The 100% Tax in Theory and Practice*

by

Casey B. Mulligan

University of Chicago

November 1998

Abstract

Taxes often discourage work effort. Is this a necessary but unfortunate byproduct of policies intended to redistribute from rich to poor? Or is discouraging work effort the primary goal of policy? I examine the case of 100 percent labor income tax rates and argue that 100% tax rates have some of the expected disincentive effects and hence that their prevalent use is inconsistent with the goal of redistributing from rich to poor. I then introduce a model where the primary policy goal is to discourage taxable income and show how 100 percent taxes are optimal over a relatively wide range of incomes, arguing that discouraging taxable income as a primary policy goal provides a better description of policies actually used -- especially policies with respect to the elderly. Such a goal may also explain high but less-than-confiscatory tax rates, minimum wage laws, maximum hours laws, and the world-wide lack of means testing of social security programs.

---

*I gratefully acknowledge comments by Fernando Alvarez, Pierre-André Chiappori, Erzo Luttmer, Xavier Sala-i-Martin, and Aaron Yelowitz.

Table of Contents

I. Introduction

Labor income is sometimes taxed at a 100 percent rate. Why? This paper explores four explanations: (1) tax policy makers just made a mistake, (2) people enjoy work and/or paying taxes at the margin, (3) a 100 percent tax over some range maximizes social welfare because it provides greater work incentives for those with incomes above that range, and (4) the primary intent of policy is to discourage taxable income. I showed that 100 percent taxes are quite common in practice and that almost no one works who is subject to a 100 percent tax, and thus am skeptical about explanations (1) and (2).

I explore the 100 percent tax in the Mirrlees model of optimal redistribution. A 100 percent tax over some range can be optimal because it provides greater work incentives for those with incomes above that range. However, I argue that the 100 percent tax ranges used in practice are much too wide to be justified on incentive grounds.

Starting with the pretty weak hypothesis that a 100 percent tax discourages taxable income, such a tax makes a lot of sense if the policy objective is to discourage taxable income. One reason for such an objective is that suppliers of fringe benefits may be politically powerful. Another possibility, suggested by Sala-i-Martin (1996), is that some people's taxable income generates a negative externality. Or, an interest group's political power may be enhanced when its members work less -- as claimed by Mulligan and Sala-i-Martin (1998).

II. 100% Taxes in Practice

II.A. 100% Tax Defined

I refer to any tax or benefit formula involving no gain in net-of-tax-and-transfer income with an increase in labor income as a "100 percent tax".[1] This obviously includes reductions in benefits with earnings as well as increases in what the layman calls "taxes". This also includes tax rates which are apparently greater than 100 percent because, in the usual models of response to taxes, there is no difference between a tax rate of $\tau_1 > 1$ on incomes in the interval $[y_1, y_2]$ together with a tax rate of $\tau_2 < 1$ on the interval $[y_2, y_3]$ and a single tax rate of exactly 100% on incomes between $[y_1, y_3]$.[2]

II.B. American Examples

In recent years, 100 percent labor income taxes have been somewhat uncommon in the U.S., with the "Medicaid Notch" as the only notable exception. The Medicaid Notch refers to the range of labor income for unmarried women with children where a small increase in income results in an elimination of Medicaid benefits (Yelowitz 1995), and may have been reduced by recent policy (House Committee p. 882). Under some conditions, recent AFDC beneficiaries faced 100% tax rates (House Committee p. 390). An even larger fraction of AFDC beneficiaries faced 100% tax rates before 1971, although it has been disputed in the literature whether effective AFDC tax rates ever reach 100% (eg., Fraker et al 1985).[3] Some American public housing programs may tax labor income at a 100% rate over a range of labor income (Currie and Yelowitz 1998).

Historically, there have been American 100 percent labor income tax rates. An important one derived from Social Security benefit formulas between 1939 and 1971. Between 1939 and 1959 retirees lost all of their Social Security benefit if their earnings exceeded a rather low earnings limit

[1]I do not analyze 100% taxes on savings. See Hubbard, Skinner, and Zeldes (1995) for one American example of a 100% tax on savings.

[2] $y_3 = \dfrac{(\tau_1 - \tau_2)\, y_2 - (\tau_1 - 1)\, y_1}{1 - \tau_2}$. The two tax schemes are not equivalent if workers do not have enough control over their incomes to make sure their income does not fall in the interval $[y_1, y_3]$.

[3]Little work has been done to bring in the costs and opportunities of obtaining deductions and earnings disregards to determine the effective AFDC tax rate, as was done for the personal income tax by Barro and Sahasakul (1983).

by even one dollar![4,5] The limit was $15/month from 1939-49, $50/mo in 1951 and 1952, $75/mo in 1952 and 53, and $80/mo for 1954-59 (Myers 1993 pp. 271-274). For a manufacturing production worker, these limits represented less than 8, 18, 25, and 23% of monthly earnings (Census Bureau 1975 series D-804), respectively, while lost benefits were typically 30% of the average worker's wages and a larger percentage of his after-tax wages (House Committee 1996, Table 1-14).

