in 1985 (4,022 taxpayers), and (5) non-single and non-joint taxpayers (3,609 taxpayers). Applying these five sample exclusions simultaneously reduces the panel to 21,903 taxpayers. The remaining sample exclusions reduce the panel to 15,348 taxpayers.

avoids the problems inherent in small samples. For example, the panel includes 3,633 returns (24 percent) in the 50 percent rate bracket in 1985 and another 628 (4 percent) in the 49 percent rate bracket. The selection of tax returns for the panel, however, depends on whether a tax return is sampled for both the 1985 and 1989 SOI Individual Tax Files. The selection of tax returns for our panel is endogenous because the annual SOI Individual Tax Files are stratified by income, the dependent variable in our model.

Following Hausman and Wise (1981) and Imbens and Lancaster (1996), we use a weighting procedure to correct for endogenous sample selection. The sampling procedure used for the SOI Individual Tax Files has the effect of including in the panel with certainty all taxpayers who remain in the same sample strata or move to a sample strata with a higher sampling rate between 1985 and 1989. Taxpayers who move to a sample strata with a lower sampling rate have a lower probability of remaining in the 1989 SOI Individual Tax File and being included in the panel. Therefore, taxpayers with declines in income between 1985 and 1989 are underrepresented in our panel. The weights are designed to compensate for this underrepresentation. Weighting the data by the maximum weight from the 1985 and 1989 SOI Individual Tax Files compensates for the underrepresentation of taxpayers with income declines and extrapolates our sample of tax returns to the population. Weighted least squares is used for estimation and both weighted and unweighted estimates are reported. A description of the weighting procedure is provided in the Appendix.

Dependent Variable

Two income concepts are used in our analysis: (1) constant law gross income, and (2) constant law taxable income, which is a variant of a taxpayer's taxable income. The gross income concept defines income broadly and is useful for analyzing to what extent changes in tax rates may explain changes in reported incomes and the income distribution. The taxable income concept is less useful for analyzing changes in the income distribution, but is more useful for analyzing the overall behavioral response of taxpayers and the revenue maximizing tax rate, because adjustments to income, personal exemptions, and the standard deduction/itemized deductions, are included. Both the gross income and taxable income measures reflect income as reported by taxpayers and are adjusted for inflation.

We define both income concepts on a post-TRA, constant law basis to remove the statutory differences in the income measures before and after TRA. We incorporate as many of the statutory changes in TRA as our data allow. The largest adjustment is to subtract capital gains realizations from taxable income in both years. Including capital gains would cause several problems. First, long-term capital gains were taxed at a different rate than other income in 1985, making it difficult to define a single tax rate appropriate for total 1985 income. Second, TRA increased effective tax rates for capital gains by repealing the 60-percent exclusion, while decreasing rates for other income. Thus, total income including capital gains is subject to two offsetting rate effects.

Because TRA disallowed certain passive losses, post-TRA taxable income reflects smaller losses. Failure to control for the change in passive loss rules would tend to overstate income growth and mistakenly attribute this portion of the income change to the tax rate response. Because of the lack of the data necessary to construct post-TRA losses absent this statutory change and to abstract from the effect of this statutory change on our results, we add all losses back to income both before and after TRA.[14] Other adjustments to gross income include adding back excluded dividends and unemployment compensation.

Additional adjustments are required to calculate constant law taxable income. The two-earner deduction and non-itemizer deduction for charitable contributions are added back. Adjustments are made for the increase in the floor under medical expenses, the new 2 percent floor under miscellaneous deductions, and the repeal of the deduction of sales taxes. Adjustments are made for changing the deductions for moving expenses and employee business expenses from above-the-line deductions to deductions available only to itemizers. The deduction for IRAs is put on a post-TRA basis by limiting it to lower income taxpayers. We also adjust for increases in the real value of the standard deductions and personal exemptions in TRA. We control for these statutory changes by adjusting 1985 taxable income by the increase in the real value of the standard deductions and personal exemptions in 1989.

Tax Price

Combined federal-state marginal tax rates were calculated using detailed tax calculators to compute each taxpayer's net-of-tax rate (i.e., $1-\tau$) or "tax price.".[15] The tax price used for estimation is the difference in the natural logarithms of the tax prices in 1985 and 1989.

A taxpayer's actual tax price, however, is likely to be endogenous with respect to the amount of income reported. As taxable income rises, a taxpayer's tax rate may also rise despite

---

[14]The assumption that all changes in reported losses were caused by statutory changes tends to understate tax-induced behavioral effects. We have not, however, controlled for changes in depreciation rules and for statutory changes limiting contributions for 401(k) and other similar retirement saving plans. Failing to control for these other statutory changes could bias the estimated effects of tax rates upward.

[15]The tax calculators were adapted for use with panel data from the U.S. Treasury Department's Individual Tax Simulation Model.

the rate reductions under TRA. An instrumental variable procedure is used to overcome this problem.[16] The instrument is constructed by first computing a taxpayer's 1989 tax rate using 1985 income inflated to 1989 levels. The instrumental variable is then calculated as the difference between a taxpayer's "synthetic" tax price for 1989 and the actual 1985 tax price. Two stage least squares estimates are obtained by first regressing the change in the actual tax price against the change in the synthetic tax price and the other exogenous variables. Consistent estimates are obtained in the second stage by replacing the change in the actual tax price with fitted values from the first stage.[17]

By construction, the instrument eliminates the effect of income changes attributable to tax-induced behavioral responses on the change in the tax price and only reflects the exogenous statutory change in tax rates due to TRA. It is this exogenous change in tax rates due to TRA that is the primary source of identification of the tax price in our model.

A taxpayer's tax price includes both state and federal marginal tax rates, allowing for the deductibility of state income taxes for taxpayers who itemize deductions. Variation in state tax rates that are independent of taxpayers' income also helps to identify the tax price.[18] The means and standard deviations for the combined federal-state tax price and the separate federal and state

[16]Following Hausman (1978), we tested for the possibility of endogeneity of the actual tax price and rejected the null hypothesis of exogeneity.

[17]Feldstein (1996) uses a somewhat different approach than described above. By distinguishing between taxpayers based on their 1985 actual tax rates, Feldstein (1996) omits the first stage in procedure described above. This has the effect of ignoring the coefficient on the change in the synthetic tax price in the first stage equation. We considered the effect ignoring the first stage equation by estimating a reduced form equation that simply includes the synthetic tax price, and found the estimated tax price elasticity to be somewhat lower than the results reported below.

[18]Feenberg (1987) and Burman and Randolph (1994) use cross-sectional differences in state income tax rates to identify tax rate effects.

tax prices are provided in Table 1.

