The Use of Panel Data in the Analysis of the Behavioral Response to Taxation

JOEL SLEMROD AND WILLIAM SHOBE

I. Introduction

The behavioral response to a tax policy is critical to assessing its effect on total revenue, the distribution of the tax burden, and the true cost of taxation to the economy, including its distortionary impact. Research on the behavioral response has, for the most part, relied on analyses of the historical evidence on tax structures and the observed choices of taxpayers. Most studies have used either cross-sectional or aggregate time-series data.

In this paper we discuss an alternative source of evidence about behavioral response to taxation - panel data, also known as longitudinal data. A panel data set follows a given sample of individuals over a period of time, and thereby contains multiple observations on each individual in the sample. Panel data sets have been utilized extensively and have provided valuable insights in other areas of economics, especially labor economics, but have received little attention in taxation. This oversight is unfortunate because the analysis of panel data potentially can overcome some of the fundamental econometric problems that plague analysis of cross-sectional and aggregate time-series data.

We proceed as follows. Section 2 briefly discusses the analytical characteristics of panel data, and emphasizes its advantages over other data sets. The hazards of using panel data are also addressed. Section 3 critically surveys the small body of existing literature that studies tax issues using panel data. A promising new source of panel data on taxpayers in the U.S. is described in Section 4, and Section 5 discusses some research on the tax elasticity of charitable contributions that uses this data. Section 6 concludes.

II. Analytic Advantages (and Disadvantages) of Panel Data

2.1. Controlling for unobserved variables

Econometric analysis can never account for every possible influence on the behavior of individuals. In some cases, the influence is mundane and measurable but not included in the data set. In other cases, the influence is inherently difficult to measure. It is well known that whenever there are omitted variables correlated with (included) explanatory variables, standard methods of estimation will yield biased estimates of the true model. The estimated coefficients will be a mixture of the true effect of the included variable and the effect of the correlated omitted variables.

Panel data can begin to resolve this problem whenever the omitted variable is an unobserved personal characteristic. As an example, consider the problem of estimating the effect of capital gains taxation on the realization of capital gains (the "lock-in" effect). One suspects that, in addition to the tax effect, realizations may be related to income, or wealth. Furthermore, taxpayers differ in their propensity to trade assets some are "churners," others are not. We have no measure of this propensity, but believe it to be a permanent characteristic of an individual. If this propensity is correlated with income, then a regression analysis using cross-sectional data will yield biased estimates of both the income and price effects. With multiple observations on the same individual, this problem can be eliminated by, for example, using first-differenced data. For any individual, the change in capital gains realizations is related to the change in income. Because, by assumption, there is zero change in the propensity to churn assets, no omitted variable bias arises.

2.2. Identifying tax and income effects

The identifiability of both price (tax) and income effects is a serious problem in the econometrics of the behavioral response to taxation. Consider the case of charitable contributions, which in the U.S. are deductible from taxable income for most upper-income taxpayers. For these taxpayers the price, or cost, of contributing one dollar is one minus the applicable marginal tax rate. Of great interest is the response of charitable giving to a change in that price arising from, for example, a change in the deductibility status or a change in marginal tax rates. The econometric problem arises because a taxpayer’s income is also likely to influence the level of giving and the fact that, in a graduated income tax system, the marginal tax rate is a function of taxable income. If the marginal tax rate were a linear function of income, then it would be impossible to separately identify the
price and income effects. The relationship is in fact nonlinear, so that a linear model can identify the two effects. In addition, it is usually assumed that the measure of income that affects charity is different from taxable income. In this way identification is assured due to the fact that different taxpayers of the same income make different use of ways to reduce taxable income, and thus their marginal tax rate. Both of these methods of identification have problems. Since we have no strong prior reason to believe that a linear model is appropriate, identification based on that assumption is problematic. Furthermore, if the reductions in taxable income, such as moving expenses, are correlated with the propensity to give charity, the price term will be biased by the effect on charity of the omitted variable.

Panel data that spans years when the tax schedule has changed provides an ideal vehicle for identification, because in this case the same taxable income implies different marginal tax rates in different years. This provides a source of price variation that is unrelated to changes in income.

