NBER WORKING PAPER SERIES

TAXATION AND THE LABOR SUPPLY DECISIONS OF THE AFFLUENT

Robert A. Moffitt
Mark Wilhelm

Working Paper 6621
http://www.nber.org/papers/w6621

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1998

An earlier version of this paper was presented at the Conference, "Does Atlas Shrug? The Economic Consequences of Taxing the Rich," October 24-25, 1997, Office of Tax Research, University of Michigan. The authors would like to thank Gerhard Fries, Arthur Kennickell, Lillian Mills, Joel Slemrod, Christopher Taber, and other participants at the conference for suggestions and comments as well as seminar participants at the University College London, Harvard-MIT, Northwestern University, and Yale University. Cristian de Ritis provided excellent research assistance. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

Taxation and the Labor Supply:
Decisions of the Affluent
Robert A. Moffitt and Mark Wilhelm
NBER Working Paper No. 6621
June 1998
JEL No. H2, J22

ABSTRACT

We examine the effect of the 1996 Tax Reform Act on the labor supply of affluent men. The Act reduced marginal tax rates for the affluent more than for other taxpayers. Using instrumental-variables methods with a variety of identifying variables, we find essentially no responsiveness of the hours of work of high-income men to the tax reduction. However, we do find hourly wage rates of such men to have increased over the period.

Robert A. Moffitt
Department of Economics
Johns Hopkins University
Baltimore, MD 21218
and NBER
moffitt@jhu.edu

Mark Wilhelm
Department of Economics
Pennsylvania State University
University Park, PA 16802

Research on the labor supply effects of taxation has a long history (Hausman,1985) but very little research has directly concerned high-income taxpayers. This is a serious deficiency in the literature because it is widely assumed that high-income taxpayers may be more responsive to tax rate changes than other income groups because their marginal tax rates are very high and because they have more opportunities for altering their behavior. In part this neglect has been the result of data difficulties because relatively few data sets have contained labor supply information on a sufficient number of high-income taxpayers--hours of work being the most traditional measure. The most widely-used data sets for tax analysis of high-income taxpayers have used information from IRS returns (e.g., Feldstein, 1995a) but these data sets contain no direct information on labor supply (other than whether earnings are positive). We address this data difficulty by using the Survey of Consumer Finances (SCF), a data set which oversampled high-income taxpayers and which was conducted at several points during the 1980s and 1990s. We use the SCF to analyze the effects of the 1986 Tax Reform Act on the labor supply decisions of the affluent.

There are a number of methodological issues that must be addressed in any study of the labor supply effects of income taxation and several special issues in a study of the affluent. Among the general issues are those dealing with how to obtain cross-sectional variation in changes in marginal tax rates; this is a problem because the same tax law, in essence, applies equally to everyone. Another, separate issue is how to deal with the nonlinearity of the tax schedule when a nonproportional tax schedule is in force. Because of its special importance, we will address the first issue in detail in our paper.

Section I reviews prior work on evaluating the effects of high-income taxpayers. Section II outlines our approach to estimation and Section III presents our data and results. A final section draws conclusions.

I. Prior Work on Taxation and Labor Supply

Empirical work on the effect of taxation on labor supply through the early 1980s is reviewed by Hausman (1985). By and large those studies suggested that male labor supply is rather insensitive to tax rates but that female labor supply, at least that of married women, is considerably more sensitive.[1] Studies of the effects of 1981 and 1986 tax legislation have found generally consistent results, with responses larger for women than for men and small, if not zero, effects for the latter (Bosworth and Burtless, 1992; Eissa, 1995, 1996a, 1996b; Mariger, 1995; Ziliak and Kniesner, 1996). With the exception of two recent studies of high-income physicians, lawyers, and managers (Showalter, 1997; Showalter and Thurston, forthcoming), these studies have not had large numbers of observations of high-

---

[1] We consider here only the uncompensated elasticity. If income elasticities are sufficiently large, compensated elasticities can be nontrivial. See Hausman (1981) for an example.

income taxpayers.[2]

Because IRS data have many more such observations, there have been more studies of the effects of the 1981 and 1986 legislation on incomes as reported to the IRS. These studies have generally revealed quite significant responses to tax rates, although the magnitude of the effect differs considerably across studies (Auten and Carroll, 1998; Feenberg and Poterba, 1993; Feldstein, 1995a; Feldstein and Feenberg, 1996; Lindsey, 1987).[3]

The issue which has preoccupied much of the recent literature has concerned how to identify the effects of the federal income tax on either labor supply or income, given that individuals with the same characteristics face the same tax schedule at a given point in time. If the social and economic characteristics which cause tax schedules to differ across individuals (marital status, family size, forms of nonlabor income, etc.) have independent effects on behavior, there is no remaining variation in tax rates to permit the identification of tax responsiveness. While many of the earlier studies reviewed in Hausman (1985) made the assumption that some of those characteristics did, in fact, not affect behavior independently, the studies since that time have eschewed that variation in favor of other forms of identification. Cross-sectionally, a few studies have used state

[2] There is also a literature on the effects of the Earned Income Tax Credit (EITC) on labor supply. Because that tax feature is aimed at low-income families, and we are concerned with high-income families, we will not review those studies.

[3] We note that a major issue in these studies is whether the changes in income reported to the IRS reflect real changes in behavior or only changes in the form of income as a means of tax avoidance (Slemrod, 1994, 1996).

variation in taxes for identification (Auten and Carroll, 1998; Showalter, forthcoming), but these studies ignore migration and income shifting across states. The more common methods of identification have used the "differences in differences" method, which uses variation over time in tax schedules for different individuals to identify tax effects (see the references above for U.S. studies; see Blundell et al., forthcoming, for a U.K. study using this method).

We devote the next section of our paper to a discussion of this method and of the assumptions underlying it, and show that it is a form of instrumental variables estimation which requires exclusion restrictions for identification. We then proceed with our empirical work and apply that method to the effect of the 1986 Tax Act on the labor supply of high-income men, using the SCF data.

II. Modeling the Labor Supply Effects of Taxation

As just noted, a major problem in estimating the effects of nationwide tax systems is that they provide no variation upon which to base estimation, at least holding constant individual characteristics. The methodology of "differences-in-differences," or fixed effects, which is employed in some of the recent studies, makes use of panel data or repeated cross-section data to address these problems. This methodology can be applied in a simple tabular fashion but can also be applied in a regression context. We shall begin by discussing this method in general and will show that, when put into a regression framework, the method can be seen to rely for identification on exclusion restrictions of a particular kind and that a leading case of the methodology is equivalent to instrumental variables estimation with panel data. We shall then briefly discuss the issues raised by using repeated cross-section data and by the nonlinearity of the budget constraint.[4]

Differences-in-Differences (Fixed Effects) with Panel Data. The differences-in-differences methodology can be viewed within the context of the treatment-effects literature (e.g., Heckman and Robb, 1985), where interest centers on the effect of some treatment d (usually defined as a dummy variable) on some outcome variable y, possibly conditional on a vector of other regressors x, which we will take to be individual socioeconomic characteristics (possibly including income amounts). However, the models with which we are concerned differ in an important respect from the standard model in the treatment-effects literature, for here it is assumed that d has no cross-sectional variation conditional on x. The federal income tax is of this type because all individuals with the same characteristics face the same schedule, and all individuals with the same characteristics and income components and amounts face the same marginal tax rate.

The fact that the tax schedule does vary with individual characteristics and income implies that the stimulus induced by the tax system is a function of x, and this is what furnishes variation

[4] See Blundell and MaCurdy (forthcoming) for another econometric discussion of the differences-in-differences method and Meyer (1995) for an earlier discussion with references to applications of the method in areas other than taxes and labor supply.

that can be used for identification. In the case of tax systems, the tax formula dictates that marginal tax rates differ for individuals with different characteristics (marital status, number of dependents, income, whether a home is owned, and other variables). Letting p denote the time period, our starting point is a linear model of the form

$$y_p = \alpha_p + \beta d_p(x_p) + \gamma_p x_p + \epsilon_p \tag{1}$$

where $d_p(x_p)$ is the treatment variable of interest, which is often the marginal tax rate faced by the individual.[5] For a particular choice of tax variable $d_p(x_p)$, the parameter of interest is the effect of that variable, which is β. For the most part, we will assume that $d_p(x_p)$ is a known parametric function because the tax formula is known and hence x_p are the variables that go into the tax formula. However, all of our important conclusions will apply as well to the case in which d_p is an observed variable, x_p are instruments, and $d_p(x_p)$ is a function to be estimated (e.g., in a first-stage regression); in this interpretation, the identification problem arises if there are no instruments that do not appear independently in the equation.

We assume that eqn(1) is an equation derived from theory and hence is assumed to be the "true" equation, or at least that it can be formally derived as an approximation to that theory. That is, we

[5] We leave aside for the moment exactly what feature of the tax formula is of interest, including the issue of "which" marginal tax rate is of interest if the tax system is nonlinear. We will discuss this issue in the next section.

assume that the theory puts no other restrictions on the equation which might furnish sources of identification. This is an issue in the labor supply case, where one element of x_p (the wage rate) interacts with $d_p(x_p)$ (the marginal tax rate) according to most theories. We address this issue below, and for now simply restrict ourselves to the class of theories, whether it be a large or small class, which generate equations of the form of eqn(1).

To illustrate the problem most cleanly, we assume in eqn(1) that the set of x that enter the tax formula is equivalent to the set that appears independently in eqn(1); in practice, the former is likely to be a subset of the latter but this merely would mean that we should add another set of variables into (1).[6] Adding such a set will not affect the identification problem so we do not do so. All variables and parameters in eqn(1) are assumed to vary over time except β, which is not allowed to vary because it is the main parameter of interest and it is generally desired to estimate only a single time-invariant response effect, at least over a short period of time.

As it stands, with a single cross-section of data, β is identifiable from nonlinearities in the $d_p(x_p)$ function because x_p appears linearly in the equation. But this source of identification is weak because slight relaxations of the linearity would result in a

---

[6] If $d_p(x_p)$ is the exact tax function, then necessarily x_p includes endogenous variables like income. In that case those variables will generally not appear on the right-hand-side of (1) but their variation will nevertheless not identify β because they will be correlated with ϵ_p.

loss of identification.[7] If instead variables can be found which affect marginal tax rates but which do not affect y_p directly, the effect of $d_p(x_p)$ on y_p would be nonparametrically identified (at least over the range of the data) and the problem would be solved; but we will assume throughout that such variables are not available.

The critical vector of variables in this model is x_p, and a number of different cases can be distinguished depending on the nature of that vector. One major distinction is whether it is time-invariant or varies over time; another is whether it is endogenous (i.e., correlated with ϵ_p) or exogenous. For an income tax application, the relevant case is clearly endogenous, time-varying x, because x includes income, which varies over time, and which is endogenous because y, if labor supply, is one determinant of income. However, we will build up to that case by first considering exogenous time-invariant x and then exogenous time-variant x_p; after which we will consider endogenous x and x_p.[8]

The case which serves as the prototype for all the others is the case of a time-invariant exogenous x. In the tax case, filing status,

---

[7] If d(x) is linear in x, identification would clearly be lost (p subscripts, which are irrelevant in a single cross-section, are ignored). If d(x) is nonlinear in x, identification is generally lost if eqn(1) is generalized to $y=\alpha+h[d(x)]+g(x)+\epsilon$ where g and h are arbitrarily nonlinear functions with unknown parameters (recall that d(x) is a known parametric function). A qualification to this statement is that some portions of g(x) can be identified if x is a vector rather than a scalar because multiple points in the support of x yield identical values of d(x). We ignore this source of identification.

[8] Under the interpretation of x_p in eqn(1) as instruments, these cases correspond to the use of different types of instruments. The discussion is thus relevant to a different, but perhaps larger, class of applications than the tax example.

if taken as exogenous and time-invariant in the short term, is one example. Assuming that panel data on a set of individuals is available for two periods (we will consider below the case of more than two waves of data) and that the law changes between the periods, we have

$$y_{p+1} - y_p = (\alpha_{p+1}-\alpha_p) + \beta[d_{p+1}(x)-d_p(x)] + (\gamma_{p+1}-\gamma_p)x + (\epsilon_{p+1}-\epsilon_p) \qquad (2)$$

With this first-differenced equation, β is identifiable (apart from nonlinearities in $d_p(x)$) if $\gamma_{p+1}=\gamma_p$, in which case x drops out of (2) as an independent determinant of the change in y. In words, it must be assumed that there is no trend in the independent effect of x on y. This is the assumption that has figured in much of the differences-in-differences analysis of tax effects (Eissa, 1995, 1996a, 1996b; Feldstein, 1995a; Blundell et al., forthcoming; and others). Thus at least one variable must be found which affects how individuals react to the program but whose independent effect is stationary; that is, an exclusion restriction is necessary for eqn(2). This will be the critical assumption in all the models to be discussed. Note that the model is equivalent to a fixed-effects model where x is the fixed effect which differences out.[9,10]

---

[9] Note that with x defined as a vector of individual or area-specific dummy variables, the model fits into the standard individual-level or state-level fixed effects models.

[10] As in all fixed effect and differences-in-differences models, an issue is the degree to which the linearity and additivity in the model can be relaxed and identification retained. Replacing the

The assumption $\gamma_{p+1}=\gamma_p$ is a nontestable, just-identifying assumption in the model as stated because estimates of β cannot be obtained if it is relaxed. However, if data on additional periods prior to p are available, the assumption can be relaxed to some degree because a time pattern of γ_p can be estimated and it can thereby be determined whether γ_p contains a time trend. While it can never be known for certain whether the independent effect of x on y would have changed from p to p+1 in the absence of a change in the d (this is the usual problem of the missing counterfactual in treatment-effects models), more history on y and x can at least assist in establishing priors on whether the effect changed between p and p+1.[11]

If x is a time-varying exogenous variable (e.g., number of dependents, if taken as exogenous), a differenced eqn(1) is

$$y_{p+1} - y_p = (\alpha_{p+1}-\alpha_p) + \beta[d_{p+1}(x_{p+1})-d_p(x_p)] + \gamma_{p+1}x_{p+1} - \gamma_p x_p + \epsilon_{p+1} - \epsilon_p \qquad (3)$$

---

additive linear γx by an additive nonlinear g(x) requires for identification only the restriction that the function g be constant over time. Relaxing additivity and permitting interactions between x and $d_p(x)$ is also possible. If we let $y_p=\alpha_p+h[d_p(x),x]$, where h is of unknown form, a nonparametric regression of the change in y on x will identify differences in response across different values of x. The fundamental restriction is that h is not indexed by p and that all non-law-related changes over time appear in an additive term (the intercept). Thus there is still a substantive differences-in-differences restriction even when additivity and nonlinearity is considerably reduced.