Non-Tax Factors

Differencing controls for many non-tax factors by eliminating individual effects. The tax data also allow us to identify taxpayer characteristics whose relationship to income may have changed over time. Taxpayer wealth is likely to influence a taxpayer's ability to alter portfolios and labor arrangements in response to tax changes. Tax return data, however, provide no direct measure of a taxpayer's wealth. Instead, the sum of a taxpayer's dividend and interest income in 1985 is used as a proxy for an individual's financial wealth.[19]

Age and age squared (in 1985) are included to control for life cycle effects. We include the number of children in 1985 (up to a maximum of 6) and a dummy variable indicating whether a taxpayer has any children away from home but still claimed as dependents in 1985. Dependent children away from home are typically in college, which may limit the ability of a taxpayer to respond to lower tax rates. We also include an "entrepreneurship" dummy variable indicating whether a taxpayer reports income in 1985 from a sole proprietorship, partnership, or subchapter S corporation. This variable may reflect business ownership and entrepreneurship skills, and the propensity for risk-taking.

Some taxpayers with very low (high) transitory incomes in 1985 are likely to have large increases (decreases) in income between 1985 and 1989, many of which would likely occur even absent the changes in tax rates under TRA.[20] The exclusion of low income taxpayers (i.e.,

[19]Both capital gains and tax-exempt interest are excluded from this variable. The tax return data only contain a taxpayers capital gains when realized, not accrued, and contain no information in 1985 on tax-exempt interest income.

[20]Transitory income is somewhat reduced because the income variable excludes capital gains income and business losses, which are often large components of transitory income.

taxpayers below the threshold for the 22 percent rate brackets) from the sample was intended to help limit mean reversion bias. Taxpayers with temporarily high incomes, however, may experience large declines in income that would be associated with their large declines in tax rates, biasing the estimated tax price elasticity downward. In order to control for reversion-to-the-mean effects, a taxpayer's 1985 income is included as an independent variable.

During the 1980s, some regions of the United States grew quickly, while others experienced stagnant economies or recessions. Based on the tax return data, income growth from 1985 to 1989 was about three times higher in the Mid-Atlantic and Northeast regions (27 percent) than in the West South Central region (9 percent). Dummy variables for Census regions based on state of residence in 1985 are included to capture the different opportunities for income growth that existed in different regions during this period.

To explore the view that increasing returns to human capital also play a role in the increased inequality in the reported income distribution during the 1980s, we add dummy variables reflecting a taxpayer's occupation in 1989 to our model.[21] These variables allow us to consider whether the relationship between income and returns to human capital may have changed between 1985 and 1989. Returns to human capital may have increased because individuals in certain occupations may have experienced relative increases in labor productivity due to technological advances or may have been in higher demand perhaps due to increasing

[21]Taxpayer reported occupation descriptions were classified using the Standard Occupation Classification (SOC) system (see Clark, Reilly, and Sailer, 1989). Nearly 60 percent of the returns in our panel were classified by the Statistics of Income Division of the IRS using information from both a taxpayer's occupation description in 1989 and the industry classification of a taxpayer's employer obtained from W-2 Forms. We classified an additional 25 percent through detailed examination of taxpayer reported occupation and sources of income. SOC classifications were aggregated into occupational groups reflecting education and skill levels, and, in some cases, type of employer (government) and industry (doctors and other health related services).

international competitiveness or technological change. Income growth, however, could also vary across occupation categories because some occupations may offer greater flexibility to alter work effort or the mix of compensation than others. For example, doctors and lawyers, the occupation groups exhibiting the greatest income growth during this period, also have considerable flexibility to alter their work schedules and compensation arrangements in response to the tax rate changes. If the occupation variables are highly correlated with this tax-related effect, inclusion of these variables could bias the estimated elasticity for the tax variable downwards (Feldstein, 1996). In recognition of this measurement issue, we present sets of estimates with and without the occupation variables in the model.

IV. Results

Estimates of equation (2) for constant law gross income are shown in Table 2. Unweighted and weighted estimates are shown for three variants of the model: (1) "taxes-only" (columns 1 and 2), (2) taxes plus non-tax factors, excluding occupation (columns 3 and 4), and (3) taxes plus non-tax factors including occupation (columns 5 and 6). The "taxes only" specification includes only a constant term, the tax price (i.e., $1-\tau$), and income. This equation is most comparable to the earlier research that excludes non-tax factors from the model (for example, see Feldstein 1995a, Navratil 1994, and Lindsey, 1987). For this specification, the estimated "tax price" elasticity is 1.10 with a standard error (s.e.) of 0.13. The estimates using weighted least squares account for the endogenous sample selection and extrapolate our results to the population of taxpayers. Because the endogenous sample selection results in underrepresentation in our panel of taxpayers with income declines and sampling rates are the highest for high income taxpayers (i.e., taxpayers with the largest decline in tax rates under

TRA), we expect the elasticities using weighted least squares to be smaller than the unweighted results. Using weighted least squares lowers the estimated tax price elasticity to 0.77 with a standard error of 0.16. Only the unweighted results are within the range of the tax price elasticities reported by Feldstein (1995a), who reports elasticities between 1.04 and 3.05. The weighted results are similar to those reported by Navratil (1995), who reports elasticities around 0.8. The substantial difference between the estimated tax price elasticities obtained from the weighted and unweighted specifications, however, suggests that the unweighted results are likely biased upwards due to endogenous sample selection.

Columns 3 and 4 show results for the second specification of the model, which includes all of the non-tax factors except a taxpayer's occupation. The coefficient for the proxy for financial wealth is positive and significant implying that income grew more rapidly for those with greater wealth. The signs of the coefficients for age and age squared are jointly significant and imply that income growth declines with age.

The number of children is positively related to income growth. The dummy variable for college age children is negative (but is statistically different from zero only in the unweighted regression). The negative sign is consistent with the view that taxpayers with college age children may have already been realizing income at a high level in 1985. The coefficient for the entrepreneurship dummy is positive, but only significantly different from zero in the unweighted regression. The results for the regional dummy variables are consistent with the view that a taxpayer's income growth is influenced by regional economic performance. Four of the eight regional dummy variables are statistically significant, with the greatest income growth in the northeast and midatlantic states.

Compared to the taxes only model, the tax price elasticity falls by 15 percent to 0.64 with a standard error of 0.16 in the weighted model (column 4). The results for this specification suggest that both non-tax-factors (excluding occupation) and changes in the tax price are important determinants of income growth and that failure to control for non-tax factors may result in an overstated tax price elasticity.