2.3. Analysis or dynamic behavioral response

Panel data have a natural advantage over cross-sectional data for the analysis of dynamic phenomena. The estimated cross-sectional correlation between behavior and tax rate has usually been interpreted as being the long-run equilibrium response, but this may be incorrect if individuals are slow to respond to changes in the environment. Panel data allows the researchers to trace the immediate impact of changes in the environment on behavior and also the lagged effect. Past behavior may also have a direct effect on current behavior. For example, ceteris paribus, an individual who has realized an unusually large amount of capital gain in one year will be less likely to realize a large gain the following year, due to the depleted stock of appreciated assets in his portfolio. Current behavior may also be influenced by expectations of future changes in the tax environment. An investor who expects the tax on capital gains to fall in future tax years has an incentive to postpone realizations in order to take advantage of the lower tax rate. Panel data offer the researcher the opportunity to trace the course of anticipated tax changes and assess its impact on behavior.

2.4. Disadvantages or analysis using panel data

The availability of multiple years of observations on individuals is always an advantage. As outlined above, it facilitates correcting for the bias of estimated response due to unobserved influences, and can help to separate out tax and income effects on behavior, and allows the investigation of dynamic behavioral response to taxation.

Although the availability of panel data can only be an advantage, the application of panel data econometric techniques does have its pitfalls. In particular, by focusing on the relationship between year-to-year changes in behavior and year-to-year changes in the econometric environment, the signal-to-noise ratio of the data may decline significantly. For example, differences across individuals in income reported for tax purposes may largely reflect differences in real permanent income. The partial effect on behavior of income reported for tax purposes may then be an accurate measure of the income effect. However, annual changes in income reported for tax purposes may reflect voluntary changes in reporting strategy, losses taken purely for tax purposes, or transitory changes in annual income. All of these are noise if permanent income is the true influence on behavior.

Thus the use of panel data to minimize the potential bias due to omitted personal characteristics may exacerbate errors-in-variables bias. Careful attention must be paid to the meaning of the variables reported on an annual basis to the tax authorities.

III. Previous Tax Research Using Panel Data

The analysis of longitudinal data sets falls naturally into three categories that reflect the extent to which the researcher takes advantage of the special characteristics of the data. At the most informal level, we can use the data to add a second dimension to descriptive analysis that would otherwise be limited to cross-section or time-series. Panel data may also be used to increase the size of the dataset by treating each observation on a given individual as a separate data point; this is called pooling. At the most general level, we can use longitudinal data to combine the advantages of cross-section and time-series datasets; accounting for individual differences between cross-section units, estimating explicitly dynamic models with large cross-section samples and correcting for certain types of model misspecification.
3.1. Descriptive analysis

3.1.1. Minarik

Since income taxes are progressive, one-time realizations of gains accrued over several years may be taxed at higher rates than if the gains were taxed in the years accrued. This is often referred to as the “bunching” problem. Minarik (1981) used the IRS Seven-Year Panel of Taxpayers for the years 1967-73 to examine the effect of the bunching of capital gains realizations on the amount of taxes paid.

Minarik uses the panel to calculate the average annual tax over a seven year period for taxpayers who realize capital gains and to measure how taxes would be affected by averaging all capital gains realizations over the same period. Averaging capital gains realizations over the seven years of the panel would have reduced tax liabilities by about five percent, an indication that bunching of gains does result in higher tax payments on accrued gains.

Minarik also informally estimated how deviations from average income net of capital gains and average itemized deductions affect deviations of capital gains realizations from the average. He found that a larger than average income or smaller than average total deductions will reduce the likelihood that the taxpayer will take a relatively large gain realization. These estimates are consistent with the hypothesis that many taxpayers tend to realize gains when their tax rates are temporarily reduced. For these taxpayers there will be a less substantial bunching penalty.

