BURCH CENTER WORKING PAPER SERIES

University of California, Berkeley
Department of Economics - #3880
Berkeley, CA 94720-3880

Capital Gains Taxation and Tax Avoidance: New Evidence from Panel Data

Alan J. Auerbach
University of California, Berkeley and NBER

Leonard E. Burman
The Urban Institute

Jonathan M. Siegel
University of California, Berkeley

December 1997

Burch Working Paper No. **B97-27**

JEL Classifications: H24, H31

This paper was presented at a conference sponsored by the University of Michigan Office of Tax Policy Research, October 24-25, 1997. We are grateful to the staff of the Tax Analysis Division of the Congressional Budget Office for help in analyzing and interpreting the data used in this paper, to our discussant Jane Gravelle and other conference participants for comments on an earlier draft, and to the Robert D. Burch Center for Tax Policy and Public Finance for financial support. Views expressed in this paper are the authors' alone and do not necessarily reflect the views or policies of the institutions with which we are affiliated.

No. B97-27

CAPITAL GAINS TAXATION AND TAX AVOIDANCE: NEW EVIDENCE FROM PANEL DATA

Abstract

Previous theoretical analyses of the capital gains tax have suggested that investors have considerable opportunity to avoid the tax. Yet, past empirical work has found relatively little evidence of such activity. Using a previously unavailable panel data set with a very large sample of high-income individuals, this paper aims to bring the theory and evidence closer together by examining the behavior of individual taxpayers over time.

Though confirming past findings that avoidance of tax on realized capital gains is not prevalent, we do observe that tax avoidance activity increased after the passage of the Tax Reform Act of 1986, and that high-income, high-wealth and more sophisticated taxpayers were most likely to avoid tax. However, the efficacy of tax avoidance strategies depends on being able to avoid tax for long periods, and we find that most tax avoidance is of relatively short duration. Thus, the effective tax rate on realized capital gains is very close to the statutory rate in all years and tax brackets.

Alan J. Auerbach
University of California, Berkeley
Department of Economics
549 Evans Hall #3880
Berkeley, CA 94720-3880

Jonathan M. Siegel
University of California, Berkeley
Department of Economics
549 Evans Hall #3880
Berkeley, CA 94720-3880

Leonard E. Burman
The Urban Institute
2100 M Street NW
Washington, DC 20037-1207

1. Introduction

In the United States, capital gains taxes long have sparked interest among economists and policy makers. The Taxpayer Relief Act of 1997 contains the latest changes in the taxation of capital gains. The Act lowers the tax rate on most gains and makes the tax rate dependent on holding period. As before, gains on assets held for at least a year qualify for long-term treatment and a maximum tax rate of 28 percent, well below the maximum rate on ordinary income. In addition, assets held for at least 18 months qualify for a maximum tax rate of 20 percent, and assets held for at least five years (and purchased after the year 2000) will face a top rate of just 18 percent. The Act also exempts from tax almost all gains from sales of owner-occupied housing.

Other provisions of the Act are aimed at reducing tax avoidance associated with the already-favorable treatment of capital gains. These include changes that lessen the favorable tax treatment on real estate investments through a change in recapture provisions, and elimination of the ability of investors to hedge open positions by "shorting against the box" (taking an offsetting short position) without realizing their locked-in gains. Such restrictions build on those introduced by the Tax Reform Act of 1986 that limited the ability of taxpayers to deduct losses associated with real estate investments and other "passive" investment activities.

This legislation, which reduces capital gains tax rates in general but also seeks to eliminate certain advantages of holding assets subject to capital gains taxation, reflects an underlying tension in how the capital gains tax is perceived. On the one hand, a low rate of capital gains tax is seen as facilitating the efficient turnover of investor portfolios and a spur to venture capital investment and entrepreneurship. On the other hand, the favorable rate of tax and the ability of investors to time realizations is understood to generate opportunities to avoid not

only capital gains taxes, but other taxes as well. The continued existence of the annual $3,000 limit on capital loss deductions reflects the perceived need to limit such activity.

The same tension is evident in the economics literature. Theoretical analysis (e.g. Constantinides 1984, Stiglitz 1983) has elucidated strategies to avoid taxes on capital gains and to generate capital losses to offset ordinary income. Much empirical research, however, emphasized the potentially large response elasticities to capital gains tax reductions (e.g., Feldstein, Slemrod and Yitzhaki 1980). Subsequent empirical work (including Auten and Clotfelter 1982, Auerbach 1988, and Burman and Randolph 1994) distinguished between short-run and long-run responses, but thus far has failed to focus on the more sophisticated avoidance strategies detailed in the theoretical literature. This remaining gap between theory and evidence has been due in part to data limitations. Complicated avoidance transactions may be difficult to discern without considerable information about the behavior of the high-income individuals who realize most capital gains. But it also seems clear that the theory offers an inadequate description of taxpayer behavior. As Poterba (1987) shows, relatively few taxpayers realizing capital gains appear to utilize the avoidance strategies that theory would predict.[1] Put simply, over $100 billion of capital gains are realized every year, and most of them face a positive rate of tax.

This paper aims to bring theory and evidence closer together, by examining more closely the behavior of individual taxpayers over time. We follow Poterba in searching for the presence of avoidance activity, but our analysis is facilitated by the use of a rich data set that tracks every capital gains realization for a large number of high-income individuals over a decade, from 1985 through 1994. Having a relatively long panel also allows us to consider changes in avoidance behavior over time, and to ask whether growing taxpayer sophistication, perhaps aided by the increasing efficiency of financial markets, has led to an increase in avoidance activity. In

addition, with information on individual transactions, we can explore the extent to which avoidance behavior is a function of portfolio composition. While the theoretical arguments made by Stiglitz, Constantinides and others assume that transaction costs are negligible – that assets are highly liquid – this may not be a good assumption for some assets, such as real estate and business property. Thus, the ability of taxpayers to shelter their gains from tax may depend on the kinds of assets they own.

In a sense, our investigation is complementary to the typical empirical investigation, in that we focus especially on a period, 1987-1994, during which there were no important changes in the treatment of capital gains taxes (other than an increasing differential created by higher tax rates on ordinary income). Our view is that further analysis of the response of aggregate capital gains realizations to changes in the capital gains tax rate requires a better understanding of the underlying behavior generating these realizations.

A useful starting point for our analysis is a simple description of the relevant capital gains tax provisions in effect during the period we analyze. Figure 1, based on one presented in Poterba (1987), shows four distinct tax regimes that apply to marginal short-term gains and losses (applicable, during the sample period, to sales of assets owned less than one year) and long-term gains and losses (on those assets held for at least one year), based on a taxpayer's overall levels of gains and losses.[2] It distinguishes between long-term gains and losses – those on assets held for more than one year – and short-term gains and losses – those on assets held for less than one year.

The "normal" situation, in which the rates on long-term and short-term gains are equal to their distinct statutory rates, applies only in the region labeled A in the figure. In region A, taxpayers have both positive long-term and short-term gains. These net short-term gains are

taxed at the same rate as ordinary income, τ. In 1987, the maximum tax rate on ordinary income was 38 percent. From 1988 to 1990 the maximum tax rate on ordinary income was 33 percent, because of the phase-out of the15-percent bracket for some moderately high-income taxpayers. In 1991, the top rate on ordinary income increased from 28 to 31 percent. In 1993, the maximum ordinary income tax rate increased again to 39.6 percent. The tax rate on net long-term gains was capped at 28 percent in 1987 and from 1991 to 1994. From 1988 to 1990, the tax rate on capital gains was the same as that on ordinary income—as high as 33 percent due to the bubble. The resulting tax rate on long-term gains is denoted τ^*.[3]

A taxpayer with net long-term losses but net short-term gains is required to net the long-term losses against the short-term gains and is taxed fully on the difference if positive, and allowed a full deduction of any net loss up to $3,000. Thus, the effective marginal tax rate on both long-term and short-term gains is τ. Similarly, a taxpayer with both long-term and short-term losses, or short-term losses in excess of long-term gains, is allowed to deduct any net loss up to $3,000. These taxpayers fall into the region denoted B in the figure, in which the effective tax rate on both short- and long-term gains is τ.

Region C includes those taxpayers with total (short-term plus long-term) losses in excess of $3,000. These taxpayers face no *current* tax on marginal short-term or long-term gains, because such gains simply reduce the amount of losses that cannot be deducted. However, because capital losses may be carried forward indefinitely and used to offset gains realized in later years, gains realized while in region C may affect a taxpayer's *future* tax liability. We return to this point later. Note also that the gains calculated in the current year are net of any losses carried forward from earlier years.

The final region in Figure 1, labeled D, includes those taxpayers with long-term gains in excess of short-term losses. In this case, the long-term gain is reduced by the short-term loss and the difference is taxed at the long-term gain rate, τ^*. Thus, on the margin, all gains are taxed at the same rate, τ^*.

Poterba (1987) shows that successful use of capital gains tax avoidance strategies should lead investors to be in the vicinity of region C and to stay there over time, but most investors he observed did not appear in region C. Recent press reports indicate that this might have changed, however. Henriques and Norris (1996) argue, for example, that by exploiting devices like "short against the box" transactions that (until 1997) allowed constructive realization of a capital gain without triggering capital gains tax liability, many high-income taxpayers had learned how to escape taxation. That is, they approached region C by reducing their taxable gains to near zero. Several prominent economists quoted by Henriques and Norris agreed that high income taxpayers employed successful tax avoidance strategies.[4]

We are interested in understanding the behavior of those who are near region C, and to see if that behavior has indeed changed over time. We stress the word "near" because there is no clear division between taxpayers who actively use avoidance strategies and those who don't. A taxpayer who annually realizes a million dollars of gross gains and $995,000 of gross losses will always be in region A, but is qualitatively similar to investors who hit the $3,000 loss limit. Indeed, there may be some taxpayers who enter and remain in region C as the result of a single, unplanned loss, who should not be included in the group identified as successful tax avoiders. Our methodology attempts to take account of these and other issues of classification.

Before discussing this approach further, we turn to a brief discussion of the data set on which this analysis is based.

2. The Data

In our analysis we use the Internal Revenue Service's 1985-based Sales of Capital Assets (SOCA) panel study.[5] This panel was initially selected as a subsample of the 1985 Statistics of Income (SOI) cross-section of tax returns. The tax returns of panel members were then collected and linked for subsequent tax years through 1994. The data include full Federal individual tax return information for approximately 13,000 filers. In addition to the Form 1040 information that is in the standard SOI file, the panel also includes extensive detail on each capital transaction reported on Schedule D, Form 4797 (sale of business property), and several other forms on which capital gains are reported.

Several features of this panel make the data uniquely appropriate for the analysis of capital gains tax avoidance. First, the sample was highly stratified by income, creating an unusually large sample of wealthy taxpayers. Additionally, the fact that the data are a true panel allows observation of persistence of gains realization behavior and changes in behavior over time. Finally, the detail on individual transactions allows an analysis of heterogeneity by asset types of investors realizing capital gains and losses.

