An IRET Publication

CAPITAL GAINS SERIES

Sponsored by The Searle Freedom Trust

October 9, 2009
No. 2

THE RELATIONSHIP BETWEEN REALIZED CAPITAL GAINS AND THEIR MARGINAL RATE OF TAXATION, 1976-2004

FOREWORD

The tax rates on long-term capital gains (gains on assets held at least one year) are scheduled to rise in 2011. The current top statutory rate of 15% would rise to 20%. President Obama has suggested at various times raising the top rate further, to as much as 28%. How much additional revenue, if any, should the government expect to get from such increases in the tax rate on capital gains? Previous research has indicated that cutting the tax rate to roughly current levels may actually have raised revenue for the Treasury. If so, hiking the rates could reduce revenue, in addition to hurting growth and job creation.

IRET asked Professor Paul D. Evans of Ohio State University to take a fresh look at how taxpayers might respond to an increase in the capital gains rate, using taxpayer data from the most recent available years. His study, *The Relationship Between Realized Capital Gains and Their Marginal Rate of Taxation, 1976-2004*, is presented below.

Evans finds that taxpayers are still sensitive to the tax rate on capital gains, and would report fewer gains if the rate were raised. Based on 2004 data, the revenue maximizing tax rate may be just under 10%. Raising the capital gains tax rate from the current 15% to 20% or more would reduce federal capital gains tax revenue. Additional revenue would be lost from other parts of the income tax and from other federal taxes due to reduced investment, employment, and income. The optimal capital gains tax rate to maximize public welfare, and to help the federal budget, is surely closer to if not zero.

Background

The tax rates on long-term capital gains (gains on assets held at least one year) are scheduled to rise in 2011. They were reduced in the Jobs and Growth Tax Relief Reconciliation Act of 2003, with the cuts effective through 2008. The reductions were extended through 2010 in the Tax Increase Prevention and Reconciliation Act of 2005. (The top tax rate on dividends was also cut to 15% to match the top rate on capital gains. It too will increase in 2011, reverting to ordinary income tax rates.)

Before the 2003 reduction, the top tax rate on long-term capital gains was 20% for taxpayers in the top four tax brackets and 10% for taxpayers in the 10% and 15% ordinary income brackets. (However, for gains on assets held five years or more, the top rate was only 18% in the top four brackets and 8% in the 10% and 15% brackets.)

In 2003, the top rate of 20% (or 18%) was cut to 15% for taxpayers in the top four tax brackets. The capital gains rate for taxpayers in the 10% and 15% brackets was reduced from 10% (or 8%) to 5% through 2007, then to zero for 2008-2010.

Under current law, in 2011, the 15% tax rate on long term gains will revert to 20% (or 18% for 5-year holdings) for the top four brackets. The zero rate will revert to 10% (or 8% for gains held 5 years or more) in 2011. The 10% (or 8%) rate will apply to taxpayers in the 15% ordinary income tax bracket. (The 10% tax bracket, which was carved out of the 15% bracket by the Economic Growth and Tax Relief Reconciliation Act of 2001, will expire and be reincorporated in the 15% bracket in 2011).

(Note: these rates exclude the effects of the Alternative Minimum Tax. For taxpayers in the phase-out range of the AMT exemption, an additional dollar of capital gains reduces their AMT exemption by $0.25, and subjects that additional amount of wages, salaries, and interest income to additional AMT. The combined effective marginal tax rate due to the additional capital gains is then 21.5% or 22% in the 26% and 28% AMT rate brackets (15% plus a quarter of 26% or 28%). If the capital gains rate were to rise, the new AMT-impacted rates would be 6.5% or 7% above whatever the new rate is set at.)

Areas of previous research

Changes in the capital gains tax rate affect taxpayer behavior, the after-tax returns on investment, capital formation, employment, and income. These impacts on behavior affect federal revenues from the capital gains tax itself (the direct "own tax" revenue effect), and revenues from all other taxes imposed on economic activity (indirect revenue effects).

There were many studies between 1980 and 2000 of the relationship between the capital gains tax rate and the amount of gains that taxpayers choose to realize. These choices affect how much capital gains tax revenue would be raised at various tax rates. If a higher tax rate results in fewer gains, then the rate hike may yield less revenue than expected, or even result in a revenue reduction. Some of the studies explicitly try to answer the question, "What tax rate on capital gains would raise the most capital gains revenue for the government?" (This is not the optimal tax rate for the economy, but it is a question often raised by government officials.)

The early studies include several Treasury efforts in the mid-1980s to gauge the effect of the capital gains tax rate reductions enacted in 1978 (a cut from 35% to 28%) and in 1981 (a cut from 28% to 20%). Other effective rate changes in 1986 (to 28%) and 1997 (to 20%) gave researchers additional data to examine. With the passage of time, and further changes in the tax law, there are now several additional years of data providing more information on how people react to changes in the taxation of gains.

These studies have explored the direct effect of changes in the tax rate on the quantity of reported gains in taxable income and the resulting income tax revenues from the tax on the gains. There have been far fewer studies of the effect on economic activity, income, and revenues from other parts of the income tax and from other taxes.

One reason that tax analysts' studies focus on the relationship between the tax rate and realizations behavior is that this relationship is one that revenue estimators for the federal budget process are willing to take into account (at least in part). By contrast, the larger effect of capital gains tax rates on the performance of the economy and on other tax revenues is one that the budget estimation protocols ignore.

Estimators are willing to take "micro-economic" behavior changes into account in the budget process (such as a reduction in smoking or gasoline consumption in the event of an increase in the excise taxes on these products, or a change in the timing of the taking of capital gains if that rate is altered). However, they assume that the excise or capital gains tax rate changes do not alter the baseline forecast of the level of total economic output, incomes, prices, and the quantity and value of capital assets for the whole economy. This eliminates the possibility that the indirect changes in other taxes that are sensitive to the level of economic activity would enter into the revenue estimate of the effect of a change in the capital gains tax rate.

Although studies that focus only on realizations and the own-tax effect tell only part of the story, the issue is important. Congress relies on these revenue estimates in deciding whether or not to raise or lower the capital gains tax rate. If the realizations effect is significant, and tells a portion of the real story, then it can nudge Congress toward adoption of a more appropriate level of tax.

