Joint Working Paper
of the
Staff of the Joint Committee on Taxation
and the
Congressional Budget Office

New Evidence on the Tax Elasticity of Capital Gains

Tim Dowd
Joint Committee on Taxation
Tim.Dowd@mail.house.gov

Robert McClelland
Congressional Budget Office
robert.mcclelland@cbo.gov

Athiphat Muthitacharoen
Congressional Budget Office
athiphatm@cbo.gov

June 2012

[Updated on August 9, 2012, with minor corrections to equation 3, to Table 6, and to references to Table 6 in the body of the paper]

JCX-56-12
CBO Working Paper 2012-09

The views expressed in this paper are those of the authors and should not be interpreted as those of the Congressional Budget Office or of the Joint Committee on Taxation. The authors would like to thank Gerald Auten, Jane Gravelle, Janet Holtzblatt, Pamela Moomau, Larry Ozanne, James Poterba, William Randolph, and participants in the Fall 2011 National Bureau of Economic Research session on Public Economics for helpful comments and suggestions.

ABSTRACT

This study uses a large panel of tax returns from 1999 to 2008 to investigate how taxes affect the decision to realize gains. The study distinguishes the persistent effect of tax changes from the transitory effect. Similar to earlier studies in the literature, we use the generalized Tobit model to address the sample selection problem and the endogeneity problem in the tax variables, but we improve the identification of the tax elasticity by using the presence of carryover loss as an exclusion restriction. We also control for the financial sophistication of taxpayers because that could be an important source of omitted variable bias. The preferred persistent elasticity estimate is -0.79, and the transitory estimate is -1.2. Those estimates are statistically significant and are robust to a number of sensitivity tests. Although we focus our examination on personal capital gains, we also compare the results of our model to results from the original model applied to contemporary data, estimate our model on subperiods, and estimate our model on other types of capital gains. We find that passthrough capital gains are highly sensitive to persistent tax changes, but gains from mutual fund distributions are extremely insensitive.

I. INTRODUCTION

The relationship between the marginal tax rate and the timing of capital gains realizations has been studied intensely. One reason for that intense interest is the apparent relationship between those two variables, as demonstrated in Figure 1. Realizations sharply increased prior to the increase in tax rates in 1987, and relatively low levels of realizations accompanied relatively high marginal tax rates, even as the S&P index rose, during the years 1987 through 1996. However, there are many factors that might explain this simple interpretation, such as the increase in the variety of financial products, the increase in realizations from partnerships and S corporations, and announcement effects.

An early econometric estimate of the response of capital gains to the tax rate is reported by Feldstein, Slemrod, and Yitzhaki (1980). Using a sample of tax returns from 1973, they estimate that, in response to a capital gains tax rate reduction, taxpayers with substantial holdings of corporate stock would increase their realizations by enough to raise their total taxes paid.[1] The study sparked a flurry of other research, some using cross-section data on individuals, others relying on aggregate time-series data. Auten and Cordes (1991) note that cross-section estimates of elasticities using data on individual observations tended to be greater than 1.0 in absolute value, but time-series estimates of elasticities using aggregate data tended to be between -0.5 and -0.9.

Although studies of how other forms of income respond to taxes find a similar range of uncertainty, the variability in capital gains estimates may also stem from the issues that complicate its study. Income from capital gains realizations may be timed much more easily than income from salary and wages; in principle, capital gains realizations may be put off indefinitely. In addition, the decision to realize gains and the amount realized may jointly depend on unobservable factors, confounding attempts to consistently estimate models of those decisions. Finally, because relatively few taxpayers realize gains, microdata from a random sample of taxpayers contain few observations with gains, and samples stratified toward high-income taxpayers require weights for consistent estimation. Applying different solutions to those problems and examining different time periods can lead to substantially different estimates.

Burman and Randolph (1994a,1994b) offer evidence that the disparity in estimated elasticities is caused by whether taxpayers view the changes in tax rates as a long-run “permanent” change or a short-run “transitory” change. Using a Type II Tobit on data for the years 1979–1983, they estimate an elasticity of “permanent” tax rates at -0.18, and an elasticity on “transitory” rates of -6.42. However, their estimates are very imprecise, so that their permanent elasticity of -0.18 is insignificantly different from both zero and -1.00. One likely cause of that imprecision is their use of the same set of explanatory variables to model both the decision to realize capital gains and the amount of gains to be realized. Since then, little additional research has been conducted.

---

[1] Feldstein, Slemrod, and Yitzhaki reported their results in a National Bureau of Economic Research Working Paper in 1978. That study and related work by those authors played a role in the enactment of the capital gains tax rate reductions of 1978. Even before that study, revenue estimators at the Joint Committee on Taxation and the Treasury Department had made smaller ad hoc adjustments to allow some response of realizations to changes in tax rates.

One exception is Auerbach and Siegel (2000), who estimate the Type II Tobit model of Burman and Randolph on individual tax data for the years 1986–1993. That study also suggests that imputation of the permanent tax variable may not capture important information about permanent rates. Using the Burman and Randolph imputation, Auerbach and Siegel find a permanent elasticity of -0.34, but with a modified formula they find a permanent elasticity of -1.72. The transitory elasticities are -4.91 and -4.35, respectively.

The purpose of this paper is to estimate the responsiveness of capital gains realizations on a panel of taxpayers followed over the period 1999–2008, the most recent period available. Those years include two major tax acts: the Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTTRA) and the Jobs and Growth Tax Relief Reconciliation Act (JGTRRA) in 2003. Because most capital gains realizations in the 1980s and early 1990s were personal realizations, rather than gains from passthrough entities or from mutual funds, we concentrate our analysis on personal capital gains realizations reported on line 8 of Schedule D. We also examine total realizations, as well as realizations from passthrough entities and mutual funds.

As in Burman and Randolph and in Auerbach and Siegel, we separately estimate the elasticities for long-run and short-run tax changes. We also use a Type II Tobit but improve the identification of the elasticities by adding a variable that affects the decision to realize gains but not necessarily the level of realization. In addition, we include variables that control for taxpayers' financial sophistication. Our approach measures the effect of an increase in tax rate that has persisted over the previous year, and is also expected to persist into the next year—thus, we use the term "persistent elasticities" to describe our results rather the term "permanent elasticities" used in other work. Using our preferred model specification, the persistent elasticity is estimated to be -0.79 with standard error of 0.11, and the transitory elasticity is estimated to be -1.20 with a standard error of 0.35. These elasticity estimates are robust to a number of sensitivity tests. In addition to personal capital gains, we find that capital gains from passthrough entities are very sensitive to long run tax changes (persistent elasticity of -1.95, and statistically significant) and that capital gains from mutual funds are quite insensitive (persistent elasticity of -.09, and statistically insignificant).

II. TAX TREATMENT OF CAPITAL GAINS

Accruing a capital gain does not itself generate liability, because gains are taxed only when they are realized through their sale. When the gain is realized, under the tax code it is considered income and is subject to taxation. The taxable amount is the difference between the price at which the asset was sold and the price for which it was purchased, minus adjustments for items such as commissions and tax depreciation deductions. The gain's taxable status (or lack thereof) and the applicable tax rate depend on factors such as how long the asset was held, whether the asset is an owner-occupied home, and whether the sale takes place after the death of the owner.[2]

[2] Starting in 1951, people selling their home and buying another were allowed to "roll over" any gain on the first home into the second, as long as the second home was of the same or greater value than the first home. Then starting in 1964, people age 65 or over were allowed a one-time exclusion of up to $20,000 on gains from the sale of their home. The exclusion amount was raised to $35,000 in 1976 and $125,000 in 1981; the age at which it became available was lowered to 55 in 1978. Those

Realizations of long-term capital gains—defined generally as those on assets held for more than a year—are taxed at rates lower than rates imposed on ordinary income. Short-term gains—those assets held for a year or less—are taxed at the same rate as ordinary income. States typically treat gains, long or short, as regular income. There are several reasons why it may be more efficient to tax long-term gains at a lower rate.

First, capital gains taxation encourages investors to "lock in" to a specific set of assets. Although the initial purchase of an asset may be optimal at the moment of purchase, unforeseen events will inevitably call for an investor to rebalance his or her portfolio of assets. Higher marginal tax rates on capital gains realizations increase disincentives to rebalance that portfolio through the sale of assets, and inevitably lead to some level of inefficiency in the portfolios of taxpayers. Second, lower tax rates on capital gains can encourage entrepreneurs to invest time and effort into starting, building up, and selling new businesses. Third, because the taxable amount of realizations is not indexed for inflation, reducing the tax on those realizations also reduces the tax paid on items that are nominal gains but real losses. Finally, lower tax rates on capital gains encourage savings, which may increase the long-term productivity of the nation.

Since 1979, tax rates on long-term capital gains have risen and fallen while the overall tax rate structure has simplified. Before the Tax Reform Act of 1986, taxpayers were allowed a 60 percent deduction of any net capital gain. As a result, the maximum tax rate on capital gains was 20 percent (0.4 x 50 percent, the maximum ordinary rate). In addition, there were 14 ordinary tax brackets for married individuals and heads of household and 15 ordinary tax brackets for single taxpayers. The multiple brackets combined with the 60 percent deduction resulted in considerable variation in tax rates on capital gains. Between 1986 and 1998, individual income was subject to ordinary tax rates up to a maximum statutory tax rate of 28 percent, and there was no special tax treatment for income from realizations of gains. Moreover, there were at most five tax brackets (declining to three in 1987 and two in subsequent years). From mid-1997 until mid-2003, most long-term capital gains were subject to rates of 10 percent and 20 percent. In mid-2003, JGTRRA reduced the tax rates on capital gains to a bottom rate of 5 percent (0 percent in 2008) and a top rate of 15 percent through December 31, 2008. Public Law 109-222, enacted in 2006, extended the 0 percent and 15 percent rates through December 31, 2010. In 2010, the Tax Relief, Unemployment Insurance Reauthorization, and Job Creation Act of 2010 extended those rates through December 31, 2012.

The tax structure has similarly simplified at the state level. For example, in 1986 California had 12 tax brackets and New York had 13 brackets, but by 1999 the numbers had been reduced to six and five brackets, respectively.