The third specification (column 5 and 6) adds the occupation dummies to the model. The results for most of the non-tax factors are similar to the results reported in columns (3) and (4). Most of the coefficients for the occupation dummies are statistically different from zero. Our results for the occupation dummies suggest that changes in returns to human capital were an important factor explaining the changes in the income distribution during the 1980s. Taxpayers in occupations with high educational content experienced the greatest income growth during this period, while taxpayers in occupations with relatively less educational content did not fare as well. Lawyers, doctors, and executives and managers, have the largest positive coefficients. The effects of occupation, however, must be interpreted with care. In addition to a taxpayer's human capital, occupation may also reflect taxpayers' flexibility for rearranging affairs in response to changes in tax rates.

After controlling for occupation, the tax price elasticity remains virtually unchanged at 0.66 (s.e.=0.16) in the weighted regression. Even though occupation is an important factor explaining income growth, including occupation in the model has little measurable effect on the estimated tax price elasticity in our preferred (i.e., weighted) model. Although the tax price elasticities are well below those reported by Lindsey (1987) and Feldstein (1995a), the results reported in Table 2 suggest that *both* non-tax factors and the tax rate reductions of the 1980s help explain the increase in measured income inequality during the 1980s.

Robustness Checks

In Table 3 we report a number of robustness checks to consider the sensitivity of the results to sample restrictions and variable definitions. The robustness checks also help evaluate to what extent the estimated tax price elasticity might vary for different taxpayer groups. These robustness checks are based on the model using weighted least squares and includes occupation along with the other non-tax factors (i.e., column 6 in Table 2).

In our original sample, we abstract from individual retirement decisions by excluding all taxpayers age 55 and above in 1985 (i.e., age 59 and above in 1989). Many individuals, however, may not retire until age 62 when they can begin to receive limited Social Security benefits or age 65 when they can receive full Social Security benefits. Of course, adding additional taxpayers close to typical retirement ages will likely include some taxpayers who experience declines in income because of retirement. In addition, the larger personal exemptions and standard deductions under TRA could have induced middle income taxpayers to retire earlier because of higher after-tax incomes in retirement. On the other hand, high-income taxpayers have may been induced to postpone retirement because of higher net-of-tax income from continuing to work. Relaxation of the age exclusion to 60 and above in 1985 (i.e., age 64 and above in 1989), adds 1,823 tax returns to our sample (n=17,171) and increases the estimated tax price elasticity to 0.84 (s.e.=0.15) from 0.66 (s.e.=0.16).

The basic sample eliminates all taxpayers whose 1985 taxable income fell below the thresholds for the 22-percent statutory tax rate in 1985. Because 1989 income is affected by a taxpayer's behavioral response, we did not impose a comparable sample criteria based on a taxpayer's 1989 income for our preferred specification (i.e., Table 3). Nevertheless, in Table 4 we report how such an additional sample criterion affects the estimated tax price elasticity. In particular, we inflate the threshold for the 22 percent statutory rate bracket to 1989 levels using the Consumer Price Index and exclude the 1,677 taxpayers whose 1989 taxable income fell below these income thresholds. Excluding these taxpayers lowers the estimated tax price elasticity to 0.47 (s.e.=0.09).

TRA made S corporations, where business income is passed through to individual owners before being taxed, more attractive relative to C corporations that pay income tax at both the business entity level and the owner level (see, for example, Carroll and Joulfaian, 1997; Ayers et al; 1996; and Plesko, 1994).[22] Our basic sample excludes taxpayers who report S corporation income in 1989, but not in 1985, to eliminate taxpayers that may have converted their C corporation(s) to S corporations in response to TRA. It is possible, however, that some taxpayers may have reported S corporation income both before and after TRA, but still converted some business activities from the C corporation form to the S corporation form. The growth of total income would be overstated for such taxpayers in 1989. To consider this possibility, we examine a sample that excludes all taxpayers with any S corporation income in either 1985 or 1989. This additional sample restriction eliminates 2,358 taxpayers from our panel, but leaves the estimated tax price elasticity virtually unchanged at 0.66 (s.e.=0.17). If we instead include the 4,324 taxpayers in our panel with S corporation income in either 1985 or 1989, the estimated tax price

---

[22]TRA also created a similar incentive for C corporations to convert to sole proprietorships and partnerships. These other business forms, however, did not receive the same advantages from limited liability protection as S and C corporations. Ayers et al (1996) find that liability protection may be a distinguishing factor between corporate and unincorporated businesses and in the selection of organizational form.

elasticity rises to 0.79 (s.e.=0.15), suggesting that income in 1989 and the tax price elasticity could be somewhat overstated without controlling for this change in organizational form.

Taxpayers with AMT in 1985 were excluded from the basic sample. If we exclude taxpayers with AMT liability in either year, rather than just in 1985 (excluding an additional 206 tax returns), the tax price elasticity for the weighted regression remains virtually unchanged (coef.=0.67, s.e. = 0.35). If we instead include in our panel the 1,486 taxpayers subject to the AMT in either year, the estimated tax price elasticity rises slightly to 0.76 (s.e.=0.15).

We test whether our results are sensitive to whether a taxpayer reported income from capital gains by including a dummy variable for the presence of capital gains in 1985. The repeal of the capital gains exclusion combined with the lower tax rates on ordinary income under TRA may have induced taxpayers to shift their portfolios from high-capital gain assets to high-dividend or high-interest paying assets to take advantage of the relatively lower marginal tax rates on ordinary income. Because of this portfolio effect, our dependent variable may overstate income in 1989 as compared to income in 1985.

In order to determine whether our results are sensitive to capital gains portfolio effects and to disentangle the effects of capital gains from the AMT, we conduct several experiments. First, we include a capital gains dummy variable and find that the estimated tax price is unchanged (coef.=0.67, s.e.=0.16) and the dummy variable is not statistically different from zero. Since prior to TRA excluded capital gains were treated as a preference for the AMT and we exclude taxpayers subject to the AMT in 1985 from our basic panel, however, we also likely exclude many taxpayers who realize large amounts of capital gains in 1985. Thus, a second capital gains test is to include the capital gains dummy variable and also include taxpayers subject to the AMT in either year. In this test the estimated tax price elasticity increases to 0.79 (s.e.=0.15) and the coefficient for the capital gains dummy variable is positive and statistically significant. These results suggest that excluding AMT taxpayers may remove some of the portfolio effects associated with the change in the relative effective tax rates on capital gains and ordinary income under TRA. Moreover, the positive and statistically significant coefficient on the capital gains dummy variable (coef.=0.055, s.e.=0.029) suggests that some taxpayers who realize capital gains in 1985 altered their portfolios in response to the higher effective tax rate on capital gains realizations and lower tax rates on income received as ordinary income under TRA.

