University of Wisconsin-Madison

IRP Discussion Papers

Robert Moffitt
Michael Rothschild

VARIABLE EARNINGS AND NONLINEAR TAXATION

DP #731-83

Institute for Research on Poverty
Discussion Paper 731-83

Variable Earnings and Nonlinear Taxation

Robert Moffitt
Department of Economics
Rutgers University

Michael Rothschild
Department of Economics
University of Wisconsin, Madison
and
National Bureau of Economic Research

July 1983

We are indebted to Gary Chamberlain and Eugene Smolensky for helpful discussions, to Peter Gottschalk for graciously performing special tabulations for us, and to the National Science Foundation and the Institute for Research on Poverty for research support. The research reported here is part of the NBER's research program in taxation. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

ABSTRACT

Taxes (and income-tested transfer programs) transform before-tax income into after-tax income. Income is variable and the tax and transfer system is nonlinear--i.e., marginal tax rates vary. As a consequence, the tax and transfer system punishes (and rewards) income variability in the following way: if before-tax income is y and the tax and transfer system is represented by the function B(•), then after-tax income is B(y). If y is random, then expected after-tax income is E[B(y)]. Define the certainty equivalent of after-tax income, y_c, as the certain income which would yield an after-tax income of E[B(y)]. We measure the extent to which the tax and transfer system penalizes (or rewards) variability by comparing y_c with $\bar{Y}$, the expected value of y. Specifically, we examine $R = y_c/\bar{Y}$. If R is less than unity, then expected after-tax income is less than it would be were income not random and equal to its expectation; in this case the tax and transfer system penalizes variability. If R exceeds unity, the tax and transfer system rewards variability. In this paper we attempt to see how the tax and transfer system in the United States treats variability by calculating R under different plausible assumptions. We find that for some individuals R is substantially less than 1; for others it exceeds 1. We conclude that the tax and transfer system punishes and rewards variability in a manner which is both substantial and capricious.

Variable Earnings and Nonlinear Taxation

I. INTRODUCTION

The earned income tax credit (EITC) is a refundable tax credit on the earnings of low-income families with dependents. It operates as an earnings subsidy and is intended to encourage labor market participation. In 1982 the subsidy paid to an eligible family with an earned income of $7500 was about $250. In this paper, we show that, at least in one extreme case, the EITC provides an incentive to leave the labor market in an amount of about $800 to a family with expected income of the same $7500.

This seems a dramatic example of Milton Friedman's dictum that government policies accomplish exactly the opposite of what they are intended to achieve. It is thus appropriate that another of Friedman's theories, the permanent income hypothesis, should provide the simplest (but not the only) explanation of how the EITC, along with other aspects of the tax and transfer system, should discourage work.

One of the main premises of the permanent income hypothesis is that income is variable. Another is that income earners are--and policymakers and social critics should be--more concerned with permanent or expected income than with particular realizations of the random process which generates annual income. Yet taxes are levied on current income. Since tax and transfer systems are nonlinear (i.e., the marginal tax rate is different at different income levels), they can substantially affect the transformation of random current income into after-tax expected income.

In this paper, we show that the current system of taxes and transfers in the United States has substantial and capricious effects on the transformation of current pretax income into permanent after-tax income. We examine not only the EITC but also the federal income tax in general and transfer programs such as AFDC (Aid to Families with Dependent Children) and Social Security. Not all of these tax and transfer programs discourage work. For example, to families with an expected income of $5,000 who are eligible for the AFDC program, the tax and transfer system provides a work subsidy of almost $2,000.

The mechanism which causes these effects is straightforward. Since income is variable, we may write

$$y = \bar{Y} + \varepsilon ,$$

where $\bar{Y}$ is expected (or, in Friedman's terms, permanent) income and ε represents the transitory variation in income.[1] Taxes are levied on current income, y, rather than expected income, $\bar{Y}$. The tax and transfer system transforms current income y into after-tax income, x. We can write this in general form as

$$x = B(y).$$

Expected after tax-income is $E[B(y)]$. If we assume individuals base their labor supply decisions on expected after-tax income, the shape of the schedule $B(\cdot)$ determines how the tax and transfer system rewards or punishes variability. If $B(\cdot)$ is concave, as it will be if the tax system is progressive, then Jensen's inequality implies that the tax system will penalize variability. If $B(\cdot)$ is convex, as it will be if the tax system is regressive, then the tax system will reward risk. The tax and transfer system in the United States is not consistently progressive or regressive. The EITC example cited above, presented in more detail in Section III, is an example of income variability occurring in a region where $B(\cdot)$ is concave. In the AFDC example, income variability occurred in a region where $B(\cdot)$ is convex.