Because taxes are paid upon realization of a capital gain rather than as accrued, taxpayers can in effect choose when they pay their capital gains taxes. For instance, the Tax Reform Act of 1986 raised the top statutory tax rate on capital gains from 20 percent to 28 percent, effective at the beginning of 1987. Probably in anticipation of that increase, investors realized substantial gains in 1986 ($327.7 billion in

---

provisions stayed in effect until legislation in 1997 replaced them with an exclusion of $250,000 (or $500,000 for joint returns) that could be claimed if the seller had owned the home for at least two years and had used it as a primary residence for two out of the previous five years.

1986, compared with $172.0 billion in 1985).[3] Then, in 1987, realizations fell by almost as much, returning to a level comparable to the level before the tax increase. These large swings in realizations suggest that investors may be very responsive to rate changes immediately before or immediately after the change.

Finally, the treatment of capital losses is different from the treatment of capital gains. Preferential capital gains rates are applied after netting out any short- or long-term losses. Net losses can be used to offset up to $3,000 of ordinary income. Any remaining loss may be carried forward to the next taxable year. Thus, net losses up to $3,000 receive a tax subsidy at ordinary tax rates. Moreover, carryover losses from the previous year and current-period losses both may be used to offset current-year gains. The ability to offset gains with losses means that some taxpayers who otherwise would be subject to a positive tax will be able to have a zero tax rate on their gains.

III. EMPIRICAL MODEL

To facilitate comparison with previous empirical studies on the capital gains tax elasticity, we first review the relationship between capital gains realizations described in Burman and Randolph (1994a,1994b) and Auerbach and Siegel (2000). This relationship is modeled as:

$$\ln g_{it} = \gamma_1(\tau_{it} - \tau_{it-1}) + \gamma_2\tau_{ip} + \gamma_3(\tau_{it} - \tau_{ip}) + X_{it}\gamma_4 + \varepsilon_{2it} \qquad (1a)$$

where *i* indexes individuals, *t* indexes years, and the γs are conformable vectors of coefficients. The dependent variable, $\ln g_{it}$, represents the natural log of capital gains (measured as the net long-term personal gains before prior-year carryover losses). The tax variables τ_{it-1} and τ_{it}, are the combined federal and state marginal tax rate on long-term capital gains for their respective time periods. The tax variable τ_{ip} is the long-run, permanent tax rate, and (τ_{it} - τ_{ip}) is the transitory rate. The control variable vector X_{it} includes a variety of wealth, income and demographic variables that will be explained below. In equation (1a), the effect on the capital gains realizations of a permanent increase in the tax rate is represented by γ_2. The effect on the capital gains realizations of a transitory increase in the tax rate this year that is expected to disappear next year is given by $\gamma_1 + \gamma_2 - \gamma_3$.

Rather than describe the tax rate as the sum of permanent and transitory components, we model the immediate, transitory response to a tax change and the long run, persistent change as:

$$\ln g_{it} = \gamma_1(\tau_{it} - \tau_{it-1}) + \gamma_2\tau_{it} + \gamma_3(\tau_{it+1} - \tau_{it}) + X_{it}\gamma_4 + \varepsilon_{2it} \qquad (1b)$$

In equation (1b), the effect on the capital gains realizations of a persistent increase in the tax rate is represented by γ_2. The coefficient γ_2 measures the effect of an increase in tax rate, holding changes relative to the previous year and the next year constant. This occurs when there has been an increase in tax rate that has persisted over the previous year and is also expected to persist into the next year. The effect on the capital gains realizations of a transitory increase in the tax rate this year that is expected to disappear next year is given by $\gamma_1 + \gamma_2 - \gamma_3$.

---

[3] U.S. Department of the Treasury (2010).

Note that both equations (1a) and (1b) can be rearranged such that the tax variables enter the equation as $\beta_1\tau_{it-1} + \beta_2\tau_{it} + \beta_3\tau_{it+1}$. In that specification, the sum of all tax coefficients ($\beta_1 + \beta_2 + \beta_3$) is equivalent to the coefficient γ_2 in (1b) and therefore it is the effect of a persistent change in tax rate. We therefore refer to persistent elasticities rather than permanent elasticities when discussing our results. The coefficient β_2 is equivalent to $\gamma_1 + \gamma_2 - \gamma_3$ in (1b) and therefore it is the effect of a transitory change in the tax rates.

We take into account both the decision to realize capital gains and the amount of capital gains realized. Our full empirical specification is:

$$I_{it}^{*} = \alpha_1\tau_{it-1} + \alpha_2\tau_{it} + \alpha_3\tau_{it+1} + X_{1it}\alpha_4 + \varepsilon_{1it} \qquad (2)$$

$$\ln g_{it} = \beta_1\tau_{it-1} + \beta_2\tau_{it} + \beta_3\tau_{it+1} + X_{2it}\beta_4 + \lambda_{it} + \varepsilon_{2it};\ \text{if } I_{it}^{*} = 1 \qquad (3)$$

where the indicator I_{it}^{*} is a latent variable representing the decision to realize long-run capital gains, the αs and βs are conformable vectors of coefficients, λ_{it} is the inverse Mills ratio and the control variable vector X_{1it} is a superset of X_{2it}. This framework allows for the possibility that the effects of regressors are different between the extensive margin of whether to realize capital gains and the intensive margin of the level of gains to be realized.

The progressivity of the individual income tax schedule makes it very likely that the capital gains tax rate variables are affected by the amount of realized capital gains. The resulting endogeneity problem requires us to find instruments that are strongly correlated with the current and future tax variables τ_{it} and τ_{it+1} but are uncorrelated with the level of realized capital gains.[4]

We use the "first-dollar" marginal tax rate variables (τ_{0it}) and the maximum combined federal and state tax rate variables (τ_{sit} and τ_{sit+1}) as instruments for the two endogenous tax variables. Because those variables do not depend on any characteristic of the taxpayer other than his or her state of residence, they are exogenous.[5] The first-dollar marginal tax rate variable is computed with the amount of realized gains set to zero. However, it is still possible that the first-dollar rate is endogenous if taxpayers time their realizations to coincide with large deductions that lower the tax rate. To guard against that problem, we calculate the first-dollar tax variable with the following elements set to zero: state income taxes, property and sales taxes, charitable contributions and passive and active losses from partnerships and S corporations.

To address the possibility of a selectivity bias caused by taxpayer decisions to realize a capital gain, we employ the generalized Tobit model developed by Lee et al. (1980). It consists of four steps. First, we regress the two endogenous tax variables τ_{it} and τ_{it+1} on instruments and other regressors to obtain their fitted values $\hat{\tau}_{it}$ and $\hat{\tau}_{it+1}$. Next, we use a Probit model to estimate the decision to realize gains on the full sample (with $\hat{\tau}_{it}$ and $\hat{\tau}_{it+1}$ replacing τ_{it} and τ_{it+1}, respectively). In this step we also use the predicted values from the Probit estimation to compute the inverse Mills ratio. Third, we use the subsample of realizers to re-estimate the fitted values of tax variables. The regression includes the inverse Mills ratio

[4] The tax rate in the previous year τ_{it-1} is a predetermined variable.
[5] We assume that taxpayers do not move to a low tax state in anticipation of realizing a large gain.

calculated from the second stage. Finally, we use Ordinary Least Squares on the subsample of realizers to estimate the level equation. The regression includes the fitted tax variables computed from the third step and includes the inverse Mills ratio. Because the asymptotic variance of this estimator is unknown, we use a standard bootstrap method to estimate the standard errors.

A taxpayer's decision to realize capital gains reflects economic factors as well as his or her life cycle savings and consumption decisions. The vector of control variables X_{it} thus includes wealth and income variables as well as variables that reflect demographic characteristics such as age, family size, marital status, and regions. In order to account for taxpayers' financial sophistication, we include in the model the number of short-term transactions that taxpayers make. Dummy variables to indicate various losses (any personal long-term loss realization, net losses from the sale of a business or business asset, net losses from passthrough entities, and net short-term losses) are also included because those losses may affect the realization decision.[6]

Technically, we do not need an exclusion restriction (a variable that affects the decision to realize gains but has no effect on the level of realization) for the estimate to be identified. However, if there is not much variation in the sample, the inverse Mills ratio could be well approximated by a linear function of X_1. When $X_1 = X_2$, this correlation can introduce severe collinearity among regressors in the second stage, leading to large standard errors of the estimate. Without the exclusion restriction, the identification is based entirely on the functional form imposed by the Probit model.[7]

Our preferred model uses a dummy variable indicating whether the taxpayer recorded a carryover loss as an exclusion restriction. A taxpayer who carries a loss can use that loss to offset any capital gains that he realized. To the extent that the amount of realized gains is smaller than the amount of loss carryover, this lowers his capital gains tax rate to zero. Taxpayers with large carryover losses have an incentive to realize large gains, while those taxpayers with small losses only have an incentive to realize small gains. The mere presence of a carryover loss tells us that the taxpayer has an incentive to realize capital gains without telling us the magnitude of that incentive. We also perform a sensitivity analysis with respect to the use of this exclusion restriction.

Following Auerbach and Siegel (2000), we calculate the persistent elasticity as:

$$\varepsilon_{pit} = \hat{\tau}_{it+1}[\beta_1 + \beta_2 + \beta_3 + (\alpha_1 + \alpha_2 + \alpha_3)\lambda_{it}] \quad (4)$$

where λ_{it} is the inverse Mills ratio evaluated at $\hat{h}_i + \hat{\sigma}_{12}$, the predicted value of the selection equation (2) plus the covariance of the error terms in equations (2) and (3).[8,9] The transitory elasticity is estimated with an analogous equation that excludes future and lagged tax coefficients. We estimate the elasticity

[6] Some of these loss variables may reflect components that are endogenous. We perform a sensitivity test of our results to exclusion of these variables.
[7] See Vella (1998) and Wooldridge (2010) for further discussion of this issue.
[8] See Burman and Randolph (1994b) for the derivation.
[9] Auerbach and Siegel (2000) differ from Burman and Randolph in that they use the fitted value of the future tax rate as the permanent variable and thus use it in their elasticity calculation. We follow their practice here in order to facilitate the comparison. We also calculate the elasticity using the average value of the past tax rate, the fitted current tax rate and the fitted future tax rate. The difference is negligible.

separately for each return and then calculate a weighted mean, using as weights the product of the population weight and the amount of gains realized.