We consider whether our results change if we include a dummy variable for itemizers. Itemization may help distinguish between taxpayers who own a home rather than rent (i.e., home mortgage interest deduction), live in high tax states (i.e., state and local property, income and sales tax deductions), have substantial investment expenses (i.e., interest expense deduction), or have above average medical expenses (i.e., medical expense deduction). Homeownership may serve as a proxy for real estate wealth. Of course, taxpayers who itemize may also have more flexibility to rearrange their affairs by changing their decisions about, for example, how much equity to hold in their residence and how much to give to charity. Including a dummy variable for itemization status in 1985 leaves the tax price elasticity for either model virtually unchanged (coef.=0.67, s.e.=0.17), and the itemization variable is not statistically different from zero.

Most of the independent variables are constructed based on 1985 values. As an additional robustness check, some of the variables can be defined symmetrically in both years. This may introduce a substantial endogeneity problem in some cases because the 1989 values are likely to be influenced by a taxpayer's tax-induced behavioral response. For example, the amount of taxable interest and dividend income reported in 1989 will partly be determined by tax-induced changes in the taxpayer's portfolio. Similarly, changes in housing consumption, homeownership, and home equity all may be partly determined by changes in the value of the home mortgage interest and property tax deductions due to the lower tax rates and repeal of the consumer interest deduction under TRA. In any case, using average financial wealth instead of the 1985 value and defining the capital gains and itemizer dummy variables to one if a taxpayer reports capital gains income or itemizes in either year increases the estimated tax price elasticity somewhat to 0.80 (s.e.=0.16).

In another symmetry test, we recompute both constant law gross income and marginal tax rates on a pre-TRA basis to the extent the data allow and use 1989 sample exclusions. In our preferred specification (i.e., Table 2), we calculated marginal tax rates primarily on post-TRA basis because the tax data allow a more complete set of adjustments. For example, TRA limited deductible IRA contributions to taxpayers below certain income thresholds. Because we have no information on the IRA contributions that would have been made had these IRA contributions not been limited by TRA, in our preferred specification (i.e., Table 2) we compute the IRA deductions on a post-TRA basis. Computing IRA contributions on a pre-TRA basis would either involve imputing which taxpayers above the income thresholds contribute and how much, or simply excluding all IRA contributions from income when computing marginal tax rates. Similar problems arise for the medical expense deduction and miscellaneous deductions above the 2 percent floor. When we estimate the tax price elasticity using marginal tax rates and income based on the tax law prior to the TRA, we find the estimated tax price elasticity to be somewhat higher than the results in Table 2 (coef.=0.91, s.e.=0.10). This rise in the tax price elasticity, however, appears to be attributable to the 1989 based sample exclusions, rather than basing the income measure and tax rates on pre-TRA tax law.

We include income as an explanatory variable in our preferred model to, at least partially, control for mean reversion bias. We find that dropping this variable lowers the estimated tax price elasticity to 0.35 (s.e.=0.13), suggesting that a model with no controls for mean reversion bias may bias the estimated tax price elasticity downwards.

After a variety of robustness checks to different sample restrictions and model specifications, our basic finding is that the results are somewhat sensitive to changes in assumptions, but the estimated tax price elasticity is still significantly above zero, and is generally between 0.4 and 0.9.

Behavioral Response by Occupation Type

As discussed above, the tax-induced behavioral responses can be expected to vary by occupation not only because of differential changes in the returns to education associated with different occupations, but also because some occupations may offer more flexibility to adjust compensation packages and labor supply in response to tax law changes. Doctors, lawyers, and tax accountants may have greater control over hours worked than other occupations, and tax accountants may have a better understanding of how individual finances can be restructured to minimize taxes in the face of a changing tax environment.

Table 4 presents results that interacts the tax price with the occupation dummies to separately measure the tax price elasticity across occupations and an F-test of the joint significance of the tax price elasticity and the interaction term for each occupation category. The dummy variable for government employees and educators is excluded. Only four of the twelve occupations listed are jointly statistically different from zero -- executives and managers, investors, self-employed and farmers, and sales. It should be noted that the tax price elasticity estimates for the occupation groups should be interpreted with some caution because some of the occupation classifications tend to be clustered within relatively narrow income ranges. The identification of the tax price elasticities for some occupation groups may, therefore, be suspect. The highest tax price elasticity is for the investors category (coef.=3.34, F-value=7.32), which includes both individuals describing their occupations as investors, venture capitalists, or stock brokers. The self-employed group has an estimated tax price elasticity of 2.40 (F-value=6.10). Taxpayers in this occupation group, consisting of 127 farmers and 380 other self-employed, are likely to have considerable flexibility in how they organize their business affairs, as well as control over their labor supply. The tax price elasticity for taxpayers classified as executives and managers, consisting of 2,399 tax returns is 1.08 (F-value=9.47). The tax price elasticity for taxpayers classifying themselves as having sales-related occupations is 0.98 (F-value=3.06).

Taxable Income and the Revenue Maximizing Tax Rate.

One issue of interest for tax policy is what the results imply for the revenue maximizing tax rate, the rate at which any further tax rate increases would produce revenues losses. For this purpose, it is more appropriate to focus on taxable income rather than the broader income concept used in the previous analysis (Feldstein 1995b). Estimation of equation (2) using constant law taxable income rather than constant law gross income yields a tax price elasticity of 0.57 (s.e.=0.226) using weighted least squares and including all of the non-tax factors (i.e., comparable to column (6) in Table 2).

The tax price elasticities based on constant law taxable income can be used directly to compute the revenue maximizing tax rate. Making the simplifying assumption that the income tax is proportional, and following Lindsey (1985), the revenue maximizing tax rate is given by,

$$\tau = \frac{1}{(1 + \beta_\tau)} \tag{3}$$

where β_τ is the tax price elasticity. Following this formulation, the revenue maximizing tax rate for our preferred model (i.e., using weighted least squares and including all of the nontax factors) would be about 64 percent. Of course, the current tax system is not proportional, but progressive. As Lindsey (1985) points out, progressive rates will imply a lower revenue maximizing tax rate because inframarginal income will be taxed at a rate unrelated to the top marginal tax rate.

The use of the tax price elasticities reported above to evaluate changes in tax policy requires several caveats. First, although the tax price elasticities represent the average response of taxpayers to the tax rate changes in TRA, as shown in the previous section some taxpayer groups are likely to have different behavioral responses. For example, the response of high income taxpayers may be larger than middle income taxpayers because higher income taxpayers may have greater flexibility to alter their labor arrangements and portfolios. Second, the behavioral response may differ for different types of tax rate changes. For example, the behavioral response for rate increases may be different than the response for rate reductions.

Third, the potential for avoidance and income shifting associated with tax changes will vary over time. Taxpayers and their advisors develop new strategies and technologies to minimize tax liability. Changes in compliance measures, audit techniques, IRS regulations, and court decisions also affect the tax environment over time. All of these factors affect both the estimates of behavioral responses and the use of such estimates to predict future taxpayers responses.