Summary of the Evans paper

There are several types of studies that deal with the realizations effect, including cross section analysis, time series analysis, and panel studies over a limited time period. Each has its own strengths and weaknesses. IRET asked Professor Evans to update the cross section literature of the 1980s and 1990s by extending that analysis to taxpayer data from the most recent available years. The chief findings in his paper are:

- Taxpayers continue to exhibit significant sensitivity to the tax rate on capital gains in deciding how much of their gains to realize over time. Behavior in recent years is broadly similar to that found in earlier studies. The sensitivity to the tax rate is not a one-time fluke related to a few large tax changes in the more distant past.

- This sensitivity appears to be more than a short-run timing issue (moving gains from this year to next year if the tax rate is scheduled to fall, or from next year to this year if it is scheduled to rise). It also appears to be more than an "unlocking effect" (the sudden taking, after a rate reduction, of accumulated gains that have been pent up over time to avoid the old, higher tax rate).

- At current tax rates, a 1 percentage point reduction in the marginal tax rates on capital gains might trigger a 10.32 percent increase in realized capital gains.

- The capital gains tax rate structure that would have brought in the most capital gains revenue in 2004 appears to be a flat rate of 9.69 percent, or just under 10 percent. More precisely, a flat rate imposed at 9.69 percent on both the first and last dollars of individuals' capital gains (raising rates below that level, and lowering rates above it) would appear to bring in the most capital gains revenue. In 2004, the average tax rate on capital gains was close to the own-tax revenue maximizing rate, but the higher marginal rates applied to incremental gains discouraged realizations and held down revenue. Converting it to a flat rate of 10% (or just under) would raise more revenue than the current structure.

- The longer term revenue effect of the scheduled 2011 rate hike from 15% to 20% is likely to a permanent revenue loss, extending beyond the (larger) initial decrease one might expect due to timing behavior. Higher revenues from increasing the zero rate would likely be more than offset by reduced revenues from raising the higher rates. By extension, the revenue loss (and economic damage) would rise further if the rates were increased further to 24% or 28%.

- The optimal marginal tax rate on capital gains is lower than the capital gains revenue maximizing rate, however, because taxation leads to economic inefficiencies that reduce the welfare and income of the public. The lower income reduces government tax revenue from other sources. A cap well below 9.69 percent may therefore have been in order.

Evans is careful to admit the limitations of cross-section analysis (as earlier authors have done). It is indicative, not definitive. Of necessity, it omits taxpayer characteristics that may affect realization decisions, but are not recorded on tax returns and therefore do not appear in the public use tax sample data.

Evans extensively tested the types of information presented on the tax returns to determine the optimal set of variables and the optimal structural form to help predict the taking and reporting of capital gains. The extensive testing was carried out using 2004 data. Once the optimal structure was determined, estimates were made for all tax years from 1976 through 2004.

The estimated sensitivity varied fairly widely across the years. Estimates for the most recent less volatile years appear to fall in the mid-range of the historical variation, and may be indicative of some basic relationship to the tax rate.

The public use files are random annual samples that, unlike panels, do not track the same people over time. Each year's sample is a different group of people, whose reactions may differ from one group to the next (even in as large a sample – over 200,000 returns – as the Treasury provides). Therefore, it is desirable for the Treasury to continue, and to expand, its panel studies of taxpayer behavior over periods of time. (However, there are limits to what can be done in that regard, as more people in the panels may die over time as the time frame is lengthened, or alter their asset holdings and sales behavior as they age.)

Cross-section annual studies are not time series studies. They cannot directly measure changes in behavior immediately before or after changes in the tax rate. Evans points out that all cross section studies must therefore employ techniques to try to separate short-run timing and unlocking effects from longer term behaviors. He describes the usual methods in the literature and explains which approach he adapted for this paper. He notes that no one can be certain how successful such techniques are at purging timing influences. (Note that time series studies of the aggregate capital gains realizations of the total taxpaying population have their own problems, including a very limited number of annual observations, and the fact that people taking capital gains in one year are not necessarily the same as the people taking gains the next.)

I believe the Evans paper provides important verification of the utility of the cross-section studies. It suggests that the scheduled capital gain tax increase in 2011 be reconsidered. That rate increase may well lose revenue, in addition to hurting employment and income. If there were to be a revenue gain, it would be far smaller than the static score would indicate. It would adversely affect economic activity, and hurt federal revenue from other sources. (The same can be said about the scheduled increase in the tax rate on dividends.)

Clearly, this is an important area for further research. It would be beneficial for the Treasury to integrate the results of cross section, panel, or time series studies with a dynamic general equilibrium model of the economy that could add the consequences for capital formation, income, and total federal revenue. Some progress in such modeling was made at the time of the President's Advisory Panel on Federal Tax Reform in 2005. (See the Panel's report: *Simple, Fair, and Pro-Growth: Proposals to Fix America's Tax System*.[1]) The Treasury modeling effort is still in a developmental state, and needs to be improved and tested against historical experience.

Stephen J. Entin
President and Executive Director

[1] President's Advisory Panel on Tax Reform, *Simple, Fair, and Pro-Growth: Proposals to Fix America's Tax System*, November, 2005, available on the internet at www.taxreformpanel.gov.

The Relationship Between Realized Capital Gains and Their Marginal Rate of Taxation, 1976-2004

Paul Evans*

October 9, 2009

---

*Professor of Economics, Ohio State University, 410 Arps Hall, 1945 N High ST, Columbus OH 43210-1172; evans.21@osu.edu. The research was supported by the Institute for Research on the Economics of Taxation. I wish to thank Steve Entin and Gary Robbins, without whom this paper could not have been written.

1. Introduction

This paper investigates the relationship between realized capital gains and the marginal rate at which they are taxed. It is far from being the first paper to do so. Lindsey (1986); Darby, Gillingham and Greenlees (1988); Jones; and Eichner and Sinai (2000) investigated the relationship between the aggregate level of realized capital gains and other aggregate variables, including aggregate and statutory measures of the marginal tax rate on realized capital gains. Feldstein, Slemrod and Yitzhaki (1980, 1984); Minarik (1984); and Gillingham, Greenlees and Zieschang (1989) investigated the relationship on cross-sectional data. Auten and Clotfelter (1982), *Report to Congress* (1985); Auten, Burman and Randolph (1989); Slemrod (1990); Bogart and Gentry (1993); and Burman and Randolph (1994) investigated the relationship on panels of longitudinal data. Their findings are decidedly mixed.

