

# The Optimal Asymptotic Income Tax Rate<sup>1</sup>

By Momi Dahan\* and Michel Strawczynski\*\*

Revised, June 2005

---

<sup>1</sup> We thank Louis Kaplow for his questions and comments that helped instigate this project. We also wish to thank Avi Shimhon and Emmanuel Saez and the seminar participants at Tel-Aviv University and Hebrew University of Jerusalem for their helpful comments.

\*Hebrew University of Jerusalem, School of Public Policy, Mount Scopus, Jerusalem 91905, Israel. Email: [momidahan@mscc.huji.ac.il](mailto:momidahan@mscc.huji.ac.il).

\*\*Bank of Israel, Research Department, Jerusalem, Israel  
Email: [michels@boi.gov.il](mailto:michels@boi.gov.il).

# The Optimal Asymptotic Income Tax Rate

## Abstract

*This paper shows that a policy maker needs only two types of information to set the optimal income tax rate at the top: the compensated elasticity of labor supply and the shape of skills distribution. While results in the literature in recent years emphasize that optimal asymptotic tax rates are high - around 60 percent- we find, by using plausible estimates for these two variables, that the optimal top marginal tax rate should be between 33% and 60%, which is in line with the existing rates in the real world.*

Keywords: Optimal Income Tax

JEL Classifications: H21

## 1. Introduction

This paper presents a simple analytical expression for the optimal asymptotic tax rate. We show that the factors that are important for the optimal asymptotic tax rate are only a subgroup of those that influence the shape of income tax rates. It turns out that focusing on the optimal asymptotic tax rate generates a very simple analytical kit for policy-makers.

The user friendly formula for the asymptotic tax rate presented here can be easily employed to calculate the optimal income tax rate at high income levels for various functional forms that are consistent with empirical evidence. It also helps to explore the reasons for the switch from a relatively low asymptotic tax rate in the old literature to a high tax rate in the more recent literature on income taxation.

Until late nineties, most simulations have shown relatively low tax rates at high income levels, varying from 15 (Mirrlees, 1971) to 40 (Kanbur and Tuomala, 1994).<sup>2</sup> These early simulations were based on a lognormal distribution of income and concave utility functions for both leisure and consumption.

One of the main innovations of Diamond (1998) was to show analytically that the optimal income tax rates are determined by three factors: efficiency effect, inequality effect and distribution effect. He has shown, using a linear utility of consumption, an example where optimal tax rates go up at high income levels reaching a high asymptotic rate.

---

<sup>2</sup> In addition to Mirrlees simulations, declining optimal tax rates at the top were found by Atkinson (1973), Tuomala (1984) and Kanbur and Tuomala (1994). Mirrlees (1971) calculated also optimal asymptotic tax rates for different cases, which obviously do not include the more recent cases analyzed in the new literature.

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 skills, the zero tax rate result was perceived to be both local and exotic.<sup>3</sup> Thus, it seems to be that the zero tax result is limited to bounded distribution.

Dahan and Strawczynski (2000) have shown that *income effect* is one additional factor that plays a role in shaping the optimal income tax rates. Using simulations they 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.

The last turn so far has been made by Saez (2001), who finds a high asymptotic tax rate even after introducing all four factors, including the income effect. 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 51 and 69 percent.<sup>4</sup>

This finding opens two separate questions. First, should the result of a high asymptotic tax rate be considered as a benchmark result? Second, what are the factors

---

<sup>3</sup> 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".

<sup>4</sup> See Saez (2001), Table 2 in Section 5.

that are responsible for the second switch from a low to a relatively high asymptotic tax rate (Saez, 2001)?

In order to isolate the factors that influence the optimal income tax at the top we focus on the asymptotic rate. Diamond (1998), Dahan and Strawczynski (2000) and Saez (2001) did not distinguish explicitly between the factors that are important to the optimal shape of income tax rates from those that are important to the optimal asymptotic income tax rate.

The paper is organized as follows: in Section 2 we present the analytical framework and a simplification of the expression for calculating optimal asymptotic tax rates. In Section 3 we use this expression for calculating optimal asymptotic tax rates under different assumptions concerning the utility functions of leisure and consumption. In this section we use empirically plausible estimates of the relevant variables to calculate the optimal marginal tax rates at the top. While the main results presented in Section 3 are based on a Pareto distribution of earnings for high income levels, in Section 4 we extend the results to a lognormal distribution, which is the distribution used in the old income taxation literature. Section 5 concludes.

