Bank of Israel   **Research Department**

The Optimal Asymptotic Income Tax Rate[1]

by

Momi Dahan[*] and Michel Strawczynski[**]

Discussion Paper No. 2004.15
October 2004

[*] Hebrew University of Jerusalem, School of Public Policy, Mount Scopus, Jerusalem 91905, Israel. Email: momidahan@mscc.huji.ac.il.
[**] Bank of Israel, Research Department, Jerusalem, Israel. Email: michels@boi.gov.il.

Any views expressed in the Discussion Paper Series are those of the authors and do not necessarily reflect those of the Bank of Israel.

Abstract

Recent works on optimal income tax found that for unbounded distributions of earnings the optimal tax rate at the top is relatively high (around 60 percent). This finding is puzzling in light of the well-known result for bounded distributions of a zero optimal tax rate at the top. Our paper shows that the more recent papers have used assumptions that favor a high asymptotic tax rate: Pareto instead of Log-normal distribution and linear instead of non-linear utility of consumption. Using these two assumptions along with a logarithmic utility of leisure leads to an optimal rate of 100%, a result that is avoided in recent literature by assuming a constant compensated elasticity of labor. We find that even when using a Pareto distribution of earnings the optimal asymptotic tax rate is about a half compared to recent literature.

1. Introduction

This paper presents an analytical expression for the optimal asymptotic tax rate that allows us to explore the conditions under which the optimal tax rate is high or low at the top of the earnings distribution. The benchmark model for non-linear taxation analysis in Public Economics was introduced by Mirrlees (1971). The main analytical result of this model is that optimal tax rates are in the range between zero and one. Consequently, in order to get a further insight on optimal tax rates the usual convention has been to run simulations.

Until late nineties, most simulations have shown declining tax rates at high income levels, with tax rates at the top varying from 15 (Mirrlees, 1971) to 40 (Kanbur and Tuomala, 1994). [1] These early simulations were based on a lognormal distribution of income and concave utility functions for both leisure and consumption. However, using a linear utility of consumption, Diamond (1998) has shown an example where optimal tax rates go up at high income levels reaching a high asymptotic rate.

The high asymptotic tax rate is in sharp contrast with the familiar result of a zero tax rate at the top of the earnings distribution (Sadka, 1976, Seade, 1977). Following Diamond (1998), that have used an unbounded distribution of earnings, the zero tax rate result was perceived to be both local and exotic.[2] Thus, it seems to be that the zero tax result is limited to bounded distribution.

---

[1] In addition to Mirrlees simulations, declining optimal tax rates were found by Atkinson (1973), Tuomala (1984) and Kanbur and Tuomala (1994).

[2] In fact, this point was previously shown by Tuomala (1984) who emphasized that zero limit of marginal tax rate at the upper end of the distribution "is really very local".

Dahan and Strawczynski (2000) – henceforth DS - used simulations to show that Diamond's increasing pattern is sensitive to the assumption of a linear utility of consumption: using a logarithmic utility of leisure, they show that the concavity of the utility of consumption makes the difference between increasing (as in Diamond) and decreasing (as in Mirrlees) optimal income tax rates. The decreasing optimal tax rates at the top of the distribution casts doubts on Diamond's result regarding the relatively high asymptotic tax rate.

In a recent contribution, Saez (2001) has extended Diamond's result of increasing optimal tax rates at the top and high asymptotic tax rates to a case where the utility of consumption is concave. He found that with a constant elasticity of labor and a logarithmic utility of consumption the optimal tax rates go up at high income levels, and reach an asymptotic tax rate between 50 and 60 percent. This result was obtained using a distribution of wages that approximates the empirical distribution of wages in the U.S., which, according to the author it resembles fairly well a Pareto distribution. Since Saez's model was based on a non-linear utility of consumption, his finding re-opens the question on whether the result of a high asymptotic tax rate should be considered as a benchmark result.