The SOI cross-section over-samples the returns of high-income individuals, and the subset of returns selected for the SOCA panel was even more top-heavy. Table 2.1 shows the resulting distribution of panel members by permanent income. Permanent income is defined here as the individual's mean over the panel years of the positive components of income expressed in 1982 dollars.[6] The top panel in the table indicates that over half of the panel members have permanent income above $200,000, and more than 2,000 members have permanent income above $1 million. Comparing the unweighted counts to the population-weighted counts reveals the extent to which these data oversample high-income taxpayers. The population weights

account for the panel's sampling stratification, and transform the panel aggregates to nationally representative levels in 1985.[7]

The importance of using a high-income sample of taxpayers for capital gains tax analysis is apparent in Table 2.2, which shows the distribution of net capital gains realizations by income for each year of the panel. In nearly every year, more than half of all net gains were realized by taxpayers with permanent income above $200,000, or the top 2 percent of all taxpayers. Note that, unlike similar results based on cross-sectional data, these conclusions do no represent transitory or timing effects. They represent, as a first approximation, the long-run relationship between capital gains and permanent income.

Table 2.3 summarizes the distribution of gains and losses over time between long-term and short-term. The dramatic timing effect of the capital gains tax rate changes in the Tax Reform Act of 1986 is clear. Realizations of long-term gains nearly doubled in 1986 and fell sharply in 1987.[8] Over the panel time frame, long-term losses grew relative to long-term gains. Short-term gains and losses both also grew substantially relative to long-term gains. These trends may reflect a lagged response to the 1987 rise in the long-term capital gains tax rate, and the end of the distinction in tax rates on long-term and short-term gains.

A drawback of this panel (and all other tax panels) is that exiting members are not replaced. Thus, the aggregate numbers may not represent the national population in later years for at least two reasons. The panel suffers from attrition, because some members die, some stop filing income tax returns because their incomes fall below the filing threshold, some taxpayers report the wrong Social Security number, and some returns are lost due to processing errors. A potentially more important source of panel non-stationarity is the aging of panel members. For those reasons, we compare the later years of the SOCA panel to SOI cross-sections from the

same years to test such panel drift. We find that attrition does affect the aggregate totals, but does not affect the qualitative conclusions in any apparent way. (See endnote 11 for an example.) There is also now a new 1993-based SOCA panel, which eventually can be used as a further check on panel drift.

For each taxpayer in the panel, the SOCA data contain detailed information on every asset with capital gain or loss that is sold, including: the type of asset by 21 classifications, the gain or loss, the sale price, and the purchase and sale dates. In order to utilize this information in a panel data set organized by individual, we summed each taxpayer's gains and losses, separately by asset type, term (long or short), and year. So, for example, we created variables for the individual's long-term stock gains, short-term stock gains, long-term stock losses, and short-term stock losses in each year of our sample.[9] Additionally, we recorded the number of transactions, and consolidated several of the asset classifications.

3. Evidence on Tax Avoidance Behavior Over Time

The panel data provide an extraordinarily detailed picture of the kinds of gains and losses people realize, and how they have changed over time. The earlier discussion suggests several working hypotheses to be examined using these data:

- Wealthier taxpayers are more likely to avoid tax on their capital gains than the less wealthy (because the former have larger, more diversified, portfolios and access to better tax advice).
- Gains on liquid assets, such as shares of corporate stock, should be more lightly taxed than gains on illiquid assets, such as real estate.
- Tax avoidance may have increased over time, because taxpayers, prodded by higher tax rates, learned successful techniques to shelter gains from tax.

To test these hypotheses, we examine how capital gain realization patterns vary by wealth or income, by asset type, and over time. We start out by examining how successful taxpayers are

at sheltering gains from tax in individual years, and then look at how such tax avoidance affects the distribution of taxes paid on capital gains.

Evidence on Tax Avoidance Activity

The perfect tax planner (in the frictionless world with complete financial markets) would have net capital losses of at least $3,000 every year. In this region, denoted C in Figure 1, both long-term and short-term capital gains are untaxed and losses have sheltered the maximum possible amount of ordinary income from tax. One simple test of whether investor behavior has been moving in this direction is to examine whether more taxpayers (or more gains) have been moving into region C over time.

Table 3.1 shows the percentage of taxpayers in each of the marginal tax rate regions over the ten years of the panel, based on three different weighting schemes.[10] The top panel of the table uses population weights. In this panel, we find, as did Poterba, that the majority of taxpayers with a capital gain or loss had both positive net short-term and long-term gains. Poterba reported that, in 1982, 64 percent of taxpayers were in that situation (region A in Figure 1). In 1985, we find that an even larger share of taxpayers – 77 percent – are in region A, but the percentage varies considerably from year to year, reaching a low of 56 percent in 1990.

There is, nonetheless, a clear break in 1987 – when the tax rate on long-term gains increased for most taxpayers. The percentage of investors in region A never approaches its level in 1985 in the subsequent years. It is tempting to conclude that this is a permanent response to the higher tax rates on capital gains, but many other factors make it hard to draw firm inferences. For example, the sharp decline in the stock market at the end of 1987 and the decline in real estate prices at the end of the 1980s both would have generated losses, although the stock market

was generally robust through most of the 10-year span. Moreover, the huge sell-off of assets in 1986 in anticipation of the increase in capital gains tax rates would have left investors with few capital gains for the years following enactment of TRA. All of these, though, were temporary phenomena.

The percentage of taxpayers in region A plummeted immediately after enactment of the Omnibus Budget Reconciliation Act of 1990 (OBRA90), which raised tax rates on ordinary income but capped rates on long-term gains. The higher tax rates on both long- and short-term gains should have deterred taxpayers from region B as well, but that percentage increased by the same amount that the percentage in A decreased. Thus, the drop may be coincidental.

The perfect tax planner should be in region C and stay there. Poterba reported that about 10 percent of investors in 1982 were in that region. We find that the percentage had fallen to 5 percent in 1985, but jumped to the levels found by Poterba after passage of TRA. Nonetheless, there does not seem to be a clear trend – the peak percentage was actually in 1990 (16 percent).[11]

Under the working hypothesis, wealthier taxpayers should be more likely to be in region C. A simple control for wealth is to weight the number of observations in each region by the average dividends earned from 1985 to 1994, which we refer to as "permanent dividends."[12] We use average dividends to smooth out transitory variations over time. Absent clientele effects, average dividends would be a good proxy for holdings of corporate stock over the ten-year period. If there is a clientele effect, this measure will cause tax-conscious investors in high marginal tax brackets to be underrepresented in the averages. In that sense, it will understate the impact of weighting by wealth.

The second panel of Table 3.1 reports the populations of the four regions, weighted by dividends. In this case, the movement into region C seems to demonstrate a clear trend that starts

in 1987, and continues for the next 7 years (peaking in 1992). The dividend-weighted percentage of investors in region C is not much different from the sample-weighted total in 1985 (4 percent versus 5 percent), but by 1992-1994, the dividend-weighted percentage is about twice the percentage using sample weights. These trends suggest that wealthier investors have become more likely to optimize their portfolio behavior over the 10-year period.

Finally, the table also shows the percentages weighted by average gross capital gains over the ten years. This weighting scheme tells us the fraction of gross capital gains in each year that fell into each of the four regions. In 1985 and 1986, only 1 percent of capital gains were in the tax-free zone (region C). In 1987, this percentage jumped to 5 percent. The percentage grew still larger in later years, though it bounced around considerably from year to year.

We would expect taxpayers with higher incomes to be more likely to be in region C, and that turns out to be the case. Table 3.2 repeats the calculations of Table 3.1, grouping investors according to their permanent incomes, as defined in Section 2. The table shows that taxpayers with incomes under $100,000 were much less likely than higher-income taxpayers to be in region C in 1985. By 1994, the percentage of lower-income people in region C increased, but the percentage of people with higher incomes in that region increased much more.

The distribution of taxpayers in region C by the level of their imputed wealth (the construction of which is described in Appendix 1) follows the same pattern. As Table 3.3 shows, wealthier people are more likely to shelter their gains from tax than less wealthy people, although the increase is not monotonic. The differences by wealth generally grow after 1986. The pattern is perhaps clearest if one compares the highest and lowest wealth categories. In 1985, 5 percent of taxpayers with a capital gain or loss and wealth below $0.5 million were in region C, compared with 8 percent of taxpayers with wealth greater than $50 million. In 1987, 8 percent of

the low-wealth taxpayers were in region C, compared with 12 percent of those with high wealth. But during each year between 1990 and 1994, from 27 to 34 percent of the wealthiest taxpayers were in region C, compared to a range of 7 to 13 percent of the least wealthy.

Because of high transaction costs, sales of illiquid assets such as businesses and real estate are likely to be motivated more by non-tax factors and harder to shelter from tax (especially for undiversified investors).[13] Thus we would expect more gains on liquid assets such as corporate stock and mutual funds to be in region C than gains on illiquid assets. Table 3.4 shows this to be the case, although the differences are not overwhelming. With the exception of 1986, a larger share of gains on stock, mutual funds, and bonds is in region C than on real estate and business assets.[14] The table also shows that gains from short sales, options, futures, and commodity contracts are much more likely to be in region C in most years. This may be because such investments are much riskier than typical assets, so a taxpayer who engages in such transactions is more likely to realize large capital losses than one who sticks with safer investments. Or, it could be that relatively sophisticated investors are more likely to engage in successful tax planning. For example, as explained earlier, short sales against the box were, until 1997, one of the major techniques for avoiding the realization of large taxable gains.

Based on these data, we created a measure of investor "sophistication" – an indicator of whether the investor ever traded options or sold short. Sophisticated investors by this measure were more than twice as likely to end up in region C as unsophisticated investors. (See Figure 2.) Both sophisticated and unsophisticated investors were much more likely to be in region C after 1986 than before, but the trends of the two groups diverge after 1987. The percentage of sophisticated investors in region C remained roughly constant at around 20 percent, a remarkable stability compared to the volatile time series of region C probabilities reported in Tables 3.2-3.4.

However, the share of unsophisticated investors in region C declines after 1990. A possible explanation for this peak is the recession of 1990-91. Losses are much more prevalent in a recession, and the pattern among unsophisticated investors seems to reflect that fact. The cyclical stability of the pattern for sophisticated investors suggests that their losses are driven by a different process – not so much by exogenous macroeconomic forces as by tax planning.

To what extent might differences in investor sophistication explain the patterns of the previous tables, which showed that investors with higher income and wealth were more likely to be in region C? Table 3.5 sheds light on this, showing that our measure of sophistication (here, based on annual participation in these markets) is strongly related to income.

Determinants of Tax Avoidance

The previous analysis suggests that successful tax avoidance is related to income, wealth, and the types of assets held in portfolio. We now bring these results together and consider the simultaneous effects of all of these factors, as well as demographic variables such as age and family status, and a time trend on the likelihood of being in region C, modeled using a probit equation. The results are presented in Table 3.6. Because of the considerable volatility in capital gains realizations associated with the Tax Reform Act of 1986, we consider only the post-reform period 1987-94. Also, we present two sets of estimates, the first based on sample-weighted observations, the second based on unweighted observations. While the former approach may seem more appropriate if we wish to characterize the behavior of the representative individual, the vast majority of capital gains are realized by people with high incomes. Thus, the unweighted data, which primarily represent higher-income taxpayers, better represent the population of those with substantial gains.