This paper investigates the relationship using cross-sectional data for each year between 1976 and 2004. Not having panel data is a significant handicap in the empirical analysis, but the long time span of the data is a mitigating factor. It permits the relationship to be estimated in several widely different tax regimes: a period of 50 percent exclusion with a maximum tax rate of 35 percent; a period of 60 percent exclusion with a maximum tax rate of 28 percent; a period of 60 percent exclusion with a maximum tax rate of 20 percent; a period of no exclusion with a maximum tax rate of 28 percent; and a period of no exclusion with a maximum tax rate of 15 percent. It is hoped that the large differences in tax regime provide a useful testing ground for the econometric specifications employed.

One novel feature of the empirical analysis is its construction and use of proxies for the timing effects that have clouded the interpretation of previous purely cross-sectional analysis. These proxies may enable the identification of the long-run effects of the marginal tax rate and not a mixture of long-run and purely transitory effects.

The rest of the paper is divided into four sections. Section 2 discusses the guidance that theory can provide in specifying the relationship between the realization of capital gains and the marginal tax rates to which it is subject. It also lays out the framework that guides the empirical work. Section 3 describes the data used in the empirical analysis, the limitations that it imposes, and the econometric framework. Section 4 presents the empirical findings of the paper, which suggest that the realized capital gains respond strongly to the marginal rates at which they are taxed. Finally, section 5 concludes with a brief discussion of the implications of the paper for tax policy.

2. Theory

Economic theory can provide powerful insights into how our economy works, especially when economic frictions can be largely abstracted from. Unfortunately, the decisions of whether to realize capital gains and if so, how much are very much affected by frictions.

To see why, consider the theoretical result of Stiglitz (1983). He laid out a model in which rational households could trade in completely frictionless financial markets, showing that they would never voluntarily realize capital gains and hence would pay no capital-gains taxes. In reality, taxes on capital gains raise appreciable revenue, indicating that economic frictions must play an important role. We therefore confine ourselves to a minimalist framework, assuming merely that the decision to realize at least some capital gains and the decision on how much to realize take the form:

$$R = \mathrm{R}(T^*, F^*, B^*, Y, X) \tag{1}$$

and

$$G = \mathrm{G}(T, F, B, Y, Z)\,. \tag{2}$$

These equations employ the following notation:

- R is a dummy variable that is one or zero according as the household chooses to realize capital gains or not;
- G is the amount of capital gains the household chooses to realize if $R = 1$;
- T is the household's last-dollar marginal tax rate on capital gains in the current year;
- F is how much the household expects its last-dollar marginal tax rate on capital gains in the following year to exceed T;

- B is how much the household's last-dollar marginal tax rate on capital gains in the previous year exceeds T;
- Y is the household's income;
- T^* is the first-dollar analogue of T;
- F^* is the first-dollar analogue of F;
- B^* is the first-dollar analogue of B;
- X is a set of other variables that affect whether the household realizes capital gains in the current year;
- Z is a set of other variables that affect how large G is, given that $R = 1$; and
- R(•) and G(•) are functions.

It is natural to model the decision of whether to realize a capital gain as occurring if, and only if, the net after-tax benefit of realizing the first dollar of capital gains exceeds some threshold. For this reason, the marginal tax rate on the first dollar of capital gains is more relevant to the decision than the marginal tax rate on the last dollar. The choice of whether to realize a capital gain in the current year is also affected by how much more the household can benefit after-tax from doing so next year rather than now. In the previous year, the household was then making a decision about whether to realize a net capital gain on the basis of its expectation of this year's B^*. This decision affects the amount of accrued capital gains available to be realized in the current year and hence affects the realization decision itself.

Given that the household has chosen $R = 1$, its decision on G equates its marginal after-tax benefit from realizing capital gains to the marginal after-tax cost of doing so. For this reason, only marginal tax rates on the last dollar of realized capital gains are relevant for how large G is; hence, the presence of T , F and B in equation (2).

Households realize more capital gains, the greater their incomes. The primary reason is simply that households with higher current incomes generally are wealthier, own more stock and have accumulated more unrealized capital gains. Ideally, we should like to model G as depending on the household's stock of unrealized capital gains. Unfortunately, the requisite data are unavailable, limiting the analysis to some measure of the household's income.

In including F and B in equation (2) and $F*$ and $B*$ in equation (1), we are attempting to control for the possibility that changes in capital-gains taxation generate large shifts in the timing and much smaller permanent changes in the level of realized capital gains. This attempt is imperfect, however, because we do not have data on the marginal tax rates actually faced by the households in the next and previous years. Rather, these variables measure the tax rates that the households would have faced in those years, had their situations been identical to their current situations.

3. The Data and Econometric Issues

Each year the Internal Revenue Service prepares a large public-use dataset from the income-tax returns filed that year. The IRS constructs these datasets using stratified random sampling of the taxpaying units—hereafter, households. The stratification is designed to greatly oversample high-income households, but the IRS provides the weights used in the sampling, enabling the estimates to be adjusted for this source of nonrandomness.

The empirical analysis employs the datasets for the years between 1976 and 2004, the most recent year for which data were available. We work with data on the variables reported in Table 1. Per force, our method must be cross-sectional since a different set of households is sampled every year. As a result, we cannot control for many effects that might bias our findings. These biases may arise because all of the data stem from decisions that households either are making or have previously made. These decisions in turn depend on unobserved characteristics of the households, characteristics that also affect whether households realize capital gains and if so, how much. Many of these characteristics may be highly persistent and hence may even have affected past decisions. Of necessity, then, we shall be relating the endogenous variables R and G