We now discuss how to assay the magnitude of the effect we are examining. The certainty equivalent after-tax income, y_c, is defined as the solution to

$$B(y_c) = E[B(y)].$$

We measure the effect of the tax system on variability by the ratio:[2]

$$R = y_c/\bar{Y}.$$

If the tax system is linear, R will be equal to unity. Under a progressive tax system, $B(\cdot)$ will be concave and R will be less than unity. Under a regressive tax, system R will be greater than unity. The tax and transfer system in the United States is neither convex nor concave (see Hausman, 1983 and Moffitt, 1982); there can be no general presumption that R is greater or less than unity. We present calculations below which suggest that for reasonable estimates of the size of the variable component of income, R ranges from .78 to 1.38 for individuals with different tax statuses and different levels of expected income.

Before we present these estimates it is worthwhile trying to think about what R measures. At high income levels it is simply a way of measuring the extent to which the tax system encourages or discourages risk-taking. However, in this paper we focus on the poor. We do so for two reasons. First, as discussed in the next section, the amount by which R deviates from unity is determined by two things; the size of the kinks (i.e., the points at which marginal tax-rates change) in the tax system and the amount by which income varies. In Section III we present evidence which suggests that in the United States the tax and transfer systems combine to produce a very kinked tax system at low levels of income and that the earnings of the poor are quite variable. These two factors interact to produce values of R which differ significantly from unity. Our second reason for focusing on the poor is that for them R may be given a more dramatic interpretation than the mere encouragement or discouragement of risk-taking. For many of the poor, welfare and equivalent activities such as participation in training programs offer an income stream that is far more certain than that obtainable from participation in the labor market. R measures the ratio between expected income available in the labor market and the equivalent certain income which would make withdrawal from the labor market attractive. A low R — associated with progressive taxes — discourages labor market participation. A high R encourages it. Under this interpretation, an R of .8 means that a welfare program which offers certain income equal to 80% of expected (before-tax) income is as attractive as participation in the labor market. Alternatively, a guaranteed public jobs program at the same certain wage would be more attractive than private-sector employment.

An unrealistically simple model underlies the discussion of the preceding paragraph. We wrote as if the labor supply decision for an individual were equivalent to deciding once a year whether or not to participate in a lottery. If he does participate in the labor market he draws his annual labor market income from a distribution with known characteristics and anticipates a net gain of $E[B(y)]$. If he chooses not to participate he gets a certain income of, say y_o. In making his labor supply decision he compares these two quantities and chooses to enter the labor market only if $E[B(y)] > y_o$. This story is wrong for at least three reasons. First, the labor supply decision is made continually through the year; individuals drop in and out of the labor market and adjust their hours of work (when they can). Second, the benefits to be gained from not participating in the labor market are not certain; eligibility criteria and program benefits change. Third, we have assumed individuals are risk neutral. Taking these factors into consideration seems quite difficult. One of the contributions of this paper is the simple analytic formula of equation (1) below, which relates R to the variability of income and the curvature of the tax system. We do not know how to modify this formula to take these factors into account. It seems likely that a more realistic model of labor supply and of welfare programs would reduce the size and significance of the effects we report; abandoning risk neutrality would increase them.

II. HOW DIFFERENT IS R FROM UNITY? THEORETICAL RESULTS

The magnitude of the effect we are looking at depends on the variability of income and the shape of the tax system. In this section we present calculations which show those parameter combinations that will produce an R which is much different than unity. Our calculations are based on the following observation, which is proved in Appendix A. For a given probability distribution of income $F(\cdot)$, the progressive tax system with maximal marginal tax rate τ which minimizes R is that which taxes all income above y_c at rate τ and leaves all income below y_c untaxed. (Of course, y_c is not independent of the tax system; it is the solution to $B(y_c) = E[B(y)]$). The R-minimizing tax system is a linear tax system with a single kink, a form of taxation often advocated because of its simplicity.[3]