V. Conclusion

Using panel data this paper finds that both tax rates and non-tax factors explain the changes in reported income (holding tax law changes constant) following TRA. After controlling for nontax factors, we estimate the tax price elasticity to be about 0.66. Including nontax factors in the model reduced our estimate of the tax price elasticity by about 15 percent. Controlling for changes in returns to human capital by adding occupation to the model has no significant effect on the overall tax price elasticity, but there is some evidence that the tax price elasticity varies considerably across occupation categories.

The estimates we present for the model that only includes a taxpayer's tax price and controls for endogenous sample selection is lower than the tax price elasticity reported by Feldstein (1995a), which relies on a somewhat different empirical approach. Nevertheless, our results indicate that tax rate changes are likely to have significant effects on taxpayer behavior, reported income, and measured income inequality.

REFERENCES

Ayers, Benjamin C, C. Bryan Cloyd, and John R. Robinson, "Organizational Form and Taxes: An Empirical Analysis of Small Businesses," *Journal of the American Taxation Association*, (1996), forthcoming.

Bound, John and George Johnson, "Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations," *American Economic Review,* 82 (June 1992), 371-392.

Burman, Leonard E. and William C. Randolph, "Measuring Permanent Responses to Capital-Gains Tax Changes in Panel Data," *American Economic Review,* 84 (September 1994), 794-809.

Carroll, Robert and David Joulfaian. "Taxes and Corporate Choice of Organizational Form." U.S. Department of the Treasury, Office of Tax Analysis, Working Paper No. 73, April 1997.

Clark, Bobby, Dodie Riley and Peter Sailer, "1979 Occupation Study/1979-1983 Mortality Study," In *Statistics of Income and Related Administrative Record Research: 1989,* Internal Revenue Service, 181-187.

Congressional Budget Office, "Effects of the 1981 Tax Act on the Distribution of Income and Taxes Paid," Staff Working Paper, August 1986.

Eissa, Nada. "Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment," NBER Working Paper Series, No. 5023, February 1995.

Feenberg, Daniel, "Are Tax Price Models Really Identified: The Case of Charitable Giving," *National Tax Journal,* 40 (December 1987), 629-633.

Feenberg, Daniel and James Poterba. "Income Inequality and the Incomes of Very High Income Taxpayers: Evidence From Tax Returns." In *Tax Policy and the Economy* 7 (1993), 145-177.

Feldstein, Martin, "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act," *Journal of Political Economy,* 103 (June 1995a), 551-572.

Feldstein, Martin, "How Big Should Government Be?" NBER Working Paper Series, No. 5868, December 1996.

Feldstein, Martin. "Tax Avoidance and the Deadweight Loss of the Income Tax," NBER Working Paper No. 5055, 1995b.

Freeman, Richard B. "How Much Has De-Unionization Contributed to the Rise in Male Earnings Inequality," In Peter Gottschalk and Sheldon Danziger, eds., *Uneven Tides; Rising Inequality in America* (New York: Russell Sage Foundation, 1993).

Goldin, Claudia and Robert Margo, "The Great Compression: The Wage Structure in the United States at Mid-Century," *Quarterly Journal of Economics,* 107 (February 1992), 1-34.

Gravelle, Jane G, "Behavioral Responses to Proposed High Income Tax Rate Increases: An Evaluation of the Feldstein-Feenberg Study," Congressional Research Service Report, April 1993.

Hausman, Jerry, "Specification Test in Econometrics," *Econometrica,* 46 (1978), 1251-1272.

Hausman, Jerry and James Poterba, "Household Behavior and the Tax Reform Act of 1986," *Journal of Economic Perspectives,* 1 (Summer 1987), 101-119.

Hausman, Jerry and David Wise, Stratification on Endogenous Variables and Estimation: The Gary Income Maintenance Experiment." In *Structural Analysis of Discrete Data with Econometric Applications,* eds. Charles Manski and Daniel McFadden. (Cambridge University Press, 1981).

Heckman, James, "Sample Selection Bias as a Specification Error," *Econometrica,* 47 (January 1979), 153-161.

Hsiao, Cheng. *Analysis of Panel Data* (New York: Cambridge University Press, 1986).

Internal Revenue Service, Statistics of Income Division, *1989 Individual Income Tax Returns*, Publication 1034, (Washington, DC: U.S. Government Printing Office, 1992).

Imbens, Guido W. and Tony Lancaster, "Efficient Estimation and stratified sampling," *Journal of Econometrics* 74 (1996), 289-318.

Katz, Lawrence F. and Kevin M. Murphy, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics,* 107 (February 1992), 35-78.

Lindsey, Lawrence, "Estimating the Revenue Maximizing Top Personal Tax Rate," NBER Working Paper Series, No. 1761, October 1985.

Lindsey, Lawrence, "Individual Taxpayer Response to Tax Cuts: 1982-1984," *Journal of Public Economics,* 33 (July 1987), 173-206.

Murphy, Kevin and Finis Welch, "The Structure of Wages," *Quarterly Journal of Economics,* 107 (1) (February 1992), 285-326.

Navratil, John, "The Economic Recovery Act of 1981," in Essays on the Impact of Marginal Tax Rate Reductions on the Reporting of Taxable Income on Individual Income Tax Returns, Ph.D. dissertation, Harvard University, 1995.

Plesko, George A, "Entity Choice: Financial Characteristics of S and C Corporations," *Public Finance Quarterly,* 22 (July 1994), 311-334.

Slemrod, Joel. "On the High-Income Laffer Curve." In Joel Slemrod, ed., *Tax Progressivity and Income Equality* (Cambridge: Cambridge University Press, 1994), 177-210.