## **2. The Analytical Framework**

### **2.1 The Model**

Throughout this paper we follow the previous literature on optimal income tax and assume a particular form of utility that is both additive and separable in leisure and consumption :

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

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:

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

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:

$$(3) \quad \frac{V_{(1-L)}}{U_C} = (1 - \tau)w, \quad \tau \equiv \frac{\partial T[wL(w)]}{\partial [wL(w)]}$$

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  $w^5$ :

(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:

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

where  $w_L$  is the bottom of the positive and continuous distribution of skills which is denoted by  $F(w)$  and its respective density function is denoted by  $f(w)$ . The budget constraint of the economy is:

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

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

---

<sup>5</sup> This assumption assures agent monotonicity; i.e., before taxes, income and consumption rise with skill (see Myles, 1995, p.140, and Stiglitz, 1987).

constraint of the economy and the differential equation for the state variable  $u$  (given by the self-selection constraint):

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

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

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

$$= \gamma(w + \frac{V_L}{U_C}) \frac{dF}{dw} + \frac{\lambda(w)V_L}{w} \varepsilon = 0, \quad \text{where } \varepsilon = 1 + \frac{LV_{LL}}{V_L}$$

Note, that  $\varepsilon=1+1/\eta$  where  $\eta$  is the compensated elasticity of labor supply.

$$(9) \quad H_u = (g - \gamma \frac{dC}{du}) \frac{dF}{dw} = \frac{d\lambda}{dw}, \quad \text{where } g \equiv \frac{dG(u)}{du}$$

The transversality conditions are:

$$(10) \quad \lambda(\infty) = \lambda(w_L) = 0$$

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

$$(11) \quad - \int_w^\infty (\frac{\gamma}{U_C} - g) \frac{dF}{dx} dx = \int_w^\infty \frac{d\lambda(x)}{dx} dx = \lambda(\infty) - \lambda(w) = -\lambda(w)$$

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

$$\gamma(w + \frac{V_L}{U_C})f(w) + \frac{\int_w^\infty (\frac{\gamma}{U_C} - g) \frac{dF}{dx} dx V_L \varepsilon}{w} = 0$$

$$\left(-\frac{V_L}{U_c(1-\tau)} + \frac{V_L}{U_c}\right)f(w) = -\frac{\int_w^\infty \left(\frac{\gamma}{U_c} - g\right) \frac{dF}{dx} dx V_L \varepsilon}{\gamma w}$$

$$-\frac{V_L}{U_c} \left(\frac{1}{1-\tau} - 1\right) = -\frac{\int_w^\infty \left(\frac{\gamma}{U_c} - g\right) \frac{dF}{dx} dx V_L \varepsilon}{\gamma w f(w)}$$

$$(12) \quad \frac{\tau}{1-\tau} = \left[\frac{\varepsilon}{w}\right] [U_c] \left[ \frac{\int_w^\infty \left[\frac{\gamma}{U_c} - g\right] f(x) dx}{\gamma(1-F(w))} \right] \left[ \frac{(1-F(w))}{f(w)} \right] \quad , f(w) \equiv \frac{dF}{dw}$$

Where  $U_c$  and  $g$  are evaluated at a particular  $w$ . 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 since  $\eta$ , the compensated elasticity of labor supply, equals  $(\varepsilon - 1)^{-1}$ .

The second term is the standard "income effect", which is dependent on the marginal utility of consumption. As explained in any basic textbook in economics, raising income tax rate works to reduce net income and as a result individuals work more. 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 determining factors of top optimal marginal tax rates

In this sub-section we ask whether all the effects that appear in equation (12) are necessary to calculate optimal tax rates at the top. We show that the analytical kit of the policy maker can be reduced to two factors: the efficiency and distribution effects. In other words, to calculate optimal tax rates at the top policy makers need to know the compensated elasticity of labor supply and the shape of skills distribution.

### **Lemma**

*The optimal asymptotic income tax rate is determined by the efficiency and distribution effects as long as the marginal utility of consumption or marginal social utility tends to zero as the wage goes to infinity.*

As wage tends to infinity Equation (12) becomes:

$$(13) \quad \lim_{w \rightarrow \infty} \frac{\tau}{1 - \tau} = \left[ \frac{\varepsilon}{w} \right] \left[ \frac{(1 - F(w))}{f(w)} \right]$$

### **Proof**

Using *L'Hopital's* rule the product of the income effect and inequality aversion effect that appears in equation (12) is equal to:

$$\lim_{w \rightarrow \infty} U_c * \lim_{w \rightarrow \infty} \frac{\int_0^{\infty} \left( \frac{1}{U_c} - \frac{g}{\gamma} \right) f(x) dx}{1 - F(w)} = 1 - \lim_{w \rightarrow \infty} U_c \frac{g}{\gamma} = \left\{ \begin{array}{ll} 1 & \text{if } \lim_{w \rightarrow \infty} U_c = 0 \text{ or } \lim_{w \rightarrow \infty} g = 0 \\ 1 - \delta \frac{g}{\gamma} & \text{if } \lim_{w \rightarrow \infty} U_c = \delta \text{ and } \lim_{w \rightarrow \infty} g > 0 \end{array} \right\}$$

Equation (13) shows that the optimal top income tax rate is determined by two factors only: the shape of the distribution of skills and the efficiency effect ( $\varepsilon/w$ ), that depends on the compensated elasticity of labor supply.