IV. DATA

Our data come from a unique 10-year panel of federal tax returns, over the years 1999–2008, created by the Internal Revenue Service's Statistics of Income (SOI) Division.[10] The data are a stratified random sample of returns selected in tax year 1999. The data include each item on the federal tax form 1040 and their attendant schedules, including Schedule D (capital gains and losses). SOI has linked the data to Social Security Administration records to determine the dates of birth of the primary taxpayer, secondary taxpayers, and the first four dependents in the file. In addition, we linked the data to a 1999 SOI study of occupation and industry. We use Jon Bakija's tax calculator (Bakija 2009) to generate federal and state marginal tax rates by year for each observation.

The panel is a stratified random sample of tax returns that oversamples high income tax returns. The set of taxpayers is taken from the 1999 cross-section sample, which contained 176,966 returns. The panel subsample contains 88,123 returns from 21 income stratifications,[11] weighted to represent 123 million tax returns. The stratification by income includes sampling rates ranging from 0.05 percent to 100 percent. Each taxpayer and his or her spouse who were on a selected tax return in 1999 or who filed late for tax year 1999 in tax-processing years 2000 or 2001 are included in the panel for each year that they filed a return. Dependents and new entrants to the panel through marriage are not followed separately from the original panel member. Because of the complexity of the weighting procedures in handling taxpayers whose weighting shifts dramatically over time, we restrict our sample to taxpayers that did not experience a change in marital status over the 10-year period (dropping about 19,000 returns in 1999) or who had a change in the value of their weighting of more than 5 percent in 1999 (dropping about 174 returns).

The lefthand panel of Table 1 shows the sample sizes and total capital gains from all sources, including short-term, long-term, and passthrough gains, for the unrestricted sample. As can be seen, there was $559.8 billion in total capital gains realizations in 1999. Realizations of capital gains fluctuate considerably over this period, from a low of $234 billion in 2002 to a high of $752 billion in 2007. The restricted sample follows the same pattern with a low in 2002 and a peak in 2007. Over the course of the panel's 10 year period, there is approximately 14 percent attrition of the unweighted number of returns in the unrestricted sample and 16 percent in the restricted sample.[12]

The final column in each panel of Table 1 shows the positive long term capital gains realizations from the sale of or exchange of a capital asset, excluding capital gains and losses reported on lines 11–14 of Schedule D. The excluded amounts reported on those lines are items from the sale or exchange of a

[10] See Weber and Bryant (2005) for a detailed description of the stratification and selection process of the 1999 edited panel.
[11] Weber (2006) reports that the subsample contains 83,434 returns. For the unrestricted sample, we report the number of returns after adjusting for split records in the case of divorce in any of the subsequent years, by creating duplicate returns in each year prior to the divorce and splitting the weights. This maintains the 1999 income totals. The restricted sample does not include these returns because of the filer's change in marital status.
[12] Bryant (2008) reports that there is 15 percent attrition of primary and secondary taxpayers over the period 1999-2005.

capital asset used in a trade or business, involuntary conversions, amounts received from passthrough entities or mutual funds, and loss carry forwards.[13] Our dependent variable is the positive value of the sum of long-term gains excluding swaps, distributions, partnerships and S corporations, and involuntary conversions. In this paper, we focus on long-term gains from the sale of capital assets that are personal in nature and are reported on line 8 of Schedule D, including stock held for investment or the gain from the sale of primary residence in excess of $250,000 ($500,000 in the case of joint return). As shown in Figure 2, personal capital gains realizations made up a significantly larger portion of total realizations in the 1980s and early 1990s than in later years. In 1984 and 1985, personal capital gains represented approximately three-quarters of total long-term realizations. Therefore, we believe that concentrating on line 8 totals provides a better comparison with earlier analyses. However, personal capital gains averaged only slightly more than a third of total realizations for the period 1999–2008, with capital gains realizations attributable to passthrough entities experiencing the largest increase – rising from approximately 10 percent of total capital gains in the late 1980s to almost 30 percent of capital gains in 2008. We therefore examine gains declared on lines 11–14 of Schedule D separately.

The restricted data have a total of 558,525 observations. We further restrict these observations for the following cases: We drop all dependent returns; we keep only those returns with the age of the primary taxpayer in 1999 between 18 and 120; we drop any return with a calculated total capital gains marginal tax rate less than zero or greater than 0.4; and we drop any observation with a missing value (as opposed to a zero) for any variable needed in the estimation process.[14] The combination of these restrictions results in a sample with 341,793 observations, with 70,377 reporting a long term capital gain on line 8 of Schedule D.

We create the marginal tax rate variables using Jon Bakija's tax calculator (Bakija, 2009). The tax calculator has detailed information on the federal tax structure and on each of the 50 states and the District of Columbia. The calculator has information for each year 1999-2007. We use the 2007 values for 2008 tax rates.

The exogenous variables that form the vector *X* include demographic and economic variables that may be correlated with capital gains.[15] Family size is the number of personal and dependent exemptions, and marital status is determined from the filing status of the taxpayer. We include dummy variables for taxpayer age brackets (such as 30–39, 40–49, and so on) created from an age variable provided from Social Security records matched to the data. The gender of the head of household is also included. We include dummy variables for region, which are derived from the taxpayer's state of residence.[16] We also

---

[13] The excluded amounts on lines 11 are from the sale of property used in a trade or business, amounts from involuntary conversions from loss due to casualty or theft, amounts from swaps and straddles, or like-kind exchanges. The excluded amounts from line 12 are amounts from partnership, S corporation, and other passthrough entities. Line 13 excluded amounts are distributions from mutual funds. The line 14 exclusion eliminates capital loss carry forwards from prior years.

[14] About three percent of the deletions were due to problems with the tax variables.

[15] In the instances in which explanatory variables in log form take a value of 0 we replace the natural log with 0. We further add to the regression a dummy variable equal to 1 for all instances when the variable equals 0, and 0 otherwise.

[16] We use four regions: Northeast (CT, ME, MA, NH, NJ, NY, PA RI, and VT); South (AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, VA, WV, and TX); Midwest (IA, ID, IL, KS, MI, MN, MO, ND, NE, OH, SD, and WI); and West (AK, AZ, CA, CO, HI, ID, MT, NM, NV, OR, UT, WA, and WY).

include dummy variables for year in order to account for the aggregate shocks that affect all taxpayers in the same way across years.

We include in our model two imputed measures of unrealized capital gains—the total amount of unrealized gains and the proportion of unrealized gains in stocks. Both measures are imputed using Survey of Consumer Finance (SCF) data for 2001, 2004, and 2007. We use SAS code provided by the Federal Reserve Board to create variables total unrealized capital gains and unrealized capital gains from stocks, bonds, and mutual funds.[17] The natural logs of these variables are regressed on demographic variables and tax variables common to both the SCF and our panel, and we use the estimated coefficients to impute the log of unrealized gains to our data. (See Table A1 in the appendix for the regression results.) We also include the lagged values of business losses and rent losses, taken from tax data, as measures of business wealth.

As in Burman and Randolph (1994a, 1994b) and Auerbach and Siegel (2000), we estimate permanent income by regressing the natural log of the average of real positive income over all years on demographic characteristics. Then the regression estimate is used to impute annual permanent income based on lagged values of the regressors. Transitory income is measured as the difference between current income (the sum of positive income from all sources except gains) and permanent income.

Because a taxpayer may use up to $3,000 of his capital loss carryover to offset against his ordinary income, we create a dummy variable for the presence of capital loss carryover in excess of $3,000. In order to account for financial sophistication of taxpayers, we include dummy variables for the number of short-term realizations that a taxpayer made in the prior year: 0, 1–34, 35–167, 168–1001, and more than 1001. These brackets are based on the sample distribution of those with positive lagged number of short-term realizations.[18] We also include dummy variables for any long-term loss realization, net losses from the sale of a business or business asset, net losses from passthrough entities, and net short-term losses.

Table 2 shows the mean values of the variables from the sample of 341,793 observations. As described above, the panel is a stratified sample that oversamples taxpayers with high incomes. Comparison of the weighted and unweighted mean long-term gains highlights the effect of the sample stratification. The average long-term gain reported on line 8 of Schedule D is $2,136. However, if we do not use the weights, the average jumps to $466,665. Similarly, for the sample of observations reporting a long-term capital gain, the weighted average among realizers is $36,885 and the unweighted average is almost $2.5 million.

Finally, note that the sample of realizers is different from the overall sample. Compared with the population of all returns, on average the sample of realizers tends to have higher income, has a higher marginal tax rate, is roughly 10 years older, is more likely to be male and married, and is less likely to have children.

---

[17] We use bulletin_macro.sas, available at http://www.federalreserve.gov/pubs/oss/oss2/bulletin.macro.txt

[18] The second group corresponds to those between 0 and the 75th percentile, the third group corresponds to those between the 75th and 90th percentiles, the fourth group corresponds to those between the 90th and 99th percentiles, and the fifth group corresponds to those above the 99th percentile.

V. RESULTS

We implemented the empirical approach described in section III using a double-hurdle model and test for the presence of sample selection in who realizes a gain. Because we find evidence of sample selection, we use the generalized Tobit model.

Estimates from the Double-Hurdle Model

We start our regression analysis with an estimate of a naive double hurdle model, in which we assume that, conditional on our explanatory variables, the decision to realize a long-term capital gain is independent of the level of realization. The estimation method is similar to the generalized Tobit method described earlier, except that the inverse Mills ratio is omitted from the third and fourth steps. Both the criterion function and the level equation include the dummy variable for the presence of carryover loss. The results are shown in Table 3.

The three tax variables as (τ_{it-1}, τ_{it}, and τ_{it+1}) constitute the main focus points of our analysis—their coefficients capture the persistent and transitory effects of changes in the capital gains tax rate. The sum of all the tax coefficients measures the effect that an increase in tax rates has on capital gains realizations, holding changes relative to the previous year and the next year constant. Using equation (3), the elasticity of capital gains realization with respect to the persistent change in the tax rate is -0.69, with a standard error of 0.10. On the other hand, the effect on capital gains realizations of a transitory increase in the tax rate this year that is expected to disappear next year is measured by the coefficients on the current tax rate $\hat{\tau}_t$. The transitory elasticity is estimated to be -1.02, with a standard error of 0.32.