Topel, Robert J. "Regional Labor Markets and the Determinants of Wage Inequality," *American Economic Review,* 84 (May 1994), 17-22.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF VARIABLES (n=15,348)

Variables	**Mean**	**Standard Deviation**
Δ log of Constant Law Gross Income	0.148	0.901
Δ log of Constant Law Taxable Income	0.122	1.098
Δ log of $(1-\tau_T)$	0.161	0.164
Δ log of $(1-\tau_F)$	0.153	0.158
Δ log of $(1-\tau_S)$	0.0083	0.0282
Log of Income	11.601	1.347
Log of Capital Income	6.894	3.086
Age	40.9	8.2
Age Squared	1,737	668
Entrepreneur Dummy	0.514	0.500
Number of Children	1.435	1.272
College Age Children	0.031	0.174
Marital Status	0.875	0.331
Regional Dummy Variables:		
Northeast	0.068	0.252
Mid-Atlantic	0.197	0.398
West-North Central	0.063	0.243
South Atlantic	0.155	0.362
East South Central	0.044	0.205
West South Central	0.101	0.300
Mountain	0.042	0.202
Pacific	0.159	0.365
Occupation Dummy Variables:		
Executives and Managers	0.255	0.436
Investors	0.034	0.181
Lawyers	0.044	0.205
Artists, Athletes, and Journalists	0.018	0.131
Doctors and Other Health-Related Occupations	0.106	0.308
Science, Engineers, and Computers	0.086	0.280
Supervisors	0.028	0.166
Blue Collar	0.109	0.312
Farming and Self-Employed	0.033	0.179
Sales	0.049	0.216
Government and Educators	0.063	0.243
Not Elsewhere Classified	0.017	0.128
Occupation Information Not Available	0.135	0.342

τ_f= Federal marginal tax rate, τ_S= State marginal tax rate, τ_T=combined Federal-state marginal tax rate.

TABLE 2

REGRESSION RESULTS: CHANGE IN GROSS INCOME

Variables	**Taxes Only**		**Add Nontax Factors**		**Add Occupation**	
	Two Stage Least Squares	Weighted Two Stage Least Squares	Two Stage Least Squares	Weighted Two Stage Least Squares	Two Stage Least Squares	Weighted Two Stage Least Squares
Intercept	0.342* (0.142)	0.694* (0.172)	0.510* (0.237)	1.226* (0.221)	0.717* (0.240)	1.762* (0.234)
Δ log of $(1-\tau_f)$	1.101* (0.134)	0.774* (0.159)	1.025* (0.136)	0.640* (0.155)	0.948* (0.136)	0.660* (0.155)
Log of Income	-0.032* (0.014)	-0.067* (0.017)	-0.050* (0.015)	-0.118* (0.021)	-0.068* (0.015)	-0.170* (0.022)
Log of Capital Income			0.0101* (0.0036)	0.0102* (0.0021)	0.0086* (0.0035)	0.0090* (0.0021)
Age			0.0037 (0.0095)	-0.00097 (0.00601)	0.0028 (0.0095)	-0.00044 (0.00598)
Age Squared			-0.00014 (0.00012)	-0.00006 (0.00008)	-0.00012 (0.00012)	-0.00006 (0.00008)
Entrepreneur Dummy			0.0652* (0.0185)	0.0077 (0.0127)	0.0627* (0.0190)	0.0168 (0.0129)
Number of Children			0.0125* (0.0068)	0.0246* (0.0050)	0.0094 (0.0068)	0.0228* (0.0050)
College Age Children			-0.1864* (0.0429)	-0.0035 (0.0326)	-0.1860* (0.0426)	-0.0009 (0.0323)
Marital Status			0.0841* (0.0252)	0.0930* (0.0157)	0.0915* (0.0253)	0.130* (0.016)
Regional Dummy Variables:						
Northeast			-0.005 (0.034)	0.067* (0.022)	-0.007 (0.033)	0.069* (0.022)
Mid-Atlantic			0.031 (0.025)	0.045* (0.017)	0.027 (0.025)	0.040* (0.017)
West-North Central			-0.088* (0.034)	-0.044* (0.022)	-0.089* (0.034)	-0.050* (0.022)
South Atlantic			-0.021 (0.026)	0.023 (0.017)	-0.024 (0.026)	0.016 (0.017)

TABLE 2 (CONT.)						
Variables	**Taxes Only**		**Add Nontax Factors**		**Add Occupation**	
	Two Stage Least Squares	Weighted Two Stage Least Squares	Two Stage Least Squares	Weighted Two Stage Least Squares	Two Stage Least Squares	Weighted Two Stage Least Squares
East South Central			-0.093* (0.040)	-0.028 (0.025)	-0.088* (0.039)	-0.029 (0.024)
West South Central			-0.240* (0.029)	-0.148* (0.020)	-0.234* (0.029)	-0.149* (0.019)
Mountain			-0.144* (0.040)	-0.039 (0.024)	-0.135* (0.040)	-0.035 (0.024)
Pacific			-0.039 (0.026)	0.007 (0.017)	-0.033 (0.026)	0.012 (0.017)
Occupation Dummy Variables:						
Executives and Managers					0.146* (0.034)	0.072* (0.020)
Investors					-0.053 (0.051)	-0.253* (0.048)
Lawyers					0.188* (0.047)	0.122* (0.041)
Artists, Athletes, and Journalists					0.026 (0.064)	-0.052 (0.049)
Doctors and Other Health-Related Occupations					0.072 (0.039)	0.144* (0.029)
Science, Engineers, and Computers					0.024 (0.039)	0.030 (0.022)
Supervisors					-0.042 (0.052)	-0.046 (0.028)
Blue Collar					-0.091* (0.037)	-0.095* (0.020)
Farming and Self-Employed					-0.261* (0.050)	-0.308* (0.034)
Sales					-0.081* (0.040)	-0.069* (0.024)
Not Elsewhere Classified					-0.017 (0.064)	-0.023 (0.032)
Occupation Information Not Available					-0.029 (0.036)	-0.065* (0.019)

Note: Standard errors appear in parenthesis. * indicates significant at the 95 percent level of confidence. τ_T= Combined federal-state marginal tax rate. Instrumental variables is implemented using two stage least squares. In the first stage the actual tax price is regressed against the synthetic tax price and the independent variables. Predicted values from the first stage are then included in the second stage. Observations=15,348.

TABLE 3

Robustness Checks for Complete Model

Robustness Check	**Sample Size**	**Tax Price Elasticity**
Modify Sample Selection Criteria		
Include taxpayers age 64 or below in 1989	17,171	0.84* (0.15)
Apply income cutoff in both 1985 and 1989	13,671	0.47* (0.09)
Exclude taxpayers with any S corporation income	12,990	0.66* (0.17)
Include all taxpayers with S corporation income	17,314	0.79* (0.15)
Exclude AMT taxpayers in both 1985 and 1989	15,142	0.67* (0.16)
Include taxpayers subject to the AMT	16,628	0.76* (0.15)
Include capital gains dummy variable	15,348	0.67* (0.16)
Include capital gains dummy variable and taxpayers subject the AMT	16,628	0.79* (0.15)
Other Robustness Checks		
Include itemization dummy variable	15,348	0.67* (0.17)
Financial wealth, capital gains, and itemization based on both 1985 and 1989 information	15,348	0.80* (0.16)
Base marginal tax rates and dependent variable on pre-TRA law	15,089	0.91* (0.10)
Drop income as independent variable	15,348	0.35* (0.13)

Estimates are for the complete weighted model, including all non-tax factors. Instrumental variables is used with generalized least squares. Standard errors appear in parenthesis. * indicates variable is significant at the 95 percent level of confidence. τ_T= Combined federal-state marginal tax rate.