Minarik’s analysis points up a problem with all models that attempt to predict capital gains realizations from other financial variables. That is, when we take into account the taxpayer’s optimization problem, we suspect that all of the variables are determined, to some extent, simultaneously. He estimated that the deductions effect on capital gains realizations was more pronounced than the income effect. The interpretation given to this result depends on whether the deductions are “voluntary” or “involuntary.” For involuntary deductions such as medical expenses this result would suggest that deviations of gains realizations from the average are used more to meet unanticipated expenses than to maintain a constant level of taxable income. With voluntary deductions, the interpretation would be that charitable contributions are planned simultaneously with gains realizations in order to minimize tax liabilities.

Even at this informal level, however, the analysis of panel data offers insights unavailable from cross-section or time-series datasets.

3.2. Pooling

Panel data is frequently used to increase the number of data points available by pooling all of the observations into one large dataset. The researcher then applies standard estimation techniques as if each observation were an independent draw from the population.\(^1\)

Economic models using pooled panel data (or cross-section data) generally assume that all of the agents have the same behavioral response function up to a random error term: \(y = f(X) + e\). The error term, \(e\), will vary among individuals but the parameters of the function must be the same for the estimation to be meaningful. If the true behavioral function parameters vary among individuals or over time, especially in some systematic way, then the estimation will suffer from possibly serious heterogeneity bias. “Heterogeneity bias” simply means that if each individual in a panel has a different set of behavioral parameters, then estimating with pooled data will give very unreliable parameter estimates, possibly of the wrong sign. Tests for heterogeneity bias are straightforward but are often not useful, as they require several more years in the panel than there are parameters to be estimated. In this case, the researcher must carefully weigh the risks of heterogeneity bias against the potential benefits of greatly increasing the number of data points by pooling. Techniques do exist for relaxing the homogeneity assumption. We will mention some of these later in this section.

3.2.1. Auten & Clotfelter

One of the most visible current issues in U.S. tax policy is the effect of taxation on capital gains realization behavior. Cross-section estimates including those of Fieldstein, Slemrod, and Yitzhaki (1980), hereafter FSY, and time series estimates such as the U.S. Treasury Department (1985) suggest that capital gains realizations are very responsive to changes in the capital gains tax rate. The response could even be strong enough to cause an inverse revenue response to changes in the tax rate.

Auten & Clotfelter (1982) used the same seven-year panel used by Minarik to estimate the tax elasticity of capital gains realizations. The authors construct a 3-year backward-looking moving average of both marginal tax rate and income to serve
as proxies for permanent tax rate and income, respectively. They define the transitory tax rate and income as the difference between current tax rate and income, respectively, and their measure of permanent tax rate and income. Testing for the influence of both transitory and permanent variables allows them to distinguish between the short-run and long-run impact of tax policy. The authors pool the data in the panel in order to increase the number of observations, implicitly assuming that parameter values do not change across individuals or across time.

The authors find that transitory changes in tax rate do have a large and significant effect on realization behavior. The estimated effect of permanent changes is smaller. The estimated elasticities are somewhat lower than that found by FSY, who could not separately identify the short-run and long-run effects.\(^2\)

Two key criticisms may be made of this study. First, the authors do not seem to consider the possibility of heterogeneity bias. There are many individual characteristics that might affect the rate of asset turnover and the pattern of gains realizations that are not observable. These differences would be less of a problem if we allow the intercept term to vary among individuals. As it stands, we have no indication of either the sign or magnitude of the bias.

Second, we would certainly expect that capital gains realizations follow definite temporal patterns. Unusually large realizations would likely be followed by smaller ones. Thus, we would expect that last year's gains would help explain this year's gains. This variable is not included in Auten and Clotfelter's study. On the other hand, last year's income (net of gains) is a regressor, as part of the income moving average. Since current income is related to the level of gains realizations, there is good reason to believe that a correlation exists between permanent income and the error term.

3.2.2. Treasury (Auten)

In a study for the U.S. Treasury Department, Auten (1982) attempted to improve on Auten and Clotfelter by more carefully measuring wealth and by using a more reliable measure of the effective tax rate. He also used group averages to construct a proxy variable for marginal tax rates, since the observed marginal tax rate on the last dollar of capital gain is determined simultaneously with income. As in his study with Clotfelter, Auten pooled the panel data.