Equation (4) can be decomposed into the elasticity for the decision to realize and the elasticity for the amount realized. Table 2 lists the weighted mean marginal tax rate for those realizing gains, and Table 3 lists the estimated coefficients. Multiplying the sum of the coefficients of the tax variables by the mean marginal tax rate yields an approximation of the elasticity of average realized gains from a persistent tax change: -0.56. The elasticity of the realization decision is therefore about -0.12. An analogous decomposition of the transitory elasticity shows that the elasticity of average amount of gains is -0.83 and the realization elasticity is -0.19. These results suggest that tax rates influence the amount of gain that taxpayers choose to realize more than those tax rates influence whether taxpayers choose to realize a gain.

Note that our estimate of transitory elasticity is substantially smaller in absolute magnitude than prior estimates. One might expect taxpayers to be less responsive to temporarily low rates for at least three reasons. First, as described previously, the tax rate structure has simplified, and for most taxpayers flattened, considerably since the time period studied by Burman and Randolph (1994a, 1994b). There are never more than two applicable brackets in our sample, as opposed to 14 to 15 brackets in 1979. Consequently, in our sample temporary changes in income are less likely to result in changes in tax rates. Second, more taxpayers are in the top bracket than in prior years, including the period studied by Auerbach and Siegel (2000). We show this in Table 4, where we compare results for 1993 (found in Table 13 and Table 14 of Burman and Ricoy 1997) with comparable calculations on total realizations of long-run capital gains for 2007. In 1993, slightly less than 60 percent of capital gains were realized by those with incomes greater than $200,000, and they faced an average tax rate of 23.9 percent—and the top rate was 28 percent. In 2007, 84 percent of capital gains were realized by those with incomes greater than $200,000, and they faced an average tax rate of 14.8 percent—and the top rate was 15 percent.[19] The table makes it clear that there are more gains being realized at the top rate in 2007 than in 1993, which is likely due to a lack of opportunity to realize them at a lower rate. Finally, unlike the 1986 rate change, the federal rate reduction in JGTRRA was not known in the year prior to its enactment, giving taxpayers fewer opportunities to delay realizations until the drop in rates.

Coefficients on the other control variables generally conform to expectations. Capital gains realizations are significantly and positively related to both permanent and transitory incomes. Imputed unrealized gains have a large and positive effect on the level of realizations, although they have a small negative effect on the criterion function. The share of wealth held in stocks is included in the model because of the ease with which stock can be liquidated relative to other assets, such as real estate. The extremely small standard error of the share in wealth variable in the Probit stage suggests that friction is an important determinant in explaining the realization decision.

The categorical variables for age show that the probability of realizing gains increases with age; however, the average size of gains is less clearly related to age. In both the criterion function and the level equation, the categorical variables for the lagged number of short-term realizations have coefficients of increasing magnitude—except, surprisingly, for the last group (with the largest number of transactions). Those magnitudes suggest that up to a point, the more actively one participates in short-term realizations, the higher the probability of realizing a long-term gain and, other things being equal, the larger the gain realized. The standard errors of the coefficients on the last group are noticeably higher than those of the other groups, possibly because there are too few observations to precisely estimate their effects.

The coefficients on dummy variables for long-term loss realization, net losses from passthrough entities, and net short-term losses are positive and significant in both equations (2) and (3). The coefficients on the dummy variables indicating net losses from business sales are positive in both stages but are significant only in only the level equation—possibly because of too few observations reporting such losses. The statistical significance of those loss variables in the level equation is surprising in that the mere presence of a loss, without regard to size, appears to be correlated with larger average realizations. The dummy variable for having carryover loss, however, is positive and significant in the Probit stage but is insignificant in the level equation. It is also predetermined, which suggests that the significance of the other loss variables may be due to endogeneity—taxpayers may be simultaneously deciding to realize relatively large gains and claim a loss. We explore this issue in the sensitivity analysis section.

The estimates in the double hurdle specification are consistent only if the assumption about the independence between the decision to realize and the level of realization holds. If those two decisions are correlated, as one might suspect, the estimates are inconsistent. To account for this sample selection problem, we estimate the generalized Tobit model as discussed earlier. We use the dummy variable that indicates carryover loss as the exclusion restriction in the Probit stage. Above we argue

---

[19] Burman and Ricoy point out that the statutory maximum rate is not reached because "some taxpayers have losses and deductions that lower their taxable income (before capital gains) below the threshold". That rate was 28 percent in 1993, and the same reasoning applies to the maximum rate of 15 percent in 2007.

that a taxpayer having a carryover loss does not affect the amount of gains realized by a taxpayer. Nevertheless, we add the carryover loss variable to the second stage. The statistical insignificance of that variable supports the idea that it does not affect the amount of gains realized. Approximately 35% of the observations in any given year used a carryover loss to offset other gains.

Estimates from the Generalized Tobit Model

Table 5 shows results from the generalized Tobit model. The coefficient on the inverse Mills ratio is positive and statistically significant, indicating that we need to account for the sample selection in order to avoid bias and inconsistent estimates. Because the coefficient is an estimate of the covariance of the error terms of the two stages, our results suggest that—conditional on the explanatory variables—the larger the gain, the greater the probability that it is realized.

Both persistent and transitory elasticities are slightly greater in absolute magnitude than those from the double-hurdle specification. The persistent elasticity is estimated to be -0.79, with a standard error of 0.11, while the transitory elasticity is estimated as -1.20, with a standard error of 0.35. Comparing the persistent elasticities between the two models, the amount of bias in the persistent elasticity estimate appears to be on the order of 0.10.

Decomposing the elasticities into average realizations and the realization decision, the persistent elasticity of average realizations is -0.70 and the elasticity of realizations is -0.09. The elasticities for transitory changes are -1.06 and -0.14. In both transitory and persistent changes, therefore, it appears that tax rates play a small role in the decision to realize capital gains. That accounts for the relatively small amount of sample selection bias we find, even though the selection effect is statistically significant.

Accounting for selection bias doubles the elasticity of gains with respect to persistent income, and it raises the elasticity of gains with respect to transitory income by about 50 percent. The proportion of gains held as stocks becomes much more pronounced and statistically significant. The effect of age on realized gains becomes clearer, and for taxpayers over 40 follows the same pattern as the effect on the probability of realizing a gain. The number of short-term transactions has the same effect on average realizations, although the increase in average gains with the number of transactions becomes even more pronounced.

VI. SENSITIVITY ANALYSES

We explore the robustness of our results to a number of alternative model specifications and assumptions. Table 6 reports the key results. More complete results are provided in the appendix.

Unweighted Estimates

As indicated earlier and highlighted in Table 2, our data come from a stratified random sample of tax returns. The sample design specifically oversamples high-income taxpayers, including those with income from capital gains. The introduction of this non-random sampling process complicates the estimation of capital gains realizations. Hausman and Wise (1981) show that the proper treatment of the non-random sample selection process (when the sampling structure within the strata is known) is to perform weighted least squares, or maximum likelihood procedures may be used to correct for the endogeneity of the sample design. Consequently, we use population weights to account for the stratification of the sample. The weights are the inverse of the sampling rate for each of the 21 substrata.

Minarik (1984) points out that consistent estimation of the effect of taxes on capital gains requires the use of weights because the stratification is based on total income and thus is correlated with realizations of capital gains. In a sensitivity analysis, he shows that weighted regressions result in substantially smaller elasticity estimates than are produced by unweighted regressions.[20] The standard errors for the weighted regressions are not calculated, but it is reasonable to suspect that the standard errors are larger in this case.

Although the estimates derived from unweighted regressions are inconsistent, it is worthwhile to understand the sensitivity of the results to the stratification of the sample. To that end, in Table 6 we report the results from the model in the previous section, but without the use of weights. The persistent elasticity increases from -0.79 to -1.13 (unweighted). The transitory elasticity changes from -1.20 to -1.60 (unweighted). In both cases, the standard errors are smaller in the unweighted regressions.

Allowing Heterogeneous Wealth Effects Across Time

Aggregate shocks, such as those from changes in the stock market, are likely to have large effects on the stock of unrealized capital gains, and these effects are likely to be heterogeneous for taxpayers with different levels of wealth. Although our model specification includes dummy year variables to control for those aggregate shocks, they are constrained to have similar effects on gains realizations for every taxpayer in the same year. To allow for the heterogeneous effects, we interact the imputed unrealized gains variable with the year dummy variables. As shown in Table 6, both persistent and transitory elasticities change very slightly from the base model without the interaction terms.

Possible Endogeneity of the Tax Variables

Gravelle (2010) suggests that the -1.72 elasticity estimated in Auerbach and Siegel (2000) may be due to transitory components in the first-dollar tax rate. Although we define the first-dollar tax rate to minimize that concern, remaining transitory components might still bias our estimates away from zero. We conduct two tests of that problem.

First, we re-estimate our model without the first-dollar tax variable, also dropping τ_{it-1} and τ_{it+1}.[21] In this case, the coefficient on τ_{it} is the persistent effect of a tax change and its instrument, the maximum state and federal tax rate, does not have the endogeneity problem that may affect the first-dollar tax variable. The resulting estimated elasticity of -0.735 is nearly identical to the estimate in Table 5. We then re-estimate our model with τ_{it-1}, τ_{it}, and τ_{it+1}, using as instruments $\tau_{\mathrm{sit-1}}$, τ_{sit}, $\tau_{\mathrm{sit+1}}$. Although the instruments are exogenous, their obvious correlation should lead to larger standard errors than in other models. In

---

[20] Feldstein et al. (1984) attribute this decline to heterogeneity of elasticities across taxpayers, because they suspect that weighted regressions emphasize the responses of low-income taxpayers, who they believe are less sensitive to tax rates than high-income taxpayers. We address the points of both Minarik and Feldstein et al. by using population weights in the regression analysis and the product of population weights and capital gains realizations in calculating mean elasticities.

[21] We would like to thank William Randolph for suggesting this approach.

this case, the elasticity is estimated to be -0.682, and, as expected, the standard error is nearly triple the standard error in our preferred model. In sum, the potential endogeneity of the first-dollar tax rate does not appear to be a problem.