If income is lognormally distributed, then there is a simple analytic relationship (derived in Appendix B) between the marginal tax rate, τ, the coefficient of variation, c, and R:

$$(1) \qquad \tau = (1-R)/[(1-\Phi((\log R - \frac{1}{2}\sigma^2)/\sigma)) - R(1 - \Phi((\log R + \frac{1}{2}\sigma^2)/\sigma))]$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable and $\sigma^2 = \log(1+c^2)$. Table 1 gives the results of calculations which use this formula. In interpreting this table it is well to keep in mind that it represents extreme cases. If income is lognormally distributed with a coefficient of variation equal to unity and if the highest marginal tax rate is 51%, then the minimum value which R can assume is .8. R will attain this value only if the tax system has a single kink, leaving all income less than y_c untaxed and taxing all income greater than y_c at a rate of 51%. Furthermore only persons with a level of expected income equal to y_c will have an R of .8; other persons with a different permanent income (but with current income lognormally distributed with the same coefficient of variation) will have an R between .8 and 1. If the tax system has several brackets (but the top bracket is 51%) then R is larger than .8.

We choose the lognormal for use in Table 1 for three reasons. First, the lognormal permits us to obtain Table 1 analytically. Second, empirical work on earnings commonly uses the log of earnings as a dependent variable; it often assumes that earnings are lognormally distributed. As a consequence, most empirical reports of the variability of earnings are estimates of the variance of log earnings, σ^2. If earnings are lognormally distributed, there is a simple analytic relationship between this parameter and the coefficient of variation, c:

$$c = (\exp(\sigma^2) - 1)^{1/2}.$$

Since we measure variability of earnings by c, this correspondence is useful. Finally, experiments with exponential, binomial, and truncated normal

Table 1

Marginal Tax Rate Required to Generate
Different Values of R

Coefficient of Variation	Required Marginal Tax Rate	
	R = .9	R = .8
.25	.66	.90
.50	.43	.70
.75	.33	.58
1.0	.28	.51
1.25	.25	.46
1.35	.24	.44

Note: For the definition of R, see text.

distributions suggest that the results of Table 1 are robust; if the mean and variance of income are held constant, the first significant digits of the entries in Table 1 do not, for the most part, change when the table is re-calculated using these distributions.

Table 1 shows that if income is variable, kinks of moderate size could make R small for some people. Kinks, that is, changes in marginal tax rates, of the order of .24 or .44 are, at minimum, necessary to produce reasonably low R's. In the next section we present calculations for the United States which suggest that for some people R is as low as .78; for others it is as high as 1.38.

III. HOW DIFFERENT IS R FROM UNITY? ESTIMATES OF THE EFFECTS OF CURRENT UNITED STATES TAX AND TRANSFER PROGRAMS

In the previous section we have shown that, in the worst case, a single-kink tax system would have to have a marginal tax rate of more than .24 to generate an R of .90. In this section we calculate the ratio R for several major tax and transfer programs in the United States to see if the kinks in those programs can generate an R more than .10 or so away from one. We shall assume that the log of earnings is normally distributed with mean μ and variance σ^2. Expected income is given by $\bar{Y} = \exp(\mu + \frac{1}{2}\sigma^2)$. As we have observed above, σ is determined by the coefficient of variation, c. We shall perform our calculations for different values of expected income, $\bar{Y}$, and c.

Since some of the largest kinks in the tax and transfer system occur near the bottom of the income distribution, we shall pick several low values of $\bar{Y}$. But which values of c we should examine is not as obvious. Cross-section earnings regressions, such as those of Mincer (1974), provide values of c (calculated from the estimated variance of log earnings) in the neighborhood of .60-.90. However, it could be argued that cross-sectional variation exaggerates the variation in transitory income for a single individual because of heterogeneity across individuals. Many authors have instead estimated earnings equations from panel data; such a procedure yields an estimate of the variance in earnings for the same individual over time. Most studies have done so with some sort of error components structure — see Lillard and Willis (1978) for one example and MaCurdy (1982) for a more recent one with a very general structure for the error term. These studies usually imply coefficients of variation (of the transitory component of the error term) of from .50 to .80, with substantial differences across studies. Unfortunately for our purposes, the variance of the transitory error is always assumed to be independent of the level of permanent income, so no information is provided on whether the (permanently) low-income population has a larger variance than the rest of the population. Studies which focus only on low-income individuals (e.g., Hausman and Wise, 1979), maintain the same independence assumption in the stochastic specification of the error term. Fortunately, Gottschalk (1980) has calculated the standard deviation of earnings over a six-year period for each individual in the National Longitudinal Survey of Men, and has tabulated the mean of the standard deviations by the level of mean earnings over the period. Professor Gottschalk graciously performed special tabulations for us which show that that coefficient of variation in the lowest quintile of the population is 1.35. It falls to about .50 in the next highest quintile and then down to about .20 for the rest of the population. Since we are interested in the low-income sector of the population, we shall therefore perform our calculations for c's ranging from .50 to 1.35.[4]