More examples of 100 percent taxes in the U.S. are found in public assistance programs prior to the 1970s. State administered public assistance programs typically (implicitly) taxed earnings at a 100 percent rate (Myers 1993 pp. 827, who also points out that some states administering old age assistance exempted the first 80 dollars of monthly income).

II.C. International Examples

100 percent taxation of elderly labor income is and has been common internationally. Elderly Spaniards and Belgians are not allowed to collect their government pension if they earn any labor income at all (Boldrin et al 1997 p. 16, SSA 1997 p. 330, Pestieau and Stijns 1997, p. 9[6]) and those benefits are typically close to or more than what the pensioner would have earned after taxes (Boldrin et al 1997). France allows pensioners to receive labor income, but not from their preretirement occupation (Blanchet and Pele 1997, p. 9, SSA 1997, p. 130). No Austrian under age 65 earning more than 3740 schillings/month may collect a public pension (= $349/month, SSA 1997 p. 21). More international examples of 100 percent taxes on elderly labor income can be found (SSA 1997).[7] Furthermore, the public pension benefits in these and other countries are nearly as much as an average worker's earnings (Gruber and Wise 1997) - and therefore the range of income to which the 100%

---

[4]Those aged 75 and over were exempt from the earnings test beginning in 1950, and those aged 72 and over exempt beginning in 1954 (Myers 1993 pp. 272-3).

[5]There were no "delayed retirement credits" before 1975 (House Committee 1996, p. 31), so foregoing benefits in one year did not affect the amount of benefits enjoyed in later years.

[6]This rule may have changed in recent years since, according to SSA 1997, p. 35, 282118 francs ($8993) could be earned in 1997 without sacrificing the pension benefit.

[7]Some of this implicit tax might be avoided by substituting work over time and taking advantage of (the country-specific version of American) delayed retirement credits, although as in the U.S. delayed retirement credits have not always been part of old age pension formulas (SSA 1967, 1997).

implicit tax applies is very large.

100 percent taxation of elderly labor income in other countries is not especially new, and probably has become less common over time. In 1967, for example, German public pension formulas taxed elderly labor income of those under age 65 at 100 percent, with no exemption amount (SSA 1967 p. 80); Canadian public pension formulas taxed elderly labor income of $900-1500/year at 50% and $500+/year at 100% (SSA 1967 p. 32; in 1997 US$ these limits are 4000 and 6667); British public pension formulas 100% taxed labor incomes (with a 1967 £5/week = $14/week = 1997 $67/week exemption for those aged 70+, SSA 1967 p. 222).

II.D. Empirical Regularities

These examples reveal several empirical regularities. First, 100 percent taxes often apply to a wide range of labor incomes; in some of the examples is probably the case that the majority of income that would have been earned by program beneficiaries in the absence of the program is subject to the 100 percent tax rate. Second, 100 percent taxes more typically apply to old people. Third, the 100 percent taxes do discourage work effort.[8] Fourth, it is often the case that some benefits are not subject to the 100 percent tax rate, such as medical benefits for the elderly. Fifth, 100 percent tax rates apply to lower, rather than higher, labor incomes. Often there is some small amount of labor income (or, in the case of France, a little bit of labor market activity) exempt from the 100 percent tax rate.

I exclude disability programs from my examples for two reasons. Doing so limits the scope of the paper. It is also possible that insurance motives (rather than, or in addition to, redistribution motives) are needed to understand design of disability programs, although I am skeptical that insurance considerations would justify 100 percent taxes any better than redistributional ones (Diamond and Mirrlees 1978, 1986).

III. Utility of Work and Taxes

If people are willing to pay to work and/or enjoy paying taxes (as in Bernheim 1986), then tax rates of 100 percent or higher might be optimal. However, this story implies that a number of

[8] See Section III below or Kotlikoff and Wise (1989) for one of the many empirical studies in the literature.

people will work even when subject to a 100 percent tax while, in fact, it is quite rare for someone subject to a 100 percent tax rate to work. For example, the 1995 Belgian and Spanish male labor force participation rates were 5% and 11%, respectively, at age 65 (the first age at which the 100% tax applies to a wide range of labor income; Pestieau and Stijns 1997, Figure 6; Boldrin et al 1997, Figure 7).

IV. 100% Taxes as Optimal Redistribution

III.A. The 100% Tax in the Mirrlees Problem

Consumers differ according to their labor productivity, w, which cannot be observed by those enforcing policy. Instead, only a person's total labor product -- the product of his productivity w and his work effort $n(w)$ -- can be observed.

The social objective is utilitarian. The fraction of consumers with labor productivity less than w is $F(w)$, and the distribution of productivity has support $[0,\bar{w}]$.

$$W(G) \equiv \max_{c(\cdot),\, n(\cdot)} \int_0^{\bar{w}} u\big(c(w),\, n(w)\big)\, f(w)\, dw \quad \text{subject to:}$$

$$\int_0^{\bar{w}} [\, w n(w) - c(w)\,]\, f(w)\, dw \;\geq\; G$$

$$\frac{d}{dw} u\big(c(w),\, n(w)\big) \;\geq\; -u_n\big(c(w),\, n(w)\big) \frac{n(w)}{w}$$

$$c(\tilde{w}) = c(w) \quad \text{for all } \tilde{w} \text{ such that } n(\tilde{w})\tilde{w} = n(w)\, w \quad , \text{ all } w$$

The first constraint says that total tax revenue must be at least G. The second constraint is an incentive compatibility constraint: a person of productivity w must not desire to (by working less) pretend to be less productive than he is.[9] The third constraint dictates that labor productivity can only be inferred from total labor product.

