MEASURING PERMANENT RESPONSES TO CAPITAL GAINS TAX CHANGES IN PANEL DATA*

by

Leonard E. Burman and
William C. Randolph
Congressional Budget Office
U.S. Congress

OTA Paper 68 **August 1994**

OTA Papers and Briefs are circulated so that the preliminary findings of tax research conducted by staff members and others associated with the Office of Tax Analysis may reach a wider audience. The views expressed are those of the authors, and do not reflect Treasury policy. Comments are invited, but OTA Papers and Briefs should not be quoted without permission from the authors. Additional copies of this publication may be purchased from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161

Office of Tax Analysis
U.S. Treasury Department, Room 1064
Washington, DC 20220

---

* **Forthcoming, American Economic Review. This paper was written while Randolph was a staff member of the Office of Tax Analysis, U.S. Treasury Department.**

Measuring Permanent Responses to Capital Gains Tax Changes in Panel Data

By LEONARD E. BURMAN AND WILLIAM C. RANDOLPH*

This paper uses panel data and information about differences in state tax rates to separate the effects of transitory and permanent tax rate changes on capital gains realizations behavior. The effect of permanent change is found to be substantially smaller than the effect of transitory change. The estimated difference is even larger than past differences between estimates from careful micro data studies, which have primarily measured the transitory effect, and time series studies, which have primarily measured (at best) the permanent effect. Our results thus resolve a longstanding conflict between micro data and time series studies of how marginal tax rates affect capital gains realizations behavior.

"Observe due measure, for right timing is in all things
the most important factor" --Hesiod (700 BC)

For more than forty years, policy analysts and economists have debated about how capital gains realizations respond to changes in capital gains tax rates. (See, e.g., Lawrence H. Seltzer, 1951) The issue has received attention in part because, if realizations of capital gains are responsive enough, the tax rate on capital gains could be cut at no cost to the Treasury. But it is also an important issue for tax reform because some argue that the welfare cost of the capital gains tax could be large relative to the tax revenues collected if realizations are very sensitive to tax rates. (Patric H. Hendershott et al., 1991)

The debate has been fueled by an array of disparate statistical estimates of the elasticity of capital gains realizations with respect to the marginal tax rate on capital gains. The evidence from time series appears entirely inconsistent with the evidence from individual tax return data. Time series studies have generally found that capital gains are relatively unresponsive to tax rates. Estimates based on micro data, however, generally suggest that realizations are highly elastic.

These empirical estimates are viewed with great skepticism by many who have studied the issue. Some authors of time series studies (Alan J. Auerbach, 1989; Jonathan Jones, 1989; Robert Gillingham and John S. Greenlees, 1992) have discounted their findings because they are subject to intractable aggregation biases and are extremely sensitive to sample period and seemingly minor changes in specification. Estimates from micro data studies have been even less robust.

---

*Burman: Congressional Budget Office, U.S. Congress, Washington, DC 20515; Randolph: Office of Tax Analysis, U.S. Treasury, Washington, DC 20220. We are grateful to B.K. Atrostic, Jerry Auten, Charley Ballard, Joe Cordes, Glenn Hubbard, Jody Magliolo, Randy Mariger, Jim Nunns, Larry Ozanne, R.P. Trost, Jenny Wahl, and seminar participants at George Washington, Georgia State, Maryland, Michigan, and Northeastern, and three anonymous referees for helpful comments and suggestions. Jim Cilke and Gordon Wilson developed the tax calculators. Views expressed do not necessarily represent the views or policies of the Congressional Budget Office or the Department of the Treasury.

More fundamentally, some authors of micro data studies have recognized that their estimates may systematically overstate the long-term response to tax changes. Indeed, the seminal empirical study of the effect of tax rates on realizations of capital gains raised this caveat:

> An individual whose tax rate varies substantially from year to year will tend to sell more when his rate is low. To the extent that low rates in 1973 are only temporarily low, our estimates will overstate the sensitivity of selling to the tax rate. *We have no way of knowing how important this is*. (Martin Feldstein et al., 1980, p. 785. Emphasis added.)

Such timing behavior is very important. The Tax Reform Act of 1986 (TRA) created a natural experiment to test the hypothesis that timing matters. TRA was passed by Congress at the end of September, 1986. It turned a "permanent" 20 percent maximum tax rate, in effect since 1981, into a temporary rate, to be replaced by a higher maximum capital gains rate of 28 percent in 1987. In response, long-term capital gains on corporate stock in December 1986 were nearly seven times their level in December 1985. (Burman et al., 1993)

Timing behavior probably explains why micro data studies have produced such large elasticity estimates.[1] As the transitory component of individuals' taxable income varies, it provides them with opportunities to time capital gains realizations in years when tax rates are relatively low. In a particular year, those with the lowest tax rates, other things constant, would be those with the largest capital gains realizations. As a result, a regression based on micro data is likely to measure a negative correlation between marginal tax rates and capital gains realizations. But, without more information, it is impossible to determine how much of the measured correlation represents purely transitory timing behavior. Much of the policy debate, however, has centered on the permanent or long-term response to statutory tax changes.[2]

To distinguish permanent from transitory tax effects, we define a "permanent tax rate" as the tax rate on long-term capital gains, purged of individual and aggregate transitory effects. We estimate the relationship between capital gains and permanent tax rates in a panel of tax returns using an instrumental variables estimator that also accounts for the endogeneity of tax rates and self-selection. Our instrument for permanent tax rates is the maximum combined federal and state tax rate on long-term capital gains. This instrument, which only varies among states, removes individual transitory effects because it is uncorrelated with transitory variations in individuals' income. We remove aggregate transitory effects by using time dummies.

Our estimates imply that the elasticity of capital gains realizations with respect to permanent tax changes is much smaller than the transitory response. Our point estimates of permanent tax effects are smaller in absolute value than most estimates from time series. Our estimates of transitory tax effects are larger than estimates from previous studies based on micro data.

---

[1]See Gerald E. Auten and Charles T. Clotfelter (1982), Joseph E. Stiglitz (1983), Jane G. Gravelle (1987), Auerbach (1988), Auten et al. (1989), Joel Slemrod and William Shobe (1990), and Donald W. Kiefer (1990) for more discussion on this point.

[2]Auerbach (1989) and Gravelle (1991) have questioned the large elasticities found in most micro data studies on conceptual grounds. They argue that such large elasticities would imply that even modest changes in tax rates could cause realizations to exceed accruals.

I. The Decision to Realize Capital Gains

There are two inherent problems in measuring taxpayer responses to capital gains tax changes: standard theoretical models do not explain why people realize significant amounts of taxable capital gains, and some variables that would enter almost any theoretical model are not observed in available data. Nonetheless, the typical empirical model may be interpreted as a reduced form, given data limitations, to test the most general implications of theory. Our analysis extends the basic empirical model to permit identification of a key policy parameter, the effect of permanent changes in tax rates.

The capital gains tax is relatively easy to avoid. Tax on an asset's gain or loss is not due until the asset is sold, and may be avoided entirely if an asset is held until death or donated to charity. Stiglitz (1983) showed that, by borrowing, hedging, accelerating losses, and deferring gains, capital gains taxes can be avoided altogether if capital markets are perfect and transactions are costless. He concluded that the existence of substantial taxable capital gains realizations implies that the underlying assumptions of his model must be violated in practice.

George M. Constantinides (1984) showed that realizing gains on stocks as soon as they qualify for preferential long-term tax rates may be optimal for very volatile stocks with low transaction costs. Yves Balcer and Kenneth L. Judd (1987) showed that, if borrowing and liquidity constraints are binding and options markets do not exist, capital gains would be realized following a LIFO strategy to maximize the benefits of deferral. However, none of these models would explain the $100 to $200 billion in taxable gains reported in a typical year.

Kiefer (1990) and Burman and Randolph (1992) developed models in which capital gains realizations occur because capital markets are limited--there are liquidity constraints and no options markets--and individuals believe they can beat the market by trading. The latter study also showed that transaction costs could be important. As well as characterizing a long-run equilibrium in which significant amounts of capital gains could be realized, these analyses also shed light on the transition path from one steady state to another after tax laws change. Conventional wisdom holds that the short-run response to a permanent cut in capital gains tax rates would be larger than the long-run response because of an immediate "unlocking" effect. Taxpayers holding assets to avoid capital gains tax suddenly flood the market with these assets when the tax rate is lowered because the tax cost of selling the assets is reduced. However, this conventional view ignores the fact that the cost of selling assets is also an increasing function of accrued gain as a fraction of asset value. On average, this fraction would be higher immediately after a capital gains tax cut than it would be in the new steady state. The high level of accrued gains will initially increase the cost of asset sales relative to the steady state, and will therefore dampen adjustment in the short run. If the initial level of unrealized accruals is high enough, the short-run increase in realizations of capital gains could actually be smaller than the long-run increase.

The response to temporary reductions in capital gains tax rates is clearer. A temporarily low capital gains tax rate provides taxpayers with an opportunity to gain from timing. A temporary tax cut reduces the tax cost of selling now, but leaves the tax cost of selling in the future unchanged. In contrast, a permanent tax cut reduces the tax cost of selling at any time. Thus, realizations of capital gains will be higher under a transitory rate cut than under a permanent cut, as illustrated by the response to TRA. (Paul J. Bolster et al., 1989; Burman et al., 1993)

In micro data, variations in capital gains tax rates include both permanent and transitory components. The permanent component results from expected differences in earnings capacity,

sources of income and deductions, and because capital gains tax rates vary across states. It may change when tax laws change. The transitory component results from tax planning and temporary changes in income and deductions. The tax law may also cause aggregate transitory changes if the statutory change is anticipated or if a new tax law is phased in over several years. An empirical model should allow for the possibility that people respond differently to changes in the permanent and transitory components.

II. Empirical Model of Permanent and Transitory Tax Effects

An individual decides whether to sell specific portfolio assets and, incidentally, whether to realize capital gains. Capital gains enter the decision because the tax price of selling an asset is the product of the capital gains tax rate and the share of asset value that is a capital gain. Assets with relatively larger accumulated gains are more costly to sell than assets with smaller accumulated gains. Unfortunately, our panel of tax returns from 1979 to 1983 only includes total capital gains and losses with no detail about sales of specific assets. Thus, like all previous empirical studies of capital gains, we estimate a reduced form relationship between total long-term capital gains and factors that may affect the decision to sell assets with capital gains.

For taxpayers who choose to realize capital gains, we model the relationship between capital gains and tax rates as follows:

$$\ln g = X\gamma_0 + \gamma_1\tau_p + \gamma_2(\tau_t - \tau_p) + \gamma_3(\tau_t - \tau_{t-1}) + \varepsilon_2 , \tag{1}$$

where g is realized capital gains by an individual at time t, X is a vector of predetermined and exogenous variables, τ_p is the permanent tax rate, τ_t is the current tax rate in year t, γ_0, γ_1, γ_2, and γ_3, are fixed parameters, and ε_2 is a random error term. This semi-log functional form has been used in most empirical capital gains research. It implies that the elasticity of capital gains realizations with respect to the marginal tax rate is an approximately linear function.[3]

The decision of an individual to realize capital gains depends on the costs and benefits of realizing gains, the size and composition of the portfolio, and preferences. Taxes affect the costs and benefits of selling. The cost of selling depends on the effective marginal tax rate on capital gains and on the size of the average accrued gain. Equation (1) separates the marginal tax rate into permanent and transitory components. The permanent tax rate is the tax rate purged of its individual and aggregate transitory components. It is the expected (normal) tax rate in a typical year given federal and state tax laws and normal levels of income for each individual. The remaining transitory component represents the tax cost of selling when the tax rate is unusually high, or holding when the tax rate is unusually low. The lagged tax rate, is also included as a proxy for the unobservable size of accrued gains. For example, if the previous year's tax rate was unusually high, then accrued gains should be larger than usual because realizations would have been postponed.