[11] See Bosworth and Burtless (1992) and Eissa (1996a) for two labor supply studies that sought to establish longer-term time trends and to determine whether there have been deviations from trends. These types of tests are common in models which permit not only fixed effects in levels but also fixed effects in trends, for example. Note too that this method is made more complicated if the law has been changing in past periods (e.g., prior changes in tax law), which may make it difficult to establish the existence of a trend.

where the effect of taxes is again unidentified if the linearity assumptions in (3) are sufficiently relaxed. Here the problem is not solved if $Y_{p+1}=Y_p$ (again, if linearities are relaxed). But a simple way of dealing with this issue is to select the subsample for which $x_{p+1}=x_p$; for that subsample, $Y_{p+1}=Y_p$ again is a sufficient condition for identification. Because both x_{p+1} and x_p are exogenous, this selection introduces no bias.

Many of the more important applications of the differences-in-differences, fixed-effects approach are cases where the excluded variable in first differences is instead endogenous. To keep this case notationally separate from the previous ones, we will use z to denote the variable instead of x, where now it is assumed that z and ϵ_p are not independent. Here the difference between time-invariant z and time-varying z will be more important, and most of the interesting cases will arise when z is time-varying. But time-invariant z is an important case as well, although examples are more difficult to imagine in practice. In the tax case where y is labor supply, selecting a subsample for whom marital status is unchanged from p to p+1 is one such z if marital status is considered to be jointly determined with labor supply.

The application of the methodology in this case can be most easily rationalized by the assumption of the panel data random-effects model. Hence we assume

$$y_p = \alpha_p + \beta d_p(z) + \mu + \nu_p \tag{4}$$

where μ is a time-invariant individual effect. In eqn(4), z is not included as a separate regressor because it is assumed to be an endogenous variable jointly determined with y and hence not too have an independent structural effect on y. The endogeneity of z can arise either from a relation to μ or to ν_p or both, but it is the former that can be addressed by first differencing. Because μ and z are both time-invariant, it follows that $E(\mu|z)$ is constant over time and therefore that the "types" (μ) of individuals associated with different values of z do not change. Hence

$$y_{p+1} - y_p = (\alpha_{p+1}-\alpha_p) + \beta[d_{p+1}(z)-d_p(z)] + (\nu_{p+1}-\nu_p) \tag{5}$$

The assumption needed in this model for consistent estimation of β is that z is uncorrelated with $(\nu_{p+1}-\nu_p)$, the trend in the unobservables in the equation (or, more precisely, that the function $[d_{p+1}(z)-d_p(z)]$ is uncorrelated with $(\nu_{p+1}-\nu_p)$).[12] This case is thus once again equivalent to a simple fixed effect model. The assumption that $\gamma_p=\gamma_{p+1}$ is equivalent in this model to the assumption that there is no time-varying coefficient on μ, as would be the case if $\psi_p\mu$ appeared in the model.[13]

---

[12] This is a case of a "balanced" bias analogous to that in randomized trials based on endogenously-selected populations discussed by Heckman (1996b).

[13] If $[d_{p+1}(z)-d_p(z)]$ interacts with μ, identification problems ensue.

If z is a time-varying endogenous variable, we have, again assuming the presence of an individual effect,

$$y_p = \alpha_p + \beta d_p(z_p) + \mu + \nu_p \tag{6}$$

$$y_{p+1} = \alpha_{p+1} + \beta d_{p+1}(z_{p+1}) + \mu + \nu_{p+1} \tag{7}$$

and, first-differencing,

$$y_{p+1} - y_p = (\alpha_{p+1}-\alpha_p) + \beta[d_{p+1}(z_{p+1})-d_p(z_p)] + (\nu_{p+1}-\nu_p) \tag{8}$$

The leading case in the tax application is that in which income or some function of income, which determines the individual marginal tax bracket, is used for z. Thus consistent estimation of β again requires that $[d_{p+1}(z_{p+1})-d_p(z_p)]$ and $(\nu_{p+1}-\nu_p)$ be uncorrelated. This is a much stronger assumption than has been needed thus far because if z is jointly determined with y (as income and labor supply are, for example), then z_{p+1} is likely to be correlated with ν_{p+1}, and z_p with ν_p. Hence the fundamental exclusion restriction necessary for the differences-in-differences approach is in jeopardy.[14]

The conventional solution to problems of endogenous regressors is to seek correlates of those regressors which satisfy exclusion and

---

[14] Note that the issue of whether z_{p+1} is affected by the change in law is irrelevant. The issue is instead whether the values of y_{p+1} and z_{p+1} are chosen together, in which case there will be a dependence between them which is independent of the law change and hence could introduce a spurious relation between $d_{p+1}(z_{p+1})$ and y_{p+1}. Of course, in many cases one would expect the law change also to affect z_{p+1} but this is not necessary for bias to occur.

other restrictions for identification. Instrumental variables (IV) is one method, among others, for consistent estimation subject to those restrictions. In the IV case we seek an instrument which is asymptotically correlated with $[d_{p+1}(z_{p+1})-d_p(z_p)]$ but not with $(\nu_{p+1}-\nu_p)$ and is excluded from eqn(8). The classes of instruments which can be sought for this purpose are precisely the three which we have already discussed--time-invariant and exogenous x, time-variant and exogenous x_p, and time-invariant but endogenous z--in each case again satisfying the exclusion and orthogonality restrictions in the first-differenced equations which we have already discussed for these three classes of variables. With $d_p(x)$, $d_p(x_p)$, and $d_p(z)$ now reinterpreted as to-be-estimated functions of instruments, all of the above analysis applies. Thus the analysis at this point comes full circle back to the original three cases, with time-invariant exogenous variables x with stationary coefficients constituting presumably the strongest instruments.

In our empirical discussion below, we will be more specific about the types of instruments in the labor-supply-tax application that might satisfy these conditions. However, here we shall discuss an approach used in a number of prior applications, namely, the use of the period-p value of z_p as an instrument (e.g., Feldstein,1995a). The variable z_p is an endogenous but time-invariant variable (if it is held constant through p+1, that is) and hence, assuming it is both correlated with the change in the tax variable and independent of $(\nu_{p+1}-\nu_p)$, it is a candidate instrument. In the two-stage-least-squares version of its application, $[d_{p+1}(z_{p+1})-d_p(z_p)]$ is regressed on z_p and its predicted value replaces the actual value in eqn(8). In an alternative version, one linearizes the tax schedule with the approximation

$$d_{p+1}(z_{p+1}) = \theta_0 + \theta_1 d_{p+1}(z_p) + \eta \tag{9}$$

and uses predicted values from estimates of this equation in place of actual $d_{p+1}(z_{p+1})$ in eqn(8). Consistent estimation requires in either case that the predicted values be asymptotically uncorrelated with $(\nu_{p+1}-\nu_p)$.[15]

The difficulty with this instrument is that z_p is unlikely to be correlated to the same degree with ν_p and ν_{p+1}, and hence is likely to be correlated with the difference $\nu_{p+1}-\nu_p$. Because z_p and y_p are jointly determined--either because z_p is equal to y_p (i.e., if the lagged dependent variable is used) or is a direct function of y_p (as income is of labor supply)--the transitory error ν_p will have a direct effect on z_p. This covariance will translate into a dependence of z_p on the differenced error, $\nu_{p+1}-\nu_p$ because z_p will almost certainly not be related to ν_{p+1} in the same way it is related to ν_p. For example, if ν_{p+1} and ν_p are independent, there will be no relation between z_p and ν_{p+1}) and the resulting bias will take the form of regression to the mean.

The influence of ν_p could be accounted for by entering z_p directly and independently into eqn(8)--essentially controlling for

---

[15] In one case it is z_p and, in the other, it is $d_p(z_p)$ that must be uncorrelated with $(\nu_{p+1}-\nu_p)$.

the lagged dependent variable--but then identification would be lost because the change in d would have no variation independent of z_p; in this sense the issue is an identification problem more than a regression-to-the-mean problem. But if z_p is entered independently in the regression, some other instrument is needed to address the initial endogeneity problem, and there one again returns to the need for one of the classes of instruments discussed previously which satisfies the same set of conditions.[16,17]

A variant of this procedure which has apparently not been reported in the published literature is the use of z_{p+1} as the instrument.[18] That instrument qualifies under the same conditions as z_p; eqn(8) is perfectly symmetrical w.r.t. periods p and p+1, and the fact that period p+1 is after the tax law change has no direct bearing on the validity of z_{p+1} as an instrument. To the contrary, there is little a priori reason to suppose that the correlation between z_{p+1} and ν_{p+1} will differ from that between z_p and ν_p. Unfortunately, if both are tested as instruments and the estimates of β are the same,

---

[16] Auten and Carroll (1998) entered the period-p value of income as a control. Identification rested in their case on other variables (state-level tax rates, composition of income, etc.).

[17] Another approach to the problem would be to utilize data for additional periods in the past. Assuming that tax rates had not changed over those periods, and that the regression-to-the-mean effect is stationary, that effect could be estimated from past periods' data and then "subtracted" off of the effect estimated from period p to p+1. The additional restriction needed for identification is that the autocorrelation is of order one and hence there is no direct additional regression-to-the-mean effect from periods prior to p. See Moffitt (1998) for a discussion of models of this type.

[18] We thank Joel Slemrod and Lillian Mills for pointing this out to us.

this can arise either because there is no bias or because the bias is the same for both. If the estimates differ, it is likely that they will be biased in opposite directions and this can indicate the presence of serial correlation in the errors. In the simple case where the instrument is y_p or y_{p+1}, which contain ν_p and ν_{p+1}, respectively, the covariances between the error term in eqn(8) and these two instruments are $[Cov(\nu_p, \nu_{p+1})-Var(\nu_p)]$ and $[Var(\nu_{p+1})-Cov(\nu_p, \nu_{p+1})]$. Assuming the variances are the same in the two periods, the estimated β using y_p as the instrument will be higher (lower) than the estimated β using y_{p+1} as the instrument if serial correlation is positive (negative).

Repeated Cross-Sections. Because our empirical work will use panel data, we will not discuss the application of the principles just outlined to data consisting of a series of repeated cross-sections. However, we provide in the Appendix a summary of the issues that arise in that case. As the analysis there shows, the models discussed above which rely on time-invariant x or z for identification can be applied to repeated-cross-section data with only small modification, and consistent estimates of β obtained under the same conditions. However, models using time-variant x require additional assumptions for identification, and models using time-variant z are very difficult, if not impossible, to use with repeated cross-section data without the imposition of implausible restrictions.

Piecewise-Linear Tax Schedules. The federal income tax creates a piecewise-linear budget constraint from which individuals can choose labor supply locations. The econometrics of this problem have been analyzed extensively in past work (Hausman, 1985; Moffitt, 1986,1990; MaCurdy et al., 1990; Blundell and MaCurdy, forthcoming). The implication of this body of literature for present purposes is that the interpretation of the coefficient on the marginal tax rate variables that we estimate, and which other investigators have estimated, must be interpreted with caution.

The object of interest in that literature has generally been the estimation of the parameters of a static utility function U(H,C)-- where H is hours of work and C is consumption. If the labor supply function is linear, those parameters are the coefficients in the equation for H if utility maximization occurs on segment s of the constraint:

$$H = \alpha + \beta W[1-t_s(x)] + \delta\tilde{N}_s(x) + \gamma x + \epsilon \tag{10}$$

where W is the hourly wage rate, $t_s(x)$ is the marginal tax rate on segment s for an individual with characteristics x, and $\tilde{N}_s$ is virtual nonlabor income for segment s. Aside from the interaction between W and the marginal tax rate, $t_s(x)$, and the presence of the virtual income variable, eqn(10) fits into the framework of eqn(1) that formed the basis for the econometric analysis above.

Unfortunately, as shown in the Appendix, eqn(10) does not correctly represent the determination of H along segment s observed in a cross-sectional data set because of segment classification error. Such error is necessarily present if the variance of ϵ is nonzero. Instead, H observed along a segment s is determined by a weighted average of marginal tax rates on all other segments of the constraint.

Further, first-differencing in the manner of the differences-in-differences, fixed effects model does not lessen this problem. In light of these problems, consistent estimates of the effect of $t_s(x)$, or of $W[(1-t_s(x)]$ on H cannot be interpreted as representing estimates of ß in eqn(10). Instead, those estimates must be interpreted as the net effect of a change in the marginal tax rate in one segment on H, including those effects arising from correlated changes in the marginal tax rates of other segments. This is the interpretation we will give to our parameter estimates.

We should also note at this point that the static labor supply theory clearly implies that an income term should be included in the equation and that the wage rate should be interacted with the marginal tax rate, regardless of nonlinear constraint issues. We will test an income term in our models and we will also test interactions of W with the marginal tax rate in our empirical work. However, we will not use the theoretically-implied interaction between W and the marginal tax ratet as a source of identifying variation on the presumption that the effects of the two variables may be different for a variety of reasons.[19]

Applying the Methodology. In the labor-supply-tax case we study the federal income tax and its effect on hours of work. There are

---

[19] One reason is that there may be omitted variables correlated with the wage rate that would bias its coefficient; another is that the theory is potentially misspecified and that some behaviorally important differences in individual responses to wage rates and tax rates are left out of the model. Reasons for, and tests of, the hypothesis that wage and tax effects are different were discussed many years ago in the negative-income-tax experiment and related literatures. See Moffitt (1979, p.480), Moffitt and Kehrer (1981,pp.106,123), and Rosen (1976).

many variables in the federal income tax code that affect the individual's marginal tax rate which constitute candidates for x or z. These variables include adjusted gross income (AGI); deductions, exemptions, and filing status, which determine taxable income; and various tax credits and adjustments for other taxes. Each of these categories include subcategories as well. However, few of these variables are direct candidates for x, x_p, z, or z_p, for most are related too closely to income and hence labor supply and hence are likely to be endogenous. Earned income is clearly in this category, but unearned income in its many forms is as well, for the majority of that income arises from investment decisions that are probably jointly made with labor supply decisions. As for the remaining variables that go into the tax formula, we are constrained by our data, which are household survey in nature (see below) to those which were obtained in the questionnaire. The only two major non-income tax-formula variables in our data are marital status, which is highly correlated with filing status, and family size, which is correlated with the number of exemptions. We will test both of these variables as instruments.