Auten's results fall neatly between those of FSY and those of Auten and Clotfelter. The tax elasticities he estimates vary between -1.16 and -2.20, adding some support to the possibility that, for capital gains, tax rates and revenues may move in the opposite direction, at least at current levels of taxation.

It is worthwhile to note that, in an attempt to eliminate the problem of simultaneity of income and tax rate, Auten is forced to use a rather diffuse proxy based on group averages for one of the main independent variables, tax rate. Finally, as before, pooling the data is subject to substantial heterogeneity bias if unobserved characteristics would cause parameters to vary across taxpayers.

Auten also used the pooled data to replicate the FSY analysis, and obtained results quite consistent with FSY. We note this result to point out that a regression with pooled data is essentially equivalent to a regression with cross-sectional data except that there are more observations.

3.3. Special panel techniques

Many techniques exist for taking advantage of the extra information in a panel dataset without assuming that individuals' response to individuals is completely homogeneous.\(^3\) We will concentrate on linear models that are stable over time, allowing individual intercepts to vary while still assuming constant slope parameters across individuals.

Under certain ideal conditions, panel estimation techniques can utilize both variation over time for a given individual and variation between individuals to fix population parameters. When these conditions hold, we use the **random-effects model**. The crucial assumption for this model is that any unknown individual characteristics are random and not correlated with any of the explanatory variables. Thus, we can treat variation due to individual characteristics as part of the error term. Under these circumstances, generalized least squares (GLS) is the best linear unbiased estimator. Ordinary least squares (OLS), because it does not account for the fact that the covariance of within-individual errors is different than the covariance of across-individual errors, will yield inefficient, though unbiased, estimates.

Often, however, we expect that certain values of the explanatory variables are more likely for some individuals than for others and that unobserved factors which affect the dependent variable are correlated with the explanatory variables. In that case, treating individual variation as part of the error term would induce a correlation between the error and the explanatory variable and would cause GLS (and OLS) to give biased results. In this case, we use the **fixed-effects model**, which treats individual variation as just another parameter to be estimated. The appropriate estimation technique is the covariance estimator (CV), which ignores variation between individuals, thereby eliminating the bias due to individual-specific error
components that are correlated with the explanatory variables. When a two-year panel is used, CV is equivalent to estimating a first-differenced model.

These models become much more difficult to estimate when the errors vary systematically over time. The maximum likelihood estimator for the fixed-effects model is biased when the number of years in the panel is small even if the number of individuals is very large. For random-effects models, the interpretation of the model depends on assumptions about how the initial observations were generated. Researchers most often assume that parameters are constant over time and allow for individual variation.

3.3.1. Clotfelter

These special panel data techniques are just beginning to find use in the analysis of tax policy questions. Clotfelter (1980) used the IRS Seven-Year Panel of taxpayers for 1967-73 to examine the effect of tax incentives on charitable giving. Data on charitable contributions are available for four years of the panel. Clotfelter did not use the CV estimator. Instead, he estimated three separate first-difference (FD) equations and reported the separate results. The FD estimation offers several advantages over earlier, cross-section, results. First, the FD model should remove any heterogeneity bias due to omitted individual-specific variables. Second, in cross-section estimation, the price and income effects are subject to a serious simultaneity bias. In the FD model this bias is reduced since part of the change in the tax price of giving is due to a change in the tax schedule. The tax schedule change is exogenous to the taxpayer’s optimization and, thus, the proportion of total variation subject to simultaneity bias is reduced. Third, Clotfelter uses prior-year income to calculate “predicted” income for a period. He then uses this predicted value to separate income into its “permanent” and “transitory” components to test for differences in how contributions respond to temporary and permanent changes in income.

Clotfelter’s results are rather dramatic. The FD estimates of the price elasticity of giving are just -0.33, compared to -1.40 for similar equations estimated with cross-section data. In fact, the FD price elasticity estimates are not significantly different from zero. He also found little difference between elasticities for the permanent and transitory income measures.