TABLE 4

Tax Price Elasticity By Occupation

	Tax Price Elasticity*Occupation$_i$ (Standard Error)	Total Effect (F-Value)
$(1-\tau_T)$	-0.275 (0.420)	NA
Occupation:		
Executives and Managers	1.356* (0.480)	1.081* (9.470)
Investors	3.612* (0.976)	3.337* (7.318)
Lawyers	0.409 (0.640)	0.134 (0.255)
Artists, Journalists, and Athletes	0.783 (0.956)	0.508 (0.392)
Doctors and Other Health-Related Occupations	0.780 (0.521)	0.505 (1.451)
Science, Engineers, and Computers	0.413 (0.552)	0.138 (0.292)
Supervisors	0.256 (0.724)	-0.019 (0.214)
Blue Collar	1.010 (0.706)	0.735 (1.050)
Farming and Self-Employed	2.399* (0.744)	2.124* (6.091)
Sales	1.253* (0.581)	0.978* (3.061)
Not Elsewhere Classified	0.026 (1.307)	-0.250 (0.233)
Occupation Information Not Available	0.700 (0.596)	0.425 (0.714)

Estimates are for the complete weighted model, including all non-tax factors, using instrumental variables with generalized least squares. The government and educators occupation category is excluded for estimation purposes. For the interaction variables, standard errors appear in parenthesis. For the total effect, F-values appear in parenthesis. * indicates variable is significant at the 95 percent level of confidence. τ_T= Combined federal-state marginal tax rate.

Appendix A

The panel of tax returns used for this paper is constructed by matching taxpayers present in the Statistics of Income (SOI) Individual Tax files for both 1985 and 1989. The population of individual income tax returns is highly skewed, with a small fraction of individuals accounting for a relatively large fraction of income and tax liability.

The SOI cross-sectional files are stratified by income, business plus farm receipts, and tax schedule type to over sample taxpayers with high incomes or business receipts and taxpayers who file particular tax schedules (e.g., Schedule C/sole proprietorship and Schedule F/farm proprietorship). Income is defined as the greater of a taxpayer's total income or total loss. Taxpayers with very high incomes are selected with certainty. In 1985, 121,480 returns were sampled. In 1989, 110,840 returns were sampled.

As noted by Imbens and Lancaster (1996) and Hausman and Wise (1981), if the sampling were random or exogenous, estimation would lead to consistent and efficient estimates. However, selection of tax returns for the panel used for this paper depends, to a large extent, on the dependent variable in our model (i.e., adjusted taxable income less capital gains realizations); that is, the sample selection is endogenous. Therefore, without appropriately controlling for endogenous sample selection, the estimates of the tax price elasiticity would be likely to suffer from sample selection bias.

The stratification of the sample by income, business receipts, and form type also means our sample does not correspond to the true proportions in the population. In order to allow inferences to be drawn about how changes in tax rates are likely to affect the population, we also need to extrapolate our sample to the population. Below we first describes the construction of the panel of tax returns used for this paper and the nature of the sample selection bias. Then we outline the weighting approach used to correct for the endogenous sample selection and produce population estimates.

Description of Panel

The sample strata used for the SOI cross-sectional files are based on the larger of total income or total loss, the size of business plus farm receipts, and the presence of particular forms and schedules. Returns are then selected from the sample strata from two methods. The first method selects all returns with the same two sets of four ending digits of their social security number (SSN). Tax returns selected under this first method contribute to a randomly drawn sample of tax returns often referred to as the Continuous Work History Survey (CWHS). The probability of any tax return being selected for the CWHS sample is roughly equal to 1/5,000.[1]

The second method uses the ending digits of numbers generated by randomly transforming each

[1]The probability of taxpayers' with the same four ending digits of their SSN being selected is 1/10,000. Since two sets of four ending digits of a taxpayer's SSN are used, the sampling probability becomes 1/5,000. Tax returns selected under this method constitute about 20 percent of the panel observations used for this paper.

tax return's SSN. Tax returns with ending digits below specified values, which define the probability of tax returns being drawn within each strata, are selected. In 1985, the sampling rates for the various strata ranged from 0.03 percent to 100 percent. From one cross-sectional file to another cross-sectional files, the random number for a particular tax return is unchanged because the seed value for the random number generator is the same across years. Provided a taxpayer remains in a strata with an equal or greater sampling rate, this sample design guarantees that the taxpayer will be selected for future cross-sectional files. This sample design also has the effect of maximizing the overlap of taxpayers sampled in different years. Of the 121,480 returns sampled in 1985 and 110,840 returns sampled in 1989, about 56,003 returns were sampled in both years.

Endogenous Sample Selection

Taxpayers selected using the SSN transformation described above will generally have an equal or greater probability of remaining in subsequent cross-sectional files if their income is unchanged or increases because they would either remain in their original sample strata or move to a strata with a higher sampling rate. Taxpayers who experience a decline in income are likely to fall into a lower sampling strata, and would, therefore, have a probability of being selected for subsequent cross-sectional samples of less than one. In fact, the probability that they will be sampled in the subsequent cross-section falls to whatever the sampling rate is for their new sample strata. The probability that they remain in the cross-section, conditional on being selected in a prior year is the ratio of the sampling rate for their new strata to their initial sampling rate (Hinkins, Jones, and Scheuren; 1988).

Tax price elasticities estimated from the sample of tax returns comprising the panel used for this paper are likely to suffer from sample selection bias because taxpayers with rising or equal incomes are likely to remain in the panel, while taxpayers with falling income are less likely to be included in the panel. That is, this paper uses a panel to study the determinants of income changes where taxpayers with declines in income are under represented. Moreover, the tax price elasticity estimated in our paper is likely to be biased upwards.[2]

Sample Weights to Correct for Sample Selection Bias and Produce Population Estimates