In this paper we explore this question by calculating asymptotic tax rates, i.e., the limit of optimal non-linear marginal tax rates when the wage tends to infinity. It is interesting to note that Mirrlees (1971) has already analyzed the optimal asymptotic income tax rates, although his results have not been emphasized in the literature. In this paper we aim at shedding light on the mechanisms that drive these results, and analyze cases that were not discussed in the literature in recent years.

The paper is organized as follows: in Section 2 we present the model that serves as a basis for the analytical expression for calculating optimal asymptotic tax rates. The basic propositions, which refer to the different cases analyzed in the literature under different assumptions concerning the utility functions of leisure and consumption and the income distribution, is presented in Section 3. Section 4 analyzes the optimum non-linear shape under a lognormal distribution of earnings, and Section 5 concludes.

2. **Non-linear optimal income tax**

2.1 The Model

Assume the following utility function:

(1) $u = U(C) + V(1 - L)$

where C is consumption, 1-L is leisure and U and V are respectively the utility of consumption and the utility of leisure. The budget constraint at the individual level is:

(2) $C(w) = wL(w) - T[wL(w)]$

where T symbolizes the income tax, which is defined on total income since the wage w and the supplied amount of labor L(w) are not observed by the government. The first order condition at the individual level is:

(3) $\dfrac{V_{(1-L)}}{U_C} = (1 - \tau)w$, $\tau \equiv T`$

where $V_{(1-L)}$ and U_C are the first derivatives of V and U, respectively. Assume also the existence of the self-selection constraint, which takes the form that utility must increase with w^3:

---

[3] This assumption assures agent monotonicity; i.e., before taxes, income and consumption rise with skill (see Myles, 1995, p.140, and Stiglitz, 1987).

(4)

$$\frac{du}{dw} = U_C[L(w) + w\frac{dL}{dw}] - V_{(1-L)}\frac{dL}{dw} = \frac{V_{(1-L)}}{w}L + V_{(1-L)}\frac{dL}{dw} - V_{(1-L)}\frac{dL}{dw} = V_{(1-L)}\frac{L}{w} > 0$$

We introduce now the government, which maximizes the social welfare function:

(5) $$SW = \int_{w_L}^{w_H} G\{U[C(w)] + V[1 - L(w)]\}f(w)dw$$

where w_L and w_H are the bottom and top of the positive and continuous distribution of skills. The budget constraint of the economy is:

(6) $$\int_{w_L}^{w_H} C(w)f(w)dw = \int_{w_L}^{w_H} wL(w)f(w)dw$$

i.e., government intervention is purely redistributive. We are now ready to write the hamiltonian (H), which is composed by the social welfare utility function, the budget constraint of the economy and the differential equation for the state variable u (given by the self-selection constraint):

(7) $$H = \{G(u) - \gamma[C(w) - wL(w)]\}\frac{dF}{dw} + \lambda(w)V_L\frac{L}{w}$$

The control variable of this problem is L. γ is the multiplier of the budget constraint and λ is the multiplier of the self-selection constraint. The F.O.C. for a maximum are:

(8) $$H_L = \frac{\partial G}{\partial u}\frac{\partial u}{\partial L}\Big|_{\bar{u}} + \gamma(w - \frac{dC}{dL}\Big|_{\bar{u}})\frac{dF}{dw} + \frac{\lambda(w)}{w}(V_L - LV_{LL}) = 0$$

$$= \gamma(w + \frac{V_L}{U_C})\frac{dF}{dw} + \frac{\lambda(w)V_L}{w}(1 + \delta) = 0, \qquad where \ \delta = -\frac{LV_{LL}}{V_L}$$

(9) $H_u = (\dfrac{dG(u)}{du} - \gamma \dfrac{dC}{du}) \dfrac{dF}{dw} = \dfrac{d\lambda}{dw}$

The transversality conditions are:

(10) $\lambda(w_H) = \lambda(w_L) = 0$

By integration of both sides of (9), and using the transversality conditions (10) we obtain:

(11) $-\displaystyle\int_w^{w_H} (\dfrac{\gamma}{U_C} - \dfrac{dG}{du}) \dfrac{dF}{dw} dw = \int_w^{w_H} \dfrac{d\lambda(w)}{dw} dw = \lambda(w_H) - \lambda(w) = -\lambda(w)$