When instruments for the endogenous earned and unearned income variables are considered, a larger number of instruments might seem to be available. Any instruments which can be thought of as determinants of permanent income or wages are candidates because they should be correlated with contemporaneous income and hence tax rates, but uncorrelated with the transitory income components which are likely correlated with the change in labor supply. In this category we consider education and broad-category occupation, which are both roughly constant over short periods of time.[20] We will also test as instruments various forms of assets which are moderately illiquid in form, such as the value of a house or the value of life insurance. Because these assets are fairly illiquid and do not generate cash income flows, they should not be directly correlated with contemporaneous income but should be correlated with permanent income.[21]

As noted previously, because the static labor supply model implies that W is interacted with (1-t), we also test interactions of our instrumented tax variable with the wage rate and with its predictors (such as education). We also test specifications which directly incorporate income effects.

We will also test the use of z_p (the pre-law-change value of income or AGI or hours of work) as an instrument, as well as various transforms of z_p. Of the transforms, that which we will use most heavily is the period-p value of the marginal tax rate, $t_p(z_p)$. This is the instrument used by Feldstein (1995a). We will also test including z_p as an independent regressor and using the other

---

[20] Eissa (1996a) and Blundell et al. (forthcoming) both used education as a instrument (both allowed education to affect labor supply in levels, but assumed that it disappeared in differences). Note that we do not test the wage rate itself as an instrument partly because it should appear explicitly in the labor supply equation, but also because we regard it as a choice variable and one which may respond to changes in tax law, as emphasized by Feldstein (1995b).

[21] We reemphasize that eliminating serial-correlation and regression-to-the-mean effects, by using instruments orthogonal to transitory errors, is necessary but not sufficient for consistent esimation; it is also required that the coefficient on the instrument (e.g., permanent income or its predictors) not change over time.

instruments we have described to identify the model; this will control for regression-to-the-mean and other serial correlation effects.

The major focus of our empirical work is specifically on the labor supply response of the rich. To maintain this focus we will will test instruments which stratify the population into groups which separate individuals in the upper tail of the distribution from the rest of the population. Thus our instruments will be variously formulated as those with very high period-p income, period-p marginal tax rates, very high education or high-earning occupations, and very high asset levels.

As we have stressed in our analysis above, the major condition needed for validity of the instruments is that their effects on labor supply be constant over time. Obtaining evidence on this question is not possible with only a two-period, before-and-after panel such as that we use, but indirect evidence can be obtained from other data sets. The Current Population Survey (CPS) contains information on income, earnings, and labor supply for a number of years and also on education, occupation, marital status, and family size. Figures 1-4 provide information on the a priori validity of education and occupation in this respect. Figure 1 shows trends in annual hours of work for prime-age men in high-earning occupations (professionals and managers) and all others, while Figure 2 shows such trends for those with high education (college degree or more) and all others.[22] Interestingly, the figures demonstrate relatively little trend in the

---

[22] The figures use all working men 25-54 in the year in question in the CPS.

hours-worked gap prior to 1986 for either variable, suggesting that they might be suitable as instruments.[23] The figures also show little evidence of a widening of the gap after 1986, as well, but this has no bearing on the validity of the variables as instruments because the true effect of the law affects the post-1986 trends. These results do not extend to family income and male earnings; those variables significantly widened prior to 1986 both between the two education groups and the two occupation groups (figures not shown).

III. Data and Results

A. The Survey of Consumer Finances

The Survey of Consumer Finances (SCF) is a household survey conducted to gather financial information from a nationally representative sample of American households (Kennickell and Shack-Marquez, 1992). Since 1983 the SCF has been fielded triennially, and in 1983 and 1989 the surveys had a panel feature for which a subsample of households appeared in both. The primary focus of the survey is on wealth information, and considerable detail is devoted to the composition of financial and nonfinancial assets and various types of debt. Because wealth holdings are very concentrated at the top of the wealth distribution, the SCF oversamples high-income households. The relatively large size of the affluent sample in the SCF has been used for estimating aggregate wealth and studying trends in wealth

---

[23] We confirmed this with simple regression tests, which yielded insignificant coefficients for the difference in trends for the two groups prior to 1986.

inequality (e.g., see Wolff 1994,1995). We will use this oversample feature of the SCF to analyze the labor supply of the rich.[24]

We use only the 1983 and 1989 waves of the SCF and the panel of individuals who appeared in both. The 1986 Tax Reform Act occurred midway between these years but not close enough to either to warrant concern about contamination due to timing responses. The fortuitous fielding of the SCF before and after the Act makes it particularly useful both for purposes of reexamining the effect of the tax changes on adjusted gross income (AGI) with data other than those available from tax returns and for investigating the response of labor supply to the Act. An additional advantage of the SCF is that it contains data which enable us to examine the sensitivity of our results—and, by extension, the previous results of others—to the use of a fairly wide array of alternative instruments. Importantly, many of these instruments are not based on income, or some function of income, from the first period.

The SCF obtained AGI information using different methods in 1983 and 1989. In 1989, respondents were queried directly about 1988 AGI in a section of the survey dealing with their Federal income taxes.[25] In 1983, there was no direct query of AGI but rather the SCF constructed two measures of respondent AGI from the responses to questions in the income and household characteristics sections. One

---

[24] Because of this stratification we use the SCF-supplied weights for all our analyses.

[25] Aggregate AGI estimated from these responses exceeds published IRS totals by $400 billion, or 13 percent (U.S. Internal Revenue Service, 1991).

measure was designed to be current law (1982) AGI and the other was constructed to include full capital gains and the dividend exclusion (unlike 1982 law). When the weighted 1983 cross-section SCF is used to generate aggregate 1982 AGI, the first measure underestimates the published IRS totals for 1982 AGI by $200 billion, or 11 percent (U.S. Internal Revenue Service, 1984), but the second measure (adjusted to current 1982 law by subtracting 60 percent of capital gains and the dividend exclusion) differs from the IRS totals by only $1 billion.[26] Hence we use the second measure. Because it already includes full capital gains and the dividend exclusion it is comparable in definition to 1989 (tax law 1988).

We use several criteria to select a sample for analysis. Appendix Table A1 provides a summary of these criteria as well as their effects on sample size. We analyze male heads of households aged 25-54 in 1983. The age restriction implies that the oldest men were no more than age 60 in the second period of the panel and, for the most part, were likely not considering retirement decisions. We select men for whom there was no ambiguity in linking 1983 data from the household record to the data from 1989. Lastly, we analyze men who had positive AGI, positive wage rates, and who worked more than 200 hours in both years; we also exclude the few observations whose labor hours were imputed in the 1989. Our final sample consists of

[26] This correspondence is in large part a result of the weight we use (called the "panel weight" in the SCF), whose construction was partially based on post-stratification to match IRS tables on AGI with full capital gains and the dividend exclusion. There is another weight on the SCF which conducts further stratification but we do not use it; our results appear not to be sensitive to which weight we utilize.

490 men.

We calculate a marginal tax rate (MTR) for each observation in each year from the data available on the SCF, using tax rules applicable in 1982 and 1988. To calculate an estimate of taxable income, AGI (in 1982, AGI is reduced by 60 percent of capital gains and the dividend exclusion) is reduced by the number of household members times the exemption amount and by an estimate of average deductions of those with similar AGI based upon published IRS tables (U.S. Internal Revenue Service, 1984, 1991).[27] This estimate of taxable income is then used with the tax tables to determine the MTR as well as the value of the tax payment. For 1983, the MTR is reduced by 5 percent if the deduction for a second worker was effective (i.e., the man's earnings had to have been under $30000 and less than his wife's).

Table 1 shows the means and standard deviations of the major variables in the analysis in 1983 and 1989. The sample for this table includes only 406 men with 1983 MTR greater than 0.20 for reasons we will discuss below (this is the sample closest to that used by Feldstein (1995a)); means for the entire sample of 490 observations as well as for those with lower 1983 MTR values are presented in Appendix

---

[27] Thus we take deductions as exogenous; see Triest (1992) and Feldstein (1995c) for a discussion of this assumption. The IRS publishes the percent of returns which itemized deductions and the average amount of those deductions by AGI category. We calculate a weighted average of the standard deduction and itemized deductions, using the percent itemizing and the amount of deductions if itemizing. For filing status, we use a direct question on the 1989 SCF but because no direct question was asked in 1983, we treat all married couples as filing jointly and all single men as filing singly in that year.

Table A2. The last four columns in Table 1 subdivide the sample into groups with midrange 1983 MTR values (from .20 to 0.44) and high 1983 MTR values (over .44). Approximately 96 percent of the sample is in the former group and 4 percent is in the latter group; thus the latter is our "high income" sample.

The first several rows in Table 1 shows mean 1983 and 1989 AGI, hours of work, and other outcome variables of interest.[28] While AGI grew for midrange-MTR men it grew more in both absolute and percentage terms for high-MTR men. Annual hours worked (calculated from the product of normal weekly hours and normal annual weeks worked) increased for both MTR groups but by approximately the same amount.[29] Total income, wage and salary income, and the latter combined with business income also increased both all men but more for those with high 1983 MTR values.[30]

A key variable in the table is the net-of-tax rate (NTR), equal to one minus the marginal tax rate. Between 1983 and 1988 the NTR increased much more for those with high inital MTR values than for

[28] All monetary values in our paper are in 1988 dollars using the personal consumption expenditure deflator. We will continue to refer to "1983" and "1989" AGI even though the SCF follows the usual survey practice of obtaining this and all other income data for the year preceding the survey.

[29] We investigated SCF measures of other measures of work effort and labor market behavior and compensation--self-employment, executive and deferred compensation, and others. Unfortunately, these measures were either not well-defined at all in the SCF or their definitions changed between 1983 and 1989 (as in the case of self employment).

[30] The income questions in the survey are separate from those for AGI and hence provide an independent measure of income. The total income measure includes wage and salary income, business income, interest and dividends, capital gains, rents, pension income, transfers, and all other sources of income in the household.

those with lower values, consistent with many prior calculations of the effect of 1986 Tax Reform Act (e.g., Hausman and Poterba, 1987). It is this differential effect that forms the basis for all the differences-in-differences, fixed effects estimates in this paper and in much recent work. The table also shows that both gross and net hourly wage rates increased over the period but more for the high-initial-MTR group.

The rest of the variables in Table 1 will be used in the subsequent analysis as control variables (particularly marital status, household size, and age) and/or instruments for the change in the NTR. In all cases only the 1983 value of the variable is used. In addition to the distribution of the observations across seven (rather than two) 1983 MTR groups, the table shows the means of several additional variables. These include a high-income ("rich") dummy, equal to one if 1983 total income exceeded $100,000; a dummy for educational experience after college; a dummy for those in professional or managerial occupations; variables for the value of a house and of life insurance plans, as well as a dummy for those either owning an expensive house (greater than $200,000 in value) or holding a large amount of life insurance (greater than $300,000).

B. Results

While our major focus is on hours of work, we initially benchmark our results against those of Feldstein (1995) both to determine whether our data give similar results to his for AGI as well as to illustrate the use of alternative instruments. We find results for AGI in our data quite similar to those of Feldstein although we also find the magnitude to be somewhat sensitive to the use of alternative instruments.

AGI Results. Table 2 shows the estimates of effects of the 1986 Tax Act on AGI using a tabular methodology similar to that of Feldstein, and Table 3 shows those estimates using a regression methodology. The instrument used in the Feldstein model is the initial-period level of the marginal tax rate (MTR) grouped into categories. We construct three groups--low (less than .20), midrange (.20-.44) and high (.44-.50)--which differ slightly from a four-group categorization used by Feldstein.[31] The first group, those with low MTRs, is omitted by Feldstein from the sample so we also omit that group for our initial analysis (but we add that subsample back in subsequently). In Table 2, the first two columns show changes in AGI and in the NTR for the midrange and high MTR groups. For linear differences (ΔAGI and ΔNTR), it can be readily seen that the high MTR group experienced both greater increases in AGI and in the NTR. The magnitudes imply that a .01 increase in the change-in-NTR is associated with a large absolute increase of $8,900 of annual AGI (1988 dollars). Converted to an elasticity at the means of the data, this yields a sizable elasticity of 1.992. Feldstein estimated elasticities in the range (1.10,3.05) for a taxable-income-related

[31] His four MTR groups were those less than .22, 22-.38, .42-.45, and .49-.50. The major difference is that we collapse his upper two groups into one for sample size reasons.

concept and (.26,.88) for a AGI-related concept.[32] While the latter is closer in concept to our income definition than the former, our estimates are closer to the former range. Given the marked differences in the way income information is obtained in the two surveys, our estimates should be judged to be reasonably consistent with those of Feldstein.

Feldstein calculates his elasticities somewhat differently, however, by first calculating mean AGI and NTR for each group, then calculating the percentage change in that mean between the years, and then using the difference in the differences of these percentages for his calculations. As shown in Table 2, when we apply this method to our data, we obtain an elasticity estimate of 1.828, quite close to the linear-difference calculation and again reasonably consistent with the Feldstein elasticities. We also show in Table 2 a third possible means of calculating an elasticity, by computing percentage changes in AGI and NTR at the individual level and by then computing an elasticity from the means of these percentage changes. This yields an elasticity of 1.757, close to our estimates from the other methods.