Table 1. Glossary of Symbols and their Definitions

Symbol	Definition
R	Dummy variable for whether the household realized capital gains
G	Long-term capital gains realized minus short-term capital losses realized
Y	Adjusted gross income
T	Marginal tax rate on the last dollar of capital gains
F	Excess over *T* that an identically situated household would expect to pay the next year on its last dollar of capital gains
B	Excess over *T* that an identically situated household would have paid in the previous year on its last dollar of capital gains
*T**	First-dollar analogue to *T*
*F**	First-dollar analogue to *F*
*B**	First-dollar analogue to *B*
*X*1, *Z*1	Dummy variable for whether the household itemized
*X*2, *Z*2	Dummy variable for whether the household is married
*X*3, *Z*3	Dummy variable for whether the household claimed children as exemptions
*X*4, *Z*4	Dummy variable for whether the household collects social security benefits
*X*5, *Z*5	Dummy variable for whether the household has positive wage income
*X*6, *Z*6	Dummy variable for whether the household has positive interest income
*X*7, *Z*7	Dummy variable for whether the household has positive dividend income
*X*8, *Z*8	Dummy variable for whether the household has positive Form-C income
*X*9, *Z*9	Dummy variable for whether the household has positive Form-E income
*X*10, *Z*10	Dummy variable for whether the household has positive Form-F income
*X*11, *Z*11	Dummy variable for whether the household has negative Form-C income
*X*12, *Z*12	Dummy variable for whether the household has negative Form-E income
*X*13, *Z*13	Dummy variable for whether the household has negative Form-F income
*X*14, *Z*14	Dummy variable for whether the household has tax-exempt interest income
*X*15, *Z*15	Dummy variable for whether the amount of tax due exceeded 0.1*Y*, yet no penalty was incurred
*Z*16	Dummy variable for whether the household paid the alternative minimum tax
*Z*17	Ratio of wage income to *Y*
*Z*18	Ratio of interest income to *Y*
Z19	Ratio of dividend income to *Y*
*Z*20	Ratio of Form-C income to *Y*
*Z*21	Ratio of Form-E income to *Y*
*Z*22	Ratio of Form-F income to *Y*
*Z*23	Ratio of the alternative minimum tax to *Y*
*Z*24	Ratio of the amount deducted to *Y*

to other variables that may well be endogenous.[1] No magic bullet can solve this problem. We shall therefore end up with estimates subject to biases of unknown size and direction. The best we can do is to estimate how *T* affects *G* using many methods in order to gauge the sensitivity of our estimates to the method used. Some credence may then attach to our estimates if they prove to be sufficiently robust.

We consider three basic specifications. The simplest specification estimates equation (2) with weighted least squares,[2] ignoring that every regressor may be endogenous and also ignoring that those who realize capital gains are a self-selected sample of the households. We estimate regressions of the form[3]

$$\log G_i = a + bT_i + cF_i + dB_i + f \log Y_i + g_1 Z_{1i} + g_2 Z_{2i} + g_3 Z_{3i} + \cdots + e_i. \tag{3}$$

In equation (3), *a* is an intercept; *b*, *c*, *d*, *f* and the *g*s are regression coefficients; the subscript *i* indicates that the *i*th household is under consideration; and e_i is the error term. The regression coefficients capture the marginal effects of the included regressors, and the error term captures the effects of all other influences on realized capital gains. The specification is linear in the logarithms of *G* and *Y* and linear in the tax rate *T* as well as the in the other regressors. As a result, the regression coefficient *b* is the semielasticity of response of the capital gains to the last-dollar marginal tax rate, and *f* is the income elasticity. Because we are primarily interested in what *b* is, we shall not report the estimates of other regression coefficients.

In our second basic specification, we estimate equation (3) with weighted two-stage least squares rather than ordinary least squares, assuming that only a few of the regressors are endogenous but again ignoring the self-selection of household in the sample. We allow first for the endogeneity of *T*, *F* and *B*. We then also allow for the endogeneity of income. To control for

---

[1] In the text, we distinguish between *jointly determined* on the one hand and *endogenous* and *exogenous* on the other hand. In our usage, endogenous variables are correlated with the error term in the regression, exogenous variables are not, and jointly-determined variables may be either endogenous or exogenous. Generally speaking, however, econometricians presume jointly-determined variables to be endogenous, absent a compelling reason for assuming them to be exogenous.

[2] The population weights provided in the dataset are used as the weights in estimation.

[3] We follow the convention in technical disciplines of using log to indicate a natural logarithm.

endogeneity, however, requires that some other variables must be assumed to be exogenous. These other variables, however, are also jointly determined with the amount of capital gains realized. We must therefore make assumptions about exogeneity, assumptions that at best are only somewhat more justified than the assumption that the regressors themselves are exogenous.

In our third basic specification, we employ the Heckit method of controlling for the self-selection of the households in the sample.[4] In this approach, each household decides to realize capital gains if, and only if, some criterion

$$C_i = hT_i^* + jF_i^* + kF_i^* + m_1 X_{1i} + m_2 X_{2i} + m_3 X_{3i} + \cdots + u_i > 0 \tag{4}$$

is satisfied, where *h*, *j*, *k* and the *m*s are unknown coefficients that are to be estimated. In equation (4), C_i can be interpreted as the net benefit that the household can receive from realizing capital gains, which is assumed to depend on T^*, F^*, B^*, a set of other observable variables X, and an error term u_i that captures the household's unobservable characteristics. The biases that may result from estimating equation (3) by either weighted least squares or weighted two-stage least squares stem from the possibility that the unobservable influences incorporated in u_i may also be incorporated in the error term e_i. James Heckman (1976) showed that under some fairly weak assumptions, these biases can be eliminated by first estimating a probit model for the dummy variable that is one if the household is in the sample and zero otherwise. Equation (3) is then augmented with the inverse Mills ratio S_i from the probit model to give

$$\log G_i = a + bT_i + cF_i + dB_i + f \log Y_i + + g_1 Z_{1i} + g_2 Z_{2i} + g_3 Z_{3i} + \cdots + pS_i + v_i, \tag{5}$$

and the resulting regression equation is estimated with whatever method would be appropriate in the absence of self-selection. In equation (5), *p* measures the effect of the sample selection on the amount of capital gains realized, v_i is an error term purged of the effects of the self-selection, and the other regression coefficients continue to have the same interpretation as in equation (3).

[4] The Tobit method is inappropriate because it assumes that there is a notional amount capital gains that is observable as the actual amount when it is nonnegative but is censored at zero when it is negative. Much more plausible is that households choose not to realize capital gains for reasons that equation (3) leaves unexplained.

4. Estimates

We begin with a preliminary econometric analysis of the dataset for 2004. We employ this dataset because it is the most recent one available and because it yields estimates that are fairly representative of those obtained for the other years. We have two objectives: to assess how sensitive the estimates are to the econometric method used; and to identify an appropriate econometric model to estimate for each of the other years.