Based on the weighted data, people with higher permanent income (net of endogenous capital gains) are much more likely to be in region C, as are people with many capital transactions, another potential measure of an investor's sophistication and portfolio liquidity. Sophisticated investors (defined as before to be those who ever have traded commodities or options or engaged in short sales) are 9 percent more likely to be in region C than others. The share of mutual fund distributions in gross capital gains has a strongly negative effect.[15] This may reflect the fact that capital gain distributions (typically from mutual funds) are involuntary, and that investors with large mutual fund holdings are less actively involved in portfolio management and tax planning. Growth in GDP and the stock market have the expected negative impact on the probability of a net loss, but neither effect is significant. Wealth, the shares of different asset types, and the demographic variables, with the exception of marital status, are insignificant. Finally, note that, with all the other factors accounted for, the probability of being in region C does not change significantly over time.

However, one should use caution in interpreting these results, because the process of weighting, while appropriate for characterizing the behavior of the overall population, gives relatively low weight to the higher-income investors who realize most capital gains. In the unweighted estimates, permanent income has a negligible and statistically insignificant effect, but the wealth effect becomes large and significant. The shares of wealth accounted for by stock and farm property now have significant negative effects on the probability of having a net capital loss. While the impact of farm wealth is not surprising, given the illiquidity of such property, the effects of the stock share are less easily explained. One possibility is that, during this period of rapid stock market growth, individuals with large stock portfolios had especially high accrued gains, some of which were realized. Although this effect should be picked up in part by the

growth rate in the Standard and Poor's index – the coefficient of which becomes large and highly significant – that index is not a perfect measure of broader stock market wealth. Municipal bond interest also has a significantly negative impact, possibly reflecting the use of a more passive strategy to avoid taxes.

One important result that does carry over from the weighted estimates is the impact of our measure of sophistication. Although the coefficient is somewhat smaller than before, the effect of sophistication on the probability at the mean of being in region C is virtually identical – approximately 9 percent. A final difference from the weighted estimates is the large and statistically significant positive effect of the time trend. The trend accounts for about a 12 percentage point increase in the probability of being in region C from 1987 to 1994.[16] The explanation for this trend appears to be the greater weight given to "sophisticated" investors in the unweighted estimates, whose region-C population remains steady after 1990 even as that among unsophisticated investors falls (see Figure 2).

To test this theory, we re-estimated the weighted probit with interaction terms for sophistication with the time trend and GDP growth (not shown in the table). After these additions, GDP growth has negligible effect for sophisticated investors, suggesting that their presence in region C is insensitive to cyclical variation, unlike unsophisticated investors. In this alternative specification, the time trend for sophisticated investors is positive and nearly as large as that reported for the unweighted specification in Table 3.6.

The Duration of Tax Avoidance

The benefit of realizing additional capital gains while in region C depends on how long a taxpayer expects to stay there. An individual who is in region C in one year and region A the

next is not really untaxed on marginal gains – he is only deferring tax for a year. That is, if the taxpayer realizes a gain of *g* while in region C, he incurs no current tax liability now, but his tax liability increases in the following year, because he has that many fewer losses to carry over. So his effective tax rate is $\tilde{\tau}/(1+r)$, where *r* is the investor's nominal discount rate, and $\tilde{\tau}$ is the tax rate applicable to gains realized in the second year. If he stays in region C for two years, his effective tax rate is $\tilde{\tau}/(1+r)^2$, and so on (assuming his tax rate stays unchanged). Since future gains and losses are uncertain, his effective tax rate is stochastic.

While this uncertainty makes a full analysis quite complex, a taxpayer's decisions presumably depends on an expected effective tax rate, $\bar{\tau}$, defined as

$$\bar{\tau} \equiv \sum_{t=1}^{\infty} f(t|X_0)\frac{\tilde{\tau}(X_0)}{(1+r)^t} \tag{1}$$

where $f(t|X_0)$ is the probability of staying in region C for exactly *t* periods conditional on being in region C in period 0 and other information known at time 0, X_0.[17]

We estimate the duration in region C using an exponential hazard function,

$$h(t|X_0) = e^{\beta_{0t} + X_0\beta}, \tag{2}$$

where β_{0t} is a parameter that varies with duration, *t*, and β is a vector of constants. The hazard function is the probability of exiting region C in period *t* given X_0.[18] This specification allows for arbitrary duration dependence, because the β_{0t} are not constrained. The parameters are estimated by maximum likelihood.[19]

The probability of a duration of *t* may be derived from the estimated hazard functions. It is the product of the probability of remaining in region C for *t* periods – the survival function,

$s(t|X_0)$ – and the hazard in period t. The survival function, in turn, is simply the probability of not exiting region C in each of the previous periods, which is

$$s(t|X_0) = \prod_{j=1}^{t-1}(1 - h(j|X_0)), \tag{3}$$

with the initial condition that $s(1|X_0) = 1$. Substituting $f(t|X_0) = s(t|X_0)\, h(t|X_0)$ into equation (1) yields

$$\bar{\tau} \equiv \sum_{t=1}^{\infty} h(t|X_0)s(t|X_0)\frac{\tilde{\tau}(X_0)}{(1+r)^t} \tag{4}$$

We estimate the hazard function $h(\cdot)$ for the period 1987-94, again using both weighted and unweighted samples and most of the same covariates (some time-varying) as those used in the probit estimation above. Table 3.7 presents the estimation results. One effect that is not surprising is that a large capital loss carryover significantly reduces the hazard rate. All else equal, the larger this loss overhang, the longer it takes an investors to use it up.

In comparing the remaining hazard model results to those in Table 3.6, one should keep in mind that variables that increase the rate of departure from region C have a positive coefficient. Thus, variables that are associated with tax avoidance not only by contributing to presence in region C but also to longer duration in region C would have a positive sign in Table 3.6 but a negative sign in Table 3.7. Among the variables in this category are wealth and our measure of investor sophistication, each of which is negative and significant in both weighted and unweighted estimates. Other variables with consistent effects across the two tables are shares of mutual funds and (for the unweighted specification) farm property and stock, which

reduce presence in and increase the rate of exit from region C. The trend over the period is positive and significant, indicating an increase in exit rates over time, perhaps reflecting the impact of stock market growth (or growth in other assets) not fully accounted for by the growth rate of the S&P 500 index.

What is the net impact of these individual effects, taken together, on the hazard rate? The answer, of course, varies across individuals, but we can get an idea of the aggregate picture by considering the hazard rates predicted at the mean values of all the covariates. Table 3.8 presents these predicted hazard rates for the weighted and unweighted estimates. For comparison, it also presents observed ("empirical") hazard rates for the unweighted sample which, not surprisingly, are quite close to the predicted values. Except for an unexplained blip at the six-year duration in the weighted sample, the two sets of estimates exhibit very strong negative duration dependence. Close to half of all investors in region C depart after one year, but hazard rates fall nearly monotonically thereafter.

One possible explanation for this apparent duration dependence is unobserved heterogeneity of the region-C population. Not all individuals in region C exercise a tax avoidance strategy. Some, perhaps, follow simpler realization strategies, but occasionally realize losses. Those investors probably have much higher exit rates than those vigorously pursuing tax-reduction strategies. The investors who remain in region C for more years are more and more likely to be the aggressive tax avoiders.

Another possible explanation for duration dependence is noise in our measure of tax avoidance. We identify individuals as being tax avoiders only if they are in region C, taking the maximum allowable deduction for capital losses. However, as noted in the introduction, it may not make sense to distinguish this behavior from that of a taxpayer who shelters all or nearly all

of his capital gains every year without hitting the exact $3,000 limit. The presence of taxpayers hovering "near" region C and randomly hitting the limit exactly could well introduce a spuriously high exit rate at short durations. In fact, as Figure 3 illustrates, the distribution of investors is bimodal. Two thirds of taxpayers with gains or losses are able to shelter less than 10 percent of their gains. (Actually, the denominator in the figure is gain plus $3,000, defined this way so that a taxpayer must reach the boundary of region C to offset 100 percent of gains.) About 12 percent of taxpayers shelter all their gains—that is, they are in region C—but only 1 percent shelter between 90 and 100 percent. The figure also shows the percentage of gains actually sheltered by losses. The bimodality remains, although considerably more gains than taxpayers are in the 0 to 10 percent sheltered category. In addition, only 6 percent of gains are fully sheltered (compared with 12 percent of taxpayers). This suggests that taxpayers who fully shelter their gains have smaller than average gross gains. This would be expected, because it is easier to generate enough losses to shelter a small gain than a large one. But it is also consistent with the idea that some taxpayers are in region C because they use the tax-avoidance strategies mentioned earlier to reduce the amount of their gross taxable gains.

We tested the sensitivity of our empirical hazard estimates by using a variety of different definitions of tax avoidance, adding to region C each taxpayer who offset at least x percent of his gross capital gains in a particular year. The results of one such specification, with $x = 100$, are given in Appendix 2. These results are quite similar to those based on the stricter definition.

Effective Tax Rates on Realized Capital Gains

Using the hazard rates presented in Table 3.7, we may use the formula given in expression (4) to calculate effective tax rates on realized gains for each individual in our

sample.[20] We focus on long-term capital gains, as these have been the subject of the greatest policy discussion over the years. Before doing so, however, we must resolve a number of technical issues.

First, as our hazard estimates do not go beyond a duration of seven years, we assume that the hazard rate (which already is very low) remains constant thereafter. Second, we must make assumptions about the values of time-varying covariates (such as GDP growth). We set values of such variables equal to their sample means, with the exception of the time trend, which we assume equals its value at the end of the sample period (7). Third, we must make some allowance for sample attrition. Our estimates simply exclude longer durations for individuals who disappear from the sample. Using these estimates to calculate tax rates implicitly assumes that attrition is uncorrelated with individuals' hazard rates. An important case in which this will not be true is when attrition is due to the death of the taxpayer. Because capital gains taxes are permanently forgiven at death, the correct treatment of a taxpayer who dies is to impose a hazard rate that is permanently equal to zero. Unfortunately, we cannot identify the reason for attrition. Therefore, we performed the calculation under two extreme assumptions: that attrition is random, and that attrition is always due to death. The results are virtually identical for the two cases (primarily because attrition while in region C is relatively unimportant). Thus, we present only those based on the former assumption. Finally, to apply expression (4), we need a discount rate and a value for $\tilde{\tau}$, the tax rate on long-term gains in the year the taxpayer leaves region C. We assume a value of 7 percent for the nominal discount rate. As for the value of $\tilde{\tau}$, it is important to keep in mind that exit can occur into one of three regions, A, B, and D. While the tax rate on long-term gains equals τ^* in regions A and D, it equals the ordinary tax rate, τ, in region B. We assume that each individual's probability of exit into each region equals the observed sample

probabilities. And further, since the value of τ in future years is not necessarily the same as in the year of a gain realization, we assign the expected future ordinary income tax rate based on permanent income.