Of the three methods of calculating elasticities reported in Table 2, only the first and third--not the second, which is the precise method used by Feldstein--can be formulated in regression terms. We do this in Table 3, where we show IV estimates of two types of equations, one for which the linear change in AGI is the dependent variable and one for which the percentage change in AGI (for the

---

[32] The first range of elasticities is taken from Table 2 (Feldstein, 1995, p.565) and the second range has been calculated by the authors from the figures in Table 1 (Feldstein, 1995, p.561).

individual observation) is the dependent variable. In both cases there is a single regressor, which is either the linear change in the NTR or the percent change in NTR. Instrumental-variables is applied by using a single dummy, in effect--whether the 1983 MTR is in the high category--as the instrument.[33] Table 3 shows, for both methods, both the reduced forms and the first-stage regressions, whose coefficients are identical to the entries in Table 2. The second-stage IV coefficient on the change in NTR is 890,000 in the linear model and 1.757 in the percentage-change model, thus replicating the estimates in Table 2; the former must be converted to an elasticity, which we showed to be 1.992 in Table 2. Table 3 shows standard errors on the estimates which are far below the coefficient magnitudes and which hence imply highly significant effects.

Table 4 shows the effect of adding additional independent variables to the model as well as, more importantly, the effect of using alternative instruments, on the second-stage estimated coefficient on the linear change in NTR in regressions for the linear change in AGI. The first row shows the coefficient that results when a number of additional regressors are included (in both the first and second stages)--family size, marital status, and age. This addition has little effect on the coefficient. The second half of the table shows the F-statistic on the instruments, the p-value for that

---

[33] We say "in effect" because, for illustration purposes in Table 3, we omit the constant term in both the reduced form and first-stage regressions and include both a high and midrange MTR dummy. This generates coefficients which are comparable to the figures in Table 2. In the remainder of the paper we include the constant and only the high MTR dummy.

statistic, and the R-squared of the first-stage regression; the high F-statistic on the instrument (i.e, the MTR dummy) shows this instrument to be strong.

The second row shows the effect of using seven separate 1983 MTR groups as instruments instead of the two used heretofore (see Table 1 for group definitions; the highest of the groups is the same as .44-and-over group). Interestingly, the NTR coefficient loses significance in this specification. The source of the difference is illustrated in Figure 5, which shows the change in AGI between the time periods for the different MTR groups. The figure includes the less-than-.20 MTR group, so it shows three and eight MTR groups instead of two and seven, respectively. For all of the eight MTR groups except the highest (MTR between .44 and .50), the relationship between initial-period MTR (and hence the change in NTR) is flat or negative; but the highest-MTR group has a very large increase in AGI. Thus it appears that it is the highest MTR group that is responsible for the positive elasticities being estimated.[34]

Because initial-period MTR is primarily a function of initial-period AGI, we examine whether using AGI itself as the instrument would alter any of these conclusions. The third and fourth rows of Table 4 show that it would not. Using a dummy for high 1983 income, a positive and significant elasticity is obtained in the same range as that obtained by using the two MTR groups. But when the log of AGI is

[34] There are five or ten very large AGI gains and losses (over $1 million in absolute value) in the data in the upper group. Deleting these extreme values reduces the magnitude of the coefficient but it remains statistically and economically significant.

used--and thereby not making a special distinction between the highest AGI group and the rest of the population--the estimated coefficient is significant but drastically reduced in magnitude.

That initial-period AGI is the implicit instrument in this approach, even if MTR groups are used, brings the two issues described in Section II into consideration. Regression-to-the-mean effects in AGI--or, more generally, serially-correlated errors--will bias estimates which use AGI as an instrument. In addition, even if a measure of permanent income or AGI were used (i.e., one purged of serially-correlated transitory components), the more fundamental issue of whether its coefficient has been changing over time arises. For these reasons we test several alternative instruments shown in Table 4. First, we test the 1983 values of marital status and household size because these both enter the tax formula independently of AGI.[35] However, as the table indicates, they are extremely weak instruments--they do not discriminate between different change-in-NTR values well--and yield insignificant results. Inasmuch as the results using AGI and MTR instruments have indicated that positive tax effects are arising only from the very top of the distribution, the fact that marital status and household size do not well discriminate between that upper group and the rest of the population makes their

[35] These instruments are time-varying but exogenous (by assumption) and hence are the type of instruments for which we recommended in Section II that only those with no change be included. Hence our estimates for these instruments only include those with no change in marital status or household size from 1983 to 1989.

insignificance not unexpected.[36]

We next test education and occupation as instruments. To discriminate to the greatest extent possible between the upper tail of the distribution and the rest of the population, we construct a dummy for whether an individual has post-college educational experience and a dummy for whether an individual is in a professional or managerial occupation, the highest paid occupations. As Table 4 shows, the occupational dummy is a very weak instrument but the education dummy is not (perhaps because the occupational dummy has a mean of only .45); nevertheless, even the latter yields an insignificant tax response estimate. However, the tax-response estimate when education is used is still positive and sizable in economic terms, even though its standard error is also quite large, indicate imprecision in the estimate.

We test in the last two rows two measures of assets which are available in our data--the value of an owned house and the value of life insurance. These variables, while financial in nature, are sufficiently loosely connected to current income flows as to increase their likelihood of exogeneity and, similarly, are less likely to be

[36] We also tested the 1989 MTR group and the log of 1989 AGI as instruments, as discussed in Section II. The estimated tax response coefficient becomes, surprisingly, negative and significant in this specification. However, when a dummy for high 1989 AGI (the top 4 percent of the distribution), the coefficient becomes positive and significant once again. A plot analogous to Figure 5, but using seven 1989 AGI groups, shows the 1989 response to occur only in the top part of the distribution. Thus the two years are consistent with each other at the top end. The top 1989 MTR group does not yield a positive response coefficient because the top bracket in that year includes almost 30 percent of the population (unlike the top 1983 MTR group).

affected by regression-to-the-mean effects than AGI. In addition, assets are less equally distributed than income or the other instruments we have tested and hence have a better chance of discriminating between the top earners and those below. However, asset values are subject to the trending-coefficient problem because asset inequality has been growing (Wolff, 1994,1995).

We test a set of instruments which include the log house value, log of life insurance value, and dummies for those with zero house value and life insurance; and a dummy for whether either is high (see discussion of Table 1 for exact definition). As the results in the table show, these instruments are strong in the first stage and also yield significant estimated tax response coefficients, albeit only about two-thirds the magnitude of those using the two MTR or top AGI-group instruments.

Finally, we show in Table 5 the effects of adding the low-MTR group back into the sample (which we have continued to exclude, for comparability with the Feldstein analysis) as well as tests for the importance of regression-to-the-mean effects. Adding the low MTR group into the sample lowers the estimated tax effect arising when the small MTR-group instruments are used.[37] We also show results from using the asset instruments because they are both strong instruments and yielded significant results in Table 4; the estimated coefficient falls slightly when the low MTR group is added as well (from .649 in Table 4 to .552 in Table 5, for example, for the log asset

[37] When adding the low MTR group into the sample, we retain only one instrument, a dummy for being in the high MTR group.

instrument). With these instruments we can also test for regression-to-the-mean effects by entering AGI into both the first and second stage equations.[38] As Table 5 shows, controlling for AGI in this way increases the estimated tax-response coefficient. This should be expected because pure regression-to-the-mean effects would tend to bias the coefficient in a negative direction (those with positive 1983 transitory errors should experience declines in AGI over time).

We thus have replicated the sizable tax elasticities for AGI found by Feldstein (1995) and have shown that those elasticities arise from behavior of the extreme upper tail of the income distribution which is quite discontinuous with that of the rest of the population. Instruments which are successful in discriminating between that top group and the balance of the population, even if they are instruments not strictly AGI-based (e.g., asset instruments), yield similarly sizable tax elasticities even if regression-to-the-mean effects are accounted for.

Results for Hours Worked. Having tested instruments for AGI, we now turn to hours of work and apply the same strategy and test the same set of instruments. Figure 6 shows the distribution of 1983 annual hours worked by the three 1983 MTR groups we used for the AGI analysis. The distribution is remarkably different for the high MTR group and the rest of the population, with about 60 percent of the high MTR group working more than 2500 hours per year and almost 30 percent working more than 3000 hours per year. For the rest of the

[38] We could also have tested AGI in the MTR-based instruments but the inference would be weak because identification would rely entirely on nonlinearities in AGI effects.

population, the mode is in the typical 1750-2250 range. Given these high hours of work, there is at least some prima facie question of whether there is much opportunity for additional work among the rich.

Table 6 shows IV estimates of the effect of NTR on annual hours worked, using the same methodology as in Table 4--exact same specification and sample for each equation but with a different dependent variable. As the table indicates, none of the effects are significant except that for the high asset group, and that effect is negative. The strength of the estimates, shown in the latter columns of the table, is necessarily the same as in Table 4; thus the insignificance of the estimated effects cannot be ascribed to weakness of the instruments.

Figure 7 shows the pattern of changes in hours worked over the period by 1983 MTR group, in analogy to Figure 5. The relative hours changes for the midrange and high MTR groups is slightly positive but small in magnitude; the coefficient in Table 6 is negative because of the addition of the other independent variables, but it is still insignificant.

Even if the lack of hours response of the upper tail of the distribution can be ascribed to hours which are already near their maximum, this is not true for the rest of the population. Indeed, the very high hours worked of the upper tail is uncontrovertible evidence that hours of work are fundamentally flexible upward in the U.S. labor market for those who are working "only" 2000 hours per year (i.e., year-round fulltime). Yet Figure 7 does not show any particular positive relationship between initial MTR (and hence the change in

NTR) and the change in hours worked. Nor do the instruments in Table 6 which treat all parts of the population distribution equally show any more positive responses than those which focus on the upper tail. Consequently, the evidence in these data is that hours of work are, as found in much previous work, inelastic for prime-age males in the U.S.

Table 7 provides several additional specifications to test alternative hypotheses for hours effects. Adding the low MTR group into the sample has no effect on the significance of the tax effects, nor does controlling for regression-to-the-mean effects in hours of work. We also show in the last two columns of Table 7 a specification which includes the change in tax payment as well as the change in the NTR. This specification approximates more closely the neoclassical labor supply function by accounting for income effects.[39] The results show insignificant income effects and do not change the insignificance of the NTR effects. The house value and life insurance variables are extremely weak instruments for the change in tax payment which could, in principle, be responsible for this result.

We turn in Table 8 to tests for whether NTR effects might be significant on hours of work for some subportions of the distribution. Our primary motivation for this exercise is that the theoretically appropriate price of leisure is W(1-t) which, unless affecting hours of work in simple logarithmic form, implies that the percentage effect of a change in NTR should vary with the value of the wage rate. For

---

[39] The tax payment in both years is evaluated at the actual AGI in those years. Use of this income variable rather than nonwage income converts the coefficient on the NTR variable to compensated form (Ashenfelter and Heckman, 1973, 1974).

completeness, we also test such interactions for AGI as a dependent variable. We test the same sets of samples and instruments shown in Table 7.

The results in Table 8 show no effects of this type to be present in the data for hours of work. When the NTR change is interacted with our post-college dummy (a predictor of the wage), the interaction coefficients are insignificant in all cases save one where the coefficient is a counterintuitive negative. When the net wage itself is treated as the endogenous variable of interest, and is instrumented accordingly, the same pattern results.

Interestingly, several positive and significant education interaction effects are found when AGI is the outcome of interest. Indeed, for some specifications the AGI effects are insignificant for the population without post-college experience. These results are consistent with the hypothesis that higher-wage taxpayers respond more heavily to changes in their marginal tax rates than those with lower wages. However, these results are sensitive to adjustment for regression-to-the-mean. As Table 8 shows, when such adjustment is made the AGI results disappear for the high-educated group but are stronger for the rest of the population; interestingly, hours of work effects appear for the latter as well.

Income Decomposition. To explore the mechanism by which the AGI tax response is occurring, we briefly decompose income into three major constituent parts--wage and salary income, business income, and other income--and apply our same methodology to estimating tax response for these three variables. We should note that the sum of these three components, or total income, is not the same as AGI in our data set. Total income is the sum of all forms of income reported on the survey, while AGI is not only not from tax records, it is the response to a specific question on the survey.

Table 9 shows the results of this exercise. The first column shows the results of applying the Feldstein methodology to total income, wage and salary income, and the sum of wage and salary income and business income. Business income is included with wage and salary because it is zero for most of the sample; business income responses are necessarily equal to the difference in the coefficients in the second and third rows of the table. The coefficients for other income are also obtainable by subtracting the third row from the first. The Feldstein methodology shows significant tax effects on wage and salary income and its sum with business income, and the larger coefficient when business income is included indicates that business income is the largest source of the response. However, this result does not hold up when asset instruments are used and regression-to-the-mean effects are allowed, as the remaining columns show. The major change occurs when regression-to-the-mean effects are permitted, which wipes out the business income effect (in fact, it turns negative). This result implies that serial correlation in business income between 1983 and 1988 was positive, not negative; those with above-average (below-average) business income in 1983 had even greater (lesser) business income in 1988. Thus the implication of the table is that the large business income responses shown in the first column are incorrectly assigning differential growth rates of such income to the tax law change.

The tax response does remain for wage and salary income, however, and it is therefore this form of income that we conclude constitutes the major source of adjustment to the act. Because we have found no hours of work response, we therefore have found implicitly that the entire response to the Tax Act of 1986 for men occurred in hourly wage rates.

IV. Conclusions

A long-standing issue in the effects of taxation on individual behavior concerns whether labor supply, most commonly measured by hours of work, responds to taxation. We have examined whether high income men--the rich--so respond. High-income taxpayers are often thought to have more opportunities to respond to tax law changes and to have a greater incentive to do so because of their high marginal tax rates. Our analysis of changes in the hours of work of such men between 1983 and 1989, in response to the marginal tax rate reductions legislated in the 1986 Tax Reform Act, find essentially no evidence of any such response. We speculate that this is partly a result of the fact that such men are already working such long hours (often over 3000 per year) that there is little remaining opportunity for response.

The major limitation of our study for learning about the behavior of the rich in response to taxation arises from the limitations of the data in yielding information about other aspects of the labor force behavior of the rich. Incentives to work as self-employed and incentives to work in jobs in which compensation is deferred or otherwise tax sheltered are just two examples. Better data on these behaviors of the rich are required before further progress can be made in investigating them.