We begin with weighted least squares, varying the regressors included in equation (3). Table 2 reports the resulting estimates. When only the marginal tax rate is included (Regression 1), we obtain an appreciably positive estimate for the semielasticity. This estimate is seriously biased, however, because the marginal tax rate is increasing in income, which positively affects the amount of capital gains realized. This finding points to the importance of controlling for income. When we augment the regression with the logarithm of income (Regression 2), we obtain an appreciably negative and highly significant estimate for the semielasticity. Next, in Regression 3, we control for *Z*1-*Z*4. The semielasticity roughly halves in magnitude, but remains appreciably negative and highly statistically significant. Then in Regression 4, we add the control variables *Z*5-*Z*16, each of which is a dummy variable indicating that the household received a type of income. We obtain an estimated semielasticity of -15.44, more than twice as negative as in Regression 3, which controls for only four variables. Finally, in Regression 5, we add the control variables *Z*17-*Z*24, each of which is a ratio of a type of income to adjusted gross income. Note that because the logarithm of *Y* is defined only for positive values of *Y*, these ratios always have positive denominators. Our estimated semielasticity of -15.25 hardly changes from its value in Regression 4, suggesting that once more than a few variables are controlled for, the estimated semielasticity is not very sensitive to what *Z*s are included in the regression. Because controlling for as many variables as possible is generally a good idea, we include in the regressions all available *Z*s[5]

[5] Some of these controls are not available for the entire period. *Z*14 and *X*14 are not available before 1987, *Z*15 and *X*15 are not available before 1986, *X*4 and *Z*4 are not available before 1985, and *Z*16 and *Z*23 are not available before 1979.

Table 2. Least Squares Estimates of the Semielasticity of Realized Capital Gains with Respect to Their Marginal Tax Rates for 2004

Regression Number	Estimate (Standard Error)	Description[*]
1	9.67 (0.35)	Least squares with no controls
2	-11.63 (0.53)	Least squares with only log*Y* as a control variable
3	-6.25 (0.58)	Least squares with log*Y* and *Z*1-*Z*4 as control variables
4	-15.44 (0.56)	Least squares with log*Y* and *Z*1-*Z*16 as control variables
5	-15.25 (0.55)	Least squares with log*Y* and *Z*1-*Z*24 as control variables
6	-15.33 (0.55)	Least squares with log*Y*, replacing *T* with *TAVE*
7	-15.34 (0.55)	Least squares with log*Y*, *F*, *B* and *Z*1-*Z*24 as the control variables
8	-11.64 (0.51)	*LYPOS*, *LYNEG*, *F*, *B* and *Z*1-*Z*24 as the control variables

[*]High-income taxpaying units were oversampled in the sample. To obtain estimates representative of the entire population of households, observations received the population weights in estimation. The standard errors are robust to heteroscedasticity.

Regression 6 attempts to estimate the long-run effect of the marginal tax rate on the amount of capital gains realized. Specifically, following Auten and Clotfelder (1982), we include in Regression 6 $TAVE_i$, the average of T_i and the values of it that would have prevailed in the adjacent years, had the household's situation in those two years been identical to its situation in the current year. The estimated semielasticity barely budges from its value in Regression 5. Note, however, that this average tax rate is an artificial construct: Unlike Auten and Clotfelder, we do not have the households' actual marginal tax rates, as we have no way of following them over time.[6]

[6] Auten and Clotfelter actually used the average marginal tax rate of the current year and two previous years, following the common practice at the time of using current and past values as proxies for current and expected future values.

Regression 7 pursues an alternative tack by including *F* and *B* in the regression. These variables are the differences between our proxies for the marginal tax rates in the adjacent years and the marginal tax rate in the current year. Doing so barely changes the estimated semielasticity from its value in Regression 5, suggesting that timing does not generate large effects.[7] The same caveat applies: Our method of controlling for timing may not have succeeded.

Regression 8 avoids excluding households with negative incomes from the sample. Instead of including just log*Y*, it also includes the two regressors *LYPOS* and *LYNEG*, which are defined to be log*Y* or log(-*Y*) according as *Y* is positive or negative and zero otherwise. The effect is to reduce the magnitude of the estimated semielasticity somewhat, from -15.25 to -11.64.

We now turn to the two-stage least squares estimates reported in Table 3. Last-dollar measures of marginal tax rates may be endogenous because they can depend on how large the household's realized capital gain is, as the additional income may have pushed the household into a higher tax bracket. Following Feldstein, Slemrod and Yitzhaki (1980), the most common approach to this problem has been to employ first-dollar measures as instrumental variables for the last-dollar measures. We point out, however, that the first-dollar measures are also jointly determined variables and hence may be just as endogenous as the last-dollar measures. With that caveat in mind, we now turn to Regression 9, which treats *T*, *F* and *B* as endogenous and treats their first-dollar analogues *T**, *F** and *B** as exogenous while continuing to maintain that log*Y* is exogenous. Interestingly, the estimated semielasticity does not even change.

Y would seem to be endogenous because it includes capital gains. Our instrumental variables for it consist of the logarithms of six components of *Y* when the component is positive and zero otherwise. Three are straightforward to define: the logarithms of wages, interest and dividends if the household has any and zero otherwise. The other six must account for the possibility that the incomes reported on Forms C, E and F can be negative. Specifically, three of the measures are the logarithms of the incomes reported on Forms C, D and F if positive and zero otherwise, while the other three are the logarithms of the negatives of these incomes if negative and zero

[7] The presence of large effects of timing is one of the two competing explanations for the large swings that have occasionally occurred in the realization of capital gains; see Auten and Clotfelter, Auten, Burman and Randolph, Slemrod and Burman and Randolph for further discussion of this point.

otherwise. Regression 10 uses these instrumental variables as well *T**, *F** and *B** in two-stage least-squares estimation of equation (3). We find that the estimated semielasticity becomes slightly more negative (-15.78) than in Regression 9 (-15.34). Again, a caveat is in order: These instrumental variables are just as jointly determined as adjusted gross income.