Federal Income Tax

Figure 1 shows the disposable income schedules for several tax and transfer programs. The first panel refers to the federal personal income tax, including the earned income tax credit (EITC). For all our calculations we have assumed a four-person family with a single earner and no other income. All tax and transfer programs are evaluated in their 1982 forms. As the panel shows, the present income tax actually subsidizes earnings up to \$5,000 of income. This feature has been present since the mid-1970s and is a result of the EITC, which is a refundable tax credit for which families with dependents are eligible. While it is an earnings subsidy up to \$5,000, it becomes an earnings tax beginning at \$6,000 because it must be phased out. The marginal tax rate begins at -.10, is 0 at \$5,000 and then becomes a positive .13 at \$6,000. A concave kink of about .23 (.10 and .13) is thereby created (note that this is close to the .24 of Table 1). A further kink, one which is convex, is created at about \$11,000 of earnings, the point at which the subsidy drops to zero and the regular marginal tax rates begin.

Our calculated values of R for the federal tax are shown in the first five rows of Table 2. As should be expected, R is inversely related to c. However, it is quadratically related to mean income because our middle ranges of income fall around the concave kink and our high ranges fall around the convex kink. (R would fall again at higher income levels.) As the results show, R does fall to .90 and a bit below when c takes on a value of 1.35 and the individual's permanent income is slightly above the concave kink (\$7,500 in the table). In this particular case y_c is \$6,685, more than \$800 below $\bar{Y}$. This is a large amount in absolute terms. It is also large relative to the effect of the tax system on after-tax permanent income--at $\bar{Y}$ = \$7500, the tax system provides about a \$250 <u>subsidy</u> to working as a result of the EITC