Using this expression and the first order conditions of both government (equation 8) and individuals (equation 3), we obtain:

$$\gamma(w + \dfrac{V_L}{U_C}) f(w) + \dfrac{\displaystyle\int_w^{w_H} (\dfrac{\gamma}{U_C} - \dfrac{dG}{du}) \dfrac{dF}{dw} dw V_L (1+\delta)}{w} = 0$$

$$(-\dfrac{V_L}{U_C(1-\tau)} + \dfrac{V_L}{U_C}) f(w) = -\dfrac{\displaystyle\int_w^{w_H} (\dfrac{\gamma}{U_C} - \dfrac{dG}{du}) \dfrac{dF}{dw} dw V_L (1+\delta)}{\gamma w}$$

$$-\dfrac{V_L}{U_C} (\dfrac{1}{(1-\tau)} - 1) = -\dfrac{\displaystyle\int_w^{w_H} (\dfrac{\gamma}{U_C} - \dfrac{dG}{du}) \dfrac{dF}{dw} dw V_L (1+\delta)}{\gamma w f(w)}$$

(12) $\dfrac{\tau}{1-\tau} = \left[\dfrac{\varepsilon}{w}\right][U_C] \left[\dfrac{\displaystyle\int_w^{w_H} [\dfrac{\gamma}{U_C} - G_u] f(w) dw}{\gamma(1-F(w))}\right] \left[\dfrac{(1-F(w))}{f(w)}\right],\ \varepsilon = 1 - \dfrac{V_{LL}L}{V_L},\ f \equiv \dfrac{dF}{dw}$

Equation 12 is the analytical expression to be used for the calculation of optimal asymptotic tax rates. The first term in the RHS is the "efficiency effect". The higher

the compensated elasticity of labor the lower the optimal marginal income tax rate. The second term is the standard "income effect", which is dependent on the marginal utility of consumption. For high income levels, the marginal utility of consumption is low, and thus the incentive to work harder as a result of net income reduction disappears. The third effect is the "inequality aversion effect", which depends both in the concavity of the utility of consumption and the social welfare function. For concave functions, this effect increases with income. The last effect is the "distribution effect": the higher the proportion of individuals above the wage level relative to the proportion of individuals at this level, the less distortionary is the marginal tax rate, since for these individuals the marginal tax rate acts like a lump-sum tax. Thus, a higher ratio of (1-F) over f implies a higher optimal tax rate.

2.2 The optimal asymptotic labor supply

In this sub-section we characterize labor supply when w tends to infinity and marginal utility of consumption equals $c^{-\mu}$.

Log Utility of leisure

F.O.C.:

$$(13) \quad \frac{L^{\mu}}{1-L} = w^{1-\mu}$$

Thus, we have three cases: 1) $\mu<1$ (which includes linear utility of consumption): in this case labor supply goes to one, and consequently ε goes to infinity. 2) $\mu=1$ (log utility of consumption): in this case labor tends to a finite number, i.e., there is an internal solution. 3) $\mu>1$ (which includes utility of the type -1/c): in this case labor goes to zero, and ε goes to one.

Constant elasticity of substitution

(14) $bL^{\mu+b-1} = w^{1-\mu}$

Again, there are three cases: 1) $\mu<1$ (linear utility of consumption): in this case labor

supply goes to infinity, but ε equals to b.[4] 2) $\mu=1$ (log utility of consumption): in this

case labor tends to a finite number, i.e., there is an internal solution. 3) $\mu>1$ (utility of

the type -1/c): in this case labor supply goes to zero, and ε equals to b.

3. Optimal asymptotic tax rates

This section explores the optimal asymptotic income tax rate utilizing Equation 12

and the results on labor supply from the previous section.

Proposition 1:

With a logarithmic utility of leisure, the asymptotic tax rate converges to one both for

a Pareto and Lognormal distribution of earnings, when the utility of consumption is

linear.