The results of these calculations are presented, for selected years, in Table 3.9, which shows the distribution of effective marginal tax rates on realized long-term capital gains broken down by the taxpayer's marginal tax rate on ordinary income. Gains realized in regions A, B, and D are assigned the rate appropriate to the taxpayer, region, and year, and those realized in region C are based on the methodology just described.

The median effective tax rate on long-term capital gains is identical to the statutory rate in every year. Though this may appear surprising, it is implied by the fact that a minority of taxpayers at each income level are in region C (for which the tax rate is lower) or in region B (for which the tax rate may be higher). Thus, in all years, more than half of the population that realizes capital gains is avoiding no tax at all on the gains they realize. Indeed, this identity holds for the 25^{th} and 75^{th} percentiles of the distribution of tax rates in every year except 1990 (when the tax rate at the 25^{th} percentile is 23 percent, not shown).

Not surprisingly, taxpayers in the lowest tax bracket (15 percent) are least likely to avoid capital gains tax. In most years, fewer than 5 percent have an effective rate below their statutory rate—and about an equal number manage to have their gains taxed at rates above the statutory rate. This occurs because some people who enter region C when their marginal tax rate is 15 percent may expect their rate to be higher when they exit region C. In addition, lower-income taxpayers who enter region C do not stay there very long. As a result, the overall average marginal tax rate for people in the 15-percent bracket is very close to 15 percent in every year.

In the higher tax brackets, a somewhat larger fraction of the population faces effective tax rates below the statutory rates. The lowest one percent of effective rates is more than 10 percentage points below the statutory rate, but the overall effective capital gains tax rate is still very close to the statutory rate in every tax bracket, with the difference never exceeding 3 percentage points. Indeed, in 1991 and 1994, more than 10 percent of taxpayers in the top bracket face effective rates above the statutory long-term capital gains rate. This is because the tax rate on long-term capital gains in region B equals the tax rate on ordinary income, which after the 1990 Act exceeded the statutory rate on long-term gains for higher-bracket taxpayers.

In part because of this factor, the overall effective tax rate on a dollar of long-term capital gains has declined only slightly between 1988 and 1994, despite the increased likelihood of being in region C. Perhaps more important is the fact that presence in region C exerts a relatively small impact on an investor's effective tax rate. Given that roughly half of all investors in region C depart in one year, and about two-thirds within two years, the typical tax rate reduction is relatively small, perhaps 5 percentage points. For, say, a 10 percent increase in the share of capital gains in region C (an upper bound, based on the numbers given in the bottom panel of Table 3.1), this would induce a mere 0.5 percentage point drop in the average marginal effective tax rate, a change small enough to be lost amid other changes occurring simultaneously over the period.

This is apparent is Table 3.10. Investors in region C in 1994 could expect an effective tax rate only slightly below their statutory rate if they are in the 15-percent bracket. Taxpayers in the higher brackets could expect greater discounts from the statutory rate, but the difference is not dramatic. The largest difference is only about 10 percentage points for taxpayers in region C in the 28 or 31 percent brackets. Taxpayers in the highest brackets, starting from region C, actually

face higher effective tax rates than those in the intermediate brackets, because if they exit into region B, they are likely to face tax rates as high as 39.6 percent.

Tax Avoidance, Progressivity, and Fairness

Our findings dispel two contentions made about the fairness of capital gains taxation. The first is that high-income people can avoid the tax at will, which subverts the slight progressivity that had been designed into long-term capital gains tax rates. The second is that the loss limitation is especially unfair to lower income taxpayers with only a single asset (e.g., ma and pa's grocery store), who therefore would never be able to fully deduct a catastrophic loss against other gains or their other income.

In fact, average effective tax rates on realized capital gains are very close to statutory rates. Furthermore, ma and pa seem to be least likely to be constrained by the $3,000 loss limit. People in the 15-percent bracket are least likely to enter region C and, when they do, they don't stay there long. That is why their effective tax rate on long-term capital gains is nearly identical to the statutory rate.

4. Evidence from the Survey of Consumer Finances

An important limitation of all the results presented thus far is that they relate only to tax avoidance associated with realization behavior. Thus, we have not focused on investors who fail to offset realized gains with realized losses. But this downplays the effect of an alternative avoidance mechanism, namely, the deferral of accrued gains, possibly until they receive favorable treatment at death or through a charitable contribution. Because such taxpayers might have little or no gross long-term gains, we will understate the effect of such strategies on the

overall effective tax rate. Even if the effective tax rate on realized gains is high, and not strongly related to income or wealth, this may not be true of the rate on accrued gains.

To gain a more complete picture of the relationship between realized and accrued gains, we look at the only available evidence on unrealized gains, from the Survey of Consumer Finances (SCF). Figures 4 and 5 compare the distribution of average realized gains from 1988 to 1994 on the SOCA (omitting 1987 to eliminate timing behavior around TRA86) with the distribution of accrued gains from the 1992 SCF.[21] If higher income people more successfully avoid realizing taxable gains, then accrued gains should be more concentrated among high-income people than realized gains. However, this pattern is not in evidence in these two data sets. For corporate stock, taxpayers with over $100,000 of income realized about 87 percent of gains in the average year of the period, whereas their accruals accounted for only 70 percent of gains. For business assets, the respective values for realizations and accruals are 76 percent and 61 percent.

These comparisons should be regarded with caution for several reasons. The SCF has relatively few very high-income respondents, and aggressive tax avoiders might have been less inclined to participate in the survey. The definitions of income are similar, but not identical. Finally, the long-run ratio of realizations to accruals need not be accurately pictured by data from a relatively short panel. Nonetheless, the lack of significant evidence of capital gains tax avoidance by high income people is consistent with our general results.

5. Conclusions

Our analysis has extended the work of Poterba (1987) to look at more recent panel data. We find evidence consistent with his general conclusion – that tax avoidance is not prevalent,

even after passage of tax reform, and that most high income people realize gains that are not sheltered by losses. Like Poterba, we also find that a minority of taxpayers – mostly with higher incomes and wealth – manage to shelter all or most of their gains with losses.

We find evidence that tax avoidance increased after 1986, and that it increased most for high-income, high-wealth taxpayers. As many as one-third of the wealthiest taxpayers were able to realize their gains without immediate tax (i.e., were in region C, as we have defined it) in the early 1990s. Moreover, we found that a subset of sophisticated investors were consistently more likely to be in region C than others. Through multivariate probit analysis, we demonstrated that this result persisted after controlling for other variables, such as income and wealth, and was robust with respect to weighting.

But the efficacy of tax avoidance strategies depends on being able to remain in region C for long periods. We found that most taxpayers exited quickly. Only about half of taxpayers are still in region C after one year; about one-quarter make it for at least three years. Combined with the small proportion of taxpayers in region C in the first place, this implies that the overall effective tax rate on realized capital gains is much closer to the statutory rates than is apparent based on a single-year's perspective. Again, however, it is the sophisticated investors who consistently remain in region C for longer than others, so a subset of taxpayers are able to shelter more of their gains from tax than most people.

Much further research is needed in this area. Our analysis focused primarily on realized capital gains. We could make only indirect inferences about the gains that are never realized, but which represent the most successful avoidance strategy. Since our analysis is inherently based on a reduced form model, it is hard to draw firm inferences about the structural parameters involved in people's decisions, and how they might have changed over time. These questions have

proved daunting because of a lack of data and difficult conceptual problems, but are still worth pursuing.

Appendix 1: Data and Methodology

This appendix describes the construction of the variables and methodology we use in our estimations reported in Section 3. Most of the data come from the 1985-based SOCA panel, which is introduced in the data section of this paper.

Nearly all of the variables in the SOCA panel are obtained directly from individual Federal tax returns for the panel members. While the data do not include every conceivable supplementary IRS form, the data include all the line-by-line entries from:

- Form 1040,
- Schedule D (Capital Gains and Losses),
- Form 4797 (Sales of Business Property),
- Form 2119 (Sale of a Residence),
- Form 6252 (Installment Sales), and
- Form 8824 (Like-Kind Exchanges).

The only data in the SOCA panel that do not originate from the Federal income tax returns of the filers are the date-of-birth for both the primary and secondary taxpayer. This information was obtained through a merger with the Social Security Administration records of the taxpayers.

We linked the panel by matching the Social Security numbers in the separate files from each year and IRS form. A complication that arose in this process was that joint filers in 1985 did not always remain as joint filers (in the same combinations) throughout the panel: a result of death, divorce, marriage, and other changes in filing status. Where possible, the IRS included in the SOCA data files the returns from both Social Security numbers. This forced us to make a

decision about what constitutes the appropriate panel observation. In general, we chose to follow the Social Security number listed as primary in 1985 when there was a conflict.

Constructed Variables

We created from the tax filings a variable for permanent income. This measure is not directly dependent on changes in tax code definitions of gross income and is not sensitive to transitory income variations. Permanent income is defined in all of the tabular results we present as the mean of real, positive income over the 10 years of the panel. For taxpayers not in the sample for all years, we use the mean over the available years. Positive income is the sum of the positive components of income from: wages, taxable and tax-exempt interest, dividends, alimony, business income, capital gains, supplemental gains, Schedule E (rental & royalty) income, IRA distributions, pension income, farm income, unemployment insurance benefits, taxable social security benefits, and other income. We use only the positive components since large business and capital losses are usually realized by only the wealthy, which, if included, would make some individuals with high lifetime income appear to have low income. We normalized prices using the Consumer Price Index (CPI), with a base period of 1982-84. In the probit and duration models, we removed capital gains from permanent income to purge that variable of a source of endogeneity.

We imputed wealth by capitalizing capital income reported on tax forms. Taxable and tax-exempt interest are capitalized based on the 3-month Treasury bill rate. Dividends are capitalized by the average dividend payout rate. We used the same rate to capitalize realizations of positive business income, Schedule E income, and farm income. To prevent transitory income shocks from causing volatility in this imputation, the average wealth over the panel years is used

for each individual. The variables for the shares in wealth of stock, business property, rental property, and farm property are the panel-year averages of the fractions of the total wealth imputation attributable to capitalized dividends, business income, Schedule E income, and farm income, respectively. The indicator variable for earning tax-exempt interest is set to one if the panel member earns such income in any year.

There are several well-known limitations to this method of imputing wealth. Even if average total returns to capital are similar across individuals, capital income payout rates may differ across individuals and capital types, and may depend on tax-related variables. For example, if there is a substantial dividend clientele effect there will be a systematic underestimation of the wealth of the high-tax-rate individuals.

Three variables are constructed from the characteristics of the taxpayers' capital gains realizations. The mutual fund distributions' share of gains is the simple ratio of mutual fund gains to total gross gains plus $3,000 (to avoid dividing by zero for taxpayers with no gains). The measure of taxpayer sophistication is set to one if the panel member ever engages in a gain or loss transaction involving a stock short sale, option, commodity, or futures contract. Our logic in defining it this way is that individuals who for at least one time have access to such markets are likely to have a permanently higher level of access. (We tested this assumption by using a sophistication measure based on only the current year's activity and it did not change any of the results significantly.) On the SOCA data set, short sales of stock are identifiable because the dates of sale and purchase are reversed. Options, commodities, and futures are coded as an asset type in the transactions data. The variable for the number of capital transactions is the total count of asset sales (both gains and losses) of all types in a year. In some cases, such as mutual fund

distributions, the number of sales was not distinguishable (and perhaps not relevant), and therefore counted as one transaction.