Possible Endogeneity of the Four Loss Variables

As mentioned previously, the significance of the four dummy loss variables (indicators for personal long-term losses, net business losses, net pass-through losses, and net short-term losses) in the level equation may be due to their endogeneity. If so, their coefficients are biased, and that endogeneity could bias other coefficient estimates as well. One method for exploring this possibility would be to find exogenous variables correlated with losses and uncorrelated with the error term, and then to use two-stage least squares to conduct a Hausman test. However, a simpler threshold test is to check the magnitude of the problem by dropping all the potentially endogenous loss variables from both equations. We include the carryover loss variable in the first stage because it is predetermined. The results show only a minor change in elasticities, suggesting that there is little problem caused by the possible endogeneity of the loss variables.

Including the Dummy Carryover Loss Variable in the Level Equation

Use of identical sets of explanatory variables in equations (2) and (3) leads to coefficient estimates being identified through the functional form imposed by the Probit model. Ideally, equation (2) will contain explanatory variables that influence the decision to realize capital gains but do not determine the amount of those gains. Those variables would reduce the collinearity and thus increase the sampling variation of the estimated coefficients. However, to further investigate the effects of the loss variables, we examine the exclusion of the carryover loss variable in the level equation.

The results of adding carryover loss in the level equation are shown in the eighth row of Table 6. Although the carryover loss variable is insignificant in the level equation of the double-hurdle model, it becomes significant when the inverse Mill's ratio is included.[22] In this model, the permanent elasticity and transitory elasticity are each slightly larger in absolute value than those in the model in Table 5.

It is possible that taxpayers with relatively large amounts of carryover losses engage in multi-year tax-minimization strategies. This may explain the significance of the carryover loss variable in the level equation. To investigate this issue, we replace the dummy carryover loss variable with two dummy variables for whether the carryover loss amount falls into the following groups: $3,000–$15,000 and above $15,000.[23] We found that both dummy variables are positive and significant in the Probit stage but only the second group's dummy variable (above $15,000) is significant in the level equation. This is consistent with our hypothesis that those with relatively large amount of carryover loss may employ multi-year strategies. When the two dummy variables are added to the model, both persistent and transitory elasticities are virtually identical to the comparable model that has only one dummy variable for carryover loss.

---

[22] The coefficient on the carryover loss variable is 0.26 with standard error of 0.07.

[23] $15,000 is approximately the 70th percentile of the weighted carryover loss distribution among taxpayers with positive amounts of carryover loss.

VII. ADDITIONAL ANALYSES

Because our model extends work by previous authors, we compare our estimates made with the previous models on contemporary data. We also examine the variation over time in the elasticity over time, and apply our model to alternative types of capital gains realizations.

Comparison with Prior Methods

In our analysis of data from 1999 through 2008 the persistent elasticity is -0.792 and the transitory elasticity is -1.196. But as described above, Burman and Randolph (1994a,b) estimate a permanent elasticity of -0.18 and a 'transitory' elasticity of -6.42 for the years 1979-1983 while Auerbach and Siegel's modification lead to permanent and transitory elasticities of -1.72 and -4.35 for the years 1986-1993. Here we apply the previous methods to our data to decompose the differences between their results and ours into differences due to methodology and differences due to data.

Table 6 provides elasticity estimates using the unweighted and weighted versions of those methods. Burman and Randolph's unweighted method leads to permanent and transitory elasticity estimates of -0.584 and -2.697. Applying weights to their method almost doubles the permanent elasticity while slightly increasing the transitory elasticity. However, Table 9 also shows that the weighted estimation method yields a standard error so large that the permanent elasticity estimate is not statistically significant. In Burman and Randolph's study the equivalent use of weights resulted in only a small change in the elasticity estimate.

Applying Auerbach and Siegel's method leads to permanent elasticities of -1.272 (unweighted) and -1.000 (weighted). In the weighted regressions, both the Burman-Randolph method and the Auerbach-Seigel methods lead to similar elasticities greater than one in absolute value. The additional variables we include in our regressions reduce the estimated elasticity to -0.792.

Estimates Using Subperiods

In our main analysis, we use the data from 1999–2008. However, the length of the panel and the large number of taxpayers in the sample allow us to precisely estimate the elasticities over shorter time periods. Here we examine shorter time periods similar to previous analyses. We break our sample into four subperiods 2000–2001, 2002–2003, 2004–2005, and 2006–2007. As shown in Table 8, estimates from the first two sub-periods (2000–2001 and 2002–2003) are similar to the main results using all available years (1999–2008). The 2004–2005 period has a much higher persistent elasticity, and the 2006–2007 period has a much lower persistent elasticity. However, the 95 percent confidence interval of the persistent elasticity for all subperiod estimates approximately contains our point estimate from the full period analysis (1999 – 2008).

Estimates Using Alternative Capital Gains Variables

Until this point, we have focused on personal capital gains. We now apply our base model specification to alternative types of capital gains realizations. First, we analyze net long-term gains from sales of businesses or business assets. The estimated elasticities of those gains are listed in column 2 of Table 9. The persistent elasticity is -0.79 and the transitory elasticity is -2.21. Although the persistent elasticity of businesses or business assets is similar to that of personal capital gains, the transitory elasticity is almost twice that of personal gains.

Net long-term gains from passthrough organizations, such as partnerships, S corporations, and estates and trusts from schedule K-1, result in a persistent elasticity of -1.93. This is a markedly higher sensitivity to tax rates than other types of capital gains show and is nearly identical to the transitory elasticity. Possibly, those results are associated with partnerships in the finance industry, such as hedge fund managers, that may be extremely sensitive to tax changes.

On the other hand, the variable for capital gains distributions from mutual funds has a persistent elasticity of -0.08, which is a markedly lower sensitivity than other types of gains show. That suggests that mutual funds use criteria other than taxes when determining the timing and amount of gains to distribute.

Finally, we consider the total net long-term gains, before carryover, which have a persistent elasticity that is slightly greater than 1 in absolute value and a transitory elasticity of -1.53.

The transitory elasticities and the inverse Mill's ratios are much less precisely estimated in the subperiod analyses than in the analysis using the full time period. The sensitivity of these results may be explained by the smaller variation associated with the shorter time frames. As we mentioned earlier, the inverse Mill's ratio could be approximated well by a linear function of the explanatory variables when there is not enough variation in the sample. Even with an exclusion restriction variable, this could lead to a severe collinearity problem and could yield estimates that are fairly sensitive and only weakly identified.

VIII. CONCLUSION

In this paper, we estimate the elasticities of long-run capital gains with respect to persistent and transitory tax changes. Adapting a model developed by Burman and Randolph (1994a, 1994b) and extended by Auerbach and Siegel (2000), we estimate elasticities of persistent tax changes in the range of -0.58 to -1.41, with most estimates about -0.80. Transitory elasticities almost always exceed 1 in absolute value, although they are much lower than some previous estimates. The decision over how much gain to realize appears to be much more sensitive to tax rates than does the decision to realize a gain. Although we focus our examination on personal capital gains, we also compare the results of our model to results from the original model applied to contemporary data, estimate our model on subperiods, and estimate our model on other types of capital gains. Two substantial differences between personal capital gains and other types of gains are worth noting: The elasticity of long-run capital gains from partnerships, S corporations, and trusts is much greater than 1 in absolute value; and the elasticity of capital gains distributions from mutual funds is nearly zero.

Although use of existing methods on new data allows for a clear comparison with previous research, there are disadvantages as well. For example, the consistency of our estimates relies on the distributional assumptions of our parametric model. In addition, our model examines the average effect of tax changes, although the effect may vary substantially across income categories, just as it varies across time periods and types of capital gains. Finally, if sophisticated taxpayers plan many years in advance, lagged variables may be endogenous rather than predetermined. In future work, we hope to address those concerns.

REFERENCES

Auerbach, Alan J. and Jonathan M. Siegel (2000), "Capital Gains Realizations of the Rich and Sophisticated," *American Economic Review,* 90(2): 276–282.

Auten, Gerald E. and Joseph J. Cordes (1991), "Policy Watch: Cutting Capital Gains Taxes," *Journal of Economic Perspectives*, 5(1), 181–192.

Bakija, Jon (2009), "Documentation for a Comprehensive Historical U.S. Federal and State Tax Calculator Program," unpublished working paper, August 11, 2009, http://web.williams.edu/Economics/ papers/bakijaDocumentation_IncTaxCalc.pdf.

Bryant, Victoria (2008), "Attrition in the Tax Years 1999 – 2005 Individual Income Tax Return Panel," Internal Revenue Service, Washington D.C., http://www.irs.gov/pub/irs-soi/08rppsbryant.pdf.

Burman, Leonard E. and William C. Randolph (1994a), "Measuring Permanent Responses to Capital Gains Taxes in Panel Data," *American Economic Review,* 84(4), 794–809.

——— (1994b), "Measuring Permanent Responses to Capital Gains Taxes in Panel Data," Office of Tax Analysis Working Paper 68, August 1994, http://www.treasury.gov/resource-center/tax-policy/tax-analysis/Documents/ota68.pdf.

Burman, Leonard E. and Peter D. Ricoy (1997), "Capital Gains and the People who Realize Them," *National Tax Journal*, 50(3), 427–451.

Feldstein, Martin, Joel Slemrod, and Shlomo Yitzhaki (1980), "The Effects of Taxation on the Selling of Corporate Stock and the Realization of Capital Gains," *Quarterly Journal of Economics* 94(4), 777–791.

——— (1984), "The Effects of Taxation on the Selling of Corporate Stock and the Realization of CapitalGains: Reply," *Quarterly Journal of Economics* 99(1), 111–120.

Gravelle, Jane G. (2010), "Capital Gains Tax Options: Behavioral Responses and Revenues," Congressional Research Service, CRS Report for Congress R41364, August 10, 2010.

Hausman, Jerry A. and David A. Wise (1981), "Stratification on an Endogenous Variable and Estimation: The Gary Income Maintenance Experiment," *Structural Analysis of Discrete Data with Econometric Applications*, Charles Manski and Daniel McFadden, eds., Cambridge, MA: MIT Press, 365–391.

Lee, Lung-Fei, G.S. Maddala, and R.P. Trost (1980), "Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity," *Econometrica*, 48(2), 491–503.