If those enforcing policy had perfect information about w or labor supply were completely

[9] Derivations of the incentive compatibility constraint can be found in the literature (eg., Mirrlees (1971), Atkinson and Stiglitz (1980), or Cooter (1978)), so I do not reproduce one here.

inelastic, then lump sum taxes would be available. For example, under separable utility, each person would be assigned the same consumption, work effort would be positive, the more productive would exert more effort and be worse off, and we would have what looks like a 100% tax over the entire range of labor incomes. These special case is obviously counterfactual - people do respond to tax rates of 100% by exerting less effort and we do not see 100% taxes levied on the entire income distribution. So I move on to the potentially more interesting case that productivity is unobservable and labor supply is somewhat elastic.

In order to show that actual policies do not solve this optimal program, suppose for the sake of argument that the qualitative features of actual and optimal policies are the same: a 100 percent marginal tax applies to persons with productivity less than w_0 and tax rates strictly less than 100 percent applies to all others.[10] Then the critical productivity for the optimal program can be described according to:

$$W(G) = \max_{c(\cdot),\, n(\cdot),\, c_0,\, w_0} F(w_0)\,u(c_0, 0) + \int_{w_0}^{\bar{w}} u\big(c(w), n(w)\big) f(w)\,dw \quad \text{subject to:}$$

$$\int_{w_0}^{\bar{w}} [wn(w) - c(w)]\, f(w)\,dw \geq G + F(w_0)c_0$$

$$u(c(w_0), n(w_0)) \geq u(c_0, 0) \quad , \quad \frac{d}{dw}u\big(c(w), n(w)\big) \geq -u_n\big(c(w), n(w)\big)\frac{n(w)}{w}$$

$$c(\tilde{w}) = c(w) \quad \text{for all } \tilde{w} \geq w_0 \text{ such that } n(\tilde{w})\tilde{w} = n(w)w \quad , \text{ all } w \geq w_0$$

For $w \geq w_0$, the objective and constraints are the same as in the previous statement of the problem. Because those with $w < w_0$ have no earnings, they must be assigned the same consumption by the planner. The program also includes the constraint that a person of productivity w_0 is indifferent between working and not working.

I assume the LaGrangian is concave in w_0 in the relevant range. Denoting the marginal social

[10] For another argument that tax rates ought to be 100% over some range, see Besley and Coate (1995). Those in their model subject to the 100% tax rate are required to work for the government.

value of revenue as λ, the marginal social incentive cost of raising $u(c_0,0)$ as q_0, and $f(w) = F'(w)$, the first order condition with respect to w_0 is (1):

$$f(w_0)\left[n(w_0)w_0 + c_0 - c(w_0)\right] + \frac{q_0}{\lambda}u_n\left(c(w_0), n(w_0)\right)\frac{n(w_0)}{w_0} = 0 \qquad \textbf{(1)}$$

The first term is the direct revenue gain associated with moving those at the very top of the 100 percent tax bracket into the labor force. This hurts incentives for those already in the labor force, the effect of which on social welfare is valued in the second term above and can be seen in Figure 1 graphing below individual utilities as a function of type.

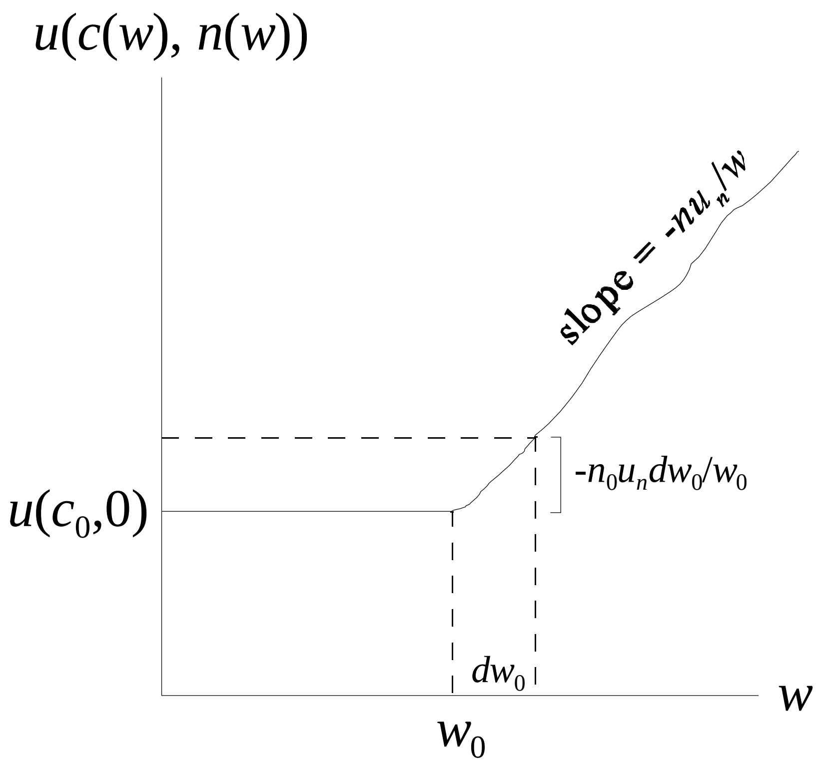

Figure 1 Individual Utility by Type

If at some w the left-hand side of the first order condition (1) is positive, then the marginal tax rate in the neighborhood of w is strictly less than 100 percent. Let w_0* be the critical productivity implied by actual policy (ie., $n(w_0^*)w_0^*$ is the smallest labor income subject to a tax rate of less than