The variable for the size of loss carryovers in the hazard model is constructed as the log of the ratio of the amount of the carryover to $3,000 plus the average size of current and recent gross gains. The specific formula used is,

(A1) $$x_t = \log\left(\frac{C_t}{3000 + \frac{1}{3}\sum_{t-2}^{t} G_t}\right)$$

where C_t is the amount of loss carryover, G_t is the amount of gross capital gains, and 3000 is added to the denominator to ensure that the variable is meaningful when gains are small (or zero).

Probit Estimation of Probability of Presence in Region C

The estimation of the probability of being in region C, reported in Section 3, was done by commonly used methods for modeling discrete choice. The data used are the pooled observations of each taxpayer in each year reporting a capital gain or loss. The dependent variable is a binary variable equal to one when the taxpayer is in region C.

We chose a probit model, which assumes the values of the independent variables, in vector **x**, relate to the probability of being in region C in the following manner:

(A2) $$\Pr(\text{Individual in Region C}) \equiv F(\mathbf{x}) = \int_{-\infty}^{\beta'\mathbf{x}} \phi(z)dz = \Phi(\beta'\mathbf{x}),$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and cumulative distribution functions of a standard normal, respectively.

The parameters were estimated by maximum likelihood. The robust standard errors are estimated by assuming that observations for the same individual are not independent (and estimating their covariance). To aid in interpretation, the partial derivatives of the probability function with respect to the independent variables, $\partial F / \partial \mathbf{x}$, are evaluated at the mean values for each independent variable. In the case of binary indicator variables, $\partial F / \partial \mathbf{x}$ reported is calculated as the change in probability associated with a change in the variable from 0 to 1 holding all other variables constant at their mean values.

Estimation of Duration in Region C and Computation of Effective Tax Rates

We used a proportional hazards model to estimate the probabilities of exiting region C at various durations. Rather than making a specific assumption about the form of the hazard function's dependence on duration, we used the semi-parametric approach of estimating a separate constant at each duration. So, our form for the hazard function is,

(A3) $$h(t \mid X) = \exp(\beta_{0t} + X\beta),$$

where β_{0t} is a constant for duration *t*.

Most of the variables included in *X* are not time-varying. Permanent income, wealth, wealth shares, and sophistication are all defined as permanent variables. We use the values only for the initial year in region C for the carryover, mutual funds distributions share, and age variables. However, we use time-varying values of the GDP and S&P-500 growth rates, and a time trend, that do not remain constant through the duration of an individual's spell in region C.

The coefficients of the hazard model, including the duration constants, are estimated by maximum likelihood. From these estimates, we can construct an estimated hazard function for

each individual. We can also look at the hazard function for any set of values for the covariates in *X*. In Table 3.8, we did so for the mean values of the covariates, $h(t \mid \bar{X})$.

We use the formula in equation (4) in Section 3 to convert hazard estimates to effective tax rates, making the assumption that the constant in the hazard function at durations beyond the reach of our sample (seven years) is equal to the constant at seven years. The survival function, $s(\cdot)$, is a function of the hazard rates,

$$s(t \mid X) = \prod_{v=1}^{t-1} [1 - h(v \mid X)] \qquad \text{for } t > 1, \text{ and } s(1) = 1. \tag{A4}$$

Since we are computing *ex ante* effective tax rates with our hazard model, we do not assume that the variation in GDP and S&P-500 growth is known, and replace the realized values for those variables with their means over the sample years. The trend variable is allowed to vary within the sample years in this calculation, but is kept at its 1994 level for subsequent years.

Appendix 2: Alternative Definitions of Region C

Theory suggests that people who successfully avoid capital gains tax should be found in region C, the area bounded by the net loss offset limitation of $3,000. However, for many investors, there may be only a small financial difference between facing that constraint, and being near it, while offsetting most or all of their gross capital gains. To examine a broader definition of region C than that used in Section 3, we redefined the region to include those taxpayers that offset high percentages of their gross gains, considering four alternative levels: 100, 90, 75, and 50 percent offset.

Tables A.1 (for unweighted and weighted samples) and A.2-A.3 (for the unweighted sample only) present results from the estimations with the region C boundary repositioned at full (100 percent) offset of gross gains, and correspond to Tables 3.6 through 3.8. A comparison of the estimates for the alternative definitions suggests that the choice does not change any of the results significantly and has no effect on our qualitative conclusions.

References

Altshuler, Rosanne, and Alan J. Auerbach, "The Significance of Tax Law Asymmeteries: An Empirical Investigation," *Quarterly Journal of Economics* **105**, February 1990, 61-80.

Auerbach, Alan J., "Capital Gains Taxation in the United States: Realizations, Revenue and Rhetoric," *Brookings Papers on Economic Activity* **19**, Fall 1988, 595-631.

Auten, Gerald E., and Charles T. Clotfelter, "Permanent Versus Transitory Effects and the Realization of Capital Gains," *Quarterly Journal of Economics* **97**, November 1982, 613-632.

Blank, Rebecca M., "Analyzing the Length of Welfare Spells," *Journal of Public Economics* **39**, August 1989, 245-73.

Burman, Leonard E., Kimberly A. Clausing and John O'Hare. "Tax Reform and Realizations of Capital Gains in 1986," *National Tax Journal* **47**, March 1994, 1-18.

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.

Congressional Budget Office, "Perspectives on the Ownership of Capital Assets and the Realization of Capital Gains," May 1997.

Constantinides, George M., "Optimal Stock Trading With Personal Taxes," *Journal of Financial Economics* **13**, March 1984, 65-89.

Czajka, John, "Income Stratification in Panel Surveys: Issues in Design and Estimation," in American Statistical Association, *Proceedings of the Section on Survey Research Methods,* (Alexandria, Va.: American Statistical Association, 1994).

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

Gravelle, Jane G., "Limit to Capital Gains Feedback Effects," CRS Report 91-250RCO, March 1991.

Henriques, Diana B., and Floyd Norris. "Wealthy, Helped by Wall St., Find New Ways to Escape Tax on Profits." *New York Times*, December 1, 1996.

Holik, Dan, "The 1985 Sales of Capital Assets Study," in American Statistical Association, *Proceedings of the Section on Survey Research Methods*, (Alexandria, Va.: American Statistical Association, 1989).

Meyer, Bruce D., "Unemployment Insurance and Unemployment Spells," *Econometrica* **58**, July 1990, 757-82.

Poterba, James M., "How Burdensome are Capital Gains Taxes?" *Journal of Public Economics* **33**, July 1987, 157-172.

Seyhun, H. Nejat, and Douglas J. Skinner, "How Do Taxes Affect Investors' Stock Market Realizations? Evidence from Tax-Return Panel Data," *Journal of Business* 67, April 1994, 231-62.

Stiglitz, Joseph E., "Some Aspects of the Taxation of Capital Gains," *Journal of Public Economics* **21**, June 1983, 257-294.

U.S. Congressional Budget Office. "Perspectives on the Ownership of Capital Assets and the Realization of Capital Gains." CBO Paper, May 1997.

Endnotes

[1] Also see Seyhun and Skinner (1994).

[2] The figure is adapted from Poterba (1987) for changes in the treatment of capital losses introduced by the Tax Reform Act of 1986. Poterba's figure reflecting pre-1986 law had seven distinct regions. A comparable figure for present law, when fully phased in, would require four dimensions to graph.

[3] Our analysis accounts for the 33-percent bubble region in effect from 1988 to 1990, but not the quantitatively less significant phase outs of itemized deductions and personal exemptions in effect after 1990, which also raised effective tax rates. We also ignore the effects of the alternative minimum tax.

[4] Henriques and Norris quote David Bradford as saying, "The simple fact is that anyone sitting on a big pot of money today probably isn't paying capital-gains taxes and the government can adopt rule after rule after rule – but the people who will get stuck paying capital-gains taxes will be the ordinary investors who own mutual funds."

[5] See Congressional Budget Office (1997) for a discussion of these data. All tabulations and estimation based on confidential tax return data were conducted by Jonathan Siegel while he was employed by the Congressional Budget Office.

[6] Income in each year is calculated independent of the tax code by summing the positive components of income, rather than using Adjusted Gross Income (AGI), which depends on the tax code.

[7] For a discussion of some of the stratification and weighting issues, see Czajka (1994) and Holik (1989).

[8] See Burman, Clausing, and O'Hare (1994).

[9] Since the SOCA transactions data come directly from the tax forms on which they originate, we excluded from this summation some gains and losses in the data set that are nontaxable or subject to ordinary income treatment. For example, nontaxable personal residence gains from Form 2119 were not included, nor were section 1231 losses and recaptured gains and losses. Furthermore, wherever detail by transaction is not needed in our analysis, we use the totals for gains and losses from Schedule D, which reflect only those realizations subject to capital gains treatment.

[10] As mentioned in endnote 2, the region definitions changed slightly after 1986. To allow comparison between 1985-86 and subsequent years, we use the post-1986 region definitions to classify taxpayers in all years. This procedure has the effect of increasing the fraction of taxpayers in region C in 1985-86, but only very slightly.

[11]As mentioned earlier, the results may be distorted by the effects of attrition in the panel. For example, if the people who leave the panel (because they do not file a tax return, die, or misreport their Social Security number) are primarily the less tax-motivated investors, these data may suggest more tax planning than really occurs in the population. We compared the estimates with data for the large Statistics on Income (SOI) sample, which is drawn every year and intended to be representative of the population of taxpayers. As in the SOCA data, about 5 percent of investors were in region C in 1985; however, the percentage in the SOI increases only to 8 percent by 1994, compared with 12 percent in the SOCA panel, which suggests that attrition may alter our numerical estimates. Nonetheless, the qualitative conclusions are the same in the representative cross-sections as in the panel.

[12] Poterba (out of necessity) uses annual dividends to weight his data.

[13] Gravelle (1991) shows why assets with high transaction costs should be less sensitive to tax rates on capital gains than more liquid assets.

[14]The large changes in year-to-year percentages for some assets, notably business property, short sales, and options, appear to be attributable to two factors. First, the denominator of the calculation, aggregate net gains for the asset class, can be quite small in any given year for such volatile investments. This magnifies small absolute fluctuations in the level of gains in region C. Second, the weights chosen in 1985 were based on a single year's income. As a result, some apparently low-income people with very high weights were actually quite wealthy with high incomes in most years, a combination that can lead to volatility.

[15]This variable equals the ratio of mutual fund distributions to gross gains plus 3,000 dollars.

[16] The trend effect is calculated by comparing the probabilities (at the mean values of the other variables) with the trend set to 0 and 7, respectively.

[17] This approach is similar to that used to calculate the effective tax rate for firms with tax loss carryforwards (e.g. Altshuler and Auerbach 1990), but is simpler in part because capital losses may be carried forward indefinitely.