Minarik, Joseph J. (1984), "The Effects of Taxation on the Selling of Corporate Stock and the Realization of Capital Gains: Comment," *Quarterly Journal of Economics*, 99(1), 93–110.

U.S. Council of Economic Advisers (2012), *Economic Report of the President*, United States Government Printing Office, Washington DC, February 2012.

U.S. Department of the Treasury (2010), *Capital Gains and Taxes Paid on Capital Gains for Returns with Positive Net Capital Gains, 1954–2008.* December 30, 2010, http://www.treasury.gov/resource-center/tax-policy/Documents/OTP-CG-Taxes-Paid-Pos-CG-1954-2008-12-2010.pdf.

Vella, Francis (1998), "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, 33(1), 127–169.

Weber, Michael E. (2006), "The Tax Year 1999-2003 Individual Income Tax Return Panel: A First Look at the Data," Internal Revenue Service, Washington D.C., http://www.irs.gov/pub/irs-soi/06asaweber.pdf

Weber, Michael E. and Victoria Bryant (2005), "The 1999 Individual Income Tax Return Edited Panel," Internal Revenue Service, Washington D.C.,. http://www.irs.gov/pub/irs-soi/05weber.pdf

Wooldridge, Jeffery M. (2010), *Econometric Analysis of Cross Section and Panel Data* (Massachusetts: MIT Press).

Table 1: Descriptive Statistics of Data Used in the Analysis

	Unrestricted Sample				Restricted Sample			
Year	Number of Returns	Weighted Number of Returns	Total Capital Gains in AGI	Positive Long-Term Capital Gains in Schedule D, Line 8	Number of Returns	Weighted Number of Returns	Total Capital Gains in AGI	Positive Long-Term Capital Gains in Schedule D, Line 8
1999	88,123	123.0	559.8	289.2	61,335	90.4	436.1	220.6
2000	84,742	116.9	619.6	375.6	58,800	85.6	476.2	289.6
2001	83,239	114.0	303.0	194.7	57,483	83.1	242.1	155.0
2002	81,710	111.3	234.4	140.2	56,283	80.8	182.5	111.8
2003	80,661	109.3	294.2	149.3	55,424	79.2	228.7	116.0
2004	79,712	107.7	440.0	225.7	54,724	77.9	355.3	182.3
2005	78,905	106.2	590.7	276.9	54,087	76.8	477.5	226.7
2006	78,550	105.7	671.9	304.9	53,874	76.4	542.9	251.1
2007	79,975	108.7	752.1	338.8	54,855	78.6	623.6	282.0
2008	75,402	101.2	314.3	175.6	51,660	73.1	265.2	150.2

Note: AGI = adjusted gross income.

Table 2: Summary Statistics of Variables Used in the Model

	All Observations		Realizers Only	
	Weighted mean	Unweighted mean	Weighted mean	Unweighted mean
Net long-term gains	2,136	466,665	36,885	2,474,404
Net long-term gains (log)	0.611	2.204	7.813	10.702
Current marginal tax rate	14.26	15.90	17.43	19.67
Imputed unrealized gains (log)	11.63	12.29	11.88	13.49
Imputed unrealized gains in stock (log)	2.632	3.771	4.179	6.281
Imputed permanent income (log)	10.79	10.89	10.93	11.08
Current income (exogenous components; log)	10.48	11.46	10.99	13.34
Business losses lagged (log)	0.133	1.062	0.507	3.335
Rent losses lagged (log)	0.410	1.028	1.003	2.094
Age	49.80	52.20	58.88	59.11
Primary taxpayer is male	0.334	0.371	0.441	0.516
Number of dependents	0.774	0.797	0.584	0.747
Primary taxpayer is married	0.516	0.621	0.683	0.833
South states	0.349	0.328	0.310	0.280
Northeast states	0.197	0.217	0.235	0.275
Midwest states	0.237	0.216	0.245	0.181
West states	0.217	0.239	0.210	0.264
Having carryover loss in excess of $3000	0.035	0.103	0.130	0.211
Lagged number of short-term realizations	0.965	16.388	5.497	53.603
Having any long-term loss realization	0.086	0.248	0.323	0.605
Having net losses from sale of a business or business asset	0.001	0.019	0.005	0.063
Having net losses from pass-through entities	0.006	0.045	0.026	0.106
Having net short-term losses	0.058	0.184	0.223	0.436

Note: All dollar amounts are in 2008 dollars.

Table 3: Double-Hurdle Model

Dependent Variable: Personal Long-Term Capital Gains Realizations (log)

	Level Equation		Criterion Function	
	Coefficient	Standard Error	Coefficient	Standard Error
$\hat{\tau}_{t+1}$	-0.024	0.018	-0.007	0.007
$\hat{\tau}_{t}$	-0.048	0.016	-0.007	0.004
$\hat{\tau}_{t-1}$	0.039	0.004	0.010	0.001
Inverse Mills ratio	n/a	n/a	n/a	n/a
Permanent income (L)	0.295	0.180	0.300	0.057
Transitory income (L)	0.225	0.026	0.075	0.007
Imputed total unrealized gains (L)	0.288	0.026	-0.027	0.007
Ratio of unrealized gains in stock (L)	0.002	0.017	0.323	0.006
Lagged business losses (L)	0.092	0.023	0.010	0.010
Lagged business losses (D)	-0.244	0.207	0.200	0.087
Lagged rent losses (D)	-0.016	0.027	0.034	0.011
Lagged rent losses (dummy)	0.277	0.235	-0.070	0.092
Male (D)	0.059	0.070	-0.068	0.018
Number of dependents	-0.051	0.026	0.016	0.007
Married (D)	-0.081	0.161	-0.169	0.053
30<=Age<40 (D)	-0.475	0.155	0.277	0.034
40<=Age<50 (D)	-0.436	0.155	0.565	0.037
50<=Age<60 (D)	-0.268	0.158	0.812	0.040
60<=Age<70 (D)	-0.017	0.163	1.106	0.040
Age>=70 (D)	-0.032	0.162	1.387	0.040
Short-term realizations group 2 (D)	0.244	0.040	0.532	0.013
Short-term realizations group 3 (D)	0.818	0.099	0.697	0.045
Short-term realizations group 4 (D)	1.598	0.130	0.687	0.084
Short-term realizations group 5 (D)	1.686	0.852	0.537	0.302
Having long-term loss realization (D)	0.945	0.041	0.167	0.014
Having net business losses (D)	0.400	0.158	0.055	0.092
Having net pass-through losses (D)	0.240	0.111	0.127	0.039
Having net short-term losses (D)	0.483	0.043	0.161	0.015
Having carryover loss (D)	0.001	0.050	0.148	0.021
Constant	1.972	1.811	-2.152	0.576
Persistent elasticity	-0.686	0.098		
Transitory elasticity	-1.019	0.318		
Number of observations	70,377		341,793	

Notes: Logarithmic variables are indicated by "L." Dummy variables are indicated by "D." Dummy variables for regions and years are included in the model but are omitted from the table. Data are weighted. Standard errors are calculated from 200 bootstrap replications. N/A = Not applicable.

Table 4: Taxes Paid on Gains as Percentage of Taxable Gains

Income	1993		2007	
	Taxes Paid as Percent of Taxable Gains	Percent of Taxpayers	Taxes Paid as Percent of Taxable Gains	Percent of Taxpayers
0 to 10,000	11.0	1.2	4.0	0.3
10,000 to 20,000	13.0	2.9	1.7	0.4
20,000 to 30,000	14.8	3.6	2.2	0.5
30,000 to 40,000	16.8	3.6	3.7	0.6
40,000 to 50,000	19.5	3.6	5.0	0.7
50,000 to 75,000	20.9	8.1	7.6	2.4
75,000 to 100,000	21.9	5.5	8.7	2.6
100,000 to 200,000	23.0	13.3	12.4	8.5
200,000 and Over	23.9	57.9	14.8	84.0

Sources: Burman and Ricoy (1997), author calculations.

Table 5: Generalized Tobit Model

Dependent Variable: Personal Long-Term Capital Gains Realizations (log)

	Level Equation		Criterion Function	
	Coefficient	Standard Error	Coefficient	Standard Error
$\hat{\tau}_{t+1}$	-0.033	0.019	-0.007	0.007
$\hat{\tau}_{t}$	-0.061	0.018	-0.007	0.004
$\hat{\tau}_{t-1}$	0.053	0.006	0.010	0.001
Inverse Mills ratio	1.666	0.291	n/a	n/a
Permanent income (L)	0.716	0.196	0.300	0.057
Transitory income (L)	0.325	0.033	0.075	0.007
Imputed total unrealized gains (L)	0.241	0.031	-0.027	0.007
Ratio of unrealized gains in stock (L)	0.416	0.073	0.323	0.006
Lagged business losses (L)	0.102	0.025	0.010	0.010
Lagged business losses (D)	-0.021	0.224	0.200	0.087
Lagged rent losses (D)	0.029	0.031	0.034	0.011
Lagged rent losses (D)	0.168	0.273	-0.070	0.092
Male (D)	-0.047	0.077	-0.068	0.018
Number of dependents	-0.028	0.028	0.016	0.007
Married (D)	-0.318	0.174	-0.169	0.053
30<=Age<40 (D)	-0.066	0.176	0.277	0.034
40<=Age<50 (D)	0.362	0.215	0.565	0.037
50<=Age<60 (D)	0.858	0.252	0.812	0.040
60<=Age<70 (D)	1.488	0.305	1.106	0.040
Age>=70 (D)	1.832	0.354	1.387	0.040
Short-term realizations group 2 (D)	0.918	0.127	0.532	0.013
Short-term realizations group 3 (D)	1.641	0.184	0.697	0.045
Short-term realizations group 4 (D)	2.360	0.200	0.687	0.084
Short-term realizations group 5 (D)	2.292	0.810	0.537	0.302
Having long-term loss realization (D)	1.159	0.055	0.167	0.014
Having net business losses (D)	0.429	0.172	0.055	0.092
Having net pass-through losses (D)	0.382	0.127	0.127	0.039
Having net short-term losses (D)	0.700	0.059	0.161	0.015
Having carryover loss (D)	n/a	n/a	0.148	0.021
Constant	-2.336	1.961	-2.152	0.576
Persistent elasticity	-0.792	0.108		
Transitory elasticity	-1.196	0.348		
Number of observations	70,377		341,793	

Notes: Logarithmic variables are indicated by "L." Dummy variables are indicated by "D." Dummy variables for regions and years are included in the model but are omitted from the table. Data are weighted. Standard errors are calculated from 200 bootstrap replications. N/A = not applicable.