Other variables summarizing individual differences are included in X. The cost of selling depends on transaction costs, so the composition of the portfolio is important. Lagged data from individual tax forms provide indirect information about whether the portfolio is likely to include real estate or business property, which is relatively costly to sell.[4] Those data, as well as lagged

[3]In estimation, we also tested the assumption that the elasticity is approximately constant. This alternative does not affect the empirical results substantially.

[4]The data are discussed in Section III.

data on sources of capital income, also allow us to create proxies for wealth, which represents the potential size of accrued unrealized gains, and for the share of wealth held as corporate stock.

Sales and purchases of assets are a part of life-cycle consumption decisions. Thus, permanent and transitory income, age, marital status, and family size may be important determinants. Panel data from tax returns allow us to estimate permanent and transitory income, and age data are matched from social security records. In addition, regional dummies are included to control for regional differences in investment preferences. Time dummies are included to control for the aggregate economic factors that affect investment opportunities. These dummies also control for the average effect of tax law changes, as occurred in 1981.

We account for the decision to realize a capital gain as well as the level of capital gains. Our full empirical specification in Equations (2)-(4) represents (1) as a generalized tobit model, and also accounts for the endogeneity of current marginal tax rates.[5] The tax terms are rearranged algebraically to simplify estimation.[6]

$$I^* = X\alpha_0 + \alpha_1\tau_p + \alpha_2\tau_t + \alpha_3\tau_{t-1} + \varepsilon_1 , \tag{2}$$

$$\ln g = \begin{Bmatrix} X\beta_0 + \beta_1\tau_p + \beta_2\tau_t + \beta_3\tau_{t-1} + \varepsilon_2 & \text{if } I^* > 0 \\ 0 & \text{otherwise} \end{Bmatrix} , \tag{3}$$

and

$$\tau_t \equiv f(Z, g) , \tag{4}$$

where I^* is a latent indicator of the decision to realize capital gains, the α and β terms are unknown parameters, and ε_1 and ε_2 are normally distributed error terms, uncorrelated with X, τ_p, or τ_{t-1}, such that $E(\varepsilon_i\varepsilon_j') = \sigma_{ij}$ for $i,j=1,2$. The combined federal and state marginal tax rate function, f, is a known nonlinear function of capital gains and Z, a vector including income items from various sources, deductions, exemptions, transfers, carried over tax losses and credits, and taxpayer filing status.

A. Estimation Procedure

We extend the instrumental variables procedure developed by Lung-Fei Lee et al. (1980) to allow for the presence of an unobserved variable, τ_p, and an endogenous variable, τ_t, in both the criterion function, (2), and the level equation, (3). The procedure is similar to the two-step regression estimation method developed by James J. Heckman (1976), except that fitted values are used in place of τ_p and τ_t. The fitted value, τ_p, is created by regressing τ_t on X, τ_{t-1}, and τ_s, the maximum combined federal and state tax rate in each individual's state. The fitted value, $\hat{\tau}_t$, is created by regressing τ_t on X, τ_{t-1}, τ_s, and τ_0, a "first-dollar" marginal tax rate on capital

[5]A separate appendix shows that correcting for endogeneity and sample selection is especially important in this type of model. Failure to properly account for these problems may explain much of the volatility in previous research on capital gains. The appendix is available upon request.

[6]The γ parameters in (1) have a simple relationship to the β's in (3): $\gamma_1=\beta_1+\beta_2+\beta_3$; $\gamma_2=\beta_2+\beta_3$; $\gamma_3=-\beta_3$.

gains. The first-dollar marginal tax rate is computed by setting g and the other sources of income and deductions that are jointly determined with g equal to zero.[7]

The parameters of (2) are estimated by probit maximum likelihood with the fitted τ_p and τ_t used in place of the actual values. The level equation, (3), is estimated by least squares using the sample of realizers, including the estimated inverse Mills ratio as a regressor to control for sample selectivity.[8] For estimation of (3), values for τ_p and τ_t are reestimated for the sample of realizers including the inverse Mills ratio as an additional variable in the fitted equations. The standard errors are corrected using a formula derived by Lee et al.[9]

B. *Consistency of the IV Estimator*

Previous micro-data studies, which lacked appropriate instruments for τ_p, could only have produced consistent estimates of tax effects if transitory and permanent responses are the same. Under this condition, our estimation procedure would produce consistent estimates using almost any exogenous instruments for the permanent and transitory tax rates. However, if transitory and permanent responses are different, then appropriate instruments for τ_p and τ_t are essential to estimate permanent and transitory tax effects consistently.

The estimation problem is unusual and interesting because we need to estimate the effects of two unobservable components of the tax rate. If τ_t is defined as

$$\tau_t = \tau_p + \mu_t , \tag{5}$$

where μ_t is transitory deviations in tax rates, then both τ_p and μ_t enter (1) as explanatory variables.[10] This problem is similar in form to an errors-in-variable model. However, in the standard errors-in-variables model, only the systematic component, τ_p, would enter the model as an explanatory variable. To consistently estimate the effect of τ_p, we need an exogenous instrument that is correlated with τ_t, but uncorrelated with μ_t, conditional on X and τ_{t-1}. Although much of the variation in τ_p is related to the other exogenous variables, especially permanent income, wealth, and the portfolio mix, differences in state tax law provide an

---

[7]The first-dollar marginal tax rate is computed by setting long-term capital gains and other income and deduction items that are likely to be endogenous equal to zero and then computing the marginal tax rate on a defined long-term capital gain. This instrument retains a substantial amount of variation independent of the other explanatory variables because marginal tax rates are a known nonlinear function of numerous exogenous factors that do not directly affect capital gains, including consumer and mortgage interest deductions, contributions to pensions and IRAs, property taxes, certain health expenses, business and employment expenses, paid alimony, and many other deductions and adjustments to income.

[8]The inverse Mills ratio is computed based on the fitted values from the probit step.

[9]The standard error estimates may be understated because the formula does not account for the use of instrumental variables in the probit equation. To check the standard errors, we randomly split the sample into 10 parts and estimated the parameters for each subsample. The standard errors of the sample mean of the 10 estimates were very close to those produced by the formula.

[10]We ignore sample selection in this discussion to focus on the key estimation problem. The IV results for the linear model, (1), are extended to the full model with truncation, (2)-(3), in a separate appendix.

exogenous source of variation that is easily measured.[11] Moreover, the variation in state taxes is closely related to an important policy question: how do realizations differ under different tax laws? Because state income taxes tend to be less graduated than the federal tax schedule, most gains are realized by taxpayers in the top state tax brackets. Thus, the top combined federal and state tax rate (τ_s) captures most of the important differences in statutes, and does not vary among individuals within a state. It is thus unlikely to be correlated with the transitory component, μ_t.[12]

To consistently estimate the transitory effect (μ_t), we need a second exogenous instrument that is correlated with τ_t, but uncorrelated with τ_p, conditional on X, τ_{t-1}, and τ_s. Our instrument has been used in various forms in most previous micro data studies of capital gains: the first-dollar tax rate (τ_0). Because marginal tax rates are a highly nonlinear function of many variables that do not directly affect capital gains (see footnote 7), this instrument captures much of the variation in τ_t, but is purged of its endogenous components. Further, τ_0 is unlikely to be correlated with τ_p, conditional on τ_{t-1}, τ_s, and the variables included in X.

The standard errors-in-variables model assumes that the random component (μ_t) is uncorrelated with the X variables--an unwarranted assumption in our model. Transitory tax differences may well be correlated with such X variables as transitory income. As a result of this correlation and the nonstandard form of our estimator, the coefficients on the X variables may be inconsistent, reflecting a combination of the direct effect on gains (γ_0) and indirect effects through correlation with μ_t. While this may make interpretation of the effects of other variables more difficult, it does not affect the estimates of permanent and transitory tax effects.

Under our assumptions, it can be shown that the estimates of permanent and transitory tax coefficients in (1) will approach the following limits:[13]

$$\text{(6)} \qquad \text{plim}(\gamma_1) = (1-\theta_1)\gamma_1 + \theta_1\gamma_2 , \qquad \text{where } \theta_1 \equiv \frac{\text{cov}(\tau_s, \mu_t \,|X, \tau_{t-1})}{\text{cov}(\tau_s, \tau_t \,|X, \tau_{t-1})} ,$$

and

$$\text{(7)} \qquad \text{plim}(\gamma_2) = \theta_2\gamma_1 + (1-\theta_2)\gamma_2 , \qquad \text{where } \theta_2 \equiv \frac{\text{cov}(\tau_0, \tau_p \,|X, \tau_s, \tau_{t-1})}{\text{cov}(\tau_0, \tau_t \,|X, \tau_s, \tau_{t-1})} .$$

---

[11]Others have used state variation to identify permanent effects of statutory changes in a panel or cross-section in different contexts. See Daniel Feenberg (1987) for an application to charitable contributions and David Neumark and William Wascher (1991) for estimating the effects of minimum wages. William T. Bogart and William M. Gentry (1993) also use state data to estimate capital gains tax effects, but their state-level aggregate data averages out individual differences.

[12]The instrument could be endogenous if the choice of state depends on capital gains tax rates. While we consider it unlikely that state tax rates on capital gains are very important to residential decisions, we test for this possible source of bias in our estimation.

[13]Probability limits for all the parameters are derived in a separate appendix. The expression cov(x,y | z) is defined to be the partial covariance between x and y given z, i.e., after the linear influence of z is removed from x and y.

Equations (6) and (7) show that the probability limits for the permanent and transitory coefficient estimates are weighted averages of the true values of the coefficients. If, as assumed, the state tax rate instrument, τ_s, is correlated with τ_p, and uncorrelated with μ_t, conditional on the other variables, (6) implies that the estimated permanent coefficient is consistent because $\theta_1=0$. Similarly, (7) implies that, if the first-dollar instrument, τ_0, is correlated with μ_t, and uncorrelated with τ_p, conditional on the other variables, then the estimated transitory coefficient, γ_2, will also be consistent because $\theta_2=0$.

Under the null hypothesis that the permanent and transitory tax coefficients are the same, (6) and (7) imply that the estimates are consistent even if θ_1 and θ_2 are nonzero. Under the alternative hypothesis that $\gamma_1 \neq \gamma_2$, the estimated difference between γ_1 and γ_2 is biased toward zero if θ_1 or θ_2 is nonzero because both must lie between zero and one. Thus, even if the assumptions for consistency are violated, our estimates provide a conservative test of the hypothesis that $\gamma_1=\gamma_2$, which is the key assumption required for the validity of previous micro data studies.

C. *Alternative Estimators*

Two studies (Auten and Clotfelter, 1982, and U.S. Department of the Treasury, 1985) attempted to measure the permanent tax rate directly from panel data by using three-year averages of marginal tax rates on capital gains. The fundamental problem with this approach is that a three-year average of federal income tax rates would be correlated with the transitory component of the tax rate. Thus, such a proxy cannot be used to estimate separately the effects of permanent and transitory tax rates since it is, itself, a combination of the two.

Slemrod and Shobe (1990) used a fixed-effects model to control for differences in permanent tax rates and other unobservable fixed effects that may affect parameter estimates. This approach can produce consistent estimates of the coefficients of transitory tax rates and other non-fixed factors, but does not allow identification of the response of capital gains to permanent tax rates, as was recognized by the authors.[14]

Bogart and Gentry (1993) use aggregate state data to estimate permanent capital gains tax effects. This approach mitigates the problem of limited sample size common to aggregate time series models, and the data set includes more years than our study. However, aggregate data precludes dealing with most of the econometric problems that we have found to be empirically important and suffers from some of the same problems that affect aggregate time series studies.