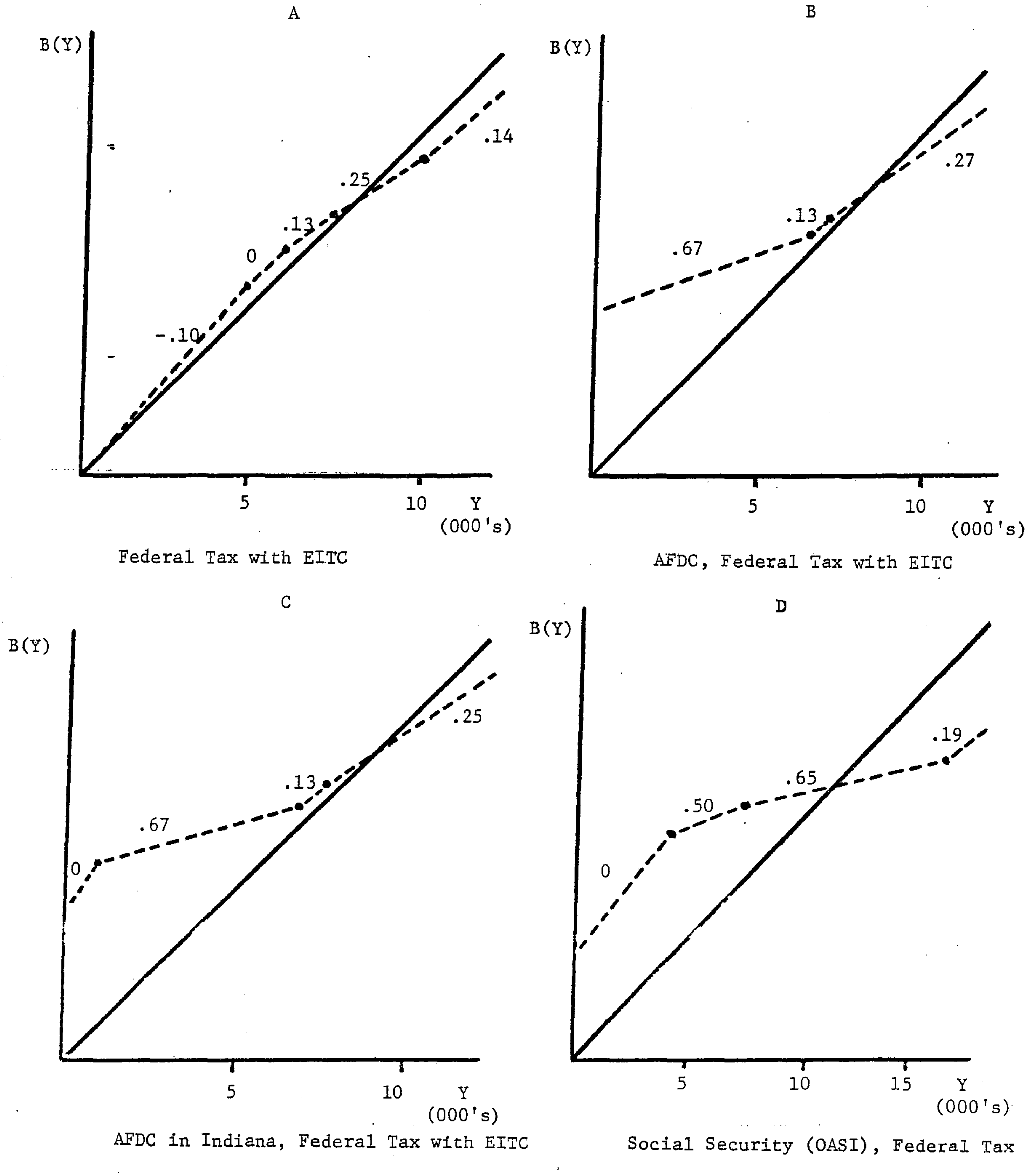

Figure 1. After-tax income, B(Y), as a function of current pretax income, Y, under four tax and transfer systems. The dotted curve is after-tax income; the solid line is at 45 degrees, representing no taxation. Numbers between the kinks in after-tax income are marginal tax rates.

Table 2

Values of R for Different Tax and Transfer Systems

Expected Income, $\bar{Y}$	Coefficient of Variation, c .50	.75	1.0	1.35
Federal Tax with EITC				
2,500	1.00	.99	.97	.95
5,000	.96	.94	.92	.90
7,500	.96	.94	.92	.89
10,000	1.00	.99	.96	.93
12,500	1.00	.98	.96	.93
AFDC, Fed. Tax with EITC				
2,500	1.01	1.04	1.10	1.19
5,000	1.12	1.23	1.32	1.38
7,500	1.08	1.12	1.16	1.18
10,000	1.04	1.06	1.08	1.09
12,500	1.01	1.02	1.02	1.03
AFDC in Indiana, Fed. Tax with EITC				
2,500	1.00	1.01	1.03	1.06
5,000	1.10	1.21	1.29	1.37
7,500	1.09	1.14	1.17	1.20
10,000	1.04	1.07	1.08	1.10
12,500	1.01	1.02	1.03	1.03
Social Security, Fed. Tax				
2,500	1.00	.98	.95	.92
5,000	.93	.89	.86	.83
7,500	.87	.82	.80	.78
10,000	.98	.97	.96	.93
12,500	1.09	1.11	1.12	1.10

(indeed the intent of the EITC is to encourage work); but the effect of risk on expected after-tax labor earnings is almost three times this amount, and negative. Thus the phenomenon we are examining in this paper seems quantitively important relative to the usual effects of tax systems discussed in certainty models —at least if the coefficient of variation is high and if an individual's permanent income falls in a certain range.

Aid to Families with Dependent Children (AFDC)

The major transfer program for poor families in the United States is AFDC. However, in most states only families without an able-bodied male present in the household can receive it.[5] Hence the female-headed household population is the largest beneficiary. In 1982 the statutory tax rate in the program in most states was .67, as shown in Panel B of Figure 1. This tax rate applies until benefits end, at which point the federal income tax system (including the EITC) takes over.[6] However, in some states--Indiana is an example (Hausman, 1983)--a range of zero marginal tax rates is imposed at low earnings levels. In Indiana, as shown in Panel C, the zero rate applies over the first $1,000 of earnings.