The striking result that emerges from equation (13) is that under plausible conditions (discussed below) the standard income effect and degree of inequality aversion do not play any role in determining the top marginal income tax rate. Those two effects are important to the optimal shape of income tax schedule but not for the optimal asymptotic tax rate.<sup>6</sup>

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, income 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. The cost of a dollar in term of social welfare is  $g$  and it goes to zero with a standard

---

<sup>6</sup> Dahan and Strawczynski (2000) and Saez (2001) emphasize the importance of income effects for characterizing the shape of optimal income taxation at high income levels.

social utility function that has some degree of inequality aversion (or to a finite number) as wage goes to infinity. The standard income effect is canceled out exactly by the inequality aversion effect, and thus the product of the inequality effect and income effect equals one.

Equation (13) implies that the asymptotic income tax rate is the same both for Utilitarian and Rawlsian social welfare function as long as the marginal utility of consumption goes to zero as wage goes to infinity. In what follows, we will assume that the marginal utility of consumption goes to zero except, for obvious reason, the case of a linear utility of consumption.

As can be seen, the top marginal tax rate is slightly different if the marginal utility of consumption converges to some positive number and  $g$  is positive. For example, with a linear utility of consumption the asymptotic tax rate depends on whether  $g$  goes to zero or to some positive number. In that case, the product of the efficiency and distribution effects should be multiplied by some constant. However, the asymptotic tax rate is the same both for Utilitarian and Rawlsian social welfare function even with a linear utility of consumption if the product of the efficiency and inequality aversion effects goes to infinity.

Equation (13) is useful to understand the difference between optimal asymptotic tax rates using bounded and unbounded distributions. It is well-known that the efficiency effect is the single factor that determines the asymptotic tax rate when the distribution of skills is bounded. Equation (13) shows that for unbounded distributions there is one additional factor: the distribution effect. This reflects the fact that with an unbounded

distribution there is always an individual with a higher wage level than the one at which we calculate the optimal marginal tax rate.

### 2.3 Optimal asymptotic labor supply

There is little consensus in the empirical literature about the size of labor supply elasticity at high income levels. Some studies have found estimates in excess of one while others have found elasticities close to zero.<sup>7</sup> In this sub-section we characterize optimal labor supply and the implied compensated labor supply elasticity when  $w$  tends to infinity and marginal utility of consumption equals  $c^{-\mu}$  for two different forms of utility of leisure.<sup>8</sup>

Each one of these two forms of utility of leisure is in line with empirical evidence in some important respects but inconsistent with empirical literature in some other dimensions.

#### a) A constant elasticity of substitution, $V(L)=-\ln(1-L)^{-1}$

Using Equation (3), the first order condition for this type of utility is:

$$(14) \quad \frac{L^{\mu}}{(1-L)^2} = w^{1-\mu}$$

There are three possible outcomes for optimal labor supply. In the first case ( $\mu < 1$ ), which includes linear utility of consumption, as wage goes to infinity labor supply goes to one. At first glance, the fact that the individual work full time ( $L=1$ ) might look counterintuitive given that the wage rate is extremely high. The individual is induced to work more up to a maximum level (which in this case equals one) because

---

<sup>7</sup> See Table 1 in Gruber and Saez (2002).

<sup>8</sup> A third case, using a logarithmic utility of leisure, is presented in the appendix.

the marginal utility of consumption is positive, even at very high wage. Note that as  $L$  goes to one,  $\varepsilon$  goes to infinity and the compensated elasticity of labor supply goes to zero (which means that the labor supply is inelastic).

In the second case ( $\mu=1$ ), which is represented by the log utility of consumption, as wage goes to infinity labor supply tends to an internal solution. The internal solution for labor supply is the result of the typical property of log utility of consumption where income effect offsets the substitution effect regardless of the wage level, including infinity. The optimal labor supply has two solutions but only one ( $L=0.38$ ) has economic significance. In this case, the compensated elasticity of labor supply goes to 0.8.