Table 3. Two-Stage Least Squares and Heckit Estimates Of the Semielasticity of Net Capital-Gains with Respect to Their Marginal Tax Rates for 2004

Regression Number	Estimate (Standard Error)	Description[*]
9	-15.34 (0.55)	Two-stage least squares using *T**, *F** and *B* as the instrumental variables
10	-15.78 (0.65)	Two-stage least squares using *T**, *F**, *B*, and the logarithms of nine components of income as the instrumental variables
11	-17.51 (0.94)	Same as Regression 8 except the Heckit procedure controlled for self-selection
12	-10.76 (1.56)	Same as Regression 9 except the Heckit procedure controlled for self-selection
13	-38.58 (6.63)	Same as Regression 10 except the Heckit procedure controlled for self-selection
14	-11.23 (1.55)	Same as Regression 11 except that TAVE replaces *T*, *F* and *B* and the first-dollar analogue of *TAVE* replaces *T**, *F** and *B**
15	-21.80 (1.18)	Same as Regression 13 except that different criterion was used for self-selection

[a]All regressions control for the variables *Z*1-*Z*24 as well as log*Y*; see Table 1 for their definitions. Population weights were used in estimation, and the standard errors are robust to heteroscedasticity.

We now turn to our Heckit estimates. In Regressions 11-14, we first estimate a probit model for the dummy variable that is one if the household realized capital gains and had a positive adjusted gross income and zero otherwise. We include the variables *T**, *F**, *B** as well as *X*1-*X*15; see Table 1 for their definitions. The inverse Mills ratio from our estimated probit model was then used as a regressor in equation (5). Regressions 11-13 redo Regressions 8-10 while controlling for self-selection; the results are, respectively, to increase the magnitude of estimated semielasticity (-17.51 vs. -11.64), to decrease its magnitude (-10.62 vs. -15.34), and to greatly increase its magnitude (-38.58 vs. -15.78). Regression 14 redoes Regression 12 with *T*, *F* and *B* replaced by *TAVE* and the instrumental variables *T**, *F** and *B** replaced by the first-dollar analogue of *TAVE*. The magnitude of the estimated semielasticity increases slightly from -10.62 to -11.23. Regression 15 is estimated using a different selection criterion, that the household not only has positive realized capital gains and a positive adjusted gross income but also receives dividends and hence owns stock. The effect of employing this more selective criterion is to appreciably increase the magnitude of the estimated semielasticity from -10.62 to -21.80.[8]

The amount of capital gains realized clearly depends on income, but that dependence is difficult to estimate without serious bias. Moreover, the estimates of the semielasticity in which we are interested are often sensitive to what the income elasticity is estimated to be. We illustrate this fact with Table 4. The estimated semielasticities range from -7.63 for an income elasticity of 1.0 to -29.78 for an income elasticity of 1.6. On the presumption that income elasticity exceeds one,[9] a value of the semielasticity near -10 is a conservative estimate.

Table 4. 2004 Estimates of the Semielasticity of Net Capital-Gains With Respect to Their Marginal Tax Rates For Four Imposed Values of the Income Elasticity

Regression Number	Estimate (Standard Error)	Description[*]
16	-7.63 (1.48)	Same as Regression 12 except that an income elasticity of 1.0 is imposed
17	-15.02 (1.77)	Same as Regression 12 except that an income elasticity of 1.2 is imposed
18	-22.40 (2.51)	Same as Regression 12 except that an income elasticity of 1.4 is imposed
19	-29.78 (3.42)	Same as Regression 12 except that an income elasticity of 1.6 is imposed

[a]All regressions control for log*Y* and *Z*1-*Z*24. Population weights were used in estimation, and the standard errors are robust to heteroscedasticity.

---

[8] The dummy variable for having positive dividend income is excluded from the criterion and the regression equation.

[9] Burman and Ricoy (1997) document that the share of realized capital gains in income is strongly increasing in income, providing prime facie evidence that the income elasticity is appreciably greater than one.

Our study of the realization of capital gains remains incomplete in one regard. In the data, household with negative adjusted gross incomes actually do realize capital gains. Indeed, 6.34 percent of the capital gains in 2004 were realized by such households. They have at least three excellent reasons to do so: their marginal tax rates are zero, the lowest they will ever be; realizing capital gains may help them to smooth their consumptions; and a bankruptcy court may be forcing the sale. For households with negative incomes, however, we cannot estimate their semielasticity since their marginal tax rates do not vary, all being zero. Nor can we identify any marginal effects on the probability of selection in a probit model. We therefore model their realization of capital gains as completely unaffected by tax policy.

We now apply the Heckit two-stage least-squares estimator to all of the datasets for 1976-2004. In the regressions, we include *T*, *F*, *B*, log*Y*, and all the available control variables *Z*1-*Z*24 and use *T**, *F** and *B** as instrumental variables, the same specification as Regression 12 employed. The resulting estimates are plotted as a histogram (Figure 1) and in a time-series plot (Figure 2). For clarity, the standard errors of the estimates have not been indicated. In all cases, they are small, reflecting the large sizes of our samples (typically well over 100,000).

According to the histogram, most of the estimated semielasticities center around -6, but in seven years the estimates are considerably more negative, indeed as negative as -26. The estimated semielasticities are therefore sensitive to the year for which they are estimated. The time-series plot indicates that the estimates are especially negative in 1978 and in the mid- to late-1980s but are only slightly negative from the end of the 1980s though the mid-1990s. Since then, they have been trending downward.

We have also examined whether the estimated semielasticities may have varied with aggregate conditions, finding some evidence that they have. Two regressions suggestive of such a relationship are

$$\hat{b}_t = \underset{(10.9)}{-28.7} + \underset{(1.40)}{2.49} \log(EQITY_t) \tag{6}$$

$$\hat{b}_t = \underset{(2.6)}{-14.6} + \underset{(2.74)}{6.14} G_t. \tag{7}$$

Figure 1. Histogram of Semielasticities of Realized Capital Gains to the Marginal Tax Rate, 1976-2004