---

[14]Because the combined federal and state tax rates vary over time as well as among individuals, we could conceivably estimate an individual fixed effect in our model. We did not do this for two reasons. First, modelling fixed effects would make it difficult or impossible to control for sample selectivity, which Auten et al. (1989) found to cause substantial biases. Second, because only a minority of mostly small states changed their tax rates on capital gains between 1980 and 1983 (Bogart and Gentry, 1993), only about 3 percent of the independent variation in the state tax instrument remains after controlling for both time and individual fixed effects. The sources of this variation are individuals who moved between 1980 and 1983, the 14 states that changed tax rates between 1980 and 1983, and the interaction effect between the change in federal tax rates in 1981 and the net capital gains tax rate. Since the precision of instrumental variables estimates depends on the correlation between permanent tax rates and the instrument, removing almost all of the variation in the instrument would yield uninformative results. Moreover, to the extent that the remaining variation corresponds to movers, who may have reasons for realizing capital gains independent of tax effects, interpreting the estimated coefficient as primarily a permanent tax effect may be unwarranted.

III. Data

The data are from a panel of individual income tax returns for about 11,000 taxpayers for the years 1979 to 1983. (U.S. Department of the Treasury, 1979-1983) In addition to detailed tax return data, the panel includes the ages of taxpayers for each return. The panel sample was stratified to oversample high-income taxpayers; thus, a much larger proportion of the sample (53.4 percent) had capital gains than in the population at large (18.5 percent). The Treasury department edited the data for consistency and developed programs to calculate marginal tax rates. (James M. Cilke and Roy A. Wyscarver, 1987) Some observations were discarded because the data were internally inconsistent.

Summary statistics for the data and instruments used in estimation are shown in Table 1. Weighted and unweighted statistics differ because the sample was stratified. We use unweighted data for estimation, but test for the possibility that endogenous stratification biases the estimates.[15]

Our data were originally prepared by Auten et al. (1989), but we have made several improvements. We created the instrument for permanent tax variation (τ_s) by computing the combined federal and state marginal tax rate on capital gains for a taxpayer with $100 million of taxable income. We also modified the first-dollar tax rate instrument (τ_0) by setting several possibly endogenous components of income and deductions equal to zero. This was done for long- and short-term capital gains and losses, capital loss carryovers, interest, dividends, business losses, charitable contributions, and the deduction for taxes paid and investment interest expense. Auten et al. did not consider the deduction items other than charitable contributions to be endogenous.

The sample period for estimation is 1980 to 1983 so that lagged values could be used. Observations on individuals were included in estimation whenever the current and lagged data were valid, which yielded a total of 42,406 observations. The dependent variable is net long-term capital gains before carryover of prior year losses as reported on Schedule D. The tax rate measure is the combined federal and state marginal rate, based on applicable tax law for each year and each taxpayer's income and deductions.[16] To smooth out kinks in the tax schedule and to represent the lumpiness of capital gains transactions, the marginal tax rate on capital gains was computed for a defined transaction, rather than for a dollar of capital gains. The capital gain on the defined transaction is the maximum of $1,000 or the square root of imputed wealth.[17]

---

[15]The sample was stratified based on income, which includes capital gains realizations and other possibly endogenous variables.

[16]Because of the way the data were coded by the IRS, state of residence is available for all returns only in 1981. In other years, we used the actual state if it was available, or the state in 1981, otherwise.

[17]As a sensitivity test, we also estimated the model using a marginal tax rate computed with a defined transaction of $1,000. This made almost no difference for the estimated effect of permanent tax rates, but increased slightly the estimated effect of transitory tax rates.

TABLE 1-DESCRIPTIVE STATISTICS FOR VARIABLES USED IN MODEL ESTIMATION

Variable Description	Population-Weighted		Unweighted	
	Mean	Coefficient of Variation	Mean	Coefficient of Variation
Net Long-Term Capital Gains	3.0	31.67	245.4	7.00
Percentage with Net Positive Gains	18.5	. . .	53.4	. . .
Marginal Tax Rate on Capital Gains	11.9	0.54	15.8	0.66
First-Dollar Tax Rate Instrument (τ_0)	11.1	0.54	13.1	0.73
Maximum Tax Rate Instrument (τ_s)	23.3	0.15	23.4	0.15
Imputed Permanent Income	28.8	0.96	125.7	1.64
Current Income (Exogenous Parts)	28.8	2.07	283.2	3.30
Imputed Wealth (Gross Assets)	125.3	0.83	286.5	6.70
Imputed Corporate Stock	11.3	4.81	118.3	8.61
Business Losses Lagged	1.6	15.19	101.8	3.96
Rent Losses Lagged	0.4	13.25	10.1	8.15

Notes: Dollar amounts in thousands of 1981 dollars. Marginal tax rates are in percentages. Statistics are for pooled years, 1980-1983.

The computed marginal tax rate is the change in tax liability divided by the amount of the defined capital gain.[18]

Other regressors are discussed above in Section II, and summarized in Table 1. Wealth was imputed by using a tobit model to regress the logarithm of total wealth, as reported in a 1982 sample of estate tax returns, on age and log capital income reported on 1981 income tax returns. The estimated wealth regression was used to impute wealth for taxpayers in our panel sample for each year based on lagged values of the regressors. Corporate stock was imputed the same way. Lagged business losses were computed as the sum of losses on rental property, losses reported on partnership returns, and losses reported by personal services corporations.

Permanent income was imputed by using the panel sample to regress the logarithm of a 5-year average (1979-1983) of real positive income on taxpayer characteristics.[19] The regression estimates were used to impute annual permanent income based on lagged values for the regressors. Current income, listed in Table 1, is defined as positive income excluding endogenous sources such as capital gains.[20] In the regression model, transitory income is the logarithm of the ratio of current to permanent income.

Family size is the number of personal and dependent exemptions claimed on the tax return. Marital status is based on tax filing status. Age was derived from social security records.

The sample period includes the Economic Recovery Tax Act (ERTA) of 1981, which reduced top tax rates on both ordinary income and capital gains by 29 percent and introduced many new tax preferences. The advantage of covering this period is that the change in tax law adds substantial variation to the statutory tax rates. The primary disadvantage is that the major tax change was far from a controlled experiment, and some of the response to the statutory change in capital gains tax rates may be transitory, although the aggregate change was controlled for by time dummies.

---

[18]Slemrod and Shobe (1990) argue that the marginal rate should be adjusted when the taxpayer has net capital losses to account for the fact that unused losses are at least partially deductible in future years. In our tax calculation, we consider only the current year, which implies a zero marginal tax rate on capital gains for taxpayers with nondeductible net losses. Given the small fraction of returns subject to the capital loss limitation, this difference is unlikely to affect empirical estimates. Moreover, since the capital loss limitation is essentially transitory, the estimates of permanent effects are unlikely to depend on the treatment of losses.

[19]Positive income includes all positive components of income (including net positive capital gains). It is an approximation of economic income used by the IRS and in several earlier studies.

[20]Current income and permanent income were scaled so that they had the same weighted population means. The unweighted mean for current income exceeds mean permanent income because the sample was stratified to oversample high-income taxpayers. Thus, people in the sample tend to have high transitory incomes. The wealth and stock variables were scaled to match the aggregates reported in the Survey of Consumer Finance for 1983, converted to 1981 dollars.

IV. Estimation Results

Estimates for equations (2) and (3) are shown in Tables 2 and 3. Table 2 shows estimates of tax rate coefficients, the corresponding elasticities, and results for three restricted models. Table 3 shows the coefficient estimates for the non-tax variables.

The first three columns in Table 2 show the marginal tax rate coefficients in the level equation (3) and the criterion function (2). To interpret these coefficient estimates, recall from footnote 6 that the total effect of changes in the permanent tax rate depends on the sum of the three tax rate coefficients, whereas the effect of transitory deviations depends only on coefficients of the current tax rate. Thus, the coefficient of the permanent tax rate in the full model is a measure of the difference in the effects of changes in the permanent and transitory tax rates.

The estimated current tax rate coefficient in the full model is negative and statistically significant at the 99 percent level in both the criterion function and level equation. The sign implies that individuals are more likely to realize capital gains (the criterion function) and realize more capital gains (the level equation) when they face temporarily low marginal tax rates. The size and significance of the coefficients imply that transitory changes in tax rates have a strong effect on taxpayer decisions about whether to sell appreciated assets and realize capital gains.[21]

The lagged tax rate coefficient is small and insignificant, which implies that lagged tax rates do not affect current capital gains decisions, holding current and permanent tax rates and other included variables constant. This result is inconsistent with the conventional wisdom that the response in the first year to a capital gains tax change is larger than the long-run response. It is, however, consistent with Kiefer's simulations, discussed in Section I.[22]

In contrast to the transitory effects, the estimated coefficient of the permanent tax rate is positive, nearly as large as the current tax rate coefficient, and significant at the 99 percent level in both the level equation and criterion function. This result implies that permanent changes in the tax rate have substantially smaller effects than transitory changes. These large and significant differences refute the basic assumption underlying the validity of previous micro data studies.

---

[21]Although our model includes measures of permanent and transitory income, the transitory tax rate component may also proxy for variation in transitory income not controlled for by other variables.

[22]Kiefer's simulations suggest a potentially larger "intermediate-term" effect, several years after a tax change. Unfortunately, there is not enough independent variation in tax rates in our data set to allow us to measure such effects with any precision.

TABLE 2–ESTIMATED COEFFICIENTS AND ELASTICITIES OF MARGINAL TAX RATE VARIABLES

Estimated Model	Marginal Tax Rate Coefficient: Current	Lagged	Permanent	Permanent Elasticity	Transitory Elasticity
Full Model					
Level Equation	−0.145	0.0013	0.116		
(Equation 3)	(0.014)	(0.011)	(0.036)	−0.18	−6.42
				(0.48)	(0.34)
Criterion Function	−0.084	0.003	0.088		
(Equation 2)	(0.005)	(0.006)	(0.016)		
Exclude Transitory and Lagged Tax Rates					
Level Equation	...	...	−0.020		
			(0.022)	−0.17	...
				(0.42)	
Criterion Function	...	...	0.007		
			(0.010)		
Exclude Permanent Tax Rate					
Level Equation	−0.144	0.039	...		
	(0.014)	(0.005)		...	−6.10
					(0.33)
Criterion Function	−0.083	0.036	...		
	(0.005)	(0.002)			
Exclude Permanent and Lagged Tax Rates					
Level Equation	−0.113	...	...		
	(0.010)			...	−4.19
					(0.22)
Criterion Function	−0.051	...	...		
	(0.003)				

Notes: Standard errors are in parentheses. Estimated coefficients of other variables included in the model are in Table 3. Elasticities are computed at an average tax rate of 18.0 and an average lambda of 2.52. See equation 8.

The last two columns show the elasticities. The elasticity (e) measures the effect of a small change in the permanent tax rate:

$$e = \tau_p \left[(\beta_1 + \beta_2 + \beta_3) + (\alpha_1 + \alpha_2 + \alpha_3) \cdot \lambda(h + \sigma\rho) \right], \tag{8}$$

where $\lambda(h+\sigma\rho)$ is the reciprocal of the Mills ratio evaluated at the mean of the systematic part of the criterion function (h) plus the covariance between the error terms in the criterion function and the level equation ($\sigma\rho$).[23] The transitory elasticity is given by a similar equation, excluding the permanent and lagged tax rate coefficients. It is interpreted as the elasticity with respect to a change in the current tax rate, holding the permanent and lagged tax rates constant.

The estimated permanent elasticity is -0.18, which implies that a 1 percent decrease in permanent tax rates would increase expected realized net long term capital gains by approximately 0.18 percent at average levels for all variables in 1983. However, the relatively large standard error implies that we cannot reject the hypothesis that permanent changes in capital gains tax rates have no long-term effect on capital gains realizations.[24] The standard error is also large enough that long-run elasticities of 0.0 and -1.0 are both included in a 95-percent confidence interval.

The estimated transitory elasticity is -6.42, which is larger in absolute value than most previous elasticity estimates from micro data.[25] The high transitory elasticity suggests that the response to a temporary tax change would be extraordinary, with realizations expected to increase by more than six times the percentage change in the tax rate. This is consistent with the dramatic increase in realizations just after passage of the Tax Reform Act of 1986, as discussed in the introduction.