100 percent). Then the following inequality implies that actual policy applies a 100 percent marginal tax to too much labor income:

$$f(w_0^*)\left[n(w_0^*)\,w_0^* \;+\; c_0^* \;-\; c(w_0^*)\right] \;+\; \frac{q_0}{\lambda}\,u_n\!\left(c(w_0^*),\,n(w_0^*)\right)\frac{n(w_0^*)}{w_0^*} \;>\; 0 \qquad \textbf{(2)}$$

I present two strategies for demonstrating this inequality. The first assumes a utility functional form and calibrates consumption and earnings from the tax system of Western European countries. The second strategy argues that the critical productivity should be greater for young workers than for old. Since actual policies typically subject more elderly income to 100 percent rates, 100 percent rates have been set suboptimally for young workers or for old workers or for both.

III.B. Calibration of the Mirrlees Problem

Although some 100% taxation is consistent with the Mirrlees model, I claim that the amount of 100% taxation used in practice is too much from the Mirrlees point of view. To support my claim, I assume parametric forms for the utility and productivity distributions, calibrate the parameters from the Belgian tax system, and show that the inequality (2) holds. But I believe the basic claim - that the revenue gain from increasing w_0* far exceeds the incentive value - would hold for many realistic functional forms and parameter values.

Suppose for the sake of argument that the qualitative features of actual and optimal policies are the same: a 100 percent marginal tax applies to persons with productivity less than w_0 and tax rates strictly less than 100 percent applies to all others. If the utility function takes the form $u(c,n) = \ln c - v(n)$ for some increasing and convex function v, then the inequality (2) becomes:

$$f\!\left(F^{-1}(\varphi_0^*)\right)\left[n(w_0^*)\,w_0^* \;+\; c_0^* \;-\; c(w_0^*)\right] \;>\; \varphi_0(1-\varphi_0)\,\frac{E\!\left(c(w)\,|\,w > w_0\right) - c_0}{c(w_0)}\,n_0(1-T')$$

Given f lognormal, five parameters are needed: (1) the fraction of the relevant population not working (φ_0), (2) the variance of log productivity, (3) the revenue gain associated with moving the most productive nonworker to employment (the term in square brackets above), (4) the percentage gap between the average consumption of workers and nonworkers, and (5) the marginal tax rate facing the least productive worker (T'). I set $n_0 = 1$, thus fixing the time units to be the amount of time worked by the least productive worker who finds it worthwhile to work rather than taking the benefit paid to those not working. This inequality is less likely to hold the closer is φ_0 to 0.5, the larger the variance, the smaller is the revenue gain, the larger the percentage gap between average consumption of workers and nonworkers, and the smaller the marginal tax rate on those working. My conservative calibration therefore takes $\varphi_0 = 0.2$ (which is larger than the fraction of those aged 65 not working in the three of the 100% implicit social security tax countries (Austria, Belgium, Spain) for which I have data[11]), a variance of log unobserved productivity of 0.25, an annual revenue gain of 3 months salary, a percentage gap between workers and nonworkers of 50%[12], and a marginal tax rate of zero. These conservative parameters still satisfy the inequality:

$$0.21 \;=\; 0.85\left[0.25\right] \;>\; 0.2(0.8)\,0.5\,(1) \;=\; 0.08$$

Since my calculations are based on Social Security programs where the fraction not working is pretty large, they do not rule out the possibility that AFDC and other 100% taxes on a small fraction of labor inomce are optimal forms of redistribution.

III.C. Mirrlees Problem with Multiple Observable Types

Most countries using 100 percent taxes use them more heavily on elderly consumers, and typically collect more revenue per capita from the young and pay more subsidies per capita to the old. I extend the Mirrlees model to evaluate the optimality of such policies.

Consider the simple case that utility functions and productivity distributions are the same for the young and old "types". In this case, the only difference between the two types is their numbers (Π and $1-\Pi$) and the weights (θ and $1-\theta$) they receive in the social welfare function:

---

[11] Gruber and Wise (1997).

[12] Although I do not have any direct data on the consumption of retirees in various countries, benefits in some of the 100% countries replace nearly all of a retiree's earnings (Gruber and Wise 1997) so c_0 is presumably fairly large.

$$\max_{G_o,\, G_y} \; \theta\, \Pi\, W(G_o) \; + \; (1-\theta)\,(1-\Pi)\, W(G_y)$$

$$\text{s.t.} \quad G_o + G_y \geq G$$

In order to obtain the most accurate description of intergenerational redistribution and to account for the fact that older consumers have greater stocks of assets, we assume that the optimal program involves $G_o/\Pi < G_y/(1\text{-}\Pi)$.