Table 6: Sensitivity Analysis

	Persistent Elasticity		Transitory Elasticity		Inverse Mills Ratio
	Coefficient	Standard Error	Coefficient	Standard Error	
Table 5 results	-0.792	0.108	-1.196	0.348	1.666
Unweighted	-1.132	0.067	-1.597	0.169	-0.814
Loss variables omitted	-0.784	0.106	-1.262	0.337	0.727
Wealth and year interactions	-0.818	0.110	-1.168	0.353	1.770
Drop $\tau_{l,t-1}$, $\tau_{l,t+1}$, τ_0	-0.735	0.264	n/a	n/a	1.904
Use instruments $\tau_{s,t-1}$, $\tau_{s,t}$, $\tau_{s,t+1}$	-0.682	0.327	-0.935	0.874	1.715
Carryover dummy in 2^{nd} stage	-0.854	0.116	-1.332	0.360	2.545
Carryover categories in 2^{nd} stage	-0.843	0.116	-1.333	0.359	2.513

Note: Standard errors are calculated from 200 bootstrap replications.

Table 7: Comparison of Methods on 1999-2008 Data

Method	Permanent/Persistent	Transitory
Current method	-1.132	-1.597
unweighted	(0.067)	(0.169)
Current method	-0.792	-1.196
weighted	(0.108)	(0.348)
Burman and Randolph	-0.584	-2.697
unweighted	(0.137)	(0.125)
Burman and Randolph	-1.116	-2.576
weighted	(2.803)	(0.702)
Auerbach and Siegel	-1.272	-1.844
unweighted	(0.076)	(0.173)
Auerbach and Siegel	-1.000	-1.733
weighted	(0.127)	(0.380)

Note: Standard errors in parentheses are calculated from 200 bootstrap replications.

Table 8: Estimates Using Subperiods

	Table 5 Results (1)	2000-2001 (2)	2002-2003 (3)	2004-2005 (4)	2006-2007 (5)
Persistent Elasticity	-0.792	-0.914	-1.001	-1.405	-0.577
	(0.108)	(0.246)	(0.419)	(0.294)	(0.165)
Transitory Elasticity	-1.196	-1.112	-1.192	1.914	-1.150
	(0.348)	(2.642)	(5.107)	(0.991)	(0.317)
Inverse Mill's Ratio	1.666	3.094	2.393	0.018	1.482
	(0.291)	(0.768)	(1.770)	(0.445)	(0.422)

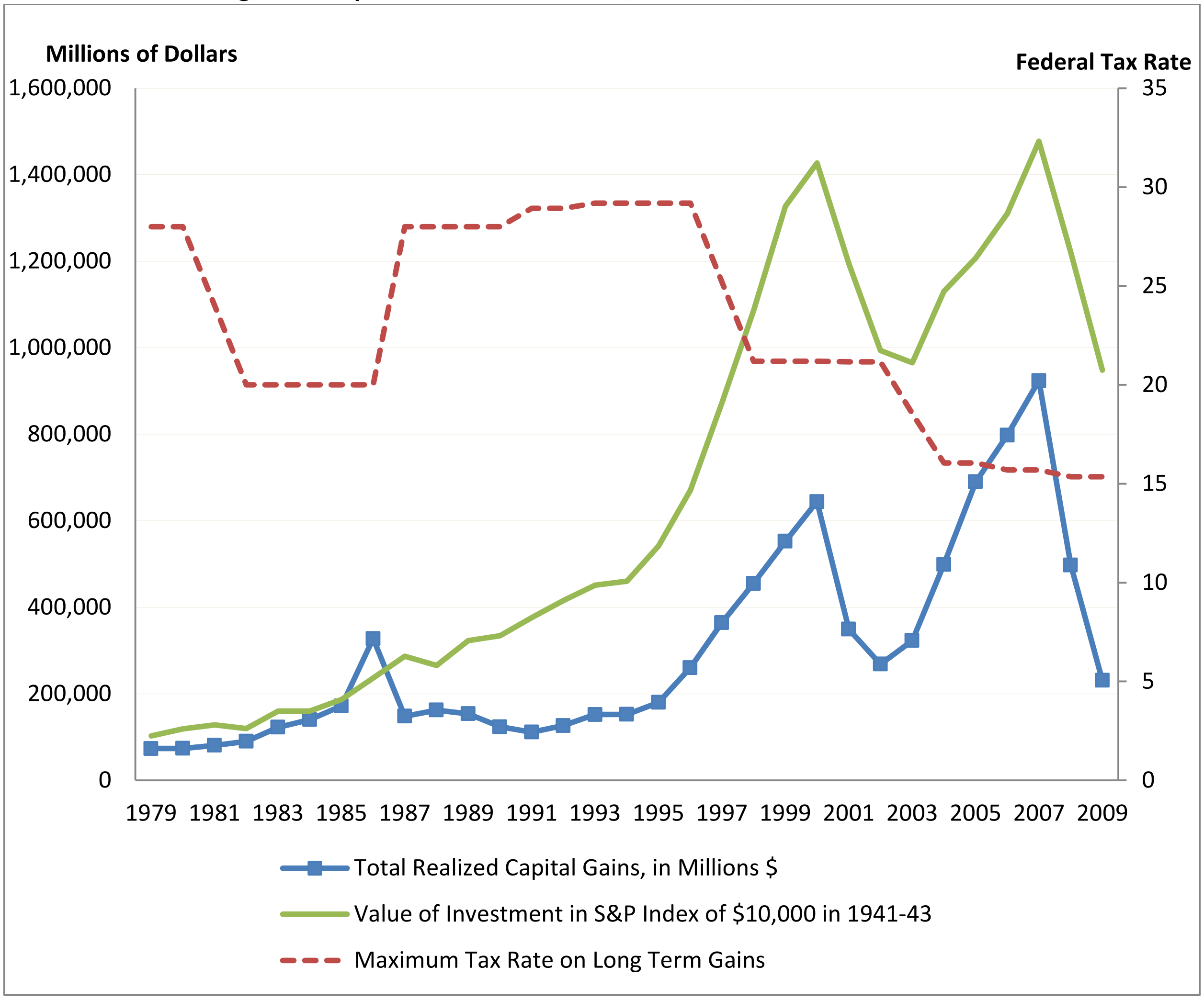
Note: Standard errors in parentheses are calculated from 200 bootstrap replications.

Table 9: Estimates Using Alternative Capital Gains Variables

	Table 5 Results (1)	Sales of Businesses /Business Assets (2)	Partnerships, S Corps, Trusts (3)	Capital Gains Distributions (4)	All Capital Gains Less Carryover (5)
Persistent Elasticity	-0.792	-0.794	-1.933	-0.079	-1.020
	(0.108)	(0.198)	(0.297)	(0.077)	(0.092)
Transitory Elasticity	-1.196	-2.206	-1.987	-0.777	-1.526
	(0.348)	(0.550)	(1.012)	(0.211)	(0.277)
Inverse Mill's Ratio	1.666	-1.879	0.921	-0.159	1.311
	(0.291)	(0.479)	(1.028)	(0.105)	(0.081)

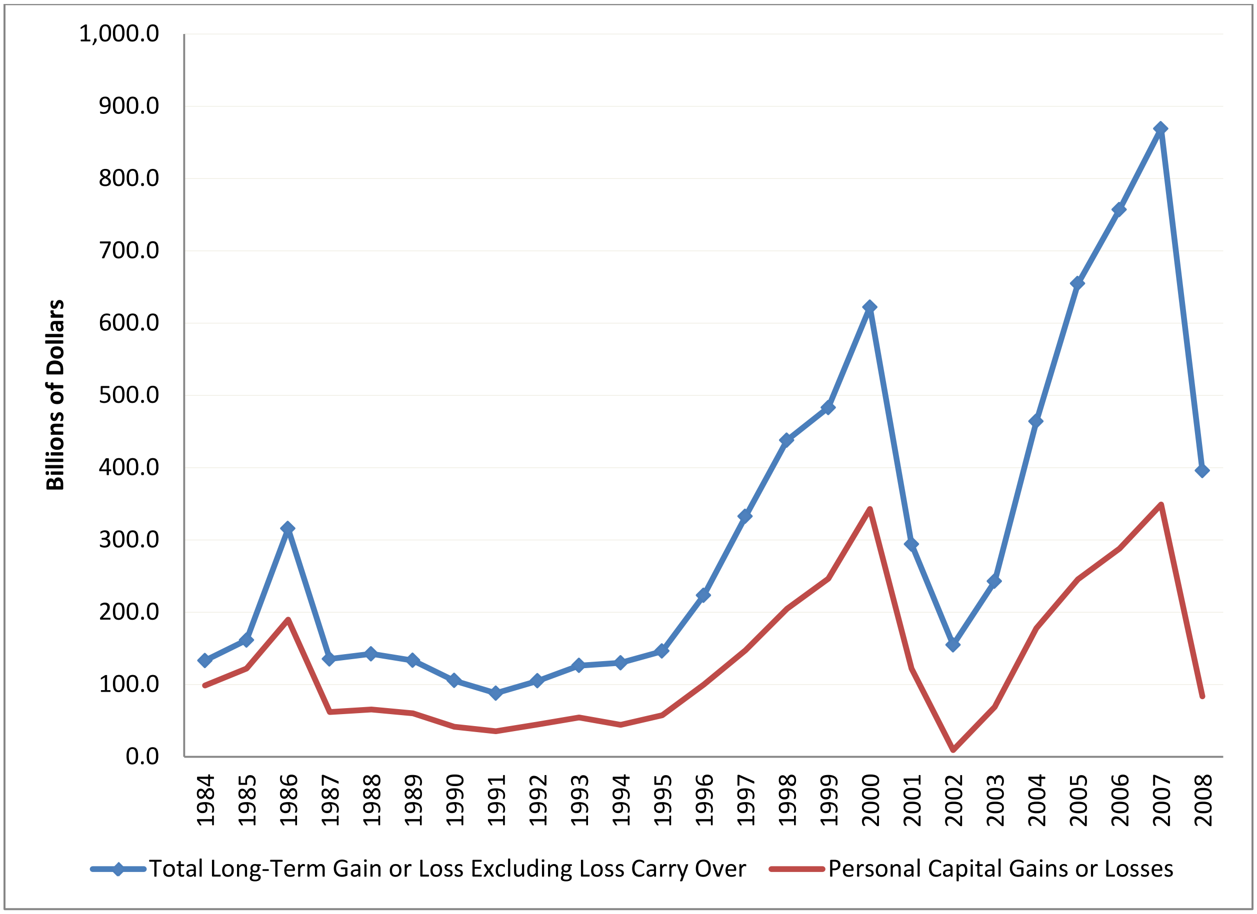
Note: Standard errors in parentheses are calculated from 200 bootstrap replications.

