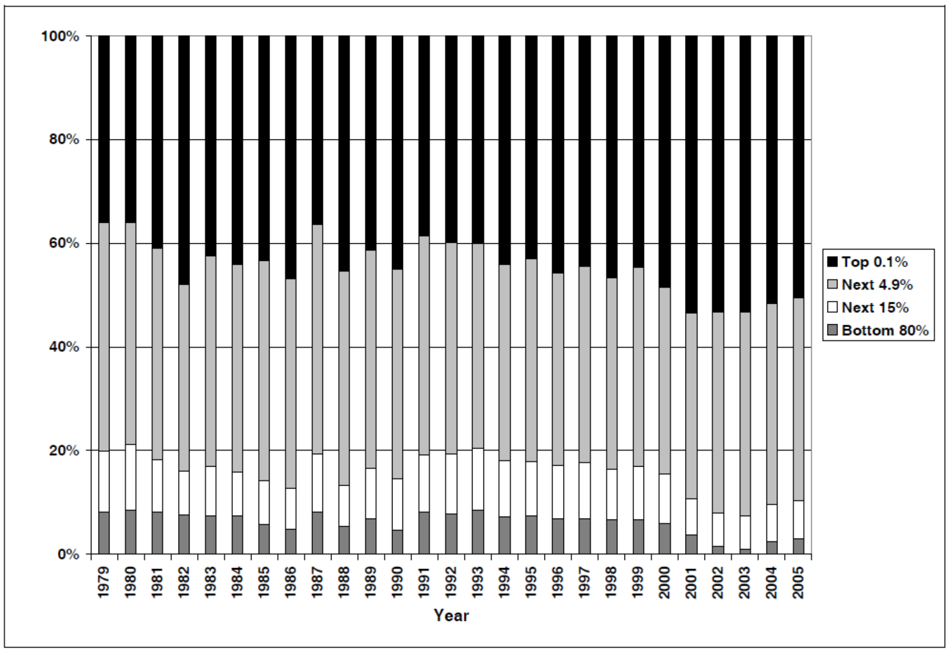
Economics 230a, Fall 2018
Lecture Note 10: Capital Gains Taxation

Capital gains taxes are of particular interest for a number of reasons, even though they do not account for a large share of revenue for a typical government, including the United States.

Figure 5. Distribution of Realized Capital Gains by Income, 1979-2005



Source: CBO, *Historical Effective Tax Rates, 1979-2005: Supplement*, December 2008.

According to Hungerford (Congressional Research Service, 2010), "Since 1954, revenue from the capital gains tax as a share of total income tax revenue has averaged 5.2%." One reason for the interest in capital gains is their concentration at the top of the income distribution, as shown in this figure (from the same paper).

Another important aspect of capital gains is that they are taxed upon *realization* rather than on accrual, which makes the tax complex and subject to a variety of potential taxpayer responses.

What does realization-based taxation do? Consider a two-period model in which an investor has an asset purchased in an earlier period for \$1, which has already appreciated in value by an amount g. The investor can either hold the asset for another period, earning an additional return r, or sell and earn the market rate of return i. Suppose all income is taxed at rate t, but only when assets are sold. Also suppose that the investor's objective is to maximize terminal wealth.

If the investor sells the asset and reinvests, terminal wealth is:

$$W_R = (1+g(1\text{-}t))(1+i(1\text{-}t)) = (1+g)(1+i) - t[g(1+i(1\text{-}t)) + (1+g)i]$$

If the investor holds the asset until the end of the second period, terminal wealth is:

$$W_H = (1+g)(1+r) - t[(1+g)(1+r) - 1] = (1+g)(1+r) - t[g + (1+g)r]$$

Comparing the terms in brackets in the second version of each expression, we can see that the "hold" strategy enjoys a tax advantage over the "realize" strategy – first period gains, g, are taxed one period earlier under the latter, and hence the tax liability has a higher accumulated value at the end of the second period because it is multiplied by $1+i(1-t)$. It follows that if $i = r$, the investor will choose to hold rather than to realize, and indeed that there is a range of values of $r < i$ for which it will still be optimal to hold rather than to sell. This phenomenon is known at the lock-in effect – in order to defer tax on previously accumulated gains, individuals have an incentive not to sell assets even when, for non-tax reasons, they would prefer to sell. In this example, the lock-in effect is associated with the investor's willingness to accept a lower before-tax rate of return, but in a realistic setting the major distortion comes from an inefficient allocation of assets across investors. That is, when an individual realizes a capital gain by disposing of an asset, that asset does not typically disappear, but instead ends up in someone else's portfolio. Thus, it is unlikely simply to have a below-market rate of return, because asset prices adjust. Rather, in a setting with risky assets, other investors may be willing to pay more for the asset than the individual currently holding it. For example, suppose there are two investors with appreciated stock, one holding Apple and the other holding Exxon. As returns on these two assets are not perfectly correlated, a combined portfolio would offer a better risk-return trade-off than either specialized position. Absent taxation, each investor could be made better off by trading with the other, but if each faces the capital gains tax, the trades may not occur.

The lock-in effect is exacerbated by two other provisions found in the US tax system and typical of others as well. First, gains on assets held for at least one year are taxed at a lower rate (in United States at present, a maximum of 20% vs. a maximum tax rate on ordinary income of 39.6%). Second, gains held until death are not taxed at all. On the other hand, the lock-in effect is reversed when an asset has gone down in value ($g < 0$ in the above example), since deferral of tax in this case means deferring a tax *refund*. Thus, individuals have an incentive to hold gains and realize losses, meaning that those with large numbers of distinct positions in different assets could, on a regular basis, achieve liquidity by "harvesting" losses without having to realize gains. This possibility, in turn, is largely responsible for another tax provision, which limits the annual value of deductible losses (in excess of realized gains) to $3,000. Unfortunately, as discussed in Lecture 9, a limit on the deductibility of losses also discourages risk-taking.

Empirical Evidence on Responses to Capital Gains Taxation

There has been a substantial literature relating capital gains realizations to capital gains tax rates. One of the key issues is the need to distinguish between short-run and long-run responses. We would expect that a change in tax rates could have a large impact on the timing of realizations, because individuals can adjust the timing of their asset sales. For example, after the October, 1986 passage of the Tax Reform Act of 1986, which increased the capital gains tax rate on high-income individuals from 20% to 28% effective January 1, 1987, there was such a surge in realizations in the remainder of 1986 that realizations for that year were approximately twice as high as those in 1985 or 1987. But that doesn't mean that we would expect realizations to be permanently twice as high under a 20% tax rate as under a 28% tax rate.

One standard approach originally developed using panel data by Burman and Randolph (*AER* 1994; hereafter BR) involves type-II Tobit estimation (for the decision to realize gains and gains realized), where the second, intensive-margin decision takes the form:

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

where g is capital gains, X is a vector of individual attributes, τ is the individual's capital gains tax rate, and τ^p is a measure of the individual's "permanent" tax rate. The intuition for including the lagged tax rate τ_{it-1} is that a higher value will mean lower past realizations, hence a large stock of gains available to be realized at time *t*. The intuition for including some permanent tax rate measure, τ_{it}^p, is that if individuals expect a higher tax rate to prevail in the future, they will (as in 1986) wish to realize more gains in period *t*. But how should one represent this permanent tax rate? BR identified τ^p using each individual's potential maximum federal plus state tax rate, but this arguably does not correctly distinguish timing and permanent responses. On the one hand, the maximum state and federal tax rates change over time, so some of the responses to the BR measure of τ^p may be timing responses. On the other hand, individual tax rates may persistently deviate from the BR measure of τ^p, meaning that some of the response classified as temporary would actually be permanent. As an alternative, Auerbach and Siegel (*AER* 2000, hereafter AS), replace τ_{it}^p in the above specification with τ_{it+1}, the tax rate the individual will face the following year, which is generally known at time *t*. This is the approach adopted by Dowd et al., using more recent administrative tax return panel data.

There is one further econometric issue that must be confronted in estimating (1): the capital gains tax rate may depend on the level of gains realized, since tax rates rise with income. To deal with this, a common problem in empirical analysis of behavioral responses to taxation, all of these papers use as an instrument for τ a so-called "first-dollar" tax rate – the capital gains tax rate the individual would face on the first-dollar of capital gains realized. In their preferred specification, Dowd et al. find a permanent elasticity (corresponding to the effect γ_2 in equation (1)) of -0.716 and a transitory elasticity (corresponding to the combined effect $\gamma_1 + \gamma_2 + \gamma_3$) of -1.194, meaning that a temporary cut in the capital gains tax rate would increase tax revenue in the current year, but that a permanent cut in the capital gains tax rate would not.