Appendix: Modeling Issues in the Use of Repeated Cross Sections and Implications of Piecewise-Linear Tax Schedules

Repeated Cross Sections

Here we discuss the application of the differences-in-differences, fixed-effect method of estimation with repeated cross-section (RCS) data instead of panel data. We assume we have two independent cross-sections of the population with information on y and x or z, but the individuals in the two are different.[40] Estimation of the models with time-invariant x or z is not difficult because the invariance of x and z implies that individuals in the two cross-sections can be matched to one another using common values of x and z; while they are not the same individuals, they are drawn from the same strata of the population. This also implies that all time-invariant error terms (like μ) will have the same mean for individuals with the same value of z in both populations. In the case of time-invariant x, eqn(1) can be pooled across periods to estimate

[40] As usual in these models, it must be assumed that there is no significant entry or exit from the population over time through immigration, birth, or mortality. See Deaton (1985) and Moffitt (1993) for more general discussions of estimation of models with RCS data and see Heckman and Robb (1985) for a discussion of estimation of the impact of interventions with RCS data.

$$y_t = \alpha_p + [\Delta\alpha_p] D_t + \beta[d_{p+1}(x) - d_p(x)] D_t + \beta d_p(x) + \gamma x + \epsilon_t \qquad t=p,p+1 \tag{A1}$$

where D_t equals 1 if t=p+1 and 0 otherwise. The coefficient on the change in law shown in brackets is identified (apart from nonlinearities) by virtue of the assumption that γ does not vary with p; if it did, then an extra term $D_t x$ would be required and the effects of the two x variables would be confounded with the effect of the change in law. Note that the separate $d_p(x)$ variable could either be allowed to have a different coefficient than that on the law-change variable or could be folded into it.

In the case of time-invariant z, eqn(4) can be pooled across periods to give

$$y_t = \alpha_p + [\Delta\alpha_p] D_t + \beta[d_{p+1}(z) - d_p(z)] D_t + \beta d_p(z) + \mu + \epsilon_t \qquad t=p,p+1 \tag{A2}$$

In this case the coefficient on $d_p(z)$ is a biased estimate of β because z and μ are not independent, but the coefficient on the change in d is asymptotically unbiased because that variable is independent of μ conditional on $d_p(z)$.[41]

Time-varying x and z raise more difficult issues because the populations with the same values of x and z in the two cross-sections are not composed of the same individuals. However, at least if the variable is exogenous (the "x" case), those with the same value of x in the two cross-sections will have the same mean y in the absence of an effect of the law. Consequently, in this case eqn(1) can be pooled across periods to obtain an estimating equation analogous to (A2), namely,

$$
\begin{aligned}
y_t = \alpha_p + [\Delta\alpha_p]D_t + \beta[d_{p+1}(x_{p+1})-d_p(x_p)]D_t \qquad\qquad & \text{(A3)} \\
+ \beta d_p(x_p) + \gamma[x_p(1-D_t)+x_{p+1}D_t] + \epsilon_t \qquad\qquad & t=p,p+1
\end{aligned}
$$

As in (A2), the separate term for $d_p(x_p)$ could be used to obtain a separate estimate of β or included in the first term in brackets for a single β estimate.[42]

On the other hand, if time-varying, endogenous z is the the variable used for identification, using RCS data is more

---

[41] This can be shown formally. The variable $[d_{p+1}(z)-d_p(z)]$ is a deterministic (though nonlinear) function of $d_p(z)$, while the variable D_t is independent of μ by the assumption of the time-invariance of μ.

[42] As in the panel-data case, identification issues arise if sufficient nonlinearities are introduced that confound the effects of an independent change in x from a change in x working through tax-law effects. A weakness of RCS data shows up in this case because, unlike the case of panel data, the sample cannot be subselected down to those with no change in x.

problematic. Pooling eqns(6) and (7) across periods, we have:

$$y_t = \alpha_p + [\Delta\alpha_p]D_t + \beta[d_{p+1}(z_{p+1})-d_p(z_p)]D_t + \beta d_p(z_p) + \mu + \epsilon_t \qquad t=p,p+1 \tag{A4}$$

Once again, the issue is whether z_p is independent of ϵ_p, and z_{p+1} of ϵ_{p+1}. It is difficult to generalize across all applications because the degree of jointness of y (and therefore of ϵ) and z will depend on the particular variables in question, but in many cases such independence will be unlikely to hold.

If the independence condition fails, the distribution of individuals with different values of z will change between the periods as will the mean of y among individuals with fixed values of z. Thus the implicit groups formed by different values of z will be endogenous, which will bias the estimated effects. The availability of lagged z_p in panel data made possible an approach which used z_p as an instrument (albeit with the regression-to-the-mean problems noted there), but this approach is not possible with RCS data.[43]

Piecewise-Linear Tax Schedules

The common approach to estimation of labor supply choice in the face of a bracket income tax system has been to specify the "marginal" labor supply function along a segment of the budget constraint--that is, labor supply as a function of the "local" marginal tax rate (or

[43] See the comment of Heckman (1996a) on Eissa (1996b).

net wage rate) and "virtual" nonlabor income (see references to this literature given in the text). Assume that the marginal tax rate in bracket s is $t_s(x)$ (s=1,...,S) and that the value of income (or a transform of income, like AGI) at the beginning of bracket s is $a_s(x)$, where x is a set of socioeconomic characteristics that affect the individual's tax position (i.e., variables affecting AGI or affecting which schedule is applied, such as filing status. These 2S parameters characterize the tax system completely for a taxpayer with characteristics x. Maximizing a utility function U(H,Y-T;x) along segment s, where where H is hours of work, Y is gross income, T is the amount of the tax payment, and x is a vector of exogenous socioeconomic characteristics that affect preferences for work, gives the 'marginal' labor supply function

$$
\begin{aligned}
H &= g\{W[1-t_s(x)],\tilde{N}_s(x);x\} + \epsilon \\
&= \alpha + \beta W[1-t_s(x)] + \delta\tilde{N}_s(x) + \gamma x + \epsilon
\end{aligned}
\qquad \text{(A5)}
$$

as given in eqn(10) in the text and with variables as defined there.

If individuals observed to locate on only one segment in a cross-section are used for estimation of (A5), the model is not identified apart from nonlinearities in the sense discussed in the text.[44] Thus the basic identification problem posed in the text is present here as

[44] This identification problem is not confined to single-segment estimation but is also the present if the model is estimated in reduced form on all observations, i.e., if H is regressed on all marginal tax rates in the schedule, as in Blomquist and Newey (1996).

well. Variation in the net wage and virtual nonlabor income can instead be obtained by pooling the data across segments, because different individuals with the same x will usually choose a variety of segments. However, this variation is endogenous because the segment upon which an individual is observed is a function of ϵ, an error term which includes heterogeneity of preferences, measurement error, and "optimization" error (i.e., deviations from optimal choice arising from the cost of fine-tuning labor supply location relative to the brackets). Further, this endogeneity cannot be eliminated for the same reason as already discussed, namely, that there are no exclusion restrictions which, apart from nonlinearities in functional form, could be used to identify the model.

Formally, let D_s be a dummy variable equal to 1 if the individual is observed on segment s and equal to 0 otherwise. Then implicitly all variables in (A5) are multiplied by D_s. Denote by V the set of variables W, N, all 2S parameters of the tax schedule, and x. Then $D_s=f(V,\epsilon)$. If instrumental variables estimation is used to address the endogeneity, then identification will not be achievable (apart from nonlinearities) because all variables in V are already in eqn(A5) and there is no variation in the tax parameters in V independent of W,N, and x.[45] Thus obtaining variation by pooling across segments does not solve the identification problem.

With it therefore established that the fundamental identification

[45] Indeed, IV is not appropriate in this model in any case because x, in addition the net wage and virtual nonlabor income, is correlated with the error term and hence should, in principle, be instrumented. The mean of ϵ conditional on $D_s=1$ is a function of all variables in V, and therefore of x.

problem discussed in the text applies as well to the model when the piecewise-linear nature of the budget constraint is accounted for, it may be asked whether the use of first-differencing and the existence of a variable in x with stationary effects on H may permit identification here as well. In a fundamental sense, the answer is affirmative because the effect of tax rates is nonparametrically identified under those conditions and hence must be here as well. If $E(H_p|x)=f[T_p(x),x]$, where T_p is the 2S vector of tax parameters which change with time (p), then the existence of an x with stationary effects is equivalent to the assumption that p does not enter the function f independently or, equivalently, that the function f is not indexed by p. Two waves of a panel thus identify the effect of $T_p(x)$ on $E(H_p|x)$.

The question instead is what parameters are identified by this strategy and here the answer is that no simple function of the parameters in (A5) are identified. This is easy to see if we consider the mean of (A5) conditional on being on segment s:

$$E(H|V,D_s=1) \quad = \quad \alpha + \beta W[1-t_s(x)] + \delta\tilde{N}_s(x) + \gamma x + E(\epsilon|V,D_s=1) \qquad \text{(A6)}$$

from which it is clear that the residual term $E(\epsilon|V,D_s=1)$ is not constant over time if the tax schedule changes and hence will not cancel out in first differencing, even if γx does.

An additional complication, which is more fundamental, is that (A5) is not consistent with a nonzero variance of ϵ in the first place because of the problem of classification error. Given the presence of measurement error and optimization error in ϵ, a sufficiently large positive or negative value of ϵ will move the individual to a segment other than s. Thus the H of some individuals observed on segment s is not generated by the net wage and virtual nonlabor income on that segment, and hence $E(H|V,D_s=1)$ is not equal to $g\{W[1-t_s(x)],\tilde{N}_s(x);x\} + E(\epsilon|V,D_s=1)$ in general. Thus the regressors are misspecified. The mean H of those observed to be on segment s is consequently not the mean of (A5) but is rather

$$E(H|V,D_s=1) \quad = \quad \sum_{s'=1}^{S} Q_{ss'}(V)\; E(H|V,D_s=1,D_{s'}^*=1) + \sum_{k=1}^{S} R_{sk}(V)\; E(H|V,D_s=1,D_k^*=1) \tag{A7}$$

where $D_{s'}^*$ is a dummy variable equal to 1 if the true segment (defined as that implied by utility maximization with no optimization costs) is segment s'; $Q_{ss'}(V)$ is the probability that an individual observed on s is optimizing on s'; D_k^* is a dummy variable equal to 1 if the true optimizing point is at the kink at the beginning of segment k; and $R_{sk}(V)$ is the probability that an individual observed on segment s is optimizing at kink k.[46] Thus observed H is a weighted average of the net wage rates and virtual nonlabor incomes on all segments, for these

[46] Although utility maximization implies that some individuals will invariably locate at kinks, assuming smoothness of preferences, the presence of measurement and optimization error implies that no observations will be precisely located at kinks. In addition, if the variance of ϵ is sufficiently large, there will be no clustering around kinks as well.

are the determinants of H on each segment. The fact that D_s and D_k are not observed implies that the conditional means in (A7) cannot be directly estimated.[47]

Equation (A7) thus represents the function whose mean can be thought to be approximated by the local net wage and virtual income. A linear projection of (A7) onto those two local variables, and x, will yield as coefficients nonlinear functions of the other parameters and variables in the model, including the other tax parameters. It is the coefficient on the net tax rate that is estimated by the models reported in the text.

---

[47] The classification problem can be eliminated by assumption if ϵ is taken to represent only heterogeneity of preferences and not measurement or optimization error. In that case, observed segment location equals optimized location. But that assumption requires that some observations be clustered at kinks. In addition to the fact that significant clustering is rarely observed, it implies that the equation is misspecified if all observations are assigned to segments.

REFERENCES

Ashenfelter, O. and J. Heckman. "Estimating Labor Supply Functions." In Income Maintenance and Labor Supply, eds. G. Cain and H. Watts. Chicago: Markham, 1973.

______________. "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply." Econometrica 42 (January 1974): 73-85.

Auten, G. and R. Carroll. "The Effect of Income Taxes on Household Income." Office of Tax Analysis Working Paper, 1998.

Blomquist, S. and W. Newey. "Taxation and Labor Supply." Mimeographed, 1996.

Blundell, R.; A. Duncan; and C. Meghir. "Estimating Labor Supply Responses Using Tax Reforms." Econometrica, forthcoming.

Blundell, R. and T. MaCurdy. "Labor Supply." Handbook of Labor Economics, Vol.3, eds. O. Ashenfelter and D. Card, forthcoming.

Deaton, A. "Panel Data from Time Series of Cross-Sections." Journal of Econometrics 30 (1985): 109-126.

Eissa, N. "Taxation and the Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment." NBER Working Paper 5023, February 1995.

________. "Tax Reforms and Labor Supply." In Tax Policy and the Economy, Vol.10, ed. J. Poterba. Cambridge: MIT Press, 1996a.

________. "Labor Supply and the Economic Recovery Act of 1981." In Empirical Foundations of Household Taxation, eds. M. Feldstein and J. Poterba. Chicago: University of Chicago Press, 1996b.

Feenberg, D. and J. Poterba. "Income Inequality and the Incomes of Very High Income Taxpayers." In Tax Policy and the Economy, Vol.7, ed. J. Poterba. Cambridge: MIT Press, 1993.

Feldstein, M. "The Effect of Marginal Tax Rates on Taxable Income: A Study of the 1986 Tax Reform Act." Journal of Political Economy 103 (June 1995a): 551-572.

__________. "Behavioral Responses to Tax Rates: Evidence from the Tax Reform Act of 1986." American Economic Review 85 (May 1995b): 170-174.

__________. "Tax Avoidance and the Deadweight Loss of the Income Tax." Working Paper 5055. Cambridge: NBER, 1995c.

Feldstein, M. and D. Feenberg. "The Effect of Increased Tax Rates on Taxable Income and Economic Efficiency: A Preliminary Analysis of the 1993 Tax Rate Increases." In Tax Policy and the Economy, Vol.10, ed. J. Poterba. Cambridge: MIT Press, 1996.