Proof

In this case, equation 12 can be written as follows:

(15) $$\lim_{w\to\infty}\frac{\tau}{(1-\tau)^2} = \lim_{w\to\infty} U_c^2 * \lim_{w\to\infty}\frac{\int_w^\infty\left(\frac{1}{U_c} - \frac{G`(u)}{\gamma}\right)f(s)ds}{1-F(w)}\lim_{w\to\infty}\frac{1-F}{f}$$

Using *L`Hopital*`s rule the second term in the right hand side is equal to:

---

[4] Henceforth we will allow labor supply to tend to infinity.

$$(16) \quad \lim_{w \to \infty} B(w) \equiv \lim_{w \to \infty} \frac{\int_w^\infty \left(\dfrac{1}{U_c} - \dfrac{G`(u)}{\gamma} \right) f(s) ds}{1 - F(w)} = \frac{1}{U_c} - \frac{G`(u)}{\gamma} = 1$$

where B(w) reflects the inequality aversion effect. The benefit of a dollar revenue worth's a dollar for a linear utility of consumption regardless of the wage level. The cost of a dollar in term of social welfare is G'(u) and it goes to zero with a standard social utility function that has some degree of inequality aversion (or to a finite number) as wage goes to infinity. Thus, at the limit of the wage distribution B(w) equals 1.

Therefore, the asymptotic tax rate converges to:

$$(17) \quad \lim_{w \to \infty} \frac{\tau}{(1-\tau)^2} = [1][1] \lim_{w \to \infty} \left[\frac{1-F}{f} \right] = \infty$$

This is true for both Pareto and Lognormal distributions. With a Pareto distribution of the form $f = \alpha k^\alpha / w^{1+\alpha}$ (where α and k are some constants), the distribution effect goes to infinity: $\dfrac{1-F}{f} = \dfrac{w}{\alpha}$ and as a result $\tau/(1-\tau)^2$ goes to infinity and the tax rate goes to one.

In the Lognormal case, the tax rate goes to one because $\dfrac{1-F}{f} = \dfrac{w}{\log w - \varpi} s$ [where s denotes the standard deviation of log(w) and ϖ denotes the average of log(w)] and it also goes to infinity as wage goes to infinity.

One way to see the intuition behind this unique result is the following. We saw that the standard income effect and the inequality aversion effect are both constant and in the case of Pareto distribution the distribution effect divided by w is also constant. Thus, the only factor that plays a role is ε, which equals 1 over the compensated labor supply elasticity plus one. Since in this case the labor supply tends to one, it is easy to see that the compensated elasticity tend to zero and the expression of ε goes to infinity as the wage goes to infinity. In other words, assuming a linear utility of consumption drives individuals to supply labor inelastically and therefore the optimal income tax rate should be 100%.

An optimal tax rate of 100% on the most able individual (loosely speaking given that we work with an unbounded distribution) is far from trivial at first glace. In particular, this result is at the opposite polar of the well known result of Sadka (1976) and Seade (1977). Our technique of driving the optimal asymptotic income tax rate helps to clarify the forces behind this unique result. Although the result of $\tau=1$ appeared in Mirrlees (1971), to the best of our knowledge it was never emphasized explicitly in the optimal income tax literature.

However, replacing the logarithmic utility of leisure with a constant compensated elasticity of labor supply and using log-normal distribution instead of Pareto distribution would yield a result that is consistent with the well known result of a zero income tax rate at the top of the distribution.

Proposition 2:

The asymptotic tax rate converges to a finite number for a Pareto distribution of earnings and constant (compensated) elasticity of labor, both with linear and non linear utility of consumption.

We assume that the utility of leisure is of the form $V(L)=1-L^b$ where b is some constant, and $1/(b-1)$ is the compensated elasticity of labor. The case of a constant compensated elasticity was not covered in Mirrlees (1971) due to the assumption that labor supply could not exceed one. However, we can use the first order condition once we relax this assumption by letting the labor supply to go to infinity as in Diamond (1998) and Saez (2001). If an upper bound of 1 was to be enforced exogenously the first order condition will not hold as equality and we would not be able to derive the optimal tax rate. In addition we assume a Pareto distribution of the form: $f=\alpha k^{\alpha}/w^{1+\alpha}$ as before.