The second panel of the table shows what happens to estimates of the permanent elasticity when the current and lagged tax rates are excluded from the estimated model. Assuming that the transitory component of the tax rate is uncorrelated with the permanent (state) tax rate instrument, the estimates of the permanent tax rate coefficient and elasticity are still consistent when the current and lagged tax rates are excluded. The permanent elasticity estimate changes very little, from -0.18 to -0.17, but the precision increases slightly. The transitory elasticity cannot be determined from this specification.

The third panel shows the effect of excluding the permanent tax rate, but including the current and lagged tax rates, as in Auten et al. (1989). The current tax rate coefficients and implied transitory elasticity decrease slightly, and the lagged tax rate coefficients increase and become highly significant. This result makes sense because the average of tax rates over two years should be positively correlated with the omitted permanent tax rate. The omission would thus positively bias the current and lagged tax rate coefficient estimates. This result suggests that the lagged tax rate partially proxies for the omitted permanent tax rate.

---

[23]Derivation is available from the authors on request. Permanent and transitory elasticities were computed for 1983 means of the permanent tax rate and λ, which were 18.0 and 2.52, respectively.

[24]We refer to long-term changes because variation in the permanent tax rate instrument represents essentially cross-section variation in state marginal tax rates. While the combined state and federal marginal tax rates changed over time during our sample period, much of the possible influence of this source of variation was removed by including time dummy variables in our model.

[25]See, e.g., Slemrod and Shobe (1991), Auten et al. (1989), and Gillingham et al. (1989) for recent estimates.

The fourth panel shows the effect of omitting both the lagged and permanent tax rates, as in Gillingham et al. (1989). The transitory elasticity estimate becomes smaller, probably because the first-dollar tax rate instrument is positively correlated with the permanent tax rate. A positive correlation would cause a positive bias in the transitory elasticity estimate, which may explain why previous micro data studies have yielded smaller transitory elasticity estimates. This result is consistent with Slemrod and Shobe's (1990) finding that elasticity estimates were biased toward zero by failure to control for unmeasured fixed effects, such as the permanent tax rate.

The effects of other variables are summarized in Table 3, which reports estimated coefficients for the level equation, the criterion function, and the combined effects of implied changes in the values of both functions on the expected value of capital gains realizations.[26] For continuous variables, the estimates in Table 3 are reported as elasticities. For dummy variables, i.e., those followed by (D) in the table, the effects are reported as percentage changes in expected capital gains realizations implied by changing each dummy variable from zero to one.

The results seem generally consistent with life-cycle motives for saving and consumption, modified somewhat by the incentive to hold assets with gains until death. Capital gains realizations are significantly positively related to permanent income, but negatively related to transitory income, suggesting a consumption motive for realizations. Wealthier people are much more likely to realize capital gains, and realize larger gains than average. The composition of wealth also matters. A larger share of stocks in the portfolio--as measured by the stock/wealth variable--makes people significantly more likely to realize gains, but the average size of a gain is smaller, *ceteris paribus*. This result may be a consequence of the lower transaction costs for stocks than for other kinds of assets, such as real estate. The positive and significant relationship between gains and lagged business losses reflects the well-known relationship between tax shelters and capital gains, although rental losses (a subset of business losses) do not seem to have a very large independent effect on realizations.

Holding wealth and other variables constant, the pattern of realizations follows the expected life-cycle profile except for the oldest cohort. The level of realizations declines steadily through the peak earning years of 50-59, and then increases. The likelihood of realizing gains steadily increases, perhaps reflecting the fact that older people are more likely to own assets that yield capital gains. The percentage change in realizations is also U-shaped through age 69. However, the oldest taxpayers realize less capital gains than the 60-69 cohort, and are slightly less likely to realize. Although this difference is statistically insignificant, it is consistent with older taxpayers avoiding realizations to take advantage of the step-up in basis at death.

The Mills ratio coefficient equals the product of the standard error of the error term in the level equation, (3), and the correlation between the error terms in equations (2) and (3). The fact that the coefficient is nonzero implies that ignoring sample selectivity would lead to biased and inconsistent parameter estimates. The negative sign implies that the error terms are negatively correlated. Thus, the tobit model used in some previous studies, which assumes a correlation of one, would be inappropriate.

[26]Recall from Section II that these coefficients may not be estimated consistently.

TABLE 3–ESTIMATED COEFFICIENTS OF NON–TAX VARIABLES INCLUDED IN MODEL

Right–Hand Variable[a]	Coefficients: Level Equation	Coefficients: Criterion Function	Elasticity or Percentage Change[b]
Intercept	3.78	−9.70	...
	(1.25)	(0.30)	
Permanent Income (L)	0.17	0.15	0.55
	(0.06)	(0.02)	
Transitory Income (L)	−0.12	−0.10	−0.37
	(0.02)	(0.006)	
Wealth (L)	0.56	0.61	2.10
	(0.08)	(0.02)	
Stocks/Wealth (L)	−0.09	0.07	0.10
	(0.03)	(0.008)	
Business Losses Lagged (L)	0.03	0.05	0.16
	(0.009)	(0.003)	
Rent Losses Lagged (L)	−0.002	0.006	0.01
	(0.005)	(0.002)	
Family Size	0.009	−0.012	−0.06
	(0.02)	(0.006)	
Married (D)	0.19	0.03	30.9%
	(0.07)	(0.02)	
Age 30–39 (D)	−0.03	0.20	62.2%
	(0.15)	(0.04)	
Age 40–49 (D)	−0.61	0.41	44.9%
	(0.15)	(0.04)	
Age 50–59 (D)	−0.87	0.47	28.7%
	(0.15)	(0.04)	
Age 60–69 (D)	−0.81	0.49	41.0%
	(0.16)	(0.04)	
Age 70 or Older (D)	−0.85	0.48	31.4%
	(0.17)	(0.05)	
Southern Region (D)	0.23	0.009	29.1%
	(0.05)	(0.02)	
Western Region (D)	0.17	−0.019	12.4%
	(0.06)	(0.02)	
Northeast Region (D)	0.35	−0.13	1.6%
	(0.06)	(0.02)	
Year 1981 (D)	0.17	0.15	72.3%
	(0.14)	(0.04)	
Year 1982 (D)	−0.51	0.13	−17.3%
	(0.11)	(0.04)	
Year 1983 (D)	−0.36	0.17	6.4%
	(0.08)	(0.03)	
Inverse Mills Ratio	−2.68	...	...
	(0.18)		
Standard error (Sigma 1)	3.43	...	...
Observations	22,635	42,406	...

Notes: Standard errors are in parentheses.

[a] Logarithmic variables are indicated by (L). Dummy variables are indicated by (D).
[b] Numbers represent elasticities for continuous variables and percentage changes in expected long–term gains for dummy variables. All elasticities and percentage changes are evaluated at unweighted sample means of right–hand variables.

Sensitivity tests for alternative specifications and segments of the data set are reported in Table 4. The results all confirm our basic finding that permanent elasticities are much smaller than transitory elasticities. The permanent elasticity is not significantly different from zero in any specification.

The log-log model tests an approximately constant elasticity specification, which we view as inferior to the semi-log form. The results are similar to those under the semi-log form, but have higher standard errors. Weighted estimates are also consistent with our basic results. This result suggests that Joseph J. Minarik's (1981) finding that weighting could substantially alter elasticity estimates was a consequence of other estimation problems rather than endogenous sample stratification. We excluded taxpayers from high- and low-tax states to test for the possibility of endogeneity bias in our state tax rate instrument. This experiment raises the standard errors significantly because much of the variation in the instrument is sacrificed, but does not alter the key conclusions. Results are similar when the sample is restricted to 1982 and 1983 (after enactment of ERTA). Even when truncation is ignored and the model is estimated by two-stage least squares, the elasticity estimates do not change much. Estimating the model by two-stage least squares based on a sample of realizers only, the transitory elasticity changed significantly, but the effect on the permanent elasticity estimate is small and insignificant.

V. Conclusion

It has long been suspected that differences between transitory and permanent responses to capital gains tax changes were at the heart of the conflicting empirical evidence from cross-section and time-series data. Using state tax rates to distinguish transitory from permanent tax effects, and correcting other econometric problems with previous studies, we find that the difference is large and statistically significant. The difference in estimated response is even larger than the differences between past empirical results from careful micro-data studies, which measured a combination of permanent and transitory effects, and time-series studies, which are likely to have measured primarily permanent effects of changes in tax rates.

Our analysis has some limitations. First, the capital gains realizations elasticity is only one of many factors that affect the proper taxation of capital gains. For example, our analysis ignores the effects of capital gains taxes on the cost of capital and the allocation of capital among kinds of investments, and it says nothing about arguments for taxing capital gains on equity grounds. Second, this paper has followed all previous empirical research in estimating a reduced form model. Although this was necessitated by data limitations, it was also important to show that permanent and transitory tax effects could be estimated separately using a model otherwise similar to previous research. Any explicit structural model would require assumptions about the nature of preferences and individuals' optimization problems and the estimation method itself would be a radical departure from all prior research. That might lay open such an analysis to the criticism that the structure of the model was generating the results. The drawback of estimating a reduced form, however, is that the estimated parameters are functions of the tax law and macroeconomic environment and may thus change over time.

The distinction between transitory and permanent tax effects may explain some other empirical anomalies. For example, the empirical evidence on the tax-sensitivity of charitable contributions seems to exhibit a similar divergence between time series and micro data estimates. The methodology developed here may help to resolve such disparities.

TABLE 4-SENSITIVITY ANALYSIS

Sensitivity Test	Permanent Elasticity		Transitory Elasticity		Sample Size
	Estimate	Std. Error	Estimate	Std. Error	
Log-log model	-0.17	0.39	-3.32	0.14	42,406
Weighted estimates	-0.06	N/A	-5.63	N/A	42,070
Exclude high and low-tax states	0.33	1.46	-5.34	0.34	24,188
Post-ERTA (1982 and 1983)	0.73	0.63	-9.28	0.37	21,062
Ignore truncation (2SLS)					
All observations	0.11	0.74	-6.91	0.36	42,406
Realizers only	-0.27	0.49	-4.63	0.32	22,635

Notes: Unless specified, all estimates are for the semi-log model.

[a] Standard errors are unknown. Excludes returns without valid weights.

REFERENCES

Auerbach, Alan J., "Capital Gains Taxation in the United States: Realizations, Revenue and Rhetoric," *Brookings Papers on Economic Activity*, 1988, (2), 595-631.

___, "Capital Gains Taxation and Tax Reform," *National Tax Journal*, September 1989, *42*, 391-401.

Auten, Gerald E., Burman, Leonard E. and Randolph, William C., "Estimation and Interpretation of Capital Gains Realization Behavior: Evidence From Panel Data," *National Tax Journal*, September 1989, *42*, 353-374.

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

Balcer, Yves and Judd, Kenneth L., "Effects of Capital Gains Taxation on Life-Cycle Investment and Portfolio Management," *The Journal of Finance*, July 1987, *42*, 743-758.

Bogart, William T. and Gentry, William M., "Capital Gains Taxes and Realizations: Evidence from Interstate Comparisons," National Bureau of Economic Research (Cambridge, MA) Working Paper 4254, January 1993.

Bolster, Paul J., Lindsey, Lawrence B. and Mitrusi, Andrew, "Tax-Induced Trading: The Effect of the 1986 Tax Reform Act on Stock Market Activity," *The Journal of Finance*, June 1989, *44*, 327-344.

Burman, Leonard E. and Randolph, William C., "Theoretical Determinants of Aggregate Capital Gains Realizations," mimeo, Congressional Budget Office and U.S. Treasury, September 1992.

___, Clausing, Kimberly and O'Hare, John, "Tax Reform and Realizations of Capital Gains in 1986," mimeo, Congressional Budget Office, Harvard University, and Joint Committee on Taxation, March 1993.

Cilke, James M. and Wyscarver, Roy A., "The Individual Income Tax Simulation Model," in *Compendium of Tax Research 1987*, Washington, DC: Government Printing Office, 1987.