The values of R for AFDC in Table 2 show that the convex kinks generated by the transfer program can have sizable effects. At c = 1.35, R reaches 1.38 if permanent income is $5,000, not far from the convex kink. The value of y_c in this case is $6,900, almost $2,000 above $\bar{Y}$--again a large amount. Moreover, it should be noted that all the R values we have calculated are greater than 1.0, even those at $\bar{Y}$ = $12,500. At higher earnings levels, of course, R would fall below 1.0. But at $12,500 and below, the c's we have used imply that the risk of falling below the eligibility point is sufficiently high that the work-encouraging effects of the program dominate the work-discouraging effects of the progressive rates at higher points. The same is true in the Indiana AFDC system — at the $\bar{Y}$'s we have calculated, all the R's are still greater than 1 (even though somewhat smaller than before because of the presence of the concave kink).[7]

Social Security

We have also calculated our R ratios for the Social Security retirement program (OASI) because it contains (to the authors' knowledge) the most severe concave kink of any major tax or transfer program in the United States.[8] As shown in Panel D of Figure 1, Social Security benefits are not reduced unless earnings rise above about $5,000, at which point they begin to be taxed at a 50 percent rate. The federal tax begins at about $7,000, whereupon the marginal tax rate rises to .65.[9] Benefits end completely at about $16,000 of earnings, at which point a large convex kink is generated.

As Table 2 shows, this program generates the lowest values of R of any program we have considered. Those with permanent income in the range $5,000 - $7,500 — who are most affected by the concave kink — have values of R that are almost always below .90 and even below .80 at high values of c. At $\bar{Y}$ = $7,500 and c = 1.35, y_c is almost $2,000 below $\bar{Y}$. Consequently, to the extent that the aged face risk in their labor earnings, this program would seem to have significant work-discouraging effects at these low earnings levels. Of course, at higher levels closer to the convex kink, values of R are greater than one, as they are for any transfer program.

IV. SUMMARY AND CONCLUSION

In this paper we have shown that the U.S. tax and transfer system has potentially large effects on risk-taking in the labor market, effects which vary greatly across individuals. Depending upon one's eligibility for different types of transfer programs and upon the mean and variance of an individual's earnings, the effect may be quite large. For many, of course, the effect discussed here will be insignificant. But the potential for large effects suggests that these effects be taken into account in the design of such programs--along with the usual efficiency and equity effects.

That those other effects--which we have ignored--are important is a point we wish to stress. A superficial reading of this paper would suggest that its main lesson is that regressive tax systems are desirable because they encourage work. This is, of course, not our conclusion. We merely urge that the effect of sharp kinks in the tax system on those with variable earnings be taken into account in the design and analysis of tax systems. This topic cries out for treatment as an optimal tax problem and for a more realistic model of the labor supply decision. Both problems must be left for future work.

FOOTNOTES

[1]As Gary Chamberlain has pointed out to us, our analysis does not depend on ε being subjectively uncertain. It could represent known variations in earnings which will occur as an individual ages and his value in the labor market changes. The analysis we give below applies to these predictable changes —with certain obvious changes to deal with discounting (see Rothschild, 1969). However, to simplify our exposition we shall refer to all variability as risk.

[2]Those familiar with Atkinson's (1970) work on the measurement of inequality will note that y_c is the same as his equally-distributed equivalent: our R is equal to 1-I where I is his measure of inequality.

[3]An obvious analogous result applies to regressive tax systems. In the next section we note that in fact convex kinks (regressive portions of the tax and transfer system) are as important in driving R away from unity as are concave kinks.

[4]Mirer (1974) presents similar estimates over a three-year period. We use Gottschalk's because they come from direct calculations of each individual's standard deviations of earnings, whereas Mirer's come from the variances of regression error terms. As the discussion in footnote 1 above indicates, both calculations underestimate variability because they disregard predictable growth of income.

Gary Chamberlain has pointed out to us that there are difficulties in inferring from panel data that the variance of income differs for people according to their average income. A complete model would explain how variables such as education and age determine average income, as well as the variability of income.

[5]Some states offer AFDC benefits to such families, but only if the male is unemployed. Participation rates in the program are consequently low. Such families can receive Food Stamps, however, but the cash equivalent values are inconsequential relative to AFDC.

[6]Moffitt (forthcoming) and others have shown that the effective tax rate in the AFDC program is considerably lower than this. We also calculated our R values for these lower tax rates, but we do not present the results because their qualitative nature is the same as those here. The convex kink merely occurs at a higher income level.

[7]We would like to note at this point that the lognormal assigns a zero probability to nonwork, possibly an undesirable feature when studying a transfer program. To see its effects we also simulated and made our calculations with a displaced lognormal with 20 percent of the distribution negative (i.e., a one-fifth chance of being unemployed). The resulting R ratios were all in the same range as those in Table 2.

[8]The case of a "notch" — where benefits are discontinuously reduced to zero and the marginal tax rate is over 100 percent — is even more concave. The R ratios would be even smaller in this case.