A further empirical finding of interest is by Ivković, Poterba and Weisbenner, who consider differences in capital gains realizations by individuals who hold both tax-favored and taxable accounts. According to standard theory, there should be no lock-in effect for assets in tax-favored accounts, so that gains should be realized sooner, and losses later, than in taxable accounts. Indeed, the authors find that, *relative to assets in their tax-favored accounts*, investors are less likely to realize gains and more likely to realize losses in their taxable accounts (Figure 3B**)**. However, they also find that investors are more likely to realize taxable gains than taxable losses (Figure 1). There are a variety of possible explanations for this latter finding, including a belief that stock prices are mean-reverting (so that those with gains are expected to fall and those with losses are expected to rise), a need to rebalance portfolios (and hence to shed those stocks that have gained and as a result occupy a larger portfolio share), and the presence of a "disposition effect," by which individuals perceive losses more fully if they are realized.

Reforming the Capital Gains Tax

Some changes in the capital gains tax (such as taxing capital gains at death) could serve to reduce the lock-in effect, but other problems remain as long as the basic approach to taxing capital gains upon realization is followed. Some arguments for keeping the capital gains tax rate lower than other capital income taxes, including the potentially higher behavioral response elasticity and the importance of capital gains in fostering venture capital investments, relate to the realization-based nature of the tax (in the latter case because risky venture-capital investments face serious limitations on their ability to deduct losses, which as discussed earlier is a necessary feature of a realization-based system).

What other alternatives exist? One approach would be to tax capital gains at death, or at least to force heirs who receive assets to "carry over" the *basis* (i.e., original purchase price) of the assets received and therefore be liable for tax on the full gain when they eventually sell the assets. This would clearly reduce the lock-in effect associated with holding assets until death.

Another frequent proposal has been to index capital gains for inflation, allowing individuals to adjust an asset's purchase price upward for changes in the price level since purchase (i.e., pay tax on the sale price V_t less the original purchase price, V_0 multiplied by the ratio of current and initial price levels, P_t/P_0). Letting π be the annual inflation rate, this would make the return to holding an asset, W_H, equal to:

$$W_H = (1+g)(1+r) - t[(1+g)(1+r) - (1+\pi)^2] = (1+g)(1+r) - t[(g-\pi)(1+\pi) + (1+\pi+(g-\pi))(r-\pi)]$$

I.e., real tax liability is independent of inflation for given real rates of return $(g-\pi)$ and $(r-\pi)$.

Perhaps the simplest idea for reform would be to tax capital gains as they accrue, rather than upon realization (perhaps combined with a reduced rate to offset the increased present value of taxes). But there are two problems with this approach: (1) taxpayers may lack liquidity to pay taxes until assets are actually sold; and (2) the government may not know the value of some assets until they are actually sold. One proposal for dealing with the liquidity problem, by Vickrey (*JPE*, 1939), amounts to keeping an account of accruing gains and the associated tax liability and charging interest on this accruing unpaid balance until asset sale. That is, the tax liability as of date s would evolve according to:

$$(2) \qquad T_{s+1} = [1+i(1-t)]T_s + tr_sA_s$$

where r_s is the rate of return at date s, A_s is the value of the asset at date s, i is the safe rate of interest and t is the tax rate. A problem with Vickrey's approach is that r_s and A_s may be unobservable, but Auerbach (*AER* 1991) argued that one can generalize Vickrey's approach to:

$$(3) \qquad T_{s+1} = [1+i(1-t)]T_s + tiA_s + t^*(r_s-i)\,A_s$$

where t^* can take on any value, since (as discussed in Lecture 9), a tax rate on a risky asset's return in excess of the safe rate has no effect on the investor's opportunities. Auerbach then showed that a tax liability of the form:

$$(4) \qquad T_{s+1} = \left[1 - \left(\frac{1+i(1-t)}{1+i}\right)^s\right]A_s$$

satisfies (3). Note that only observable information is needed to assess the tax in (4): the sale price, A_s, the holding period, s, the safe rate of interest, i, and the tax rate, t. Auerbach and Bradford generalize this result and show how it can be implemented using a tax system based exclusively on observed cash flows, without keeping track of individual assets and holding periods.