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

Feenberg, Daniel, "Are Tax Price Models Really Identified? The Case of Charitable Giving," *National Tax Journal*, December 1987, *60*, 629-633.

Feldstein, Martin, Slemrod, Joel and Yitzhaki, Shlomo, "The Effects of Taxation on the Selling of Corporate Stock and the Realization of Capital Gains," *Quarterly Journal of Economics*, June 1980, *95*, 777-791.

Gillingham, Robert and Greenlees, John S., "The Effect of Capital Gains Tax Rates on Capital Gains Tax Revenues: Another Look at the Evidence," *National Tax Journal*, June 1992, *45*, 167-178.

___, ____ and Zieschang, Kim D., "New Estimates of Capital Gains Realization Behavior: Evidence from Pooled Cross-Section Data," Office of Tax Analysis Paper 66, U.S. Department of the Treasury, May 1989.

Gravelle, Jane G., "A Proposal for Raising Revenue by Reducing Capital Gains Taxes," Congressional Research Service Report 87-562E, Library of Congress, June 30, 1987.

___, "Limits to Capital Gains Feedback Effects," Congressional Research Service Report 91-250 RCO, Library of Congress, March 15, 1991.

Heckman, James J., "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement*, Fall 1976, *5*, 475-492.

Hendershott, Patric H., Toder, Eric and Won, Yunhi, "Effects of Capital Gains on Revenue and Economic Efficiency," *National Tax Journal*, March 1991, *44*, 21-40.

Jones, Jonathan D., "An Analysis of Aggregate Time Series Capital Gains Equations," Office of Tax Analysis Paper 65, U.S. Department of the Treasury, May 1989.

Kiefer, Donald W., "Lock-In Effect Within a Simple Model of Corporate Stock Trading," *National Tax Journal*, March 1990, *43*, 75-95.

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

Minarik, Joseph J., "Capital Gains," in H.J. Aaron and J.H. Pechman, eds., *How Taxes Affect Economic Behavior*, Washington, DC: The Brookings Institution, 1981, 241-277.

Neumark, David and Wascher, William, "Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws," *Industrial and Labor Relations Review*, October 1992, *46*, 55-81.

Seltzer, Lawrence H., *The Nature and Tax Treatment of Capital Gains and Losses*, New York: National Bureau of Economic Research, 1951.

Slemrod, Joel and Shobe, William, "The Tax Elasticity of Capital Gains Realizations: Evidence from a Panel of Taxpayers," National Bureau of Economic Research (Cambridge, MA) Working Paper 3237, January 1990.

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

U.S. Department of the Treasury, Internal Revenue Service, Statistics of Income Division, Special Panel of Tax Returns, 1979-1983.

U.S. Department of the Treasury, *Report to the Congress on the Capital Gains Tax Reductions of 1978*, Washington, DC: U.S. Government Printing Office, September 1985.

APPENDICES

These appendices derive conditions for consistency of the IV estimator for permanent and transitory tax effects discussed in Section II, derive a consistent estimator for the generalized tobit model with an endogenous regressor in both the criterion and level equations, discussed in Section III, derive elasticities for the simultaneous selection model, examine the simultaneous equations bias induced by using actual tax rates or proxy variables as regressors to estimate tax effects, as has been done in several influential earlier studies and examine the variation in tax rates in the panel.

Appendix A. Consistency of the IV Estimator

The linear capital gains model can be written as

$$g = H\Gamma + \epsilon \tag{A1}$$

where $H=[X \vdots \tau_p \vdots \mu]$, $\Gamma' = [\gamma_0' \vdots \gamma_1 \vdots \gamma_2]$, $\mu = \tau - \tau_p$, and it is assumed for simplicity that all variables are expressed as deviations from means. Note that the model has been expressed for convenience in terms of the permanent tax rate, τ_p, and the transitory component of the current marginal tax rate, μ. It is also assumed that there are n i.i.d. observations, and that H is of full column rank, conformable with Γ.[1]

The estimator, $\hat{\Gamma}$, can be written as

$$\hat{\Gamma} = (\hat{H}'\hat{H})^{-1}\hat{H}'g, \tag{A2}$$

where $\hat{H}=[X \vdots \hat{\tau}_p \vdots \hat{\mu}]$, $\hat{\tau}_p = W_1(W_1'W_1)^{-1}W_1'\tau$, $W_1=[X \vdots \tau_s]$, $\hat{\mu}=\hat{\tau}-\hat{\tau}_p$, $\hat{\tau}=W_2(W_2'W_2)^{-1}W_2'\tau$, and $W_2=[W_1 \vdots \tau_0]$. Note that $\hat{\mu}$ is orthogonal to X and $\hat{\tau}_p$, from which it follows that $\hat{\gamma}_0$ and $\hat{\gamma}_1$ are instrumental variables estimates, where $\hat{\tau}_p$ is the instrument for τ_p in equation (A1). As shown below, consistency of $\hat{\gamma}_0$ and $\hat{\gamma}_1$ depends on the covariance between X and τ_s with μ and ϵ. Consistency of $\hat{\gamma}_2$ depends on the covariance of τ_0 with τ_p and ϵ, conditional on X and τ_s.

---

[1] It is also assumed that the variables have finite moments up to at least the third order. In this appendix, the matrix X is not necessarily defined in the same way as in the main body of the paper. In the context of second-stage estimation of the generalized tobit model, for example, the estimated inverse Mill's ratio can be treated as a column of X.

A. *The Probability Limit of* $\hat{\Gamma}$

By using the fact that $\hat{\mu}$ is orthogonal to X and $\hat{\tau}_p$, and applying well known rules for a partitioned inverse, $\hat{\Gamma}$ can be rewritten as

$$\hat{\gamma}_0 = (X'M_pX)^{-1}X'M_pg \tag{A3}$$

$$\hat{\gamma}_1 = (\hat{\tau}_p'M_X\hat{\tau}_p)^{-1}\hat{\tau}_p'M_Xg \tag{A4}$$

$$\hat{\gamma}_2 = (\hat{\mu}'\hat{\mu})^{-1}\hat{\mu}'g, \tag{A5}$$

where $M_p=[I_n-\hat{\tau}_p(\hat{\tau}_p'\hat{\tau}_p)^{-1}\hat{\tau}_p']$ and $M_X=[I_n-X(X'X)^{-1}X']$. In the remainder of this appendix, equations (A3), (A4), and (A5) are used to derive the probability limits of $\hat{\gamma}_0$, $\hat{\gamma}_1$, and $\hat{\gamma}_2$ as $n\to\infty$. The derivations that follow also use the fact that $\hat{\tau}_p$ can be rewritten as

$$\hat{\tau}_p = X(X'X)^{-1}X'\tau + M_X\tau_s(\tau_s'M_X\tau_s)^{-1}\tau_s'M_X\tau,$$

and $\hat{\mu}$ can be rewritten as

$$\hat{\mu}=M_{W_1}\tau_0(\tau_0'M_{W_1}\tau_0)^{-1}\tau_0'M_{W_1}\tau,$$

where $M_{W_1}=[I_n-W_1(W_1'W_1)^{-1}W_1']$.

1. *The probability limit of* $\hat{\gamma}_0$

By combining (A1) and (A3), $\hat{\gamma}_0$ (the coefficient on X) can be written as

$$\hat{\gamma}_0 = (X'M_pX)^{-1}X'M_p(H\Gamma+\epsilon) = \gamma_0+(X'M_pX)^{-1}[\gamma_1X'M_p\tau_p+\gamma_2X'M_p\mu+X'M_p\epsilon].$$

Because $\tau_p=\tau-\mu$ and $M_p\tau=0$, $\hat{\gamma}_0$ can be rewritten as

$$\hat{\gamma}_0 = \gamma_0 + (\gamma_2-\gamma_1)(X'M_pX)^{-1}X'M_p\mu + (X'M_pX)^{-1}X'M_p\epsilon. \tag{A6}$$

Under the assumptions that $var(X)$ is nonsingular, $var(\tau_s \mid X) > 0$, and by repeated application of Khintchine's theorem to the terms in equation (A6), it can be shown that as $n \to \infty$,

$$\begin{aligned} Plim(\hat{\gamma}_0) &= \gamma_0 + k_1 var(X)^{-1}[(\gamma_2-\gamma_1)cov(X,\mu) + cov(X,\epsilon)] \\ &\quad + k_2 var(\tau_s \mid X)^{-1}[cov(\tau_s,\mu \mid X) + cov(\tau_s,\epsilon \mid X)], \end{aligned} \tag{A7}$$

where k_1 and k_2 are known (messy) nonzero functions of the second order moments. Furthermore, using the expressions for k_1 and k_2, it can be shown that if $cov(\tau_s,\mu \mid X) = cov(\tau_s,\epsilon \mid X) = 0$, and $cov(X,\epsilon) = 0$, equation (A7) reduces to

$$Plim(\hat{\gamma}_0) = \gamma_0 + (\gamma_2-\gamma_1)var(X)^{-1}cov(X,\mu). \tag{A8}$$

(We define the expression $\mathrm{cov}(x,y \mid z)$ to be the partial covariance between x and y given z, i.e., after the linear influence of z is removed from x and y.)

Thus, if τ_s is uncorrelated with μ and ϵ, conditional on X, and X is uncorrelated with ϵ, then the probability limit of $\hat{\gamma}_0$ differs from γ_0 by the product of the difference between the transitory and permanent tax effects and a familiar term for omitted variable bias, where μ is the omitted variable.

2. *The probability limit of* $\hat{\gamma}_1$

By combining (A1) and (A4), $\hat{\gamma}_1$ (the coefficient of τ_p) can be written as

$$\hat{\gamma}_1 = (\hat{\tau}_p' M_X \hat{\tau}_p)^{-1} \hat{\tau}_p' M_X (H\Gamma + \epsilon) = (\hat{\tau}_p' M_X \hat{\tau}_p)^{-1} (\gamma_1 \hat{\tau}_p' M_X \tau_p + \gamma_2 \hat{\tau}_p' M_X \mu + \hat{\tau}_p' M_X \epsilon).$$

Under the assumption that $cov(\tau_s,\tau \mid X) \neq 0$, repeated application of Kintchine's theorem can be used to show that

$$Plim(\hat{\gamma}_1) = \gamma_1 + (\gamma_2-\gamma_1)\frac{cov(\tau_s,\mu \mid X)}{cov(\tau_s,\tau \mid X)} + \frac{cov(\tau_s,\epsilon \mid X)}{cov(\tau_s,\tau \mid X)}. \tag{A9}$$

Thus, $\hat{\gamma}_1$ is a consistent estimate for γ_1 if τ_s is uncorrelated with μ and ϵ conditional on X.

Note that under the assumption, $cov(\tau_s,\epsilon \mid X) = 0$, (A9) can be rewritten as a weighted average of γ_1 and γ_2:

$$Plim(\hat{\gamma}_1) \;=\; \gamma_1(1-\theta_1) \;+\; \gamma_2\theta_1, \tag{A10}$$

where $\theta_1 = \dfrac{cov(\tau_s,\mu|X)}{cov(\tau_s,\tau|X)}$.

3. *The probability limit of* $\hat{\gamma}_2$

By similar reasoning, under the assumption that $cov(\tau_0,\tau|X,\tau_s) \neq 0$, it can be shown that the probability limit of $\hat{\gamma}_2$ (the coefficient of μ_t) is

$$Plim(\hat{\gamma}_2) \;=\; \gamma_2 \;+\; (\gamma_1-\gamma_2)\frac{cov(\tau_0,\tau_p|X,\tau_s)}{cov(\tau_0,\tau|X,\tau_s)} \;+\; \frac{cov(\tau_0,\epsilon|X,\tau_s)}{cov(\tau_0,\tau|X,\tau_s)}. \tag{A11}$$

Note that under the assumption, $cov(\tau_0,\epsilon|X,\tau_s)=0$, (A10) can be rewritten as a weighted average of γ_2 and γ_1:

$$Plim(\hat{\gamma}_2) \;=\; \gamma_2(1-\theta_2) \;+\; \gamma_1\theta_2, \tag{A12}$$

where $\theta_2 = \dfrac{cov(\tau_0,\tau_p|X,\tau_s)}{cov(\tau_0,\tau|X,\tau_s)}$.