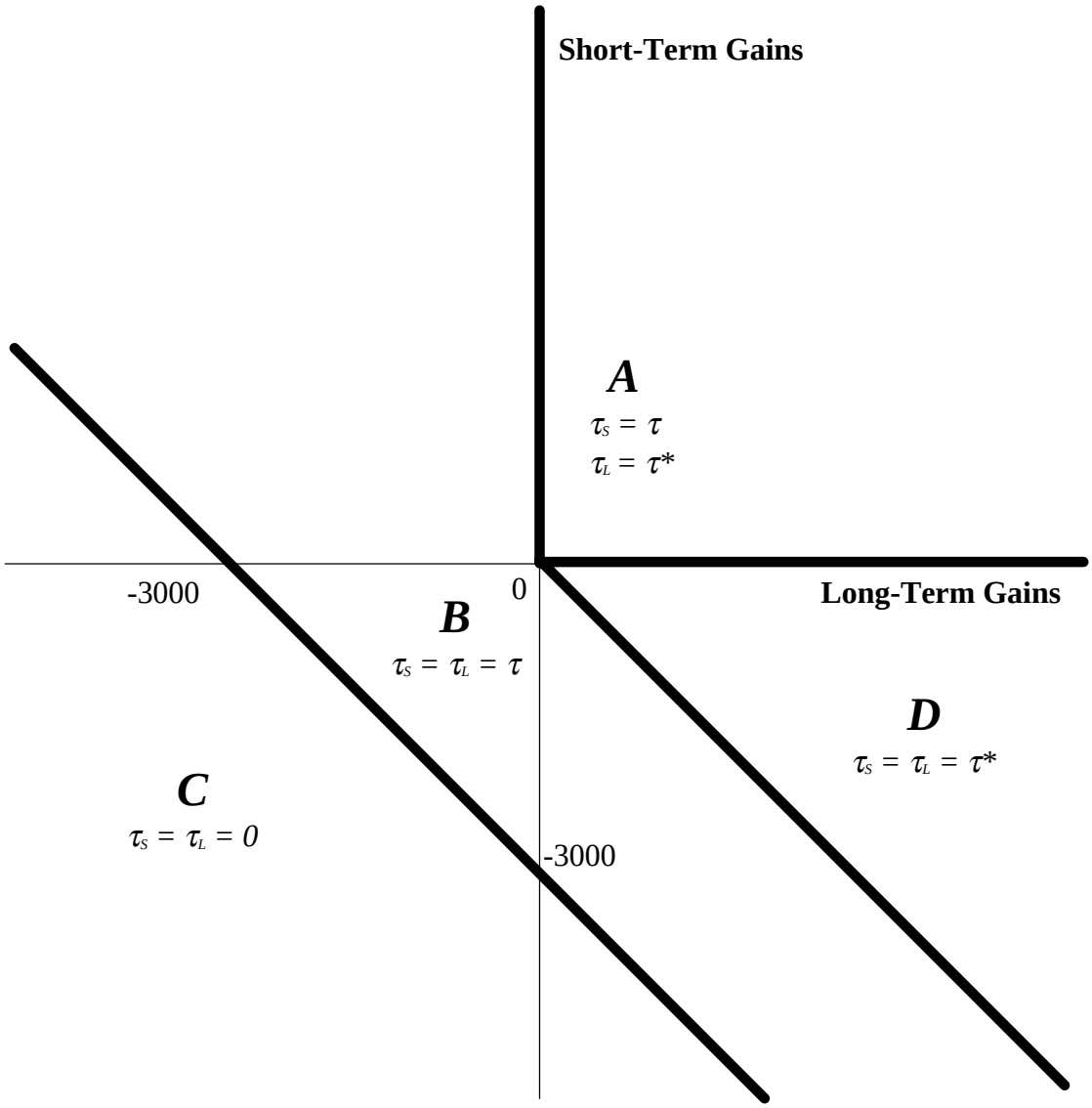
[18] See Appendix 1 for more discussion of the hazard model.

[19] Note that this is a continuous time model, but is often used in the economic literature as an approximation for discrete data. See, e.g., Blank (1989) and Meyer (1990).

[20]Keep in mind that the effective tax rates computed in this section do not account for the non-taxation of accrued but unrealized capital gains held at death. This issue is addressed in a subsequent section.

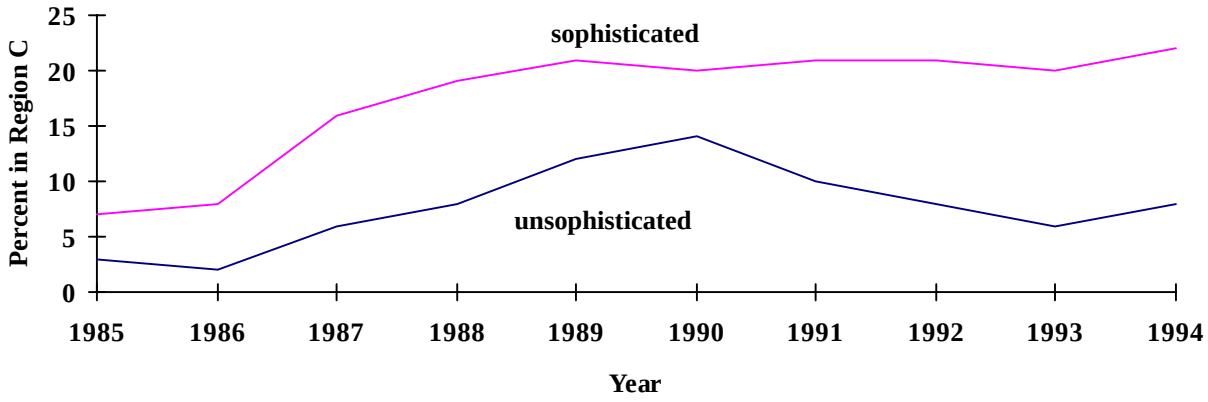
[21] We're grateful to Jeff Groen, formerly of the Tax Analysis Division of CBO, for providing the tabulations from the SCF as well as useful advice about how to interpret them.

Figure 1. Regions of Taxpayer Behavior



Note: τ_S and τ_L are the applicable tax rates in each region on short-term gains and long-term gains, respectively. τ and τ^* are the statutory tax rates on ordinary income and long-term gains.

Figure 2. Sophisticated versus Unsophisticated Investors



Note: Sophisticated investors are those who ever trade in options or future contracts or sell short.

Figure 3. Percent of Gross Gains Offset: Distributions of Taxpayers and Gross Gains

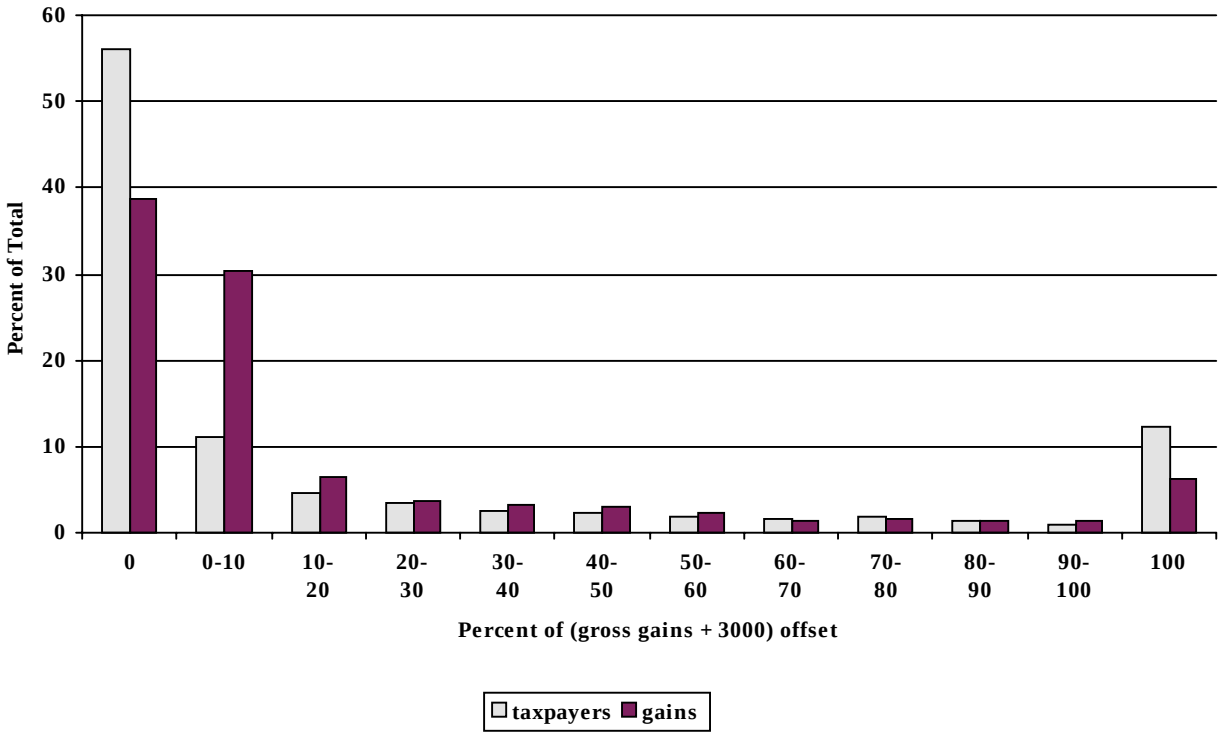


Figure 4. Percentage of Gains in Each Income Class: Stock

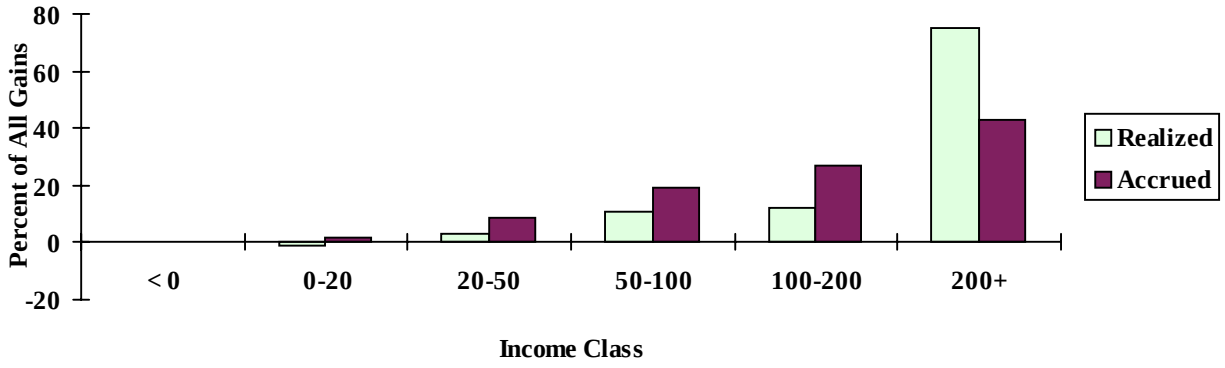
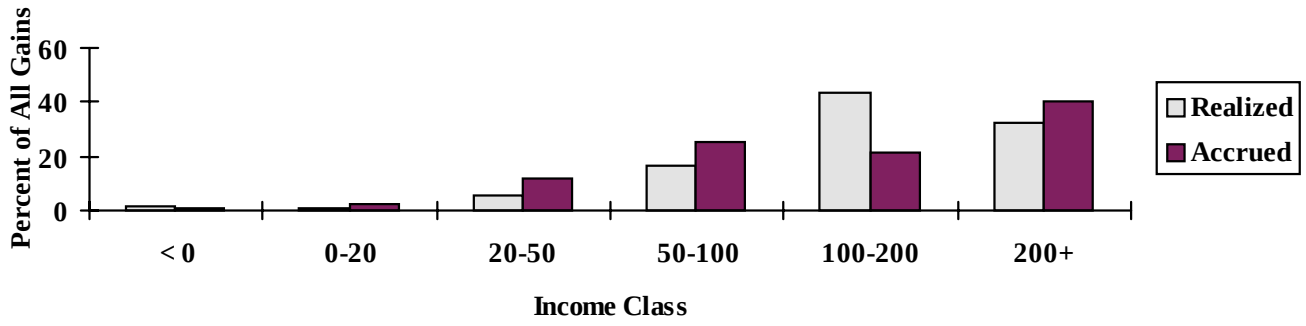