What can explain the differences between cross-section and panel estimates? The FD estimator will eliminate much bias that results from unobserved individual variation. For example, we might expect that “social status,” although unobservable to the researcher, is an important determinant of charitable giving. Although status and income are no doubt correlated, status remains quite stable between any two years. If status is absent from the estimating equation, the income measure will be correlated with the error term in a cross-section sample, causing the estimates of the effect of income to be biased upward. Because status is by assumption stable, its influence would be eliminated in a FD model; the estimates would be smaller due to elimination of the bias.

Clotfelter himself suggested that the FD model may be measuring an incomplete, short-run response and that the full adjustment takes several years to complete. To investigate this hypothesis, he added prior year giving to a non-differenced model to estimate the coefficient of adjustment. He found the adjustment to be 50% over two years. From this, he concluded that it would be reasonable to assume that adjustment delays could explain much of the difference between the FD model and the earlier estimates. Because Clotfelter did not estimate a FD model with an adjustment term, an evaluation of his results is difficult since it is unclear what effect the addition of an adjustment term to the biased non-differenced model will have on the direction and magnitude of bias.

3.3.2. Broman

Broman (1986) followed up Clotfelter’s study using the six-year panel of taxpayers from the Continuous Work History File discussed in more detail in Section 4.1. She estimated a FD equation containing an instrument for lagged giving and found that slow adjustment accounted for only a small fraction of the difference between the cross-section and FD models. Broman’s estimates for price elasticity are quite close to those of Clotfelter’s FD model.

Broman also used the panel to investigate timing effects due to the announcement of future changes in the tax rate. Although her results were quite sensitive to assumptions made about the formation of expectations, she concluded that the evidence suggests a substantial level of inter-temporal shifting of charitable contributions due to anticipated changes in the price of giving.

3.3.3. General observations

It is worth noting that, in the examples cited, using panel techniques for controlling heterogeneity bias has reduced the
estimated tax elasticities toward zero or eliminated their significance altogether when compared with cross-section results. It is not yet clear whether this pattern is due solely to the reduction of heterogeneity bias or whether some other factors are also playing a part. It is quite conceivable that the bias could work in the opposite direction. This would appear to be a fruitful area for further research.

3.4. Has previous research exploited the advantages of panel data?

The availability of panel data on U.S. taxpayers has opened up significant new opportunities for examining the behavioral response to taxation. Unfortunately, no one has yet fully exploited the power of these datasets. Two significant gaps remain in the literature. First, none of these studies cited actually apply estimation techniques that utilize all of the information available in the panels. Some studies use the panel to generate new variables to be used in single year estimations. Even the first difference models estimate a separate FD model for each period rather than using the power of the covariance estimator or, if appropriate, the feasible GLS estimator. Much information available in the structure of the panel may be wasted in this way.

The second omission is the lack of dynamic specifications of behavior in these models. For example, in the capital gains literature, no attempt has been made to model realizations behavior as a dynamic problem. Indeed, some of the specifications tested imply substantial irrationality on the part of investors. The research on charitable giving has made a first step in the direction of dynamic specification by recognizing the possibility of incomplete adjustment, but the models actually estimated are static.

IV. A New Panel of Tax Returns

Since 1979, the U.S. Internal Revenue Service has been collecting information from the tax returns of a randomly selected group of taxpayers. This panel, known as the Continuous Work History File, was developed for internal use but the IRS has made this longitudinal dataset available to academic researchers through a special arrangement with the Office of Tax Policy Research at the University of Michigan, in conjunction with the Arthur Young Tax Research Database Project. The panel now spans 1979 to 1984 with 1985 and 1986 expected soon. This panel is an unusually rich source of individual level data for those researchers interested in the effects of taxes on behavior.

4.1. Technical characteristics

The panel is a non-stratified random sample chosen on the basis of the last four digits of the social security number (SSN). Of those numbers chosen, anyone filing a return is included in the sample. The first three years of the panel each contain in excess of 45,000 returns, though the last three years of the panel show a substantial drop in the number of observations (9,000 in 1982 and 1984, 19,000 in 1983) due to budgetary limitations at the IRS. Pooling all observations in the panel gives a sample size of 177,177. Due largely to the small number of observations in 1982 and 1984, the number of individuals present in all six years of the panel is limited to 6,152 taxpayers. The information contained in each observation is a subset of the information on the standard forms filed by the taxpayer, and varies slightly from year to year.