Appendix B. Consistent Estimation of Generalized Tobit Model With Endogenous Regressors in Both the Probit and Level Equations

The generalized tobit model in the paper is the following three equations

$$I^* = X\alpha_0 + \alpha_1\tau_p + \alpha_2\tau_t + \alpha_3\tau_{t-1} + \epsilon_1 , \tag{A13}$$

$$g = \begin{cases} X\beta_0 + \beta_1\tau_p + \beta_2\tau_t + \beta_3\tau_{t-1} + \epsilon_2 & \text{if } I^*>0 \\ 0 & \text{otherwise} \end{cases} , \tag{A14}$$

and

$$\tau_t \equiv f(X,Y,g) . \tag{A15}$$

To simplify exposition, we begin by treating τ_p as if it is known and exogenous, and focus on the endogeneity of τ_t. Assume that, for purposes of the first-stage probit estimates, the reduced form for τ_t is approximately linear in X, τ_p, τ_{t-1}, and τ_0.[2] This constructed instrument, τ_0, is a "first-dollar" marginal tax rate, computed with the endogenous sources of income, g and Y, set equal to 0:

$$\tau_0 \equiv f(X,0,0) \tag{A16}$$

(recalling that f is the tax function in equation (A15)).

Let Z be the matrix:

$$Z \equiv \left[X \vdots \tau_p \vdots \tau_{t-1} \vdots \tau_0 \right] . \tag{A17}$$

Then the linearized reduced form for τ_t is

$$\tau_t = Z\pi + u , \tag{A18}$$

where π is a parameter vector and u is an error term that is assumed to be uncorrelated with Z.

---

[2]The assumption of linearity is only for convenience. The IV procedure could produce consistent estimates even if the tax rates were highly non-linear.

Substituting the reduced form for τ_t in (A13) yields

$$I^* = X\alpha_0 + \alpha_1\tau_p + \alpha_2(Z\pi) + \alpha_3\tau_{t-1} + (\epsilon_1 + \alpha_2 u) . \tag{A19}$$

Following Lee, et al (1980), the conditional expectation of capital gains is

$$E(g|I^*>0) = X\beta_0 + \beta_1\tau_p + \beta_2 E(\tau_t|I^*>0) + \beta_3\tau_{t-1} + E(\epsilon_2|I^*>0) . \tag{A20}$$

Using the standard formula for truncated means of normal random variates, the conditional expectations may be written as:

$$E(\tau_t|I^*>0) = Z\pi + (\sigma_{u1} + \alpha_2\sigma_{uu})\frac{\phi}{\Phi} , \tag{A21}$$

and

$$E(\epsilon_2|I^*>0) = (\sigma_{12} + \alpha_2\sigma_{u2})\frac{\phi}{\Phi} , \tag{A22}$$

where ϕ and Φ are the standard normal density and distribution functions, respectively, evaluated at $X\alpha_0+\alpha_1\tau_p+\alpha_2(Z\pi)+\alpha_3\tau_{t-1}$.

The difference from the model of Lee, et al (1980), appears in the σ_{uu} term that appears in the conditional mean for τ_t. This term is unambiguously positive, which implies that the conditional mean for τ_t is almost surely different from the unconditional mean, even if the covariances between different error terms are zero. This implies that correcting for selectivity is essential to finding consistent parameter estimates even if the error terms in equations (A13) and (A14) are uncorrelated, i.e., even if $\sigma_{12}=0$.

A similar potential bias from ignoring selectivity would occur because τ_p is unobserved. The IV estimator described in section II replaces τ_p with $\hat{\tau}_p$ in both the probit and level equations. Thus, the preceding analysis applies, using $\hat{\tau}_p - \tau_p - X(X'X)^{-1}X'\mu_t$ in place of u and with α_1 substituted for α_2. The potential selectivity bias discussed above is thus compounded. Fortunately, the solution is the same in both cases.

Appendix C. Computing Permanent Elasticities in the Selection Model

Let g_i be net long-term capital gains for individual i, and let E be the expectation operator (conditional on $\tau_{i,p}$ and X_i, and $\tau_{i,t}=\tau_{i,t-1}=\tau_{i,p}$). Then the long-term elasticity of gains with respect to permanent tax rates, e_i, is

$$e_i \equiv \frac{\partial Eg_i}{\partial \tau_{i,p}} \frac{\tau_{i,p}}{Eg_i} . \tag{A23}$$

For the semi-log model, it may be shown that the expression in (A23) is:

$$e_i = \tau_{i,p} \left[(\beta_0 + \beta_1 + \beta_2) + (\alpha_0 + \alpha_1 + \alpha_2) \cdot \lambda(h_i + \sigma\rho) \right] , \tag{A24}$$

where $\lambda(\cdot)$ is the reciprocal of the Mills-ratio function (ϕ/Φ), h_i is the systematic part of the criterion function, equation (A13), and $\sigma\rho$ is the covariance between the errors in the two equations.[3] The first part of (A24) is the response of the level of capital gain conditional on realizing a capital gain. The second part is the sum of the direct effect of tax rates on the probability of realizing and the indirect effect through the covariance in errors between the criterion and level equations.

For the log-log model, the elasticity is:[4]

$$e_i = \frac{\tau_{i,p}}{1 + \tau_{i,p}} \left[(\beta_0 + \beta_1 + \beta_2) + (\alpha_0 + \alpha_1 + \alpha_2) \cdot \lambda(h_i + \sigma\rho) \right] . \tag{A25}$$

The elasticity of aggregate realizations with respect to the permanent tax rate is the weighted sum of the individual elasticities, evaluated at the permanent tax rates.[5] The correct weights are the sample weights multiplied by the amount of capital gains. In practice, the results are virtually identical

---

[3]See Appendix D for the proof.

[4]Recall that the tax terms have the form $\log(1+\tau_{i,p})$ and that $\tau_{i,p}$ is measured in percents.

[5]One might be tempted to compute the individual elasticities at the actual tax rates rather than the expected permanent tax rates. This procedure would bias estimates because tax rates are endogenous and correlated with the level of capital gains. There is also a selectivity bias.

if (A24) and (A25) are evaluated at the gains-weighted means of permanent tax rates and λ.[6] Because the first part of each expression is conditional on realizing a capital gain, the appropriate average tax rate is the average estimated permanent tax rate for realizers, which is 19 percent in 1983. For the second part, we use the unconditional average permanent tax rate, which is 18 percent. The weighted average values of λ, which depend on estimated parameters, are 3.1 for the semi-log model and 2.9 for the log-log model.

Appendix D. Derivation of Elasticity Formula

The formula for elasticity is complicated somewhat because the dependent variable in the selection model is in logarithms. This appendix derives the general formula for the log-selection model, which was applied in Appendix C.

A. Expectation in a Generalized Tobit Model with Log Dependent Variable

The model may be written in general form as:

$$\begin{aligned} \ln Y_i &= f(X_i) + v_{1i} \quad && \textit{if } h(X_i) + u_{2i} \geq 0 \\ &= 0 && \textit{otherwise}, \end{aligned}$$

for $i=1,\ldots,N$. Assume that, conditional on X_i, both v_{1i} and u_{2i} are independent, identically distributed random variables such that $v_{1i} \sim N(0,\sigma_1^2)$, $u_{2i} \sim N(0,1)$, and $E(v_{1i}u_{2i}) = \sigma_{12}$. The correlation between v_{1i} and u_{2i} is $\rho = \sigma_{12}/\sigma_1$. We can rewrite the nonstandard normal random variable, v_{1i}, in terms of the standard normal, u_{1i}, as $v_{1i} \equiv \sigma_1 u_{1i}$.

[6]Computing the elasticities for the main specifications using micro-simulation changed only the third significant digit of the estimates. We report elasticities at the mean because they are easier to reproduce, requiring only a calculator and the parameter values reported below.

The expectation of Y_i conditional on X_i is:

$$EY_i = E\left[e^{f(x_i)+\sigma u_{1i}} \mid u_{2i} \geq -h(x_i)\right] \cdot (1 - \Phi(-h(x_i))), \tag{A26}$$

(treating the value of EY_i as 0 if $h(x)+u_{2i} < 0$).

The complication in equation (A26) is the conditional expectation. The following lemma derives its value.

Lemma:

$$E(e^{\sigma u_1} \mid u_2 \geq c) = e^{\sigma^2/2} \frac{1-\Phi(c-\sigma\rho)}{1-\Phi(c)}. \tag{A27}$$

Proof:

Let u_1 and u_2 be jointly standard normal with correlation ρ. Their joint density is:

$$g(u_1,u_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left(u_1^2 - 2\rho u_1 u_2 + u_2^2\right)\right\}. \tag{A28}$$

The conditional mean of $e^{\sigma u_1}$ given $u_2 \geq c$ is

$$E(e^{\sigma u_1} \mid u_2 \geq c) = \frac{\int_c^{\infty}\int_{-\infty}^{\infty} e^{\sigma u_1} g(u_1,u_2)\, du_1\, du_2}{1-\Phi(c)}, \tag{A29}$$

where Φ is the univariate normal distribution function.

Finding the solution to (A29) involves integrating out u_1 and then recognizing that the remaining terms in u_2 are proportional to a normal density. The numerator of (A29) may rewritten as:

$$N = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_c^{\infty} \int_{-\infty}^{\infty} e^{q(u_1,u_2)}\, du_1\, du_2, \tag{A30}$$

where

$$q\,(u_1,u_2) \equiv \sigma u_1 - \frac{1}{2(1-\rho^2)} \left(u_1^2 - 2\,\rho u_1 u_2 + u_2^2\right).$$

Rearrange q so that it is in the form

$$q\,(u_1,u_2) = (u_1 - \mu)^2 + r(u_2),$$

where μ does not depend on u_1 and $r(u_2)$ depends only on u_2. Then

$$q\,(u_1,u_2) = -\frac{1}{2(1-\rho^2)} \left(u_1^2 - 2\,\rho u_2 u_1 - 2\sigma(1-\rho^2)u_1 + u_2^2\right)$$

$$= -\frac{u_1^2 - 2\left[\rho u_2 + \sigma(1-\rho^2)\right]u_1 + \left[\rho u_2 + \sigma(1-\rho^2)\right]^2 - \left[\rho u_2 + \sigma(1-\rho^2)\right]^2 + u_2^2}{2\,(1-\rho^2)}$$

$$= -\frac{1}{2(1-\rho^2)} \left[u_1 - \rho u_2 - \sigma(1-\rho^2)\right]^2 - \frac{1}{2(1-\rho^2)} \left[u_2^2 - (\rho u_2 + \sigma(1-\rho^2))^2\right].$$

Substitute into (A30) and rearrange:

$$N = \int_c^{\infty} \frac{\exp\left[-\frac{1}{2(1-\rho^2)} \left(u_2^2 - (\rho u_2 + \sigma(1-\rho^2))^2\right)\right]^2}{\sqrt{2\pi}} \times$$

$$\left(\int_{-\infty}^{\infty} \frac{\exp\left[-\frac{1}{2(1-\rho^2)} (u_1 - \rho u_2 - \sigma(1-\rho^2))^2\right]}{\sqrt{2\pi(1-\rho^2)}} du_1 \right) du_2 .$$

The integrand inside the parentheses is (by design) the normal density for mean $\left[\rho u_2 + \sigma(1-\rho^2)\right]$ and variance $1 - \rho^2$. Thus, the integral is 1.