[9]Since the EITC can be received only by families with dependents, we assume that the aged do not receive it.

REFERENCES

Atkinson, Anthony B. 1970. "On the Measurement of Inequality." Journal of Economic Theory, 2, 244-263. Reprinted in A.B. Atkinson, Social Justice and Public Policy (Cambridge, Mass.: MIT Press, 1983).

Gottschalk, Peter. 1980. "Earnings Mobility: Permanent Change or Transitory Fluctuations?" Discussion Paper 604-80, University of Wisconsin, Madison, Institute for Research on Poverty.

Hausman, Jerry. 1983. "Taxes and Labor Supply." Working Paper No. 1102, National Bureau of Economic Research. Cambridge, Mass.

Hausman, Jerry, and David Wise. 1979. "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment." Econometrica, 47, 455-475.

Johnson, Norman I. and Samuel Kotz. 1970. Distributions in Statistics: Continuous Univariate Distributions--1. New York: John Wiley.

Lillard, Lee, and Robert Willis. 1978. "Dynamic Aspects of Earnings Mobility." Econometrica, 46, 985-1012.

MaCurdy, Thomas E. 1982. "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis." Journal of Econometrics, 18, 83-114.

Mincer, Jacob. 1974. Schooling, Experience, and Earnings. New York: Columbia University Press.

Mirer, Thad. 1974. "Aspects of the Variability of Family Income." In Five Thousand American Families--Patterns of Economic Progress, ed. James N. Morgan, Vol. 2. Ann Arbor, Mich.: Institute for Survey Research, University of Michigan.

Moffitt, Robert. 1982. "The Econometrics of Piecewise-Linear Budget Constraints: An Exposition." Unpublished paper, Department of Economics, University of Wisconsin, Madison.

Moffitt, Robert. 1982. Forthcoming. "An Economic Model of Welfare Stigma." American Economic Review.

Rothschild, Michael. 1969. "Changing Demand, Its Costs and Consequences." Chapter 3 in "Essays in Economic Theory," unpublished Ph.D. dissertation, MIT.

APPENDIX

A. We represent a tax system by a function B(y) which gives the amount of after-tax income retained by a person with a pretax income of y. A progressive tax system is one in which B(•) is increasing, concave, and satisfies

(2) $B(0) = 0.$

Let τ be the maximum marginal tax rate of a tax system. Then

$$\tau = 1 - \lim_{y \to \infty} B'(y).$$

We say a tax system is τ-admissible if it is increasing, concave, satisfies (2) and its maximum marginal tax rate is less than or equal to τ.

An example of a τ-admissible tax system is the piecewise-linear system which does not tax income which is less than or equal to k and taxes income greater than k at rate τ. We denote such a tax system $B_{\tau,k}(\cdot)$ and define it as follows:

$$B_{\tau,k}(y) = \begin{cases} y & \text{if } y \leq k \\ k + (1-\tau)(y-k) & \text{if } y > k \end{cases}$$

Throughout this appendix we consider the effect of a tax system B(•) on expected after-tax income when pretax income y is a random variable with support [0, M]. For a tax system B(•), define y_c as the solution to

(3) $B(y_c) = E[B(y)];$

y_c is a certainty equivalent. Under a tax system B(•), a certain income of y_c gives the same expected after-tax income as the random income y.

For any $\tau\varepsilon(0,1)$ there is a unique solution to the equation

$$(4) \qquad \hat{y} = E[B_{\tau,\hat{y}}(y)].$$

To see this, let $g(x) = x$ and $h(x) = E[B_{\tau,x}(y)]$. We seek a solution, $\hat{y}$, to $g(\hat{y}) = h(\hat{y})$. Now $g(0) = 0$, and $h(0) = (1-\tau)E[y] > 0$. However, for $\varepsilon > 0$, $g(M+\varepsilon) = M+\varepsilon > h(M+\varepsilon)$. Thus there is at least one solution to (4). There is only one since $g'(x) = 1$ and $h'(x) = \tau[1-F(x)] < 1$.

Proposition: Let $\hat{y}$ be the solution to (4); then $B_{\tau,\hat{y}}(\cdot)$ is the τ-admissible tax system which minimizes y_c where y_c is defined in (3).

The proof is a consequence of the following simple result.