Hausman, J. "Labor Supply." In How Taxes Affect Economic Behavior, eds. H. Aaron and J. Pechman. Washington: Brookings, 1981.

__________. "Taxes and Labor Supply." In Handbook of Public Economics, eds. A. Auerbach and M. Feldstein. Amsterdam: North-Holland, 1985.

Hausman, J. and J. Poterba. "Household Behavior and the Tax Reform Act of 1986." Journal of Economic Perspectives 1 (Summer 1987): 101-119.

Heckman, J. "Comment." In Empirical Foundations of Household Taxation, eds. M. Feldstein and J. Poterba. Chicago: University of Chicago Press, 1996a.

__________. "Randomization as an Instrumental Variable." Review of Economics and Statistics 78 (May 1996b): 336-341.

Heckman, J. and R. Robb. "Alternative Methods for Evaluating the Impact of Interventions." In Longitudinal Analysis of Labor Market Data, eds. J. Heckman and B. Singer. Cambridge: CUP, 1985.

Kennickell, A. and J. Shack-Marquez. "Changes in Family Finances from 1983 to 1989: Evidence from the Survey of Consumer Finances." Federal Reserve Bulletin (January 1992): 1-18.

Lindsey, L. "Estimating the Behavioral Responses of Taxpayers to Changes in Tax Rates: 1982-1984." Journal of Public Economics (July 1987): 173-206.

MaCurdy, T.; D. Green; and H. Paarsch. "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." Journal of Human Resources 25 (Summer 1990): 415-90.

Mariger, R. "Labor Supply and the Tax Reform Act of 1986: Evidence from Panel Data." Discussion Paper 95-34. Washington: Federal Reserve Board, 1995.

Meyer, B. "Natural and Quasi-Experiments in Economics." Journal of Business and Economic Statistics 13 (April 1995): 151-161.

Moffitt, R. "The Labor Supply Response to the Gary Experiment." Journal of Human Resources 14 (Fall 1979): 477-487.

__________. "The Econometrics of Piecewise-Linear Budget Constraints." Journal of Business and Economic Statistics 4 (July 1986): 317-328.

__________. "The Econometrics of Kinked Budget Constraints." Journal of Economic Perspectives 4 (Spring 1990): 119-139.

__________. "Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-Sections." Journal of Econometrics 59 (1993): 99-123.

__________. "Identification and Estimation of Treatment Effects with Panel Data." Mimeographed, Johns Hopkins University, 1998.

Moffitt, R. and K. Kehrer. "The Effect of Tax and Transfer Programs on Labor Supply: The Evidence from the Income Maintenance Experiments." In Research in Labor Economics, ed. R. Ehrenberg, Vol.4. Greenwich, Conn: JAI Press, 1981.

Rosen, H. "Tax Illusion and the Labor Supply of Married Women." Review of Economics and Statistics 58 (May 1976): 167-172.

Showalter, M. "An Investigation of the Labor Suply Response of High-Income Individuals to the Tax Changes of the 1980s." Mimeographed, Brigham Young University, 1997.

Showalter, M. and N. Thurston. "Taxes and Labor Supply of High-Income Physicians." Journal of Public Economics, forthcoming.

Slemrod, J. "A General Model of the Behavioral Response to Taxation." Mimeo, Univ of Michigan, 1995.

__________. "High-Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response." In Empirical Foundations of Household Taxation, eds. M. Feldstein and J. Poterba. Chicago: University of Chicago Press, 1996.

Triest, R. "The Effect of Income Taxation on Labor Supply When Deductions are Endogenous." 74 Review of Economics and Statistics (February 1992): 91-99.

U.S. Internal Revenue Service. "Individual Income Tax Returns." Statistics of Income: 1982. Washington, D.C.: U.S. Government Printing Office, 1984.

__________. "Individual Income Tax Returns." Statistics of Income: 1988. Washington, D.C.: U.S. Government Printing Office, 1991.

Wolff, E. "Trends in Household Wealth in the United States, 1962-83 and 1983-89." Review of Income and Wealth (June 1994): 143-174.

__________. Top Heavy: A Study of the Increasing Inequality of Wealth in America. New York: Twentieth Century Fund Press, 1995.

Ziliak, J. and T. Kniesner. "Estimating Life Cycle Labor Supply Tax

Effects." Mimeographed, University of Oregon, November 1996.

Table 1. **Means and Standard Deviations in the 1983 and 1989 SCF Panel (Men 25-54 in 1983)**

	Midrange or High 1983 MTR		Midrange 1983 MTR		High 1983 MTR	
Variable	1983	1989	1983	1989	1983	1989
Adjusted Gross Income (AGI)	49720 (49475)	61913 (128805)	44723 (16117)	52470 (41861)	168899 (199575)	287115 (563590)
Annual Hours Worked	2340 (617)	2380 (546)	2336 (611)	2375 (536)	2434 (745)	2501 (726)
Total Income	55723 (60118)	69773 (154809)	49213 (21364)	58850 (38337)	210967 (231976)	330285 (699832)
Wage and Salary Income	40945 (25715)	53409 (43663)	39107 (18008)	50519 (28461)	84808 (81810)	122342 (152001)
Wage and Salary and Business Income	49857 (39190)	63138 (114499)	44909 (17726)	55053 (32208)	167861 (127126)	255021 (512512)
Net-of-Tax Rate (NTR=1-t)	0.692 (0.074)	0.776 (0.064)	0.700 (0.065)	0.778 (0.064)	0.506 (0.006)	0.709 (0.038)
Hourly Wage Rate (W)	17.60 (16.34)	21.81 45.44	15.79 (8.92)	18.98 (21.58)	60.78 (52.81)	89.44 (188.28)
W*NTR	11.57 (8.33)	16.42 32.60	10.77 (5.21)	14.44 (15.51)	30.66 (26.45)	63.53 (135.50)
Tax Payment	8480. (13802.)	9690. (29360.)	6861. (3975.)	7389. (10132.)	47099. (52940.)	64577. (125829.)
Married	0.849 (0.358)	0.846 (0.361)	0.846 (0.361)	0.846 (0.361)	0.930 (0.254)	0.845 (0.362)
Pct. with No Change in Marital Status	.	0.828 (0.377)	.	0.828 (0.378)	.	0.848 (0.359)
Household Size	3.207 (1.424)	3.089 (1.300)	3.198 (1.426)	3.096 (1.299)	3.407 (1.373)	2.918 (1.319)
Pct. with No Change in Household Size	.	0.473 (0.499)	.	0.476 (0.499)	.	0.421 (0.494)
Age 30-34	0.226 (0.418)	.	0.232 (0.422)	.	0.063 (0.243)	.
Age 35-39	0.180 (0.384)	.	0.180 (0.384)	.	0.168 (0.374)	.
Age 40-44	0.133 (0.339)	.	0.128 (0.334)	.	0.244 (0.430)	.

Age 45-49	0.186 (0.389)	.	0.181 (0.385)	.	0.304 (0.460)	.
Age 50-54	0.101 (0.301)	.	0.096 (0.295)	.	0.206 (0.404)	.
Distribution of 1983 MTR:						
0.20 < t ≤ 0.22	0.191 (0.393)	.	0.199 (0.393)	.	0.000 (0.000)	.
0.22 < t ≤ 0.25	0.160 (0.366)	.	0.166 (0.366)	.	0.000 (0.000)	.
0.25 < t ≤ 0.29	0.188 (0.391)	.	0.196 (0.391)	.	0.000 (0.000)	.
0.29 < t ≤ 0.33	0.193 (0.395)	.	0.201 (0.395)	.	0.000 (0.000)	.
0.33 < t ≤ 0.40	0.162 (0.369)	.	0.169 (0.369)	.	0.000 (0.000)	.
0.40 < t ≤ 0.44	0.066 (0.249)	.	0.069 (0.249)	.	0.000 (0.000)	.
0.44 < t ≤ 0.50	0.040 (0.197)	.	0.000 (0.000)	.	1.000 (0.000)	.
High 1983 Income (rich dummy)	0.075 (0.260)	.	0.038 (0.190)	.	0.914 (0.281)	
Post-college	0.175 (0.380)	.	0.153 (0.360)	.	0.705 (0.456)	.
Professional-Manager	0.453 (0.498)	.	0.434 (0.496)	.	0.887 (0.317)	.
Log 1983 House Value	8.618 (4.891)	.	8.497 (4.915)	.	11.485 (3.092)	.
Zero 1983 House Value (dummy)	0.241 (0.428)	.	0.249 (0.432)	.	0.064 (0.245)	.
Log 1983 Life Insurance Value	10.268 (3.095)	.	10.181 (3.111)	.	12.344 (1.668)	.
Zero 1983 Life Insurance (dummy)	0.073 (0.260)	.	0.076 (0.265)	.	0.010 (0.010)	.
High 1983 House Value or Life Insurance Value (dummy)	0.123 (0.328)	.	0.104 (0.305)	.	0.578 (0.494)	
Observations (unweighted)	406	406	277	277	129	129

Notes: All values are weighted. Midrange 1983 MTR values those greater than 0.20 and less than or equal to 0.44 and high 1983 MTR values are those greater than 0.44 and less than or equal to 0.50. All monetary amounts are in 1988 dollars. Standard deviations appear in parentheses.

Table 2. Difference-in-Difference Estimates of the Effect of NTR on AGI

Variable	1989-1983 Differences by 1983 MTR Group: High 1983 MTR	1989-1983 Differences by 1983 MTR Group: Midrange 1983 MTR	Difference of Differences	Implied Elasticity
AGI				
Average Linear Difference	118,000	7,747	110,000	1.992 [a]
Percentage Change in Average	0.700	0.173	0.527	1.828
Average of Percentage Changes	0.706	0.213	0.494	1.757
NTR				
Average Linear Difference	0.203	0.079	0.124	
Percentage Change in Average	0.401	0.113	0.288	
Average of Percentage Changes	0.401	0.120	0.281	

Notes: [a] The absolute differences-in-differences estimate is 890,000 [=(118,000-7747)/(.203-.079)]. We convert to an arc elasticity by multiplying by [(.203+.079)/(118,000+ 7,747).
N=406

Table 3. Regressions to Generate Difference-in-Difference Estimates of NTR on AGI

	Linear Differences			Individual Percentage Changes		
	AGI (Reduced -Form)	NTR (First Stage)	AGI (2SLS)	AGI (Reduced -Form)	NTR (First Stage)	AGI (2SLS)
Midrange 1983 MTR Group Dummy	7,747 (5,934)	0.079*** (0.004)	.	0.213*** (0.044)	0.120*** (0.006)	.
High 1983 MTR Group Dummy	0.118***a (0.029)	0.203*** (0.018)	.	0.706*** (0.216)	0.401*** (0.029)	.
Change in NTR	.	.	890,000*** (289,000)	.	.	1.757** (0.882)
Constant	.	.	-62,464** (25,250)	.	.	0.002 (0.126)
R^2	0.033	0.100	-0.428	0.012	0.185	-0.257

Notes: In columns 1 and 3 the dependent variable is the linear change in AGI; in column 2 the dependent variable is the linear change in the NTR; in columns 4 and 6 the dependent variable is the percentage change in AGI; in column 5 the dependent variable is the percentage change in the NTR. All regressions are weighted. Standard errors appear in parentheses. N=406.

* Significant at .10 level. ** Significant at .05 level. *** Significant at .01 level.

Table 4. 2SLS Estimates of the Effect of NTR on AGI with Alternative Instruments

	Second-Stage Equation[a]		First-Stage Equation		
Instrument(s)	Coefficient on Change in NTR	Standard Error	F-statistic	P-value	R^2
Two 1983 MTR Groups	0.969***	0.320	38.578	0.000	0.126
Seven 1983 MTR Groups	0.162	0.126	42.742	0.000	0.420
High 1983 Income Dummy	0.839***	0.297	41.756	0.000	0.132
Log 1983 AGI	0.345***	0.140	202.16	0.000	0.365
1983 Marital Status (married dummy)	0.123	1.343	0.672	0.328	0.065
1983 Household Size	-0.152	1.128	0.683	0.317	0.075
Post-college	0.341	0.435	14.336	0.000	0.074
Professional-Manager	1.974	2.714	0.662	0.338	0.043
Log 1983 House Value and Log 1983 Life Insurance Value[b]	0.649**	0.289	9.824	0.000	0.128
High 1983 House Value or Life Insurance Value (dummy)	0.660**	0.327	30.465	0.000	0.109

Notes: Sample of men with midrange and high 1983 MTR (N=406). All regressions are weighted two-stage least squares, using linear differences in AGI and NTR in second and first stages, respectively. Each line in table shows results from a different model with a different set of instruments. The F-statistics test zero restrictions on the instruments in the first-stage and the p-values associated with those statistics as shown along with the R^2 from the first-stage regression. Each model contains in both the first and second stages a constant term and independent variables for 1983 age, marital status (dummy for whether married), and household size; the estimates on these control variables are presented in Appendix Table A3 for two of the models. When marital status and household size are used as instruments. these variables are omitted from the second stage. When these two instruments are used, only those with no change in marital status or household size are included.

[a] Coefficients and standard errors divided by 10^6.
[b] Instruments also include dummies for zero house value and life insurance.

* Significant at .10 level. ** Significant at .05 level. *** Significant at .01 level.

Table 5. 2SLS Estimates of the Effect of NTR on AGI Using Alternative Instruments and Samples and Controls for Regression to Mean

	Instrument Set				
	Two 1983 MTR Groups	Log 1983 House Value Log 1983 Life Insurance			High 1983 House Value or Life Insurance Value
Change in NTR[a]	0.815*** (0.255)	0.552** (0.224)	0.885** (0.430)	1.006** (0.443)	0.977** (0.510)
1983 AGI	.	.	-0.414* (0.249)	-0.694** (0.329)	-0.674** (0.371)
Low 1983 MTR Group Included?	y	y	n	y	y
First stage:					
F-statistic	44.408	12.036	5.201	4.432	13.033
p-value	0.000	0.000	0.000	0.002	0.000
R^2	0.128	0.135	0.173	0.211	0.204
Observations (unweighted)	490	490	406	490	490

Notes: All regressions are weighted two-stage least squares using linear differences. AGI. Standard errors appear in parentheses.