Proof

Plugging these two assumptions into equation 12 we get:

$$(18) \quad \lim_{w\to\infty} \frac{\tau}{1-\tau} = \frac{b}{\alpha} * \lim_{w\to\infty} U_c * \lim_{w\to\infty} B(w)$$

Therefore, the asymptotic tax rate converges to:

$$(19) \quad \lim_{w\to\infty} \frac{\tau}{1-\tau} = \lim_{w\to\infty} \frac{b}{\alpha}\left[U_c\left(\frac{1}{U_c} - \frac{G`(u)}{\gamma} \right) \right] = \frac{b}{\alpha}$$

This proof shows the forces behind the optimal income tax. The interaction between the standard income effect and inequality aversion effect is of particular type. We can see that the standard income effect drives the optimal tax rate to zero. Taxing the very rich will not induce them to work more because the income effect at those levels of income fades away already. At the same time, the income effect works in the opposite direction through the inequality aversion effect. Thus, taxing the income of the very rich produces an extremely large additional social welfare.

Zero marginal utility means that taking money away from the very rich does not alter their welfare but the government has more resources to improve the welfare of others. One can see that the standard income effect is canceled out exactly by the inequality aversion factor. Thus, the product of the inequality effect and income effect equals one. Hence, the asymptotic tax rate depends on the efficiency effect multiplied by the distribution effect which converges to a finite number (b/α) both with linear and non-linear utility of consumption. Note that in this particular case (i.e., Pareto distribution and constant labor elasticity) this result holds for any form of non-linear utility of consumption.

The asymptotic income tax rate is the same both for Utilitarian and Rawlsian social welfare function. In the case of Utilitarian social welfare the asymptotic rate is the same either G`(u) equals to zero or any positive number. However, in the case of linear utility of consumption the asymptotic tax rate depends on whether G`(u) goes to zero or to some positive number. So, the asymptotic tax rate is lower if G`(u) goes to

a positive number. For the values assumed by Saez (a compensated elasticity of 0.25 and 0.5), the optimal asymptotic tax rate varies from 0.714 to 0.6, respectively.[5]

Proposition 3:

For a Pareto distribution of earnings and log utility of leisure, the asymptotic tax rate converges to a finite number both in the case of log-utility of consumption and for utility of consumption of the type -1/c.

Proof

In the case of log-utility of leisure the optimal tax rate formula of equation 12 is the following:

$$(20) \quad \lim_{w \to \infty} \frac{\tau}{1-\tau} = \lim_{w \to \infty}\left[\frac{\varepsilon}{w}\right]\lim_{w \to \infty}[U_c]\lim_{w \to \infty}[B(w)]\lim_{w \to \infty}\left[\frac{w}{\alpha}\right],$$

The efficiency effects collapses to U_c because $\varepsilon = -V_L$ for log-utility of leisure and we know from the first order condition that $V_L = U_c(1-\tau)w$. Thus, the efficiency effect component goes down as wage increases in the case of log-utility of leisure. Unlike the previous case the income effect works through three channels: efficiency effect, standard income effect and inequality aversion effect. A non-linear utility of consumption implies that the efficiency effect is not constant as in the linear case. The compensated elasticity of labor is affected by the marginal utility of consumption, and the efficiency effect component (ε/w) decreases as wage goes up where $\varepsilon^* = \varepsilon^{-1} - 1$ and ε^* is the compensated elasticity.

Again, the product of the standard income effect and inequality aversion effect equals to one regardless of the exact type of utility of consumption. Hence, the asymptotic

---

[5] When the compensated elasticity is 0.25 and 0.5, Saez's parameter k equals, respectively, 2 and 4. Our parameter b equals (1+k).

tax rate is determined by the particular form of the distribution effect and efficiency effect:

$$(21) \quad \lim_{w \to \infty} \frac{\tau}{1-\tau} = \lim_{w \to \infty} \frac{\varepsilon}{\alpha}$$

In the case of log-utility of consumption the value of ε equals 2 because the optimal labor supply tends to 0.5 as wage goes to infinity, and the right hand side of equation 21 goes to 2 over α. Assuming a Pareto distribution creates a force that drives the tax rate up. It turns out that the distribution effect is completely neutralized by the efficiency effect. Using simple algebra to compute the optimal asymptotic income tax rate one can get that the asymptotic tax rate equals to $2/(2+\alpha)$ For example, the optimal asymptotic tax rate equals to 50% if $\alpha=2$.