Suppose for the sake of argument that the qualitative features of actual and optimal policies are the same: a 100 percent marginal tax applies to type k persons with productivity less than w_0^k and tax rates strictly less than 100 percent apply to all others of type k ($k = o$, y). The first order conditions with respect to w_0^o and w_0^y, respectively, take the same form (1). They differ only because the old are less willing to work (either because they are richer, because they receive government transfers from the young, or both) and therefore work less at any given after-tax wage. Subtracting the first order condition for the old from the first order condition for the young and, for simplicity, assuming that individual utility functions are separable in consumption and work, we have:

$$f(w_0^y)\left[n^y(w_0^y)\,w_0^y + c_0^y - c^y(w_0^y)\right] - f(w_0^o)\left[n^o(w_0^o)\,w_0^o + c_0^o - c^o(w_0^o)\right] =$$

$$F(w_0^y)[1 - F(w_0^y)]\left[E\left(\frac{1}{u_c^y}\middle| w > w_0^y\right) - \frac{1}{u_c(c_0^y, 0)}\right] - F(w_0^o)[1 - F(w_0^o)]\left[E\left(\frac{1}{u_c^o}\middle| w > w_0^o\right) - \frac{1}{u_c(c_0^o, 0)}\right]$$

The fractions F and 1-F are measured as the nonparticipation rates and participation rates, respectively. Since, in countries with social security programs with 100% tax rates, the young participation rate is roughly 0.9 and the old participation rate roughly 0.1, the product F(1-F) is not too different for young and old. The terms on the RHS in large square brackets are the difference between the average inverse marginal utility of wealth for those working and not working and, since those young who do not work consume quite a bit less than those young who work are likely to be larger for the young. Thus the term on the RHS is positive, or at least not very negative.

If governments were choosing w_0 optimally for young and old, then the equality above holds and the left-hand side is positive, or at least not too negative. But the left-hand side is the revenue gained by inducing the best among the nonworkers to work -- which we saw in the previous section to be much larger for the old. Hence, governments are not using 100 percent tax rates on the young and old in a way that maximizes a utilitarian objective.

V. 100% Taxes as Subsidizing Untaxed Activities

V.A. Subsidizing Fringes

Obviously, a person subject to a 100 percent marginal tax rate has a high willingness to pay for untaxed fringes. Hence, one potential explanation for 100 percent tax rates is that suppliers of untaxed fringes enjoy substantial political power. I am unaware of a study that works out this idea more systematically and do not attempt to do so here. However, there are some related studies of the labor supply of the elderly, which I introduce in the next section.

V.B. Subsidizing Leisure

The Mirrlees model implicitly assumes that nearly everyone works too little from his point of view. However, depending on who the planner is and on the economic environment, many people may work too much from the point of view of the planner. One reason this may occur is that technology is of the kind supposed by Lucas (1988) and Sala-i-Martin (1996) where, through an external affect, a person's labor productivity depends on the average human capital of other workers. In this case, the social marginal product of a person's work can be negative if his human capital is low enough (even while his private marginal product is positive). In order to align private incentives and social marginal products, 100 percent taxes may be optimal.

In a related model, Mulligan and Sala-i-Martin (1998) suppose that an interest group's political power may be enhanced when its members work less. From a group's point of view, 100 percent implicit taxes may be optimal. To see this, consider an environment like Mirrlees', where labor productivity differs across individuals and is not observable to those enforcing policy. The external funds obtained by an interest group depend on the work decisions of its members, so the group's feasibility constraint is:

$$\int_0^{\bar{w}} [\, w n(w) - c(w) \,] \, f(w) \, dw \;\geq\; G \;-\; \phi \int_0^{\bar{w}} [\, 1 - n(w)] \, f(w) \, dw$$

where G is the amount owed to the government by the interest group if each member works full-time. ϕ is a constant and $f(w)$ and $F(w)$ denote the density and cumulative density, respectively, of interest group members with productivity w. The integral on the right-hand-side is the extra revenue the interest group enjoys from external sources as a result of the leisure taken by its members. Notice that we have assumed linearity in n so that the leisure of any two group members are perfect substitutes and that the first unit of work effort hurts to the group the same as the last unit. This matters for the quantitative results; I comment more on this below.

Nonlinear labor income taxes can be designed by a group to encourage members to contribute to the group benefit $\phi[1-n]$. In order to see how nonlinear taxes might be used, an objective for a heterogeneous group must be specified. Perhaps a social welfare function is a natural objective, but my goal here is to distinguish the implications of Mirrlees-type redistribution from the implications of policies primarily designed to discourage work. So imagine that there is a political entrepreneur, an individual (or small group of individuals) who enjoy the "profits" of an interest group, which consist of funds obtained externally from the group $\phi\int[1-n(w)]f(w)dw - G$ net of nonnegative benefit payments $\int b(w)f(w)dw$ to interest group members designed to enhance the external funds.[13,14] Since a member's work effort cannot be observed, the profit maximization by the entrepreneur is a mechanism design problem:

$$\max_{c(\cdot),\, n(\cdot)} \int_0^{\bar{w}} \left(\phi[1-n(w)] + wn(w) - c(w)]\right) f(w)\,dw \quad \text{subject to:}$$

$$u\big(c(w), n(w)\big) \equiv V(w) \geq V^*(w) \quad , \quad V'(w) \leq v'\big(n(w)\big)\frac{n(w)}{w}$$

$$c(\tilde{w}) = c(w) \quad \text{for all } \tilde{w} \text{ such that } n(\tilde{w})\tilde{w} = n(w)w \quad , \text{ all } w$$

[13]The entrepreneur(s) must be a sufficiently small group, because I assume that his (their) concern for benefit payments is only according to its reduction of entrepreneurial profits (which would not be the case if benefit payments were another sufficiently important source of income for the entrepreneur because he is both an entrepreneur and an interest group member).