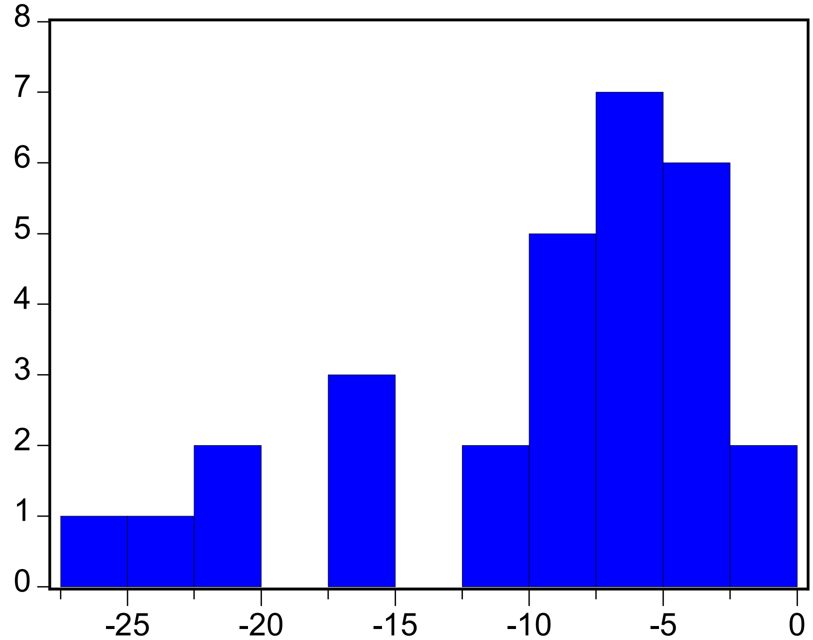
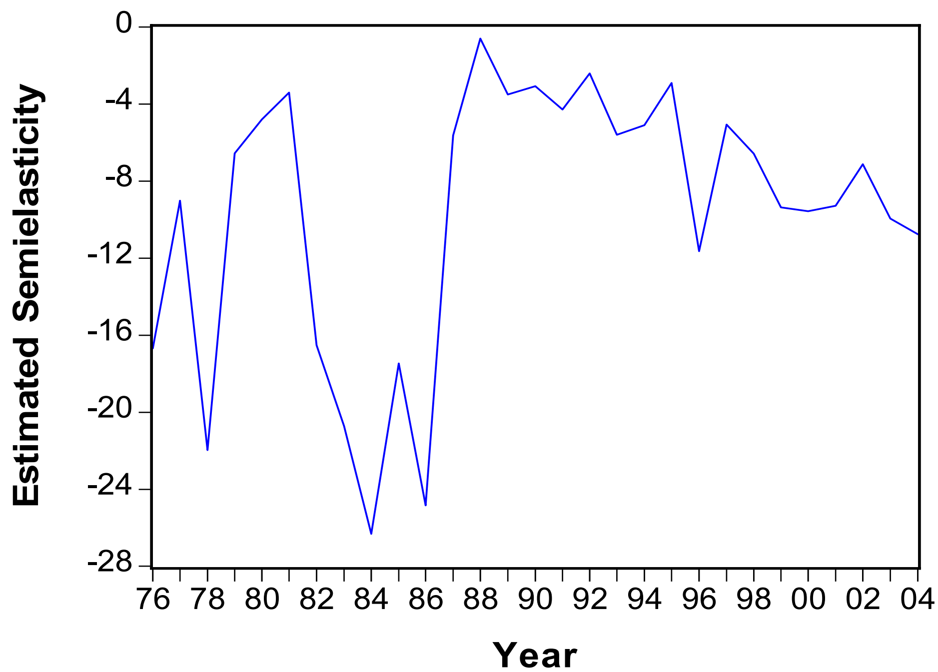


Figure 2. Plot of Semielasticities of Realized Capital Gains to the Marginal Tax Rate, 1976-2004



In these regressions, b_t is the estimated semielasticity for year t, the circumflex indicates a fitted value, $EQUITY_t$ is the real market value of stock in year t, G_t is the growth in the real Standard and Poor's 500 index over the previous 10 years, and the figure in parentheses beneath each estimate is its standard error.[10] The estimates suggest that the semielasticity is less negative in years when stockholders either have large holdings of stock or have experienced an extended bull market.

5. Implications for Tax Policy

At first blush, a highly negative semielasticity would seem to suggest that low tax rates on capital gains are desirable and a barely negative semielasticity would seem to suggest that a high tax rate is desirable. The issue is more complicated, however. A complete analysis requires attention to be paid to how the taxation of capital gains affects other decisions in the economy. On the one hand, suppose that the semielasticity is hugely negative largely because it greatly intensifies tax avoidance on other types of income. An increase in the tax rate on realized capital gains might then be justified even if it lowers the tax collected on realized capital gains.[11] On the other hand, suppose that the semielasticity is barely negative, but a lower tax on realized capital gains generates much higher stock prices, a much reduced cost of capital, and an enormous boom in economic activity. In that case, total tax collections might well increase, and even the tax on realized capital gains might increase.

It is hard to quantify such effects. For this reason, we shall assume that nothing else changes even though this assumption is certainly erroneous. Such an assumption allows the analyst to avoid making judgments about how large these effects might be, thereby making the analysis more objective.

---

[10] The nominal market value of stock was downloaded from the Flow of Funds Accounts on the website of the Board of Governors of the Federal Reserve System, and the Standard and Poor's 500 index was downloaded from the website of the Council of Economic Advisors. The series were then deflated with the consumer price index, which was downloaded from the website of the Council of Economic Advisors.

[11] Gillingham, Greenlees and Zieschang (1989), however, find no evidence for such effects.

Suppose that the estimates of Regression 12 and its associated probit equation adequately characterize how households behaved in 2004 and in subsequent years. How would the amounts of capital gains and tax revenue change in response to a permanent decrease in each household's first- and last-dollar marginal tax rates by one percentage point? Alternatively, how much would they change if these two marginal tax rates increased by one percentage point? Given our assumptions, these questions can be answered.

On the assumption that only T and T^* change and that both change by ΔT, equation (5) implies that $\log G_i$ changes by the sum of $b\Delta T$ of Regression 12 plus the household's selection effect, which is approximately $phS'\Delta T$, where b and p are the regression coefficients on T and S in equation (5), h is the coefficient on T^* in the criterion (4) and S' is the derivative of the inverse Mills ratio with respect to its argument.[12] Given ΔT = -0.01, our estimates for Regression 12 imply that the weighted average of the selection effect is 0.0026 percent over the population of households with positive realized capital gains and positive adjusted gross incomes. The total effect is therefore 0.1102. However, a fraction 0.0634 of the realized capital gains accrued to households with negative adjusted gross incomes, whose marginal tax rates do not change. Correcting for them yields a population-weighted average semielasticity of realized capital gains of (1–0.0634)·0.1102 + 0.0634·0, or 0.1032. Hence, this tax change would increase realized capital gains by approximately 10.32 percent. If the tax change also reduces the average tax rate on realized capital gains by one percentage point, tax revenue would increase by approximately 10.32 – 10.77, or -0.45 percent, a small revenue loss.[13,14,15] On the other hand, if ΔT = 0.01, a

[12] Differentiating both members of equation (5) while imposing the condition that T^* changes by the same amount as T, we obtain $\partial \log G/\partial T = b + p(\partial S/\partial T) = b + p\, S'(\partial C/\partial T^*)(\partial T^*/\partial T) = b + pS'h\cdot 1 = b + phS'$.