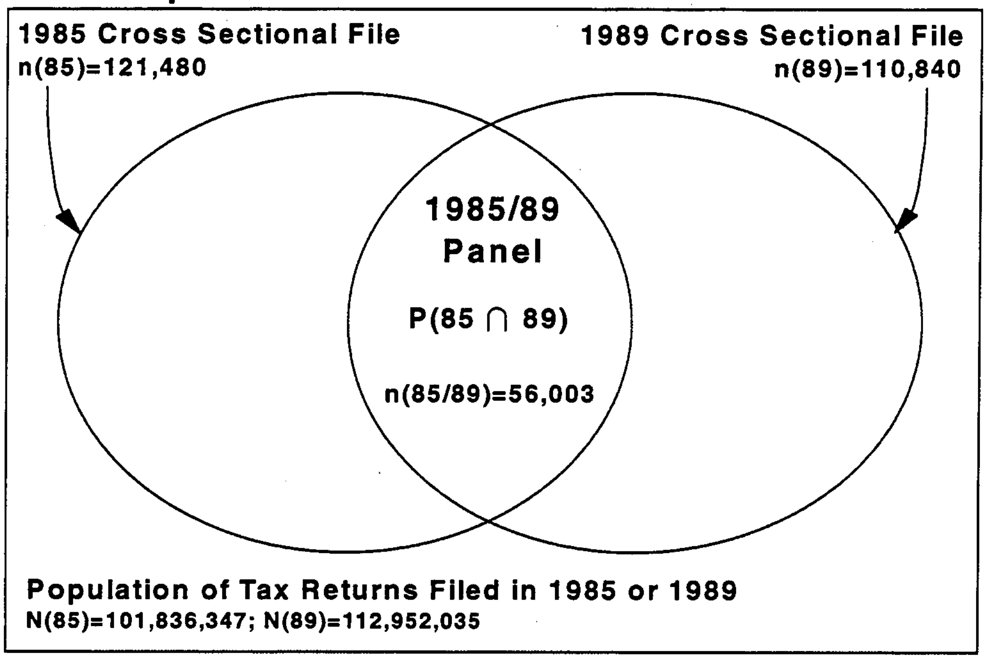
Figure 1 depicts the 1985 and 1989 SOI files and the panel constructed from the overlap between these two files. Tax returns are sampled in 1985 depending on their characteristics (i.e., income, business receipts, and form type). The sampling rate for the ith stratum in 1985, $P_i(85)$, is determined by these characteristics. Whether a tax return is sampled in 1989 will also depend on the taxpayer's 1989 characteristics. The taxpayers sampling rate, independent of whether the

[2]Although we obtain similar results when only using the randomly selected CWHS returns, we do not have a great deal of confidence in these results because the CWHS has few high income taxpayers -- the taxpayers with the greatest reduction in tax rates due to the 1986 Tax Reform Act. For example, among the CWHS returns in our panel, there are only 50 taxpayers who faced the top 50 percent statutory marginal tax rate and 49 taxpayers who faced the 49 percent statutory marginal tax rate in 1985.

taxpayer was sampled in 1985, is $P_i(89)$.

The panel used for this paper, however, only contains taxpayers that were sampled in *both* 1985 and 1989; that is, the overlap between the two cross-sectional files shown on Figure 1 as the

Figure 1
Relationship of Panel to Cross Sections of Tax Returns



intersection between the 1985 and 1989 cross sectional files.

The probability that a tax return is sampled in both 1985 and 1989, and is, therefore, included in our panel of tax returns, can be constructed in two steps depending on whether a taxpayer is sampled at all in 1985 and whether the taxpayer is again sampled in 1989. The probability that a taxpayer will be sampled in 1985 is given by a tax return's sampling rate associated with the taxpayer characteristics, $P_i(85)$. We can then consider the probability that a taxpayer is sampled in 1989 conditional on having already been sampled in 1985. This formulation follows from the identity,

$$P_i(85 \cap 89) \equiv P_i(85) \cdot P_i(89 \mid 85) \qquad \text{A.1}$$

where $P_i(85 \cap 89)$ represents our panel of tax returns (as shown in Figure 1) and $P_i(89 \mid 85)$ represents the probability that tax returns are sampled in 1989 conditional on already being sampled in 1985. A key feature of this formulation is that the probability that taxpayers are selected in 1989 and included in our panel depends on the change in the sampling rates for their respective sample strata in each year. In order to correct for the endogenous sample selection, the probability of being selected in 1989 and being included in our panel must depend on both 1989 and 1985 characteristics.

Recall that because of the sample design of the cross-sectional file, a taxpayer will be selected in subsequent cross-sectional files with certainty (i.e., $P_i(89|85) = 1$) if a taxpayer remains in the same sample strata or moves to a sample strata with a higher sampling rate. However, if a taxpayer moves to a sample strata with a lower sampling rate (e.g., because of a decrease in income), the probability that the taxpayer continues to be sampled is less than one. In this case, the probability that a taxpayer is selected, conditional on having been selected in 1985, is given by $P_i(89)/P_i(85)$. Combining these two possible outcomes, the probability of a taxpayer remaining in the panel in 1989 conditional on being sampled in 1985 is given by,

$$P_i(89|85) = \min\left(1, \frac{P_i(89)}{P_i(85)}\right) \qquad \text{A.2}$$

which will be equal to one if a taxpayer remains in the same strata or moves to a strata with a higher sampling rate, and less than one otherwise.

Consider, for example, a taxpayer selected in 1985 with a sampling probability of 1/400 whose sampling probability fell to 1/1000 in 1989. The probability that the taxpayer would be selected in 1989 and be included in our panel is 4/10. Alternatively, for a taxpayer whose 1985 sampling probability was one in 1985, but fell to 1/1000 in 1989, would have a probability of 1/1000 of being selected in 1989 and being included in our panel. Although the probability of being selected in 1989 is the same for these two taxpayers (i.e., 1/1000), the probability that they are included in our panel differs because they are sampled at different rates in 1985. Finally, a taxpayer who had a sampling probability of 1/1000 in 1985 that increased to 1 in 1989 would have a probability of 1 of being included in our panel.

Substituting equation A.1 into A.2, the probability that a taxpayer is sampled in both years and included in our panel, is given by,[3]

---

[3]Of course, this formulation does not account for the presence of CWHS returns in our sample. Since CWHS returns are sampled with certainty, the probability that they are selected in 1989 conditional on being selected in 1985 is one.

$$P_i(85)\, P_i(89|85) = P_i(85)\left[\min\left(1, \frac{P_i(89)}{P_i(85)}\right)\right] = \min\left(P_i(85), P_i(89)\right) \tag{A.3}$$

The weight is simply given by the inverse of the sampling probability, or,

$$w_i = \frac{1}{P_i} = \frac{1}{\min\left(P_i(85), P_i(89)\right)} = \max\left(\frac{1}{P_i(85)}, \frac{1}{P_i(89)}\right) = \max\left(w_i(85), w_i(89)\right) \tag{A.4}$$

where $w_i(85)$ and $w_i(89)$ are the ith taxpayer's 1985 and 1989 sampling weights, respectively. We then use weighted least squares, which is consistent and asymptotically efficient, to both correct for the endogenous sample selection and to extrapolate our sample to the population.[4] Both weighted and unweighted results are reported in the paper.

[4]Although, with our data some observations may have $P_i(85)$ or $P_i(89)$equal to zero, but are obviously not in our panel, we make no attempt to account for this in our estimation.