In the case of -1/c utility of consumption, labor supply goes to zero and consequently ε goes to one. Once again, we know that the product of the standard income effect and inequality aversion effect equals to one regardless of the exact type of utility of consumption. So the RHS of equation 21 is equal to 1 over α. When α equals 2, the optimal asymptotic tax rate is 0.33. This figure is about a half compared to the asymptotic rate found in recent literature.

In the next propositions we explore the optimal asymptotic tax rate using a lognormal distribution. There is a growing consensus that the Pareto distribution fits reasonably well the empirical earnings distribution at high income levels (Poterba and Feenberg, 1993). However, the whole distribution is best characterized by lognormal distribution (Aitchison and Brown, 1957). Moreover, almost all simulations (until the nineties)

were based on a lognormal distribution.[6] Deriving the optimal asymptotic tax rate for a lognormal distribution helps to link the more recent literature with previous results. In Section 4 we will characterize the shape of the optimal income tax structure for the case of a lognormal distribution.

Proposition 4:

With a lognormal distribution of earnings and a non-linear utility of consumption the optimal asymptotic tax rate converges to zero both for a constant (compensated) elasticity of labor and a log utility of leisure.

Proof

In this case equation 12 is as follows:

$$(22) \quad \lim_{w \to \infty} \frac{\tau}{1-\tau} = \lim_{w \to \infty} \frac{\varepsilon}{w} * \lim_{w \to \infty} U_c * \lim_{w \to \infty} B(w) * \lim_{w \to \infty} \frac{ws}{\log w - \varpi} ,$$

We know from the previous proofs that the product of the standard income effect and inequality aversion effect equals one both for linear and non-linear utility of consumption. Thus, we can rewrite the last equation:

$$(23) \quad \lim_{w \to \infty} \frac{\tau}{1-\tau} = \lim_{w \to \infty} \frac{\varepsilon s}{\log w - \varpi}$$

It is well known for normal distribution that $\phi(w)$ which denotes 1 over the hazard rate of normal distribution goes to zero when w goes to infinity. This emphasizes once more the importance of the assumption on the distribution of earnings.

---

[6] Mirrlees (1971), Tuomala (1984), Kanbur and Tuomala (1994).

In particular, with a constant compensated elasticity, the term (1-F)/fw dictates the optimal asymptotic tax rate. This term is constant for a Pareto distribution and goes to zero as wage goes to infinity using a lognormal distribution. Thus, a lognormal distribution pulls the optimal tax rate down at high levels of income compared to Pareto distribution. Note that when using a constant compensated elasticity of labor, this result holds regardless of the type of utility of consumption that is employed.

The optimal asymptotic tax rate is zero with a log utility of leisure and a non-linear utility of consumption. Again, ε tends to a finite number (the exact finite number depends on the type of non-linear utility of consumption we use) and the RHS of 23 goes to zero as wage goes to zero. This case is particularly important because it replicates Mirrlees (1971) baseline simulation where he has used a log-utility of leisure, log-utility of consumption, lognormal distribution of skills and Utilitarian social welfare function. In his main simulation Mirrlees have got a relatively low tax rate (15%) at very high levels of income (at the 99 percentile). Since then, many attempts were made to "correct" that low rate so as to be closer to actual tax rate at the top.

Our analytical expression shows that Mirrlees simulation result is in fact a bad approximation for the optimal asymptotic tax rate. More important, deriving the optimal asymptotic income tax rate analytically uncovers the forces that lie behind it. The efficiency effect and distribution effect are the two forces that dictate the optimal asymptotic tax rate since, as before, the product of the standard income effect and

inequality aversion effect equals one. In this case, the efficiency effect drives the optimal asymptotic tax rate to zero as long as the utility of consumption is concave.