[14]It is not important that the entrepreneur earn substantial profits. Results would be the same if is profits were taxed at a nearly 100 percent rate and distributed in a lump sum fashion to interest group members.

where $b(w) = c(w) - wn(w)$ and $V^*(w)$ is utility attained by the consumer of productivity w if benefits were not available for him. The condition $V(w) \geq V^*(w)$ says that interest group members must voluntarily accept the labor supply and benefit level chosen for them. The binding incentive compatibility constraint $V(w) \leq v'(n)n/w$ is, other than the direction of the inequality, the same as the binding constraint in the Mirrlees problem. The direction of the inequality is different because the interest group wants to redistribute towards those who work less rather than those who are less productive and pretending to be *more* productive is a way to hide one's work effort from the group (the group is trying to infer work from earnings). For simplicity, I have assumed that each member's utility function is separable in consumption and work ($u(c) - v(n)$).

The first order condition describing the optimal marginal benefit reduction rate T' takes the relatively simple form:

$$T' \equiv 1 - \frac{v'(n)}{w\,u'(c)} = \begin{cases} \dfrac{f(w)\phi + q(w)\frac{\sigma+1}{\sigma}u'(c(w))}{f(w)w + q(w)\frac{\sigma+1}{\sigma}u'(c(w))} & \text{if } n(w) > 0 \text{ and } V(w) > V^*(w) \\ 1 & \text{if } n(w) = 0 \text{ and } V(w) > V^*(w) \\ 0 & \text{if } V(w) = V^*(w) \text{ and } w < \bar{w} \end{cases} \quad \textbf{(3)}$$

where σ is the Frisch labor supply elasticity and $q(w)$ is the marginal willingness of the interest group entrepreneur to pay to relax the incentive compatibility constraint for a person of productivity w. When the participation constraint is not binding, the first line of (3) equates the marginal cost of raising the marginal implicit tax rate on type w (he works less and costs the group $(wT' - \phi)\sigma/(\sigma+1)$) with the marginal benefit (less work by type w makes w's pretending to be more productive less attractive, thereby relaxing the incentive compatibility constraint).

Assume that ϕ is large enough that it is optimal to pay some benefits but small enough that it is optimal to set $n(w) > 0$ for some w on $[0,\bar{w}]$. Clearly, there is no net marginal cost of raising the tax rate when $w < \phi$. If the first order conditions are sufficient then, from the first two lines of (3), the optimal policy must set $n(w) = 0$ for any $w < \phi$. Then the marginal tax rate applying to those with productivity $w < \phi$ is 1 (unless perhaps they would have chosen $n(w) = 0$ without any benefits) because, according to the constraints of the problem, different people with the same earnings (zero in this case) must be given the same benefit.

To describe the optimal policy $w > \phi$, consider the first order condition for the incentive compatibility shadow price $q(w)$:

$$q'(w) = \begin{cases} -\dfrac{f(w)}{u'(c(w))} & \text{if} \quad n(w) > 0 \text{ and } V(w) > V^*(w) \\ \lambda(w) - \dfrac{f(w)}{u'(c(w))} & \text{if} \quad n(w) > 0 \text{ and } V(w) = V^*(w) \end{cases} \qquad \textbf{(4)}$$

It can be shown that if the participation constraint binds for w near $\bar{w}$, then q is rising near $\bar{w}$ and the transversality condition $q(\bar{w}) = 0$ is violated - so the participation constraint cannot bind except exactly at $w = \bar{w}$. Thus some benefits are paid to everyone - even those with the highest productivity - and the participation constraint binds only for the person of productivity $w = \bar{w}$. From (4), the optimal q is falling for those exerting positive effort and, from the first line of (3), the marginal tax rate as a function of w approaches the function ϕ/w from above and intersects that function at exactly $w = \bar{w}$.