[a] Coefficients and standard errors divided by 10^6.

* Significant at .10 level. ** Significant at .05 level. *** Significant at .01 level.

Table 6. 2SLS Estimates of the Effect of NTR on Annual Hours Worked with Alternative Instruments

Instrument(s)	Second-Stage Equation[a]		First-Stage Equation		
	Coefficient on Change in NTR	Standard Error	F-statistic	P-value	R^2
Two 1983 MTR Groups	-0.010	0.135	38.578	0.000	0.126
Seven 1983 MTR Groups	-0.013	0.063	42.742	0.000	0.420
High 1983 Income Dummy	-0.246	0.135	41.756	0.000	0.132
Log 1983 AGI	-0.038	0.069	202.16	0.000	0.365
1983 Marital Status (married dummy)	-2.011	2.211	0.672	0.328	0.065
1983 Household Size	-0.936	1.088	0.683	0.317	0.075
Post-college	-0.320	0.229	14.336	0.000	0.074
Professional/Manager	-1.137	1.627	0.662	0.338	0.043
Log 1983 House Value and Log 1983 Life Insurance Value[b]	0.072	0.134	9.824	0.000	0.128
High 1983 House Value or Life Insurance Value (dummy)	-0.488***	0.173	30.465	0.000	0.109

Notes: Sample of men with midrange and high 1983 MTR. All regressions are weighted two-stage least squares, using linear differences in AGI and NTR in second and first stages, respectively. Each line in table shows results from a different model with a different set of instruments. The F-statistics test zero restrictions on the instruments in the first-stage and the p-values associated with those statistics as shown along with the R^2 from the first-stage regression. Each model contains in both the first and second stages a constant term and independent variables for 1983 age, marital status (dummy for whether married), and household size; the estimates on these control variables are presented in Appendix Table A3 for two of the models. When marital status and household size are used as instruments, these variables are omitted from the second stage. When these two instruments are used, only those with no change in marital status or household size are included.

[a] Coefficients and standard errors divided by 10^4.
[b] Instruments also include dummies for zero house value and life insurance.

* Significant at .10 level. ** Significant at .05 level. *** Significant at .01 level.

Table 7. 2SLS Estimates of the Effect of NTR on Annual Hours Worked Using Alternative Instruments and Controls for Regression to Mean

	Instrument Set					
	Two 1983 MTR Groups	High House Value or Life Insurance Value	Log 1983 House Value Log 1983 Life Insurance			
Change in NTR[a]	0.025 (0.127)	-0.044 (0.117)	0.164 (0.111)	-0.024 (0.122)	-0.015 (0.300)	0.129 (0.231)
Change in Tax Payment	.	.	.	.	-0.001 (0.030)	-0.001 (0.024)
1983 Hours Worked	.	-0.609*** (0.041)	-0.609*** (0.042)	.	.	-0.613*** (0.075)
Low 1983 MTR Group Included?	y	n	n	y	y	y
First stage NTR Eqn						
F-statistic	44.408	32.329	9.824	12.036	12.036	12.536
p-value	0.000	0.000	0.000	0.000	0.000	0.000
R^2	0.128	0.113	0.128	0.135	0.135	0.140
First stage Tax Payment Eqn						
F-statistic	.	.	.	.	0.747	1.241
p-value	.	.	.	.	0.253	0.293
R^2	.	.	.	.	0.012	0.014
Observations (unweighted)	490	406	406	490	490	490

Notes: All regressions are weighted two-stage least squares using linear differences. Standard errors appear in parentheses.

[a] Coefficients and standard errors divided by 10^4.

* Significant at .10 level. ** Significant at .05 level. *** Significant at .01 level.

Table 8. Estimates of the Effects of NTR on AGI and Annual Hours Worked with Education and Wage Interactions Using Alternative Instruments and Samples and Controls for Regression to Mean

	Instrument Set					
	Two 1983 MTR Groups	Log 1983 House Value Log 1983 Life Insurance				High 1983 House Value or Life Insurance Value
Hours Worked: I						
NTR	-0.077 (0.260)	0.272 (0.169)	0.042 (0.155)	0.356*** (0.135)	0.236** (0.122)	0.274* (0.156)
NTR*(Post-college)	0.186 (0.294)	-0.528 (0.358)	-0.063 (0.348)	-0.263 (0.287)	-0.033 (0.278)	-0.407* (0.246)
Hours Worked: II						
NTR*W	-4.344 (4.042)	-3.996 (3.395)	-3.752 (3.746)	-3.908 (2.684)	-2.860 (3.012)	-7.311* (4.405)
AGI: I						
NTR	1.202** (0.499)	0.452 (0.328)	0.393 (0.250)	0.823* (0.427)	1.047*** (0.398)	0.972** (0.475)
NTR*(Post-College)	-0.149 (0.565)	1.448** (0.694)	1.298** (0.560)	0.645 (0.484)	0.423 (0.328)	0.167 (0.315)
AGI: II						
NTR*W	2215.6*** (655.2)	1606.0** (559.7)	1679.7*** (519.2)	1753.0*** (679.8)	1963.9*** (633.3)	947.1 (986.4)
Low 1983 MTR Group Included?	n	n	y	n	y	y
Control for Regression to Mean?	n	n	n	y	y	y
Observations (unweighted)	406	406	490	406	490	490

Notes: All regressions are weighted two-stage-least-squares using linear differences and all include the aforementioned control variables in both stages. The regression-to-the-mean specifications include the 1983 value of the respective dependent variable. The coefficients in the AGI models are divided by 10^6 and those in the hours worked models are divided by 10^4.

* Significant at .10 level. ** Significant at .05 level. *** Significant at .01 level.

Table 9. 2SLS Estimates of the Effect of NTR on Other Dependent Variables Using Alternative Instruments and Samples and Controls for Regression to Mean

Dependent Variable	Instrument Set: Two 1983 MTR Groups	Instrument Set: Log 1983 House Value, Log 1983 Life Insurance				Instrument Set: High 1983 House Value or Life Insurance Value
Total Income	0.983*** (0.347)	0.583* (0.316)	0.504** (0.243)	0.475 (0.432)	0.500 (0.398)	0.249 (0.486)
Wage/Salary	0.230** (0.106)	0.293*** (0.110)	0.287*** (0.092)	0.494*** (0.153)	0.608*** (0.174)	0.398*** (0.124)
Wage/Salary and Business Income	0.701*** (0.256)	0.452* (0.235)	0.377** (0.180)	0.153 (0.367)	0.137 (0.396)	-0.145 (0.481)
Low 1983 MTR Group Included?	n	n	y	n	y	y
Lagged Dependent Variable Included?	n	n	n	y	y	y
Observations (unweighted)	406	406	490	406	490	490

Notes: All regressions are weighted two-stage least squares using linear differences. Each row shows the NTR coefficients for a different dependent variable. All models include a constant term and independent variables for age, marital status, and household number in 1983 in both stages.

[a] Coefficients and standard errors are divided by 10^6.

* Significant at .10 level. ** Significant at .05 level. *** Significant at .01 level.

Table A1. Sample Inclusion Criteria and Sample Size

Inclusion Criteria	Sample Size Remaining
Full SCF 1983-1989 Panel	1479
Including only:	
Male heads of households	1214
Aged 25-54 in 1983	695
No ambiguity in tracing individuals between 1983 and 1989	628
AGI in both 1983 and 1989 greater than zero	563
Wages in both 1983 and 1989 greater than zero	498
Annual hours worked in both 1983 and 1989 greater than or equal to 200	496
1989 hours worked not imputed	490

Table A2. Means and Standard Deviations in the 1983 and 1989 SCF Panel (Men 25-54 in 1983): All Men and Men with Low 1983 Marginal Tax Rates

	All Men		Men with Low 1983 Marginal Tax Rates	
Variable	1983	1989	1983	1989
Adjusted Gross Income (AGI)	43129 (45408)	57082 (115515)	20084 (5450)	40193 (40294)
Annual Hours Worked	2325 (621)	2371 (584)	2272 (631)	2337 (702)
Total Income	48995 (54645)	63285 (138570)	25470 (8839)	40602 (43383)
Wage and Salary Income	35390 (25383)	48342 (41926)	15964 (9981)	30625 (28881)
Wage and Salary and Business Income	44000 (36468)	57474 (102962)	23524 (8407)	37669 (36405)
Net-of-Tax Rate (NTR=1-t)	0.723 (0.088)	0.786 (0.065)	0.832 (0.024)	0.823 (0.051)
Hourly Wage Rate (W)	15.75 (14.91)	19.82 40.62	9.31 (3.67)	12.86 (11.62)
W*NTR	10.71 (7.65)	15.07 29.12	7.71 (2.96)	10.36 (8.27)
Tax Payment	6898. (12529.)	8587. (26377.)	1365. (694.)	4727. (9758.)
Married	0.867 (0.340)	0.858 (0.349)	0.926 (0.261)	0.899 (0.301)
Pct. with No Change in Marital Status	.	0.861 (0.346)	.	0.973 (0.162)
Household Size	3.405 (1.492)	3.255 (1.374)	4.100 (1.515)	3.833 (1.467)
Pct. with No Change in Household Size	.	0.471 (0.499)	.	0.461 (0.499)
Age 30-34	0.213 (0.409)	.	0.168 (0.374)	.

Age 35-39	0.175 (0.380)	.	0.160 (0.367)	.
Age 40-44	0.140 (0.347)	.	0.165 (0.370)	.
Age 45-49	0.183 (0.387)	.	0.172 (0.378)	.
Age 50-54	0.093 (0.290)	.	0.066 (0.247)	.
Distribution of 1983 MTR:				
$0.00 \leq t \leq 0.20$	0.222 (0.416)	.	1.000 (0.000)	.
$0.20 < t \leq 0.22$	0.148 (0.356)	.	0.000 (0.000)	.
$0.22 < t \leq 0.25$	0.124 (0.330)	.	0.000 (0.000)	.
$0.25 < t \leq 0.29$	0.146 (0.353)	.	0.000 (0.000)	.
$0.29 < t \leq 0.33$	0.150 (0.357)	.	0.000 (0.000)	.
$0.33 < t \leq 0.40$	0.126 (0.332)	.	0.000 (0.000)	.
$0.40 < t \leq 0.44$	0.051 (0.221)	.	0.000 (0.000)	.
$0.44 < t \leq 0.50$	0.031 (0.174)	.	0.000 (0.000)	.
Post-college	0.150 (0.357)	.	0.062 (0.241)	.
Professional-Manager	0.413 (0.492)	.	0.273 (0.445)	.
Log 1983 House Value	8.279 (5.010)	.	7.096 (5.236)	
Zero 1983 House Value (dummy)	0.266 (0.442)	.	0.351 (0.478)	
Log 1983 Life Insurance Value	10.103 (3.158)	.	9.527 (3.306)	
Zero 1983 Life Insurance (dummy)	0.079 (0.270)	.	0.100 (0.300)	

High 1983 House Value or Life Insurance Value (dummy)	0.101 (0.302)	.	0.026 (0.158)	.
Observations (unweighted)	490	490	84	84

Notes: All values are weighted. Low 1983 marginal tax rates are those less than 0.20. All monetary amounts are in 1988 dollars. Standard deviations appear in parentheses.

Table A3. **2SLS Estimates of the Effect of NTR on AGI and Annual Hours Worked Using Alternative Instruments: Full Coefficient Estimates**

Dependent Variable	Change in AGI		Change in Hours	
Instrument	Two 1983 MTR Groups	Log 1983 House Value and Life Insurance Value	Two 1983 MTR Groups	Log 1983 House Value and Life Insurance Value
NTR	0.969***	0.649**	-0.010	0.072
	(0.320)	(0.289)	(0.135)	(0.134)
Age 30-34	-0.064	-0.044	0.202**	0.197*
	(0.245)	(0.224)	(0.103)	(0.104)
Age 35-39	-0.331	-0.223	0.255**	0.227*
	(0.282)	(0.257)	(0.119)	(0.119)
Age 40-44	-0.329	-0.206	0.499***	0.467***
	(0.312)	(0.284)	(0.131)	(0.132)
Age 45-49	-0.334	-0.216	0.161	0.130
	(0.284)	(0.259)	(0.120)	(0.120)
Age 50-54	-0.282	-0.169	0.223*	0.194
	(0.318)	(0.290)	(0.134)	(0.134)
Married	-0.070	-0.077	-0.413***	-0.411***
	(0.264)	(0.242)	(0.111)	(0.112)
Household Size	0.015	0.012	0.056*	0.056*
	(0.069)	(0.063)	(0.029)	(0.029)
Constant	-47155.	-26065.	11.014	-43.281
	(30180.)	(27430.)	(127.0)	(127.1)
Observations (unweighted)	406	406	406	406

Notes: All regressions are weighted two-stage least squares using linear differences. The coefficients and standard errors in the AGI models are divided by 10^6 and those in the hours worked models are divided by 10^4.

* Significant at .10 level. ** Significant at .05 level. *** Significant at .01 level.