4. Characterization of the optimal shape with a lognormal distribution

In this section we aim at characterizing the optimal shape of non-linear taxation when the distribution of earnings is lognormal. The lognormal distribution is the benchmark assumption for the whole distribution.

For simplicity, the characterization is performed under the following assumptions: a log-linear (in leisure and consumption respectively) separable utility function, a utilitarian social planner (G'=1) and a lognormal distribution of skills. We characterize the optimal shape by using key values of the distribution - the mode, the median and the mean.

A. The mode ($e^{\mu-\sigma^2}$)

Marginal taxes decline at the mode.

To show this claim we write the distribution effect D(w) and its first derivative D'(w):

$$D(w) = \frac{1 - F(w)}{f(w)}$$

$$D'(w) = \frac{-F'(w)f(w) - f'(w)[1 - F(w)]}{[f(w)]^2} = -1 - \frac{f'(w)[1 - F(w)]}{[f(w)]^2}$$

Since by definition f'(w)=0 at the mode, we conclude that D'(w) at the mode is -1, i.e., marginal tax rates decline.

B. The Median (e^{μ})

If σ≈0.8 marginal taxes decline until the median and since then they rise (i.e., the minimum marginal tax is at the median). If σ>(<)0.8 then the minimum point is at the right (left) of the median.

To show this claim we use the formula for the minimum point as given in proposition 1. This minimum is given where h(x)= σ + x, where h(x) is the standard normal hazard function. At the median x=0, so we look for the value of σ where h(0)= σ; i.e., 0.7978 (see Lancaster, 1990, p. 48) . In order to see that a larger (lower) σ implies that the minimum point is to the right (left) of the median all we need is to characterize the minimum point for values of x lower (higher) than zero, by using the standard normal hazard table.

C. The Mean ($e^{\mu+0.5\sigma^2}$)

If σ≈2 marginal taxes decline until the mean and since then they rise (i.e., the minimum marginal tax is at the mean). If σ>(<)2 then the minimum point is at the right (left) of the mean.

This claim can be shown by the same method, taking into account that in this case x=0.5 σ.

5. Summary and Conclusions

Our paper provides a simple analytical expression for the optimal asymptotic tax rate that allows us to analyze the interactions between efficiency, income distribution and inequality aversion effects under the different cases analyzed in the literature following the seminal work by Mirrlees (1971). In particular, we explore the conditions under which the optimal tax rate is high or low at the top of the distribution.

We found that in general the more recent literature on the optimal income tax at high levels of wage is based on assumptions that drive up the optimal tax rate in comparison to previous literature. First, the more recent works have used Pareto distribution, instead of log-normal, and it drives up the optimal income tax rate. Second, the more recent works have used a linear (instead of non-linear) utility of consumption and it also pulls up the asymptotic tax rate. Our paper shows that these two major changes would have been translated into an optimal tax rate of 100%. However, the more recent works have introduced a third change: a constant compensated elasticity instead of log-utility of leisure. That assumption ensures that the optimal tax rate is less than 100%.

It is shown that with a Pareto distribution of earnings and a linear utility of consumption, the optimal asymptotic tax rate converges to a high tax rate. Surprisingly, the optimal income tax rate is in fact 100% for the most able individuals when the utility of leisure is logarithmic. The optimal tax rate is lower than 100% with a constant compensated labor elasticity but it is still relatively high- 60 percent.

However, when the utility of consumption is of the type -1/c, the optimal asymptotic tax rate tends to a rather low finite tax rate – around 30 percent.

In general, a zero optimal asymptotic tax rate was obtained when using a lognormal distribution of earnings. With constant (compensated) elasticity of labor the optimal asymptotic tax rate is zero both for a linear and non-linear utility of consumption. With logarithmic utility of leisure, the optimal asymptotic tax rate converges to zero if the utility of consumption is non-linear.

Bibliography

Atkinson, Anthony, "How progressive should income tax be?", in: M. Parkin and A.R. Nobay, Eds., *Essays in Modern Economics*, Logman, London, 1973.