From here, the form of the optimal policy can be described. If any benefits are to be paid, they are a function of earnings, starting at b_0 for someone with no earnings and reduced dollar-for-dollar with earnings for those with earnings less than $\bar{e} < \phi$. For those with earnings less than $\bar{e}$, benefits are reduced less than dollar-for-dollar, at a rate which declines with earnings. Figure 2 displays an optimal benefit schedule as a solid dark curve:

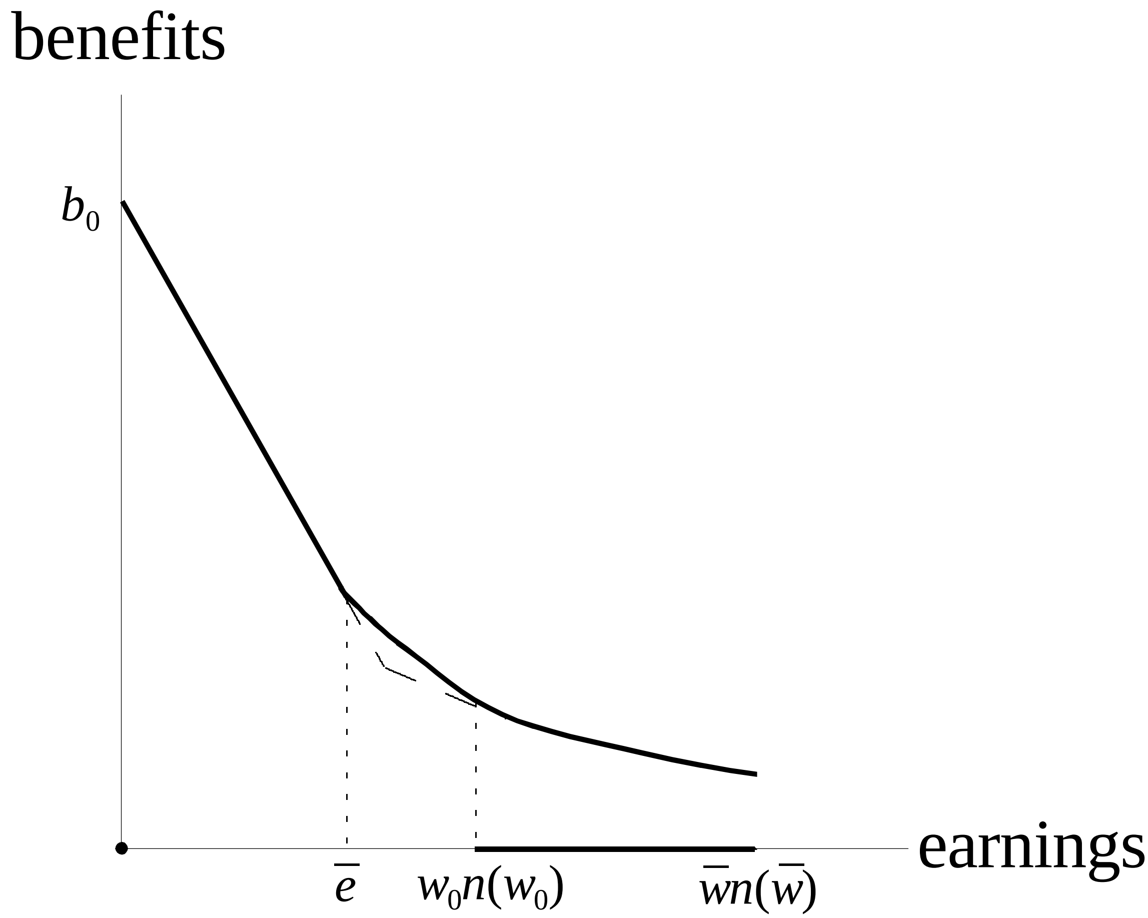

Figure 2 Interest Group Optimal Benefit Schedules

The resulting distribution of earnings is shown by darkening the horizontal axis where there is positive density. Earnings are concentrated at 0 and there are no earnings between $w_0n(w_0)$ and $\bar{w}n(\bar{w})$, where $w_0 > \phi$. Since there is a gap in the optimal earnings distribution, there is not a unique way to link optimal benefits to earnings. A dashed curve shows another optimal benefit schedule where the benefit reduction rate falls from 1 to another constant until it meets the smooth dark curve at $w_0n(w_0)$.

The group revenue loss from raising w_0 is less than in the Mirrlees model and is actually a revenue gain in the neighborhood of the optimal w_0. As in the Mirrlees model, the quantitative effect on incentives of raising w_0 cannot be observed directly, but we have the added problem that ϕ and thereby the magnitude of the revenue gain is unobservable in this model. Hence, it is more difficult to quantitatively describe the range w_0 of the 100% tax than in the Mirrlees model, although this model is more precise in its quantitative predictions for the marginal tax rates used: only 100% are used over some range of earnings. We can also say that some benefits are paid to all, implicit marginal tax rates eventually fall below one, and rates continue to fall with earnings. In these regards, the model of "discouraging taxable income" closely mimics the public pension and medical benefit systems of Spain, Belgium, France, Austria, Germany (historically), the U.K. (historically), the U.S. (before 1960), and other countries. The reduction of pensions in these countries coincides with the 100% portion of the schedule shown in Figure 2. The provision of medical benefits regardless of earnings together with capped payroll taxes (which are paid by the elderly who work)[15] coincides with a benefit (net-of-tax) schedule which falls with earnings at a declining rate. Income tax deductions and credits, and property tax deductions and credits enjoyed by the elderly may also serve this purpose.