Figure 5. Percentage of Gains in Each Income Class: Business Assets



Sources for Figures 4-5:

Accrued Gains: 1992 Survey of Consumer Finances

Realized Gains: 1988-94 Average, Sales of Capital Assets

2.1. Filers by Year and Permanent Income, SOCA Panel (1985-94)

Permanent Income in thousands of dollars	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
	Unweighted									
less than 20	1,184	1,041	1,008	1,000	984	966	932	893	885	837
20 - 50	2,051	2,014	2,007	2,004	1,995	1,991	1,982	1,962	1,943	1,904
50 - 100	1,367	1,348	1,342	1,338	1,337	1,330	1,319	1,313	1,311	1,293
100 - 200	1,497	1,489	1,482	1,472	1,469	1,461	1,452	1,439	1,428	1,393
200 - 500	2,821	2,799	2,791	2,792	2,784	2,771	2,753	2,747	2,723	2,663
500 - 1,000	1,764	1,751	1,740	1,740	1,732	1,720	1,713	1,711	1,692	1,663
greater than 1,000	2,381	2,371	2,359	2,358	2,351	2,335	2,313	2,302	2,280	2,240
All	13,065	12,813	12,729	12,704	12,652	12,574	12,464	12,367	12,262	11,993
	Population Weighted (in thousands)									
less than 20	50,897	46,954	44,995	44,168	42,992	42,837	41,178	39,214	38,436	36,115
20 - 50	40,639	40,285	40,067	40,091	39,864	39,802	39,649	39,279	39,109	38,213
50 - 100	7,918	7,893	7,880	7,876	7,875	7,836	7,804	7,794	7,756	7,698
100 - 200	1,481	1,475	1,458	1,456	1,456	1,464	1,458	1,446	1,439	1,424
200 - 500	553	551	550	550	492	549	490	543	541	536
500 - 1,000	110	108	104	104	104	103	104	103	103	98
greater than 1,000	38	38	38	38	38	38	38	38	37	37
All	101,637	97,304	95,093	94,283	92,821	92,630	90,719	88,418	87,422	84,121

Table 2.2. Distribution of Net Gains by Permanent Income
(in percents)

Permanent Income (in thousands of dollars)	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
less than 20	2.4	1.5	3.0	2.5	3.0	3.5	1.5	4.3	3.5	4.0
20 - 50	11.9	12.0	8.8	11.2	20.3	21.5	10.6	9.5	23.4	17.0
50 - 100	16.9	16.0	23.8	14.4	15.2	14.3	16.6	21.9	17.0	16.7
100 - 200	16.7	11.2	11.0	13.3	9.2	11.9	12.8	9.5	10.4	11.9
200 - 500	17.8	13.9	22.0	23.3	19.4	11.1	27.0	10.7	10.8	12.0
500 - 1,000	11.9	18.8	8.3	12.1	10.7	10.4	7.5	9.2	9.3	12.7
greater than 1,000	22.5	26.7	23.1	23.3	22.3	27.3	23.9	34.9	25.7	25.8

Table 2.3 Capital Gains and Losses by Term and Year
(in millions of dollars)

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
Long-term Gains	166,354	308,121	137,576	160,675	163,998	105,312	100,961	115,611	151,630	129,074
Long-term Losses	13,589	16,586	34,726	24,438	26,054	26,017	30,430	21,972	25,319	35,441
Short-term Gains	7,608	11,631	18,704	16,252	21,916	14,606	20,501	20,990	25,586	22,073
Short-term Losses	6,058	14,138	20,023	9,127	22,187	20,559	13,916	21,688	15,415	28,748
Net Gains	154,314	289,028	101,531	143,361	137,672	73,343	77,115	92,940	136,482	86,959

Table 3.1. Distribution of Taxpayers with Capital Gain or Loss, by Marginal Tax Rate Region
(in percents)

	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
Population Weighted										
Region A ($\tau_S=\tau$, $\tau_L=\tau^*$)	77.4	78.7	70.3	65.5	65.9	56.2	63.9	65.5	69.8	64.5
Region B ($\tau_S=\tau_L=\tau$)	12.3	11.0	15.8	19.1	15.0	23.5	18.8	19.3	15.4	18.8
Region C ($\tau_S=\tau_L=0$)	5.4	4.3	9.2	11.4	14.6	15.8	13.8	11.9	9.5	11.6
Region D ($\tau_S=\tau_L=\tau^*$)	4.9	6.1	4.8	4.1	4.6	4.5	3.5	3.3	5.4	5.1
Permanent Dividend Weighted										
Region A ($\tau_S=\tau$, $\tau_L=\tau^*$)	68.8	75.7	62.8	62.5	67.3	47.3	58.5	53.8	65.5	43.3
Region B ($\tau_S=\tau_L=\tau$)	6.5	5.6	7.5	12.9	11.0	13.2	10.7	11.8	6.8	11.3
Region C ($\tau_S=\tau_L=0$)	5.3	4.1	9.6	13.3	13.1	21.8	17.4	22.5	17.2	22.1
Region D ($\tau_S=\tau_L=\tau^*$)	19.5	14.7	20.1	11.3	8.6	17.6	13.4	11.9	10.5	23.3
Gross Gains Weighted										
Region A ($\tau_S=\tau$, $\tau_L=\tau^*$)	75.4	78.5	61.6	72.0	75.0	67.4	69.6	64.6	73.1	55.3
Region B ($\tau_S=\tau_L=\tau$)	0.7	1.4	3.3	2.1	3.3	3.6	3.0	5.7	3.8	3.7
Region C ($\tau_S=\tau_L=0$)	1.3	0.9	5.2	3.2	4.5	7.6	6.6	11.0	5.5	8.6
Region D ($\tau_S=\tau_L=\tau^*$)	22.6	19.3	29.8	22.7	17.3	21.5	20.9	18.7	17.6	32.5

Table 3.2. Percentage of Taxpayers with Gains or Losses in Region C, by Permanent Income

Permanent Income (in thousands of dollars)	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
less than 20	1.8	1.5	2.3	5.5	10.1	8.1	3.3	3.4	3.9	1.9
20 - 50	5.5	4.2	8.3	9.1	15.8	14.4	11.6	11.8	7.5	9.5
50 - 100	5.1	5.0	10.9	17.5	14.2	18.2	20.1	13.5	12.2	15.5
100 - 200	9.0	8.3	20.0	14.2	15.1	27.4	25.1	22.7	20.6	32.3
200 - 500	12.6	6.3	19.5	16.5	16.2	26.3	20.8	18.5	22.7	23.7
500 - 1,000	10.4	5.7	21.0	16.1	17.3	30.3	24.7	34.0	23.5	27.3
greater than 1,000	13.7	4.9	17.5	19.0	19.9	32.1	29.7	27.7	21.4	29.9
All	5.4	4.3	9.2	11.4	14.6	15.8	13.8	11.9	9.5	11.6

Table 3.3. Percentage of Taxpayers with Gains or Losses in Region C, by Imputed Wealth Level

Imputed Wealth (in millions of dollars)	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
less than 0.5	4.5	4.1	7.7	9.5	13.5	13.3	11.5	8.4	6.7	8.5
0.5 - 1	6.6	2.0	11.5	18.5	14.8	18.9	13.6	22.1	18.9	21.5
1 - 5	7.9	8.8	15.8	16.4	21.5	27.3	30.2	26.9	19.6	28.5
5 - 10	13.4	4.4	11.8	14.1	14.0	30.3	20.2	32.4	25.3	28.0
10 - 50	9.0	9.3	25.9	18.9	20.1	30.9	25.2	15.1	13.7	15.8
greater than 50	8.0	6.4	11.8	15.2	22.0	27.3	31.3	33.8	31.4	27.9
All	5.4	4.3	9.2	11.4	14.6	15.8	13.8	11.9	9.5	11.6

Table 3.4. Percentage of Gross Gains in Region C, by Asset Type

Asset Type	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
Corporate Stock	1.9	0.8	7.7	4.0	7.1	11.3	8.0	14.5	6.9	8.4
Mutual Funds	1.4	0.5	7.1	5.9	6.4	21.1	8.0	15.8	9.5	15.9
Bonds	4.0	0.5	15.3	18.2	6.7	19.2	14.9	14.4	12.5	21.5
Real Estate	1.3	0.7	3.0	0.7	4.0	2.8	3.8	3.4	3.4	5.9
Business Property	0.7	1.2	1.8	0.7	0.9	6.7	3.0	13.5	1.8	2.5
Short Sales	5.3	3.1	28.0	29.5	31.5	21.0	15.5	36.9	12.7	2.0
Options, Futures, Etc	8.5	16.8	18.1	14.7	41.8	12.7	28.5	35.3	17.1	37.9

Table 3.5. Percentage Taxpayers with Gains or Losses with "Sophisticated" Transactions, by Permanent Income

Permanent Income (in thousands of dollars)	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
less than 20	3.1	1.7	0.0	0.6	1.7	2.2	2.7	9.4	1.8	0.8
20 - 50	3.5	8.2	8.9	9.7	11.5	5.3	10.8	5.8	3.9	4.8
50 - 100	5.3	9.8	12.1	10.5	13.6	7.3	13.4	13.6	12.1	11.2
100 - 200	13.6	17.4	18.1	19.8	21.5	18.7	23.9	16.8	17.8	22.4
200 - 500	14.8	16.2	28.0	23.3	20.1	16.9	21.3	19.5	18.0	29.1
500 - 1,000	14.1	23.7	32.0	24.3	27.1	25.6	31.2	24.1	37.6	33.0
greater than 1,000	21.0	27.7	31.8	31.8	29.9	33.2	36.2	37.5	30.6	35.3
All	5.3	8.6	9.5	10.0	11.7	6.9	11.6	9.7	7.1	7.7

Table 3.6. Probit Estimates of the Probability of Being in Region C (1987-94)