Aitchison, J. and J.A.C. Brown. *The Lognormal Distribution*. Cambridge University Press, London, 1957.

Dahan Momi and Michel Strawczynski. "Optimal income taxartion: an example with a U-shaped pattern of optimal marginal tax rates: comment". *American Economic Review*, June 2000, 90 (3), 681-686.

Diamond, A. Peter. "Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review*, March 1998, 88 (1), pp. 83-95.

Feenberg, Daniel R. and Poterba, James M. "Income Inequality and the Income of the Very High-Income Taxpayers: Evidence from Tax Returns." in James Poterba, ed., *Tax Policy and the Economy*. Cambridge, MA: MIT Press, 1993, pp. 145-77.

Kanbur, Ravi and Matti Tuomala. "Inherent inequality and the optimal graduation of marginal tax rates." *Scandinavian Journal of Economics*, June 1994, 96 (2), pp. 275-282.

Lancaster, Tony. "The econometric analysis of transition data." *Econometric Society Monographs,* 1990, no 17, Cambridge University Press.

Mirrlees, James. "An exploration in the theory of optimal income taxation." *Review of Economic Studies*, April 1971, 38 (2), pp. 175-208.

Myles D. Gareth, *Public Economics*, Cambridge Unoversity Press, 1995

Sadka, Efraim. "On income distribution, incentive effects and optimal income taxation." *Review of Economic Studies*, June 1976, 43 (2), pp. 261-268.

Saez, Emmanuel. "Using elasticities to derive optimal income tax rates", *Review of Economic Studies*, 2001, no 68, pp. 205-229.

Seade, Jesus. "On the shape of optimal tax schedules", *Journal of Public Economics*, 1977, 7, pp. 203-236.

Stiglitz, Joseph. "Pareto efficient and optimal taxation and the new new welfare economics", *Handbook of Public Economics*, edited by Auerbach and Feldstein, Elsevier (North Holland), 1987.

Tuomala, Matti. "On the optimal income taxation - some further numerical results." *Journal of Public Economics,* April 1984, 23 (3), pp. 351-366.

Table 1 – The Optimal Asymptotic Tax Rate

	Paper	Marginal tax at high income levels (rate)			
		Schedule	0.99th	asymptotic	Our paper
Linear Utility of Consumption					
Constant Labor Elasticity					
Pareto	Diamond, Saez	Rising		0.51[a]	0.6[b]
Lognormal	-				0
Log utility of leisure					
Pareto	DS, Mirrlees	Rising	-[c]	1[d]	1
Lognormal	DS, Mirrlees	Rising	-[e]	1[d]	1
Log Utility of Consumption					
Constant Labor Elasticity					
Pareto	Saez	Rising	-	0.69[f]	0.6[b]
Lognormal	-				0
Log utility of leisure					
Pareto	DS, Mirrlees	Declining	-[g]	Finite[d]	0.5[h]
Lognormal	DS, Mirrlees	Declining	0.15[i]	0[d]	0
U = -1/C					
Constant Labor Elasticity					
Pareto	-				0.6[b]
Lognormal	-				0
Log utility of leisure					
Pareto	Mirrlees			Finite[d]	0.33[h]
Lognormal	Mirrlees			0[d]	0

Remarks:

a. Based on Saez (2001), Table 2. Assumptions: 0.5 constant compensated elasticity of labor and US calibrated income distribution.

b. Assuming that the elasticity of labor is 0.5 and that the Pareto Distribution parameter is equal to 2 .

c. Not reported by DS; However, in Figure 2a optimal tax rates rise steeply towards one.

d. As reported by Mirrlees (1971).

e. Not reported by DS; However, in Figure 2b optimal tax rates rise slowly.

f. Assuming that the elasticity of labor is 0.5 and using a calibrated distribution of income in the US.

g. Not reported by DS; However, in Figure 2a optimal tax rates decline slowly.

h. Assuming that the Pareto Distribution parameter is equal to 2.

i. Based on Mirrless (1971), Table IV. Assumptions: a 0.39 standard deviation of log income and a Utilitarian Social Planner.