Figure 1: Capital Gains Realizations, Tax Rates, and the S&P Index



Sources: U.S. Department of the Treasury (2010), and U.S. Council of Economic Advisers (2012), Tables B-95 and B-96.

Figure 2: Long-Term Capital Gains or Losses Reported on Tax Returns



Source: Author's calculations from tabulations of the 1984-2008 Statistics of Income individual income tax cross-section data.

Appendix
Table A1: Wealth and Accrued Gains in Stocks

	Wealth		Accrued Stock	
	Coefficient	Standard Error	Coefficient	Standard Error
Age	0.116	0.003	-0.105	0.020
Age Squared	-0.001	0.000	0.001	0.000
Single and Female (D)	-0.129	0.031	-1.940	0.191
Married (D)	0.177	0.041	-1.440	0.198
Wages (D)	-3.825	0.082	-10.409	0.530
Wages (L)	0.347	0.007	0.976	0.045
Taxable Interest (D)	-0.714	0.041	1.361	0.269
Taxable Interest (L)	0.138	0.005	0.116	0.032
Tax-Exempt Interest (D)	-0.629	0.088	1.294	0.528
Tax-Exempt interest (L)	0.083	0.009	0.102	0.052
Dividend Income (D)	-1.288	0.046	1.638	0.293
Dividend Income (L)	0.220	0.005	0.973	0.034
Alimony (D)	-2.270	0.378	-8.942	2.355
Alimony (L)	0.235	0.044	0.973	0.274
Schedule E Income (D)	-2.093	0.071	0.409	0.458
Schedule E Income (L)	0.304	0.006	0.046	0.039
Schedule C or F Income (D)	-2.033	0.083	-2.285	0.545
Schedule C or F Income (L)	0.272	0.007	0.299	0.047
Itemizer (D)	0.679	0.019	3.078	0.133
Home Mortgage Int Deduction (D)	0.180	0.017	0.556	0.112
SCF 2004 Observation (D)	0.112	0.019	0.270	0.125
SCF 2007 Observation (D)	0.209	0.019	1.785	0.124
2 Dependents (D)	0.203	0.033	0.853	0.154
3 Dependents (D)	0.336	0.037	0.152	0.186
4 Dependents (D)	0.441	0.039	-0.118	0.191
5 Dependents (D)	0.584	0.044	-0.218	0.235
6 Dependents (D)	0.565	0.057	-0.407	0.341
7 Dependents (D)	0.680	0.099	-1.391	0.652
8 Dependents (D)	1.047	0.151	2.549	0.932
Dependents>=9 (D)	1.143	0.216	-0.601	1.447
Married Filing Jointly (D)	0.189	0.032	n/a	n/a
Pension, Annuities, SS Income (D)	-0.799	0.134	n/a	n/a
Pension, Annuities, SS Income	0.064	0.014	n/a	n/a
Constant	5.986	0.088	-10.748	0.535
Number of observations		44920		62083
Pseudo R-Square		0.6621		0.1511

Notes: Logarithmic variables are indicated by “L”. Dummy variables are indicated by “D”.

Table A2: Current Marginal Tax Rate, τ_{it}

	Entire Population		Realizers	
	Coefficient	Standard Error	Coefficient	Standard Error
Permanent Income (L)	-0.017	0.055	2.486	0.339
Transitory income (L)	0.186	0.007	0.712	0.047
Imputed total unrealized gains (L)	-0.170	0.009	-0.047	0.041
Ratio of unrealized gains in stock (L)	-0.055	0.006	1.030	0.124
Lagged Business losses (L)	-0.270	0.032	-0.112	0.046
Lagged Business losses (D)	1.937	0.259	1.315	0.399
Lagged rent losses (D)	0.058	0.019	0.122	0.047
Lagged rent losses (D)	-0.066	0.164	-0.166	0.415
Male (D)	0.082	0.017	-0.199	0.129
Number of dependents	0.016	0.005	-0.043	0.038
Married (D)	0.093	0.049	-1.480	0.292
30<=Age<40 (D)	-0.012	0.018	0.906	0.274
40<=Age<50 (D)	0.036	0.024	1.750	0.343
50<=Age<60 (D)	0.160	0.027	2.617	0.416
60<=Age<70 (D)	0.340	0.030	3.969	0.505
Age>=70 (D)	0.466	0.034	5.230	0.591
Short-term realizations grp. 2 (D)	0.270	0.031	1.671	0.210
Short-term realizations grp. 3 (D)	0.447	0.166	2.004	0.321
Short-term realizations grp. 4 (D)	1.245	0.260	2.231	0.374
Short-term realizations grp. 5 (D)	1.565	0.411	1.982	0.646
Having long-term loss realization (D)	-0.022	0.035	0.828	0.094
Having net business losses (D)	-0.139	0.254	-0.363	0.364
Having net passthrough losses (D)	0.107	0.103	0.384	0.200
Having net short-term losses (D)	-0.041	0.042	0.541	0.100
Lagged Marginal Tax Rate	0.072	0.001	0.264	0.009
Maximum Federal, State and Local Tax Rate, τ_{sit}	0.018	0.004	0.209	0.022
First Dollar Marginal Tax Rate , τ_{0it}	0.892	0.002	0.505	0.010
Future Maximum Federal, State and Local Tax Rate , τ_{sit+1}	-0.013	0.004	-0.034	0.022
Having carryover loss (D)	-0.804	0.050	n/a	n/a
Inverse Mills ratio	n/a	n/a	4.259	0.489
Constant	1.952	0.559	-25.819	3.463
Number of observations		341793		70377
R-Square		0.9049		0.7029

Notes: Logarithmic variables are indicated by “L”. Dummy variables are indicated by “D”. Dummy variables for regions and years are included in the model but are omitted from the table.

Table A3: Future Marginal Tax Rate , τ_{it+1}

	Entire Population		Realizers	
	Coefficient	Standard Error	Coefficient	Standard Error
Permanent Income (L)	0.932	0.111	2.959	0.409
Transitory income (L)	0.489	0.013	0.753	0.056
Imputed total unrealized gains (L)	0.145	0.016	0.034	0.052
Ratio of unrealized gains in stock (L)	0.286	0.011	0.439	0.148
Lagged Business losses (L)	-0.149	0.037	-0.112	0.056
Lagged Business losses (D)	1.036	0.309	0.843	0.479
Lagged rent losses (D)	0.109	0.028	0.094	0.060
Lagged rent losses (dummy)	-0.836	0.238	-0.549	0.525
Male (D)	-0.341	0.043	-0.450	0.162
Number of dependents	0.114	0.014	-0.039	0.050
Married (D)	-0.621	0.099	-1.578	0.352
30<=Age<40 (D)	-0.134	0.061	0.357	0.399
40<=Age<50 (D)	-0.368	0.067	0.150	0.472
50<=Age<60 (D)	-0.421	0.070	0.290	0.553
60<=Age<70 (D)	-0.109	0.073	0.982	0.654
Age>=70 (D)	-0.279	0.073	1.390	0.754
Short-term realizations grp. 2 (D)	0.174	0.043	0.393	0.253
Short-term realizations grp. 3 (D)	0.149	0.176	0.200	0.389
Short-term realizations grp. 4 (D)	0.524	0.259	0.287	0.425
Short-term realizations grp. 5 (D)	-0.171	0.743	-1.301	0.929
Having long-term loss realization (D)	-0.058	0.044	0.343	0.114
Having net business losses (D)	-0.310	0.337	-0.108	0.371
Having net passthrough losses (D)	0.210	0.132	-0.005	0.237
Having net short-term losses (D)	-0.144	0.053	-0.121	0.126
Lagged Marginal Tax Rate	0.187	0.003	0.237	0.010
Maximum Federal, State and Local Tax Rate, τ_{sit}	-0.341	0.010	-0.394	0.027
First Dollar Marginal Tax Rate , τ_{0it}	0.538	0.003	0.436	0.011
Future Maximum Federal, State and Local Tax Rate , τ_{sit+1}	0.468	0.010	0.590	0.029
Having carryover loss (D)	-0.291	0.062	n/a	n/a
Inverse Mills ratio	n/a	n/a	0.741	0.588
Constant	-7.918	1.140	-29.262	4.204
Number of observations		341793		70377
R-Square		0.5696		0.5788

Notes: Logarithmic variables are indicated by “L”. Dummy variables are indicated by “D”. Dummy variables for regions and years are included in the model but are omitted from the table.

Table A4: Log of Permanent Income

	Coefficient	Standard Error
Age	0.049	0.001
Age Squared	-4.80E-4	5.00E-06
Single and Female (D)	0.187	0.004
Married Filing Jointly (D)	0.962	0.003
Married Filing Separately (D)	0.295	0.014
Male (D)	0.123	0.004
SCF 2004 Observation (D)	0.023	0.004
SCF 2007 Observation (D)	-0.022	0.005
2 Dependents (D)	0.038	0.004
3 Dependents (D)	0.082	0.004
4 Dependents (D)	0.126	0.005
5 Dependents (D)	0.050	0.010
6 Dependents (D)	-0.044	0.021
7 Dependents (D)	-0.075	0.028
8 Dependents (D)	-0.291	0.057
Dependents>=9 (D)	-0.032	0.092
Constant	9.019	0.013
Number of observations	497290	
R-Square	0.293	

Notes: Logarithmic variables are indicated by "L". Dummy variables are indicated by "D".