N, therefore, simplifies to:

$$N = \int_c^{\infty} \frac{\exp\left[-\frac{1}{2(1-\rho^2)} \left[u_2^2 - (\rho u_2 + \sigma(1-\rho^2))^2\right]\right]}{\sqrt{2\pi}} du_2.$$

Complete the square on the bracketed expression:

$$N = \int_c^{\infty} \frac{\exp\left[-\tfrac{1}{2}((u_2 - \sigma\rho)^2 - \sigma^2)\right]}{\sqrt{2\pi}} du_2$$

$$= e^{\sigma^2/2} \int_c^{\infty} \frac{\exp\left[-\tfrac{1}{2}(u_2 - \sigma\rho)^2\right]}{\sqrt{2\pi}} du_2$$

$$= e^{\sigma^2/2} \left[1 - \Phi\,(c - \sigma\rho)\right] .$$

Substituting this for the numerator of (A29) completes the proof. ■

Substituting (A27) into (A26), the unconditional expectation of Y may be rewritten as:

$$E\,Y_i = e^{f(x_i)}\, e^{\sigma_1^2/2}\, \frac{1-\Phi(-h(x_i)-\sigma_1\rho)}{1-\Phi(-h(x_i))}\,(1-\Phi(-h(x_i)) \tag{A31}$$

$$= e^{\sigma_1^2/2}\, e^{f(x_i)}\, \Phi\,(h(x_1) + \sigma_1\rho).$$

B. Elasticity

The elasticity is defined as

$$\epsilon_i = \frac{\partial EY_i}{\partial x_i} \cdot \frac{x_i}{EY_i}. \tag{A32}$$

The partial derivative, from (A31), is:

$$\frac{\partial EY_i}{\partial X_i} = \frac{\partial f}{\partial x_i}\, EY_i + \frac{\partial h}{\partial x_i}\, \frac{\phi(h(x_i)+\sigma_1\rho)}{\Phi(h(x_i)+\sigma_1\rho)}\, EY_i\,.$$

Substituting into (A32) yields

$$\epsilon_i = \left[\frac{\partial f}{\partial x_i} + \frac{\partial h}{\partial x_i} \cdot \lambda_i\right] x_i\,,$$

where λ_i is the inverse Mills ratio:

$$\lambda_i = \frac{\phi(h(x_i)+\sigma_1\rho)}{\Phi(h(x_i)+\sigma_1\rho)}.$$

Appendix E. Identification and Bias in a Simplified Model of Endogenous Taxable Income

While virtually all past studies of capital gains have recognized the endogeneity of tax rates, the corrections for endogeneity have not followed standard econometric procedure. This appendix shows how these unconventional approaches are likely to bias estimates of capital gains responses.[7]

A. Identification

A simple model of capital gains realizations is used to illustrate problems with previous solutions to the endogeneity problem. The model ignores sample selection and the excluded permanent tax rate, which are discussed in the paper. Suppose that the marginal tax rate on capital gains is a non-linear function of the exogenous variables, X, but that it is approximately linear in capital gains, g, and other endogenous income, which is grouped together in one variable, Y. The system of equations may be written as

$$g = \beta_1\tau_t + \beta_2 X_1 + \eta_1 = Z_1\beta + \eta_1 , \tag{A33}$$

$$Y = \gamma_1\tau_t + \gamma_2 X_2 + \eta_2 = Z_2\gamma + \eta_2 , \tag{A34}$$

where $Z_i \equiv [\tau_t \;\; X_i]$, for i=1,2, and

$$\tau_t \equiv f(X,Y,g) \approx \tau_0(X) + \delta \cdot (g + Y) . \tag{A35}$$

Suppose that g and Y are always observable, and that permanent tax rates are observed and are part of the vectors X_1 and X_2, which are subsets of X. The variables in (A33)-(A35) are conformable data matrices with N observations, so g, Y, τ_t, η_1, η_2, and τ_0 are $N\times 1$, X_1 is $N\times k_1$, X_2 is $N\times k_2$, and X is $N\times k$, where $k \geq \max(k_1,k_2)$. The error terms have asymptotic covariance

[7]Technically, we are not examining bias but the difference between the probability limit of alternative estimators and the actual parameter values: the magnitude of the inconsistency. We use the terms "bias" or "asymptotic bias" as convenient shorthand.

$\sigma_{ij} = E(\eta_i \eta_j)$, for i=1,2, and j=1,2, and are uncorrelated with X. The parameter vectors have conforming dimensions. In addition, define a matrix, Z_0 as

$$Z_0 = [\tau_0(X) \ \vdots \ X] \ , \tag{A36}$$

where τ_0 is the first-dollar marginal tax rate.

Suppose also that the moments of the Z matrices converge to positive definite matrices, i.e.,

$$\operatorname*{plim}_{N \to \infty} \frac{Z_i' Z_i}{N} = Q_i \ , \quad \text{for } i = 0,1,2 \ . \tag{A37}$$

Under these conditions, consistent estimation is straightforward. Because $\tau_0(X)$ is a known non-linear function and is correlated with τ_t, but not with η_1 or η_2, τ_0 is a good instrument for τ_t. All of the model's parameters are identifiable for two reasons. First, non-linearity of τ_0 would be sufficient by itself. Second, many of the elements of X that enter the tax calculation (i.e., exogenous factors that determine transitory tax rates) do not enter equations (A33) and (A34), so the parameters of the model could be identified even if tax rates were linear.

The key coefficients, β_1 and β_2, can be estimated by two-stage least squares using τ_0 as an instrument for τ_t. This estimator is analogous in this simple model to the estimator described in the paper.

B. Previous Studies

Instead of two-stage least squares, many past studies have used a proxy variable in place of τ_t. The proxy tax rate was either a "first-dollar" tax rate, which is a marginal tax rate computed by setting realized capital gains, but not other income, equal to zero, or a marginal tax rate computed by setting capital gains equal to an estimated value, conditional on exogenous taxpayer characteristics.

The class of proxy variables used in previous studies may be written as

$$\hat{\tau}_t \equiv f(X,Y,h(X)) \ . \tag{A38}$$

When $h(X)$ equals 0, $\hat{\tau}_t$ is the first-dollar tax rate used in previous studies. When $h(X)$ is a function whose expected value equals actual capital gains, g, $\hat{\tau}_t$ is a "fitted last-dollar tax rate." Because parts

of other income, Y, are endogenous, $\hat{\tau}_t$ is correlated with the error term, η_1, in equation (A33), it follows that such proxies will produce inconsistent estimates regardless of the choice of $h(X)$.[8]

1. *Least squares bias*

The problems in using the proxy variables can be illustrated by examining the naive least squares estimator of the tax rate coefficient in the capital gains equation, β_1. The asymptotic bias in this estimator, $\hat{\beta}_1$, is

$$\operatorname*{plim}_{N\to\infty} \hat{\beta}_1 - \beta_1 = q_1 \operatorname*{plim}_{N\to\infty} \left(\frac{\tau_t' \eta_1}{N} \right), \tag{A39}$$

where q_1 is the top left element of Q_1^{-1} This scalar is positive because Q_1 is positive definite.

The reduced form for τ_t is, from equations (A33)-(A35),

$$\tau_t = \frac{k\tau_0(X)}{\delta} + k(X_1\gamma_1 + X_2\gamma_2) + k\eta_1 + k\eta_2 , \tag{A40}$$

$$\text{where } k \equiv \frac{\delta}{1 - \delta(\beta_1 + \gamma_1)} .$$

The constant factor, k, is positive because δ is non-negative (marginal tax rates are assumed to be a non-decreasing function of income) and β_1 and γ_1 are negative (endogenous income falls with tax rates). Substituting this reduced form into (A39) yields the asymptotic covariance between τ_t and η_1:

[8]There are other problems with such proxies. First, the use of a proxy variable in place of actual tax rates causes coefficient estimates to be biased toward zero. Second, the use of first-dollar tax rates in place of actual tax rates would cause coefficient estimates to be biased away from zero. However, if estimation were otherwise appropriate (i.e., the first-dollar rate were really exogenous, selectivity were properly accounted for, and the model were properly specified), then elasticity estimates based on first-dollar tax rates would be consistent (except for the proxy-variable bias just noted).

$$\operatorname*{plim}_{N\to\infty} \frac{\tau_t' \eta_1}{N} = k(\sigma_{11} + \sigma_{12}) \ . \tag{A41}$$

The bias is non-zero almost surely. If the absolute value of the variance of η_1 exceeds the covariance between the two equations' errors (which seems likely), the bias will be positive (i.e., towards zero).

Figure 1 illustrates this bias for the case of $\sigma_{12}=0$ (or capital gains is the only endogenous source of income). Suppose that two taxpayers were identical, except that they had different unobserved η_1 values. On the figure, the taxpayer with the lower η_1 corresponds to gain function, GG, and taxable income function II. The taxpayer with the high η_1 has higher gains at every marginal tax rate, represented by the shift from GG to G′G′; taxable income is correspondingly higher represented by I′I′. Because the tax schedule is progressive (marginal tax rates are upward sloping with respect to income), the equilibrium tax rate is higher for the taxpayer with higher η_1 than for the other. In this example, equilibrium gains actually increase from G to G' as equilibrium tax rates increase from τ to τ'. A regression line drawn through the two points would slope upward. It is clear that this positive correlation has nothing to do with the behavioral response of taxpayers to capital gains tax changes. In fact, gains are negatively related. Thus, the least squares bias reverses the sign of the relationship in this example.

2. *Estimates based on proxy tax rate*

Most studies of capital gains realization behavior have used a proxy tax rate, such as $\hat{\tau}_t$ defined in (A38). From (A35), the proxy tax rate is

$$f(X,Y,h(X)) = \tau_0(X) + \delta \cdot (h(X) + Y) \ . \tag{A42}$$

Substituting in the reduced form for Y, $\hat{\tau}_t$ may be written as

$$\hat{\tau}_t = \tau_0(X)(1+\gamma_1 k) + \delta[h(X)+\gamma_1 k X_1 \beta_2 + X_2\gamma_2(1+\gamma_1 k)] + \delta k \eta_1 \gamma_1 + \delta(1+\gamma_1 k)\eta_2 \ . \tag{A43}$$

The asymptotic bias of an estimator based on proxies from its "true"[9] value is proportional to

$$\operatorname*{plim}_{N\to\infty} \frac{\hat{\tau}_t' \eta_1}{N} = \delta[\gamma_1 k(\sigma_{11} + \sigma_{12}) + \sigma_{12}] \ . \tag{A44}$$

If σ_{12} is zero (or negative, but small), the asymptotic bias will be negative because γ_1 is assumed to be negative. Thus, using the first-dollar tax rate is likely to bias realization elasticity estimates away from zero causing estimated elasticities to be too large in absolute value.

3. *Average tax rate as measure of permanent tax rate*

As a proxy for the permanent marginal tax rate, Auten and Clotfelter used a 3-year moving average of actual capital gains tax rates, including the actual current year tax rate. This procedure results in biased and inconsistent estimates of permanent tax effects for the same reason that "naive" least squares, discussed above, results in inconsistent estimates of the transitory tax effect. Assuming, as above, that the other regressors are not correlated with the disturbance terms, the Auten-Clotfelter procedure is roughly equivalent to using $\tau_t/3$, $\tau_{t-1}/3$, and $\tau_{t-2}/3$ as regressors. Because the lagged tax rates are assumed to be predetermined, they would be part of X_1, and the analysis of bias simply proceeds substituting $\tau_t/3$ for τ_t in Equations (A33)-(A35). It can be shown that the resulting bias would be proportional to $k(\sigma_{11} + \sigma_{12})/3$. Thus, the bias would have the same sign as the "naive" least squares bias that results from ignoring endogeneity. Under plausible assumptions, the magnitude of the bias would be about a third of the bias in the naive model. Because the least squares bias can be severe, reducing the bias by two-thirds could still result in seriously flawed estimates, even if estimation were otherwise appropriate.

In a long enough panel, if tax law remained constant, the endogeneity bias could be limited or avoided--for example by excluding the current year from tax rate averages. However, with currently

---

[9]The "true" parameter is not β_1 from Equation (1) if a first-dollar tax rate is used (h(X)=0). In that case, the parameter would reflect the relationship between capital gains and the first dollar rate. However, this, by itself, does not bias elasticity estimates. (See footnote 8.)

available data sets, using average tax rates to proxy for permanent effects would result in potentially serious biases because of the endogeneity of current tax rates just discussed and because averages of a few years' tax rates are insufficient to identify the separate effect of permanent from transitory tax changes.