Lemma: For any z in the support of y, $B_{\tau,z}(\cdot)$ is the τ-admissible tax system which minimizes

$$\frac{E[B(y)]}{B(z)}.$$

Proof of Proposition: Suppose there is a τ-admissible tax system $\tilde{B}(\cdot)$ other than $B_{\tau,\hat{y}}(\cdot)$ such that $\tilde{B}(\tilde{y}) = E[\tilde{B}(y)]$ and $\tilde{y} < y_c$ where y_c is the solution to (3) for $B(y) = B_{\tau,\hat{y}}(y)$. Then, $E[\tilde{B}(y)] = \tilde{B}(\tilde{y}) < \tilde{B}(y_c)$. Thus,

$$\frac{E[\tilde{B}(y)]}{\tilde{B}(y_c)} < 1.$$

But $E[B_{\tau,y_c}(y)] = B_{\tau,y_c}(y_c)$, so

$$\frac{E[B_{\tau,y_c}(y)]}{B_{\tau,y_c}(y_c)} = 1.$$

Thus $$\frac{E[\tilde{B}(y)]}{\tilde{B}(y_c)} < \frac{E[B_{\tau,y_c}(y)]}{B_{\tau,y_c}(y_c)},$$

which contradicts the Lemma.

Proof of Lemma: Let

$$R = \frac{E[B(y)]}{B(z)}$$

$$= \frac{E[B(y)|y<z]}{B(z)} + \frac{E[B(y)|y>z]}{B(z)}$$

$$= R_1 + R_2.$$

We show that $B_{z,\tau}$ is a τ-admissible tax system which minimizes both R_1 and R_2. Clearly the τ-admissible tax system which does not tax income $\leq z$ and taxes all income $> z$ at rate τ minimizes R_2. Such a system minimizes the numerator of R_2 while it maximizes the denominator. Consider the problem of finding a τ-admissible tax system which minimizes R_1. We need only concern ourself with how income is taxed on the interval $[0, z]$. The numerator of R_1 is (proportional to) expected after-tax income when income is distributed on $[0, z]$; the denominator of R_1 is after-tax income on income of z. Any progressive tax system will decrease the denominator of R_1 proportionately at least as much as it decreases the numerator, for under a progressive system a person with income less than z is taxed no more heavily than a person with income z. Thus the progressive tax systems which minimize R_1 are linear over the range $[0, z]$. Since $B_{z,\tau}(\cdot)$ is linear over $[0, z]$ it is τ-admissible and minimizes both R_1 and R_2.

B. The following equation is used to compute Table 1:

$$(5) \qquad \tau = (1-R)/[(1-\Phi((\log R - \tfrac{1}{2}\sigma^2)/\sigma)) - R(1 - \Phi((\log R + \tfrac{1}{2}\sigma^2)/\sigma))]$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable. To derive (5) observe first that if $B(\cdot)$ is the tax system which minimizes R and if y is distributed with density $f(\cdot)$ and cumulative distribution function $F(\cdot)$, then

$$y_c = B(y_c) = E[B(y)] = \int_o^\infty yf(y)dy - \tau \int_{y_c}^\infty (y-y_c)f(y)dy.$$

Thus,

$$(6) \qquad y_c = Ey + \tau y_c[1 - F(y_c)] - \tau \int_{y_c}^\infty yf(y)dy.$$

Now if log y is normally distributed with mean μ and variance σ^2, then

$$(7) \qquad Ey = \bar{Y} = \exp(\mu + \tfrac{1}{2}\sigma^2).$$

Furthermore, it can be shown (Johnson and Kotz, 1970, p. 129) that

$$(8) \qquad \int_{y_c}^\infty yf(y)dy = \bar{Y}\ [1 - \Phi((\log y_c - \mu - \sigma^2)/\sigma)].$$

Since

$$(9) \qquad R = y_c/\bar{Y},$$

we can substitute in (8) to obtain

$$(10) \quad \int_{y_c}^{\infty} yf(y)dy = \bar{Y}[1 - \Phi((\log R + \log \bar{Y} - \mu - \sigma^2)/\sigma)]$$

$$= \bar{Y}[1 - \Phi((\log R - \frac{1}{2}\sigma^2)/\sigma)].$$

The last step follows from (7). A similar substitution shows that

$$(11) \quad 1 - F(y_c) = 1 - \Phi((\log R + \log \bar{Y} - \mu)/\sigma)$$

$$= 1 - \Phi((\log R + \frac{1}{2}\sigma^2)/\sigma).$$

Substituting (9), (10) and (11) into (6) and rearranging, we obtain (5).