Figure 1
Annual Hours of Work by Occupational Category: Male Heads

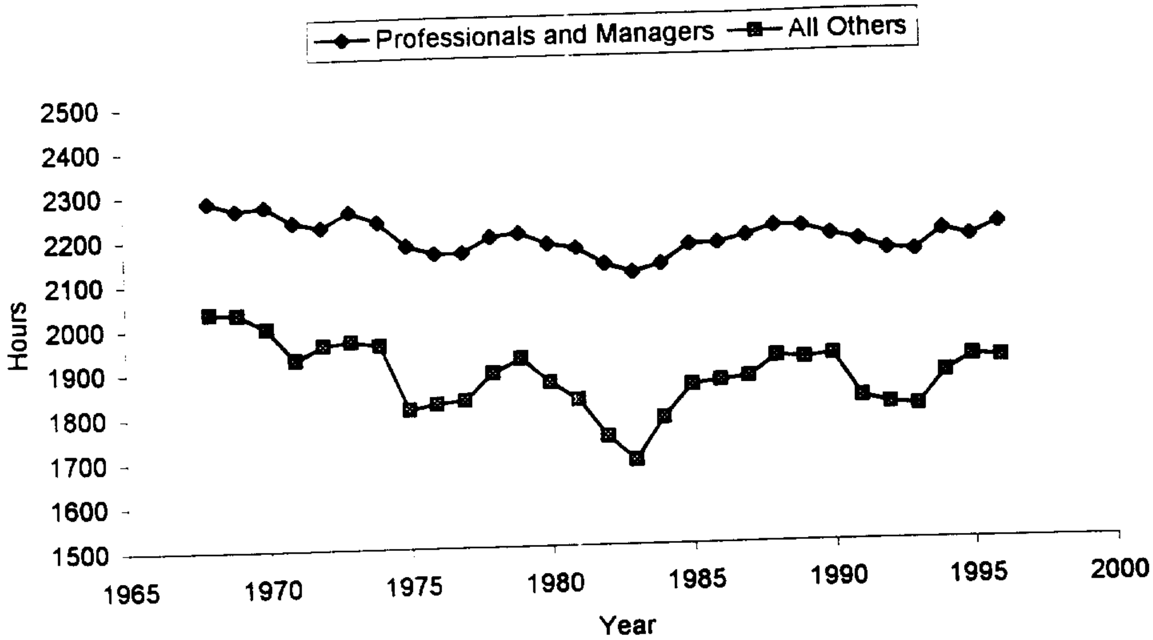

Figure 2
Annual Hours of Work by Educational Cateogory: Male Heads

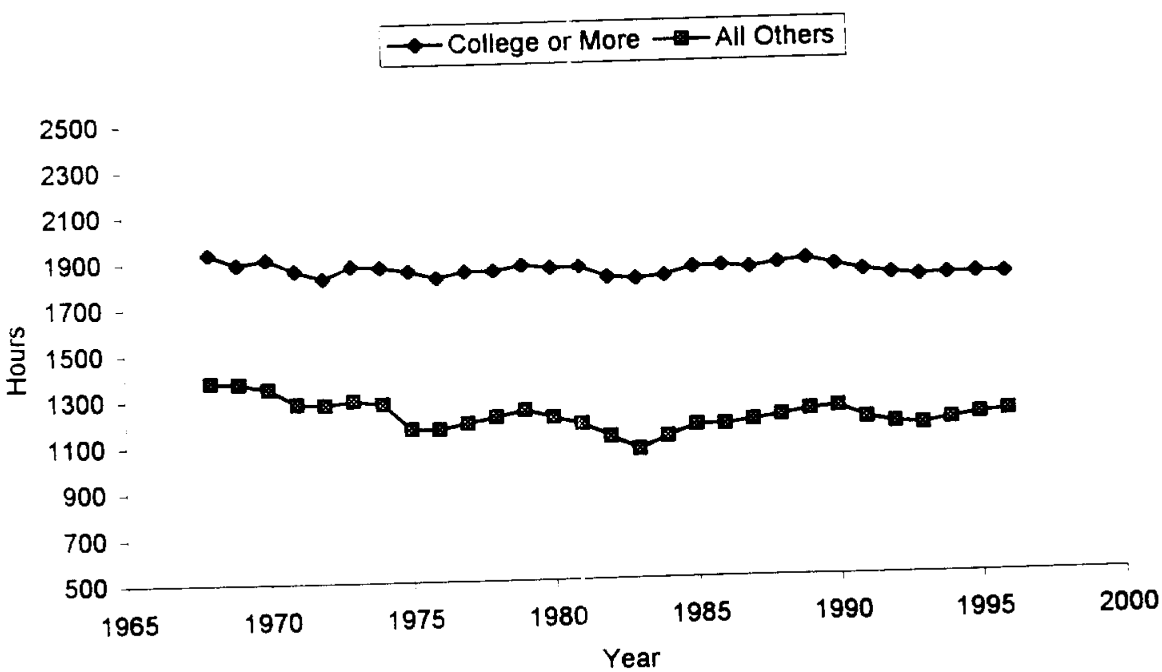

Figure 3
Family Income by Occupational Category: Male Heads

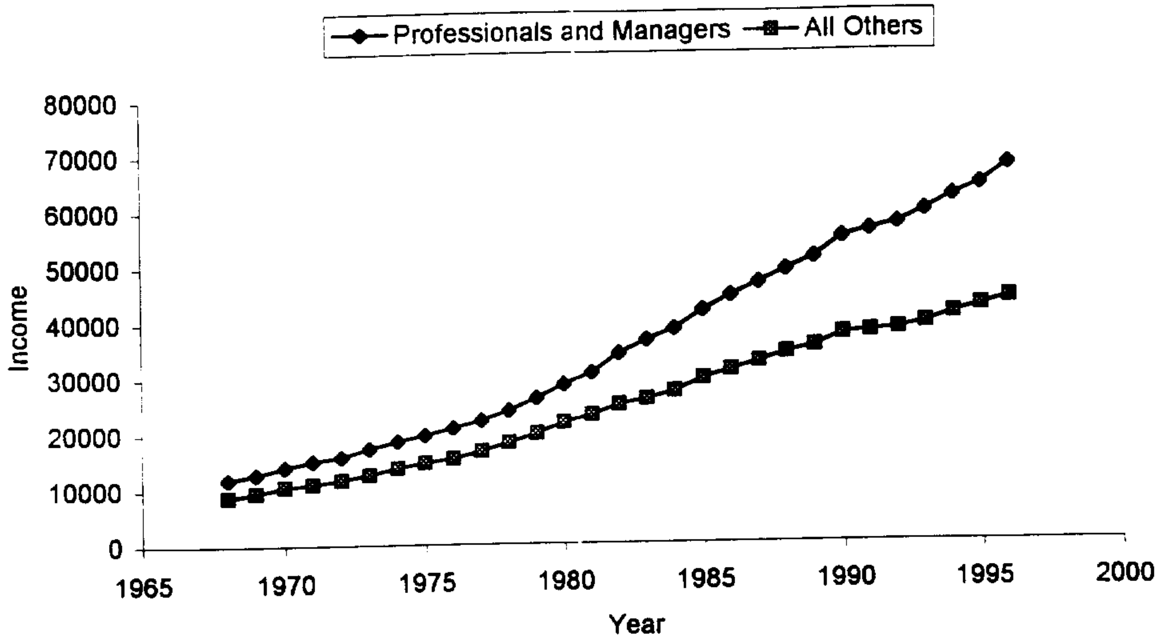

Figure 4
Individual Earnings by Occupational Group: Male Heads

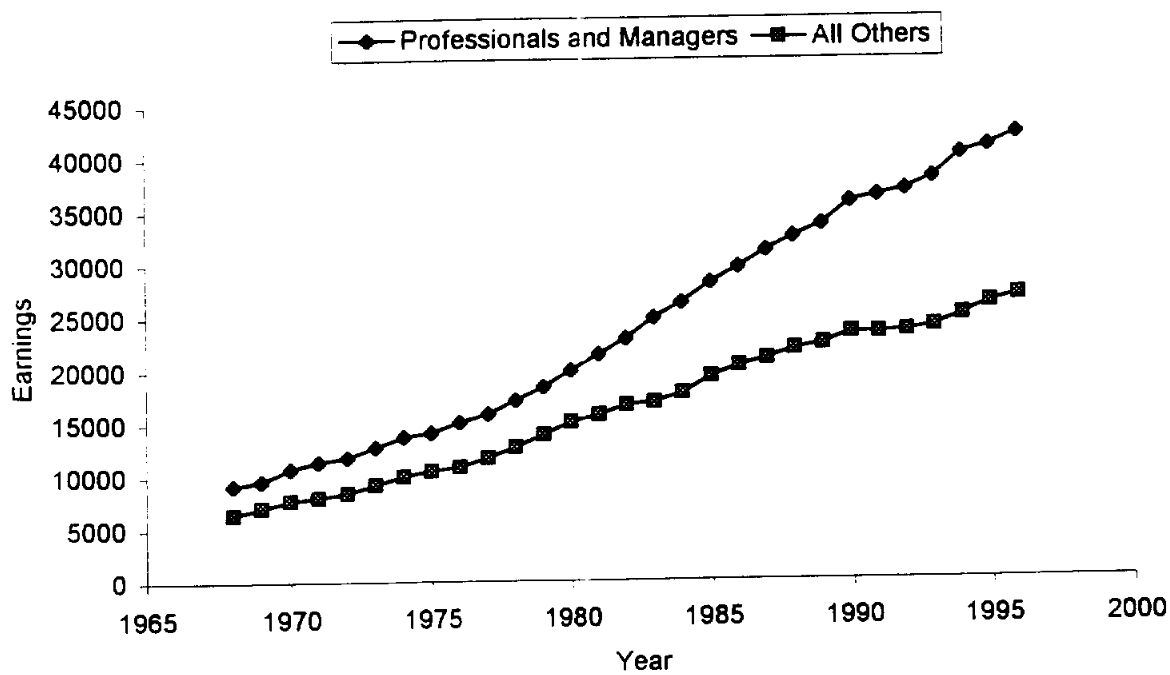

Figure 5
Change in NTR by Change in AGI

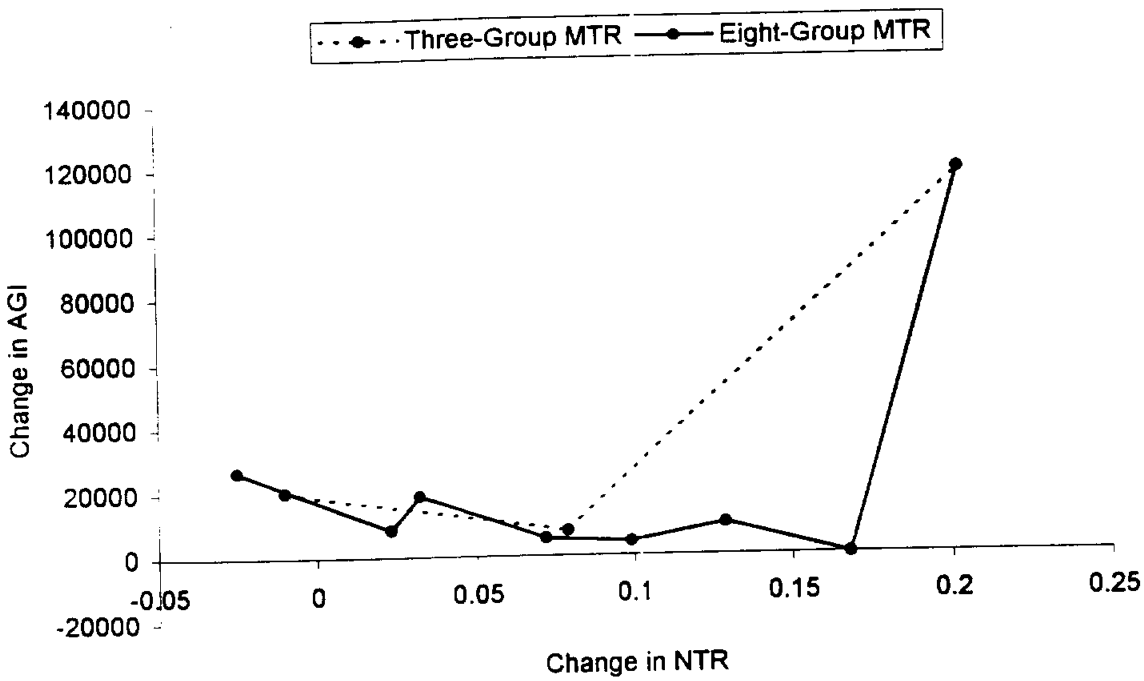

Figure 6
Annual Hours Worked By Marginal Tax Rate: 1983

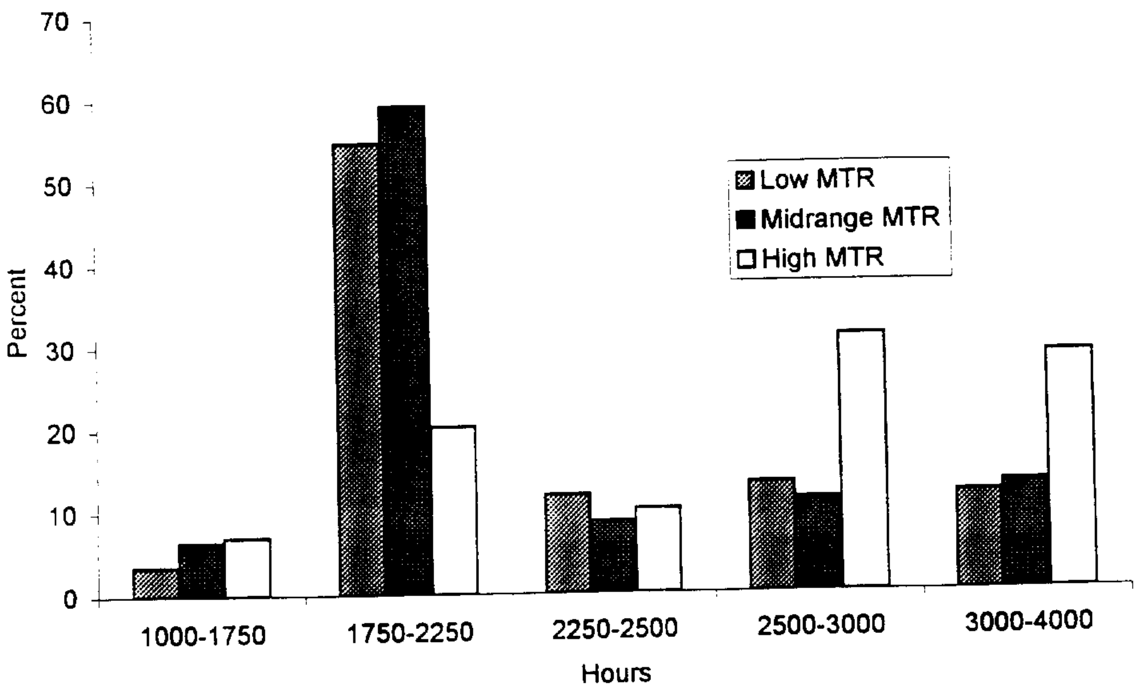

Figure 7
Change in NTR by Change in Hours Worked