Unlike a number of actual policies, optimal policies in our model do not involve exempting any labor income from the implicit tax. Perhaps, since actual exemption levels are quite small, this is just a detail. And perhaps our model could be modified to allow the interest group revenue to be nonlinear in n for very small n so that the optimal policy would tax the first bit of earnings at 0 percent, some additional earnings at 100 percent, and the highest earnings at zero percent. Or a small exemption might be optimal in a world where some earnings might accrue without the worker applying any effort during a tax and benefit accounting period -- perhaps because earnings were owed from employment in a previous period or because some past investments pay what accountants call "earnings", possibilities which may be consistent with some of the early discussions of exemptions by American policymakers (Myers 1993, p. 271).

VI. Conclusions

The key empirical result for showing that 100 percent taxes are not solve in an equity-efficiency problem is that so much revenue could be gained by inducing the best among the nonworkers to work, especially among the old. My first computation makes the case by comparing the revenue gain with a computation of the value of increased incentives. My second computation makes the case by comparing the very different use of 100 percent taxes on young and old, and

[15] Austria, France, and Spain are some of the governments currently reducing pension benefits at 100% rates which also cap payroll taxes (Social Security Administration, 1997).

requires fewer functional form assumptions and fewer estimated parameter values.

I conclude that discouraging work may be a primary objective of policy rather than an undesired byproduct of redistribution to the poor. The primacy of the former objective seems most obvious with the prevalent use of 100% tax rates, but is probably also consistent with high but less-than-confiscatory tax rates, minimum wage laws, maximum hours laws, and the frequent lack of means testing of social security programs (Sala-i-Martin 1996).

VII. References

Atkinson, Anthony B. and Joseph E. Stiglitz. *Lectures on Public Economics*. New York: McGraw-Hill, 1980.

Barro, Robert J. and Chaipat Sahasakul. "Measuring the Average Marginal Tax Rate from the Individual Income Tax." *Journal of Business*. 56, October 1983: 419-52.

Bernheim, B. Douglas. "On the Voluntary and Involuntary Provision of Public Goods." *American Economic Review*. 76(4), September 1986: 789-93.

Besley, Timothy and Stephen Coate. "The Design of Income Maintenance Programmes." *Review of Economic Studies*. 62(2), April 1995: 187-221.

Blanchet, Didier and Louis-Paul Pele. "Social Security and Retirement in France." NBER working paper no. 6214, October 1997.

Boldrin, Michele, Sergi Jimenez-Martni, and Franco Peracchi. "Social Security and Retirement in Spain." NBER Working Paper No. 6136, August 1997.

Cooter, Robert D. "Optimal Tax Schedules and Rates: Mirrlees and Ramsey." *American Economic Review*. 68(5), December 1978: 756-68.

Currie, Janet and Aaron Yelowitz. "Are Public Housing Projects for Kids?" NBER Working Paper #6305, November 1997.

Diamond, Peter A. and James A. Mirrlees. "A Model of Social Insurance with Variable Retirement." *Journal of Public Economics*. 10(3), December 1978: 295-336.

Diamond, Peter A. and James A. Mirrlees. "Payroll-Tax Financed Social Insurance with Variable Retirement." *Scandinavian Journal of Economics*. 88(1), 1986: 25-50.

Fraker, Thomas, Robert Moffitt, and Douglas Wolf. "Effective Tax Rates and Guarantees in the

AFDC Program, 1967-82." *Journal of Human Resources*. 20, Spring 1985: 251-63.

Gruber, Jonathan and David Wise. "Social Security Programs and Retirement Around the World." NBER Working paper #6134, August 1997.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes. "Precautionary Saving and Social Insurance." *Journal of Political Economy*. 103(2), April 1995: 360-99.

Kotlikoff, Laurence J. and David A. Wise. "Employee Retirement and a Firm's Pension Plan." *The Economics of Aging*. Chicago: University of Chicago Press, 1989: 279-330.

Lucas, Robert E. "On the Mechanics of Economic Development." *Journal of Monetary Economics*. 22(1), July 1988: 3-42.

Mirrlees, James A. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*. 38(114), April 1971: 175-208.

Myers, Robert J. *Social Security*. 4th edition. Philadelphia: University of Pennsylvania Press, 1993.

Mulligan, Casey B. and Xavier Sala-i-Martin. "Gerontocracy, Retirement, and Social Security." Working paper, Columbia University, March 1998.

Pestieau, Pierre and Jean-Phillippe Stijns. "Social Security and Retirement in Belgium." NBER working paper #6169, September 1997.

Sala-i-Martin, Xavier X. "A Positive Theory of Social Security." *Journal of Economic Growth*. (1)2, June 1996: 277-304.

United States Congress, House Committee on Ways and Means. *Overview of Entitlement Programs, 1996 Green Book*. Washington, D.C.: U.S. Government Printing Office, 1996.

United States, Department of Commerce, Bureau of the Census. *Historical Statistics of the United States, Colonial Times to 1970.* Bicentennial edition. Washington: U.S. Dept. of Commerce, Bureau of the Census, 1975.

United States Social Security Administration. *Social Security Programs Throughout the World - 1967*. Washington, DC: Government Printing Office, 1967.

United States Social Security Administration. *Social Security Programs Throughout the World - 1997*. Washington, DC: Government Printing Office, 1997.

Yelowitz, Aaron S. "The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions." *Quarterly Journal of Economics*. 110(4), November 1995: 909-39.