Appendix F. Variation in Tax Rates in the Panel

The central identification problem--estimation of the permanent effect--would disappear if there were no transitory variation in capital gains tax rates. In this case, all differences in tax rates would represent permanent tax effects and econometric evidence from micro data would only reflect the effects of permanent tax changes. The effect of transitory tax changes could not be identified, but it would be irrelevant.

The data, however, show that there is considerable transitory variation. Tax rates for individual taxpayers traced over five years (1979 to 1983) vary substantially, and the variance increases over time, even after controlling for the effect of the major tax legislation in 1981, as illustrated in Figure 2.

To create Figure 2, taxpayers were divided into ten groups that correspond to deciles of the unconditional sample distribution of first-dollar capital gains tax rates in 1979.[10] Each decile group was then followed through 1983 to examine how closely the group's conditional distribution in following years corresponded to the unconditional distribution of first-dollar capital gains tax rates for each year. In the figure, the distribution of each group is represented by its quartiles. A tendency for the conditional quartiles to approach the unconditional quartiles quickly would indicate a high degree of intertemporal variation in first-dollar tax rates for each taxpayer. However, if the conditional distributions remained relatively fixed, this would suggest that most variation in tax rates represents permanent differences among taxpayers.

For each of the ten decile groups, the figure displays three lines that show how the first, second (median), and third quartiles of their conditional distributions are related to the percentiles of the unconditional distribution in each of the five years from 1979 through 1983. For example, consider taxpayers in the first decile of the unconditional distribution in 1979, represented by the left-most panel of Figure 2. By construction, the median for the first decile in 1979 corresponds to percentile 5 of the unconditional distribution; the first and third quartiles of the conditional distribution correspond to percentiles 2 and 7 of the unconditional distribution. However, after five years, the first, second, and third quartiles of the conditional distribution for the first group equaled percentiles 5, 13, and 28 of the unconditional distribution. Similarly, the figure shows that the three conditional quartiles for taxpayers

[10]The sample used to construct Figure 2 only includes taxpayers who realized net positive long-term capital gains at least once between 1979 and 1983. The first-dollar tax rate is the marginal tax rate on the first dollar of long-term capital gains, i.e., computed with capital gains set to zero.

who were in the 5th decile in 1979 equaled percentiles 42, 45, and 47 of the unconditional distribution, but changed to equal percentiles 30, 50, and 63 of the unconditional distribution by 1983.[11]

The graph shows that there was substantial intertemporal variation in first-dollar capital gains tax rates between 1979 and 1983. In all ten decile groups, the conditional distributions noticeably increased in dispersion over the five year period. Further, the tendency of conditional medians to drift toward the unconditional medians shows that taxpayers with low tax rates in 1979 were likely to have had tax rates below their permanent levels; taxpayers with high rates in 1979 were likely to have been experiencing unusually high rates in that year.

While providing strong evidence of transitory volatility in tax rates, Figure 2 also illustrates that there are systematic differences in the permanent tax rates of different taxpayers. The distribution of the bottom five deciles remains below the population distribution for all fives years, and the top five deciles remain above the population distribution. Were tax rates purely random (transitory), the conditional distributions would have equalled the population distribution in 1980 through 1983.

[11]By comparing quartiles of conditional distributions to percentiles of the unconditional distributions, we have deliberately abstracted from intertemporal variations that would have resulted from general shifts in the marginal tax-rate schedule due, for example, to statutory changes in marginal tax rates in 1981 and 1982. As a result, Figure 2 provides a better picture of the degree to which dispersion of marginal tax rates in a cross section sample results from intertemporal variation because general shifts that are not represented by the figure would not result in cross-section variation in tax rates.

Figure 1
Least Squares Endogeneity Bias

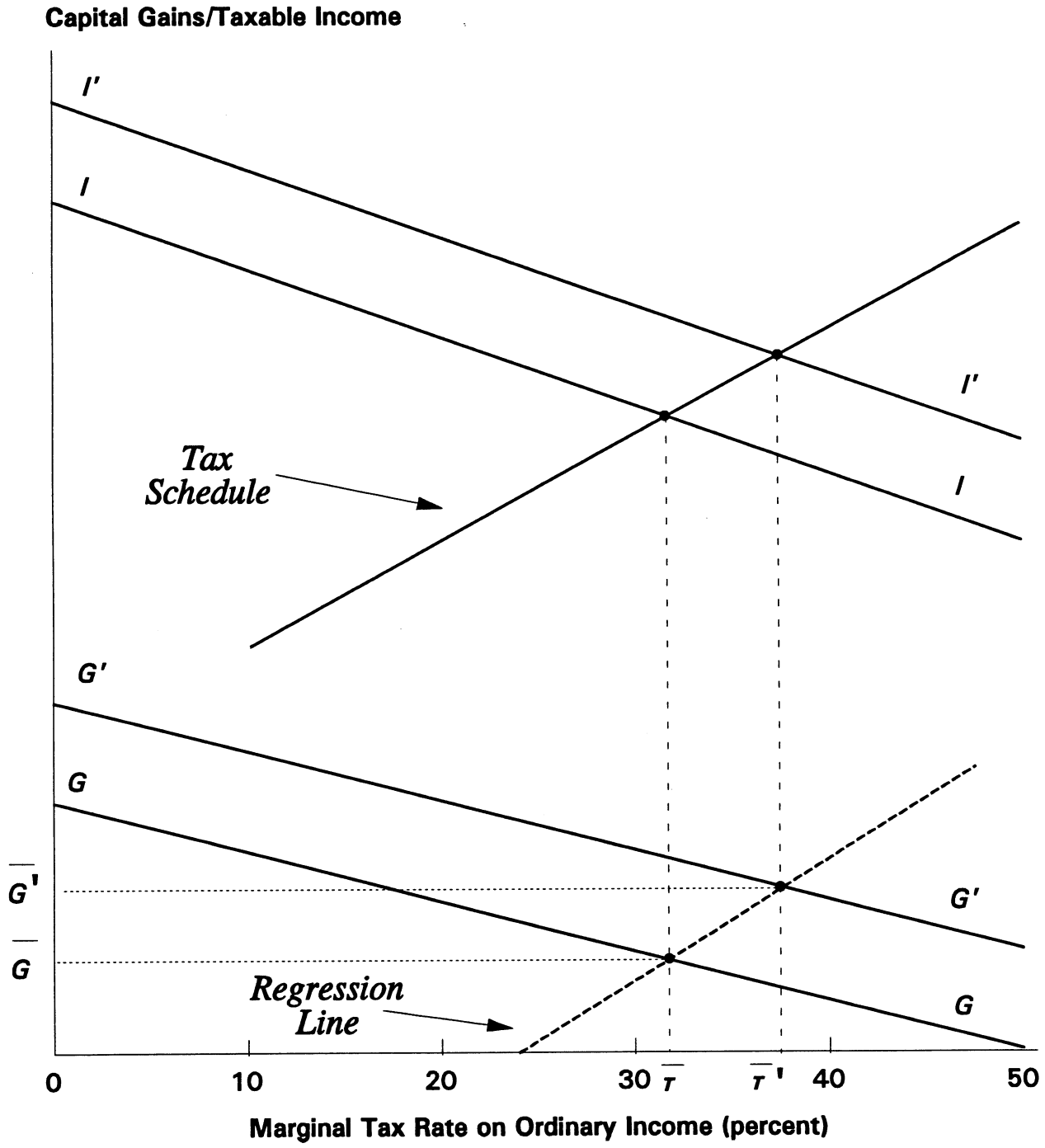
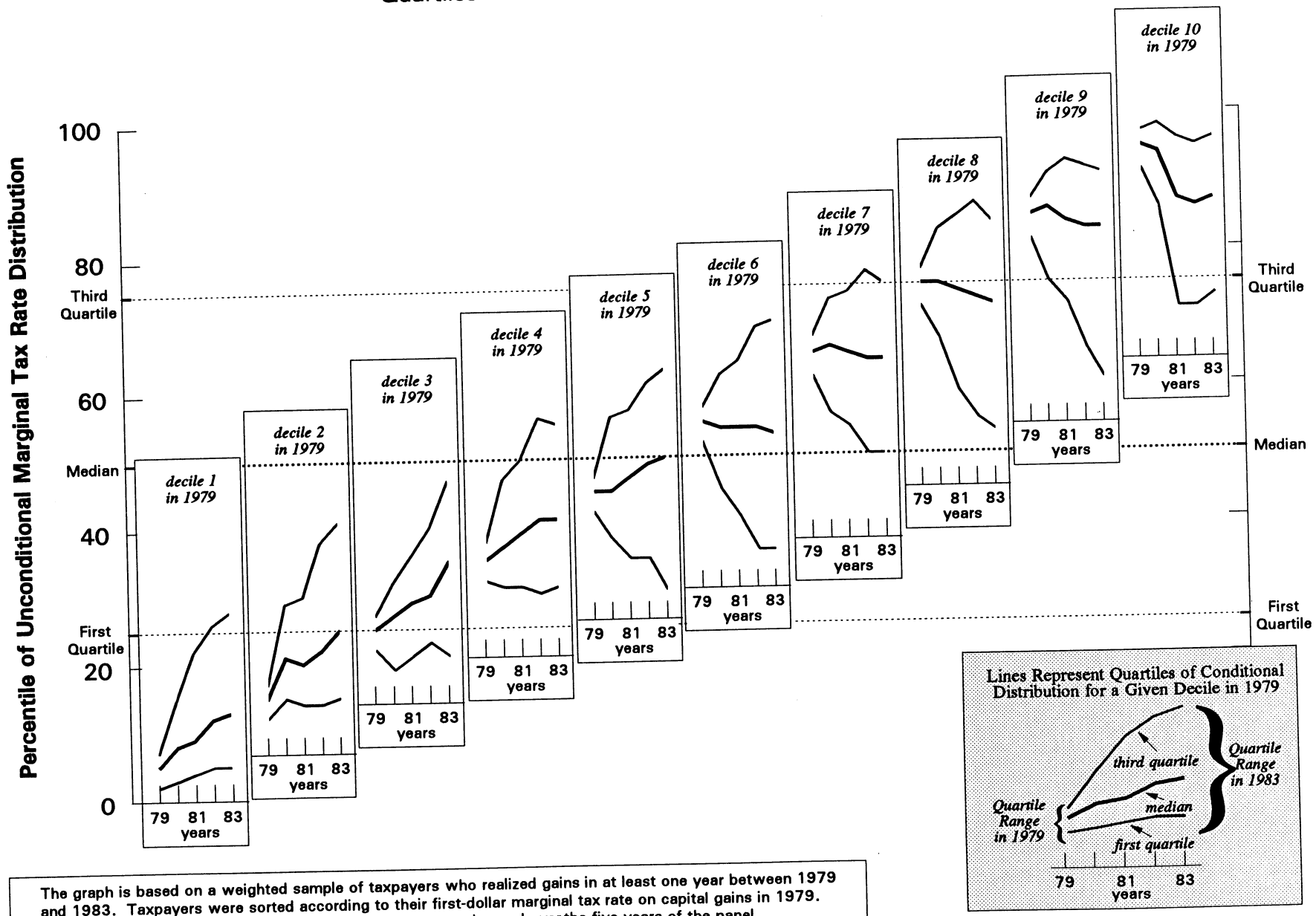


Figure 2. Intertemporal Variation of First-Dollar Marginal Tax Rates on Gains

Quartiles in 1979 to 1983 Conditional on Decile in 1979



The graph is based on a weighted sample of taxpayers who realized gains in at least one year between 1979 and 1983. Taxpayers were sorted according to their first-dollar marginal tax rate on capital gains in 1979. The lines show how the distribution of marginal tax rates changed over the five years of the panel.