[13] The revenue from that taxation of realized capital gains is $\overline{T}G$, where $\overline{T}$ is the average tax rate on them. It follows that the proportional change in the revenue is $\Delta \log G - \log(\overline{T}_0 / \overline{T}_1)$, where $\overline{T}_0$ and $\overline{T}_1$ are the initial and final average tax rates.

[14] The average tax rate on capital gains lies somewhere between T^*, the marginal tax rate on the first dollar of capital gains and T, the marginal tax rate on the last dollar. On the assumption that the average tax rate is half way between the two, the population-weighted average of this measure of the average tax rate across the households realizing capital gains and with positive adjusted gross incomes is 0.09798. After the tax change, the average tax rate becomes 0.08798. The figure 10.77 percent in the text is then log(0.09798/0.08798).

similar calculation reveals that tax revenue increase by approximately 9.72 – 10.32, or -0.60 percent, again a small revenue loss.[16] By this calculation, then, the average tax rate in 2004 was close to the revenue-maximizing rate. Indeed, the revenue-maximizing rate was 1/10.32, or 9.69 percent.[17] This finding implies that capping the first- and last-dollar marginal tax rate at 9.69 percent would unambiguously increase tax revenue. The optimal marginal tax rate is lower than the revenue-maximizing rate, however, as taxation leads to inefficiencies. A cap well below 9.69 percent may therefore have been in order.

---

[15] This calculation ignores that the fact that the increased realization of capital gains depletes the stock of unrealized capital gains in subsequent years, thereby reducing capital gains realizations then. This depletion is a small effect, however, as realized capital gains are small relative to the stock of unrealized capital gains; see Gravelle and Lindsey (1988). It is beyond the scope of this paper to account for this effect.

[16] 0.0972 = log(0.10798/0.09798).

[17] With a little calculus, one can show that the derivative of tax revenue with respect to T is $G + T\partial G/\partial T$. At the revenue-maximizing T, this derivative is zero, implying that $(T/G)\partial G/\partial T = -1$ or $T = -G/(\partial G/\partial T) = 1/(\partial \log G/\partial T)$.

References

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

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

Bogart, William, and Gentry, William M. National Bureau of Economic Analysis, Working Paper 4254, January 1993.

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

Burman, Leonard E., and Ricoy, Peter. "Perspectives on the Ownership of Capital Assets and the Realization of Capital Gains," Congressional Budget Office, May 1997.

Darby, Michael; Gillingham, Robert; and Greenlees, John S. "The Direct Revenue Effects of Capital Gains Taxation: A Reconsideration of the Time-Series Evidence," *Treasury Bulletin* (Spring 1988).

Eichner, Matthew, and Sinai, Todd. "Capital Gains Tax Realization and Tax Rates: New Evidence from Time Series," *National Tax Journal* 53 (2000): 663-681.

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): 777-791.

Feldstein, Martin; Slemrod, Joel; and Yitzhaki, Shlomo. "The Effects of Taxation on the Selling of Corporate Stock and the Realization of Capital Gains: Reply," *Quarterly Journal of Economics* (February 1984): 111-120.

Gillingham, Robert, and Greenlees, John S. "Evaluating Recent Evidence on Capital Gains Realization Behavior," US Treasury Department, Office of Economic Policy, August 4, 1989.

Gillingham, Robert, and Greenlees, John S. "The Effect of Marginal Tax Rates on Capital Gains Revenue: Another Look," *National Tax Journal* 45 (June, 1992): 167-178.

Gillingham, Robert; Greenlees, John S.; and Zieschang, Kimberly D. "New Estimates of Capital Gains Realization Behavior: Evidence from Pooled Cross-Section Data," US Treasury Department, OTA Paper 66, May 1989.

Gravelle, Jane, and Lindsey, Larry. "Capital Gains," *Tax Notes* (January 25, 1988): 397-405.

Heckman, James. “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*, Volume 5 (1976): 475-492.

Jones, Jonathan D. “An Analysis of Aggregate Time Series Capital Gains Equations,” US Treasury Department, OTA Paper 65, May 1989.

Lindsey, Larry. “Capital Gains: Rates, Realizations and Revenue,” National Bureau of Economic Analysis, Working Paper 1893, April 1986.

Minarik, Joseph J. “The Effects of Taxation on the Selling of Corporate Stock and the Realization of Capital Gains: Comment,” *Quarterly Journal of Economics* (February 1984): 93-110.

Report to Congress on the Capital Gains Tax Reductions of 1978, US Treasury Department, 1985.

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

Stiglitz, Joseph E. “Some Aspects of the Taxation of Capital Gains.” *Journal of Public Economics* 21 (1983): 257-294.

About the Author

Dr. Paul D. Evans is a Full Professor in the Department of Economics, Ohio State University. He earned his Ph.D. in economics at the University of Chicago in 1976. He has taught economics at Stanford and the University of Houston. Dr. Evans has written more than 50 published articles in textbooks and leading professional journals, including the American Economic Review, the Journal of Political Economy, the Review of Economics and Statistics, Economic Inquiry, the International Economic Review, and the Journal of Monetary Economics, among others. His areas of expertise include monetary policy, inflation, interest rates, banking, fiscal policy, capital flows, and growth theory. He is a member of the board of editors of Studies in Economics and Finance and the Journal of Financial Stability. He has served as referee for articles and research for over twenty-five leading professional journals, the National Science Foundation, and the National Academy of Sciences. Dr. Evans is a member of the American Economic Association, the Econometrics Society, and the Western Economics Association International. He is an academic advisor to the Buckeye Institute for Public Policy Solutions and the Institute for Research on the Economics of Taxation.