	Population Weighted			*Unweighted*		
	Coefficient	Standard Error	$\partial F/\partial X$	Coefficient	Standard Error	$\partial F/\partial X$
Constant	-3.4932	0.6554		-1.3584	0.1790	
Log(Permanent Income less Capital Gains)	0.1937	0.0623	0.0350	-0.0182	0.0151	-0.0058
Log(Imputed Wealth)	0.0179	0.0448	0.0032	0.0320	0.0108	0.0102
Stock Share of Wealth	-0.1531	0.2562	-0.0276	-0.2761	0.0524	-0.0884
Business Property Share of Wealth	0.3024	0.2910	0.0546	0.0352	0.0441	0.0113
Rental Property Share of Wealth	0.3582	0.2978	0.0647	0.0027	0.0488	0.0009
Farm Property Share of Wealth	-0.2926	0.5806	-0.0528	-0.9813	0.2275	-0.3141
Earns Tax-exempt Interest	-0.1038	0.1035	-0.0184	-0.1528	0.0249	-0.0499
Log(No. of Capital Transactions)	0.1090	0.0401	0.0197	-0.0057	0.0075	-0.0018
Sophisticated	0.4363	0.1251	0.0881	0.2908	0.0231	0.0906
Mutual Fund Share of Cap. Gains	-0.9908	0.2357	-0.1789	-0.3729	0.0590	-0.1194
Single	0.2226	0.1129	0.0431	0.1291	0.0268	0.0425
Married Filing Separate	0.0365	0.1751	0.0068	0.0535	0.0497	0.0174
Head of Household	-0.0436	0.3179	-0.0077	0.1511	0.0602	0.0506
No. of Dependents	0.0101	0.0401	0.0018	0.0103	0.0090	0.0033
Age of Primary Taxpayer	-0.0025	0.0182	-0.0005	0.0134	0.0053	0.0043
(Age of Primary Taxpayer)^2	0.0000	0.0002	0.0000	-0.0001	0.0000	0.0000
Real GDP Growth Rate	-2.6139	1.5694	-0.4721	-3.0170	0.3132	-0.9657
S&P 500 Growth Rate	-0.1213	0.2661	-0.0219	-0.2636	0.0438	-0.0844
Calendar Time Trend	0.0024	0.0141	0.0004	0.0335	0.0028	0.0107

Note: $\partial F/\partial x$ is the partial derivative of the probability of being in region C with respect to the covariate. It is evaluated at mean values, $\bar{x}$, for non-binary data, and evaluated for a discrete change in dummy variables from 0 to 1.
Italics indicate dummy variable.

Table 3.7. Exponential Hazard Estimates of the Duration in Region C (1987-94)

	Weighted		*Unweighted*	
	Coefficient	Standard Error	Coefficient	Standard Error
Log(Permanent Income Less Capital Gains)	-0.0838	0.0779	0.0570	0.0175
Log(Imputed Wealth)	-0.0995	0.0462	-0.0608	0.0135
Stock Share of Wealth	-0.0330	0.2621	0.3967	0.0604
Business Property Share of Wealth	0.0950	0.2835	0.2599	0.0545
Rental Property Share of Wealth	0.1583	0.2662	0.2777	0.0597
Farm Property Share of Wealth	0.1876	0.4093	0.7585	0.2554
Earns Tax-exempt Interest	0.3348	0.1116	0.1625	0.0290
Size of Initial Loss Carryover*	-0.2533	0.0326	-0.1415	0.0068
Log(No. of Capital Transactions)	0.0130	0.0748	-0.0368	0.0105
Sophisticated	-0.4566	0.1355	-0.0908	0.0273
Mutual Fund Share of Capital Gains	0.2721	0.2749	0.1886	0.0808
Age of Primary Taxpayer	0.0140	0.0212	-0.0041	0.0066
(Age of Primary Taxpayer)^2	-0.0001	0.0002	0.0000	0.0001
Real GDP Growth Rate	-0.1399	2.5710	1.2777	0.7129
S&P 500 Growth Rate	0.2861	0.5203	0.2628	0.1358
Calendar Time	0.0356	0.0271	0.0320	0.0060
constant	1.1200	0.7507	-0.7990	0.2210
duration=2	-0.3699	0.1153	-0.2846	0.0292
duration=3	-0.6009	0.1873	-0.4924	0.0411
duration=4	-0.5217	0.2515	-0.8053	0.0624
duration=5	-1.5208	0.4134	-0.8931	0.0800
duration=6	-0.2631	0.4653	-0.8172	0.0950
duration=7	-1.7515	0.5835	-1.2528	0.1573

Size of initial loss carryover variable is defined in Appendix 1.
Italics indicate dummy variable.

Table 3.8 Typical Hazard Functions (Evaluated at Mean Values of Covariates) and Empirical Hazard Function

Duration in Region C (years)	*Weighted Hazard Model*		*Unweighted Hazard Model*		*Kaplan-Meier Empirical Hazard*	
	Hazard	Survival	Hazard	Survival	Hazard	Survival
1	0.50	1.00	0.43	1.00	0.45	1.00
2	0.34	0.50	0.32	0.57	0.33	0.55
3	0.27	0.33	0.26	0.39	0.26	0.37
4	0.29	0.24	0.19	0.29	0.19	0.27
5	0.11	0.17	0.17	0.23	0.18	0.22
6	0.38	0.15	0.19	0.19	0.20	0.18
7	0.09	0.09	0.12	0.16	0.13	0.14

Table 3.9. Effective Marginal Tax Rates on Long-term Capital Gains in Selected Years, by Ordinary Income Tax Rate
(Weighted by population)

Year	Ordinary Income Tax Rate (percent)	*Percentiles of Marginal Tax Rate Distribution* 1	10	50	90	99	*Mean* By Tax Rate	Overall (Long-term Gains Weighted)
1988	15	7.6	15.0	15.0	15.0	18.8	14.8	
	28 (low)	14.5	19.6	28.0	28.0	28.0	26.6	26.9
	33	12.6	18.8	33.0	33.0	33.0	30.6	
	28 (high)	13.4	21.3	28.0	28.0	28.0	26.9	
1991	15	7.0	15.0	15.0	15.0	19.7	14.9	
	28	8.9	17.8	28.0	28.0	28.0	26.1	25.0
	31	13.5	18.7	28.0	31.0	31.0	26.8	
1994	15	8.8	15.0	15.0	15.0	20.3	15.0	
	28	15.1	21.5	28.0	28.0	28.0	26.9	
	31	12.8	23.7	28.0	28.0	31.0	27.2	25.3
	36	16.6	19.1	28.0	36.0	36.0	27.5	
	39.6	16.7	19.9	28.0	39.6	39.6	27.5	

Table 3.10. Statutory Versus Ordinary Tax Rates on Long-Term Capital Gains in 1994
(Percent)

Ordinary Income Tax Rate	Tax Rate on Long-Term Gains	*Effective Tax Rate* In Region C	All Taxpayers
15	15	14.8	15.0
28	28	17.9	26.9
31	28	17.4	27.2
36	28	19.4	27.5
39.6	28	20.5	27.5

Table A.1. Results of the Estimation, Probability of Offsetting 100% of Gross Gains (1987-94)

	Population Weighted			*Unweighted*		
	Coefficient	Standard Error	$\partial F/\partial X$	Coefficient	Standard Error	$\partial F/\partial X$
Constant	-0.6605	0.5236		-0.0360	0.1443	
Log(Permanent Income less Capital Gains)	0.0800	0.0494	0.0257	0.0106	0.0123	0.0036
Log(Imputed Wealth)	-0.0289	0.0299	-0.0093	-0.0480	0.0087	-0.0165
Stock Share of Wealth	0.0885	0.1597	0.0284	-0.0924	0.0436	-0.0318
Business Property Share of Wealth	0.2584	0.1926	0.0830	0.3112	0.0368	0.1072
Rental Property Share of Wealth	0.3061	0.2028	0.0983	0.2199	0.0418	0.0758
Farm Property Share of Wealth	-0.5808	0.4369	-0.1865	-0.6660	0.2076	-0.2295
Earns Tax-exempt Interest	0.0245	0.0807	0.0079	-0.0907	0.0205	-0.0315
Log(No. of Capital Transactions)	0.1375	0.0294	0.0442	-0.0581	0.0063	-0.0200
Sophisticated	0.2219	0.0858	0.0732	0.2838	0.0189	0.0957
Mutual Fund Share of Cap. Gains	-1.9102	0.3091	-0.6134	-0.5402	0.0580	-0.1861
Single	0.2398	0.0916	0.0798	0.1251	0.0227	0.0441
Married Filing Separate	-0.0603	0.2170	-0.0190	-0.0035	0.0458	-0.0012
Head of Household	-0.0367	0.1739	-0.0116	0.1155	0.0504	0.0409
No. of Dependents	0.0041	0.0330	0.0013	0.0064	0.0073	0.0022
Age of Primary Taxpayer	-0.0193	0.0130	-0.0062	-0.0013	0.0042	-0.0004
(Age of Primary Taxpayer)^2	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000
Real GDP Growth Rate	-3.9212	1.5585	-1.2592	-3.9947	0.3115	-1.3764
S&P 500 Growth Rate	-1.0099	0.2224	-0.3243	-0.5156	0.0447	-0.1776
Calendar Time Trend	0.0234	0.0121	0.0075	0.0452	0.0022	0.0156

Note: $\partial F/\partial x$ is the partial derivative of the probability of being in region C with respect to the covariate. It is evaluated at mean values, $\bar{x}$, for non-binary data, and evaluated for a discrete change in dummy variables from 0 to 1.
Italics indicate dummy variable.

Table A.2. Results of the Estimation, Duration of Remaining in State of Offsetting 100% of Gross Gains (1987-94)

	Unweighted	
	Coefficient	Standard Error
Log(Permanent Income Less Capital Gains)	0.0600	0.0135
Log(Imputed Wealth)	-0.0256	0.0094
Stock Share of Wealth	0.1701	0.0494
Business Property Share of Wealth	0.0405	0.0422
Rental Property Share of Wealth	0.0405	0.0483
Farm Property Share of Wealth	0.2255	0.2029
Earns Tax-exempt Interest	0.0886	0.0230
Size of Initial Loss Carryover*	-0.0672	0.0021
Log(No. of Capital Transactions)	-0.0050	0.0087
Sophisticated	-0.0585	0.0216
Mutual Fund Share of Capital Gains	0.2406	0.0727
Age of Primary Taxpayer	-0.0010	0.0051
(Age of Primary Taxpayer)^2	0.0000	0.0000
Real GDP Growth Rate	0.8132	0.5663
S&P 500 Growth Rate	0.3086	0.1060
Calendar Time	0.0174	0.0049
constant	-1.4173	0.1704
duration=2	-0.2195	0.0247
duration=3	-0.4546	0.0371
duration=4	-0.6894	0.0527
duration=5	-0.7862	0.0718
duration=6	-0.6783	0.0841
duration=7	-1.1433	0.1487

Size of initial loss carryover variable is defined in Appendix 1.
Italics indicate dummy variable.

Table A.3. Typical Hazard Functions (Evaluated at Mean Values of Covariates) for Exiting State of Offsetting 100% of Gross Gains

Duration State of 100% Gain Offset (years)	*Unweighted Hazard Model*	
	hazard	survival
1	0.45	1.00
2	0.36	0.55
3	0.28	0.35
4	0.22	0.25
5	0.20	0.29
6	0.23	0.16
7	0.14	0.12