Attrition from the panel may occur for a number of reasons unrelated to deliberate change in the sample size, including death, a change in marital status, income below the minimum that would trigger filing, late filing, or simply the choice of which spouse (between two married, joint filers) is listed first on the tax form. A taxpayer who files sufficiently late will also escape inclusion.

It is not unreasonable to suspect that a panel of this sort may exhibit some drift relative to the population as a whole. Although each year’s taxpayers in the panel may be representative of the population as a whole, the sampling method may cause a “survivorship bias” among those observations present in more than one year of the panel Christian and Frischmann (1988) have analyzed the panel for survivorship bias and have concluded that the “balanced panel” of 6,152 taxpayers present in all six years shows statistically significant variation from population averages. Income is higher and married couples are more numerous. Also, the average age of members is increasing relative to the population as a whole. Researchers should take warning from these results when making inferences about population characteristics from the sample.

4.2. Points of interest

The years of coverage of the panel, 1979 to 1984, are of special interest to researchers because this period covers significant changes in U.S. tax law. Tax rates fell substantially between 1980 and 1983. The top marginal tax rate on all income fell from 70% to 50% while the top rate on capital gains income fell from more than 28% to 20%.
This same period covers a deep recession and the subsequent recovery. The inflation rate dropped from 10% to 4% over the six years. These policy changes and business cycle events provide a significant source of exogenous variation for identifying behavioral parameters of taxpayers in the sample.

V. An Exercise in Panel Data Econometrics

In a recent paper, Daniel (1988) has used this panel data to examine the tax: elasticity of charitable contributions. Daniel first re-estimated the standard charitable giving model (using a subsample of 1,527 taxpayers who itemized deductions that are present in all six years) for each year of the panel as an independent cross-section sample. For each of the six years of the panel, the cross-section estimates are very close to cross-section estimates made by previous investigators. Price elasticity varies between -1 and -2, although none is significantly different from -1. Daniel also estimated a pooled regression on the full 9,162 observations. The estimated price elasticity is -1.21, also very similar to cross-section results.

Next, Daniel estimated a fixed-effects version of the contributions model. This procedure may be thought of as the multi-year equivalent of a first-differenced model in that it controls for time-invariant characteristics of the individuals and thus corrects for any bias due to the omission of these variables. The results are dramatically different. The estimated price elasticity becomes -0.03, which is not significantly different from zero and is significantly different from -1 at greater than 99% confidence.

Finally, Daniel estimated the random effects model using GLS. The estimated price elasticity under this specification is -0.19 and lies between 0 and -1 at the 99% confidence level. This elasticity, although greater in absolute value than the fixed effects estimate, suggests a tax responsiveness of charitable contributions that is much lower than the cross-section estimates indicate.

Unfortunately, the six-year span of the panel is not long enough to test which specification is statistically preferred to the others. Such tests do exist but have no power with such a short panel. These results do, though, cast further doubt on the cross-section estimates. Our economic models give us good reason to suspect that cross-section estimates suffer from misspecification bias, and when we correct for that bias the estimates change dramatically. Further research must now be directed toward discovering the source of the divergence of the estimates. As longer panels are made available, the data may be able to select the preferred specification.

VI. Conclusion

Panel data that follow the same taxpayers over a period of years are a potentially rich source of evidence about the behavioral response to taxation. Only recently have researchers begun to exploit this richness, at the same time that a new panel of tax return information is available in the U.S. Initial analyses of this data suggest that the behavioral responsiveness of certain activities to taxation may be lower than is indicated by studies of cross-sectional, or even aggregate time-series, data. Further work is needed to either corroborate these results or to suggest alternative explanations of the apparent behavioral insensitivity to tax changes.

Footnotes

1 This discussion follows closely that of Hsiao, 1986.
2 See Minarik (1982) and Feldstein and Slemrod (1982) for a discussion of some still controversial issues concerning the appropriateness of various estimation techniques used in these studies.
3 For two recent surveys see Hsiao, 1986, and Chamberlain, 1984.

References


Hsiao, Ch., 1983, Analysis of Panel Data (Cambridge).