In the third case ( $\mu>1$ ), which includes utility of the type  $-1/c$ , labor supply goes to zero. In this case the marginal utility of consumption (income effect) goes to zero faster than the substitution effect as wage goes to infinity. The income effect works to reduce labor supply at a faster rate than the opposite impact of the rising price of leisure (substitution effect) that works to raise labor supply. Note, that as  $L$  goes to zero  $\varepsilon$  goes to one and the compensated elasticity of labor supply goes to infinity.

This case in which the utility of consumption is  $-1/c$  is of particular interest because the elasticity of substitution between consumption and leisure is exactly 0.5. Following Stern (1973), who found an estimate of 0.406 for this elasticity, this case was considered in the literature as one of the benchmark cases (Tuomala, 1984, and Kanbur and Tuomala, 1994). However, as mentioned above, the compensated

elasticity of labor supply goes to infinity in this case, a feature that is unrealistic in view of the empirical literature on labor supply elasticities.

b) A constant compensated elasticity of labor supply:  $V(L)=1-L^b$

This type of utility of leisure is more popular in recent literature (Saez, 2001) and it has a nice property: the compensated labor supply elasticity is the same for all workers, regardless of their wage. However, this property is not in line with empirical literature. Gruber and Saez (2002) show that estimates of labor elasticity increase with income.

The first order condition in this case is:

$$(15) \quad bL^{\mu+b-1} = w^{1-\mu}$$

There are three cases for constant compensated elasticity of labor supply. The first case ( $\mu < 1$ ) produces a different result with respect to labor supply compared to the previous functional form of utility of leisure and log-utility of leisure (discussed in the appendix). The labor supply goes to *infinity* as wage tends to infinity and  $\epsilon$  equals to  $b$ . The fact that the marginal utility of consumption is positive even at very high levels of wage induces the individual to work more up to the maximum level which in the case of  $V(L)=1-L^b$  is unbounded.

Note, that when the utility of leisure is of the form  $V(L)=-\ln(1-L)$  or log-utility-of-leisure  $V$  tend to  $-\infty$  as  $L$  goes to one, and therefore  $L$  is bounded in these types of utility. In contrast, with a constant compensated elasticity of labor supply  $L$  is

unbounded since  $V$  is finite as  $L$  goes to one. It is important to note that the first order condition collapses if labor supply is restricted by an upper bound of one.<sup>9</sup>

In the second case ( $\mu=1$ ) labor tends to a finite number and in the third case ( $\mu>1$ ) labor supply goes to zero. These results are similar to the log-utility-of-leisure (see appendix). Note, that in all cases  $\varepsilon=b$ .

### 3. Optimal tax rates at the top

This section explores the optimal asymptotic income tax rate utilizing Equation (13) and the results on labor supply from the previous section.

#### Proposition 1:

*With a utility of leisure of the form  $V(L)=-\ln(1-L)$ , the asymptotic tax rate converges to one both for a Pareto and Lognormal distribution of skills, when the utility of consumption is linear.*

#### Proof

In this case, equation (15) can be written as follows:

$$(16) \quad \lim_{w \rightarrow \infty} \frac{\tau}{1-\tau} = \left[ \frac{\varepsilon}{w} \right] \left[ \frac{(1-F(w))}{f(w)} \right] \left[ 1 - \frac{g}{\gamma} \right] = \infty$$

With a Pareto distribution of the form  $f=\alpha k^\alpha/w^{1+\alpha}$  (where  $\alpha$  and  $k$  are some constants), the distribution effect  $\frac{1-F}{f}$  goes to  $\frac{w}{\alpha}$ . Thus,  $\tau/(1-\tau)$  goes

to  $\left[ \frac{\varepsilon}{\alpha} \right] \left[ 1 - \frac{g}{\gamma} \right]$ . As shown in the previous section, with a linear utility of

---

<sup>9</sup> Henceforth we will allow labor supply to tend to infinity.

consumption  $L$  approaches one and  $\varepsilon$  goes to infinity as wage goes to infinity, and therefore the optimal asymptotic tax rate goes to 100%.

In the Lognormal case, the distribution effect goes to  $ws/(\log w - \varpi)$  [where  $s$  denotes the standard deviation of  $\log(w)$  and  $\varpi$  denotes the average of  $\log(w)$ ].  $\tau/(1-\tau)$  goes to

$\left[ \frac{\varepsilon s}{\log w - \varpi} \right] \left[ 1 - \frac{g}{\gamma} \right]$  and it also goes to infinity as wage goes to infinity and the tax rate goes to 100%.

One way to see the intuition behind this rather surprising result is the following. In the case of Pareto distribution the distribution effect divided by  $w$  is also constant. The only factor that plays a role is  $\varepsilon$ , 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 tends to zero and the expression of  $\varepsilon$  goes to infinity as the wage goes to infinity. In other words, assuming a linear utility of consumption drives individuals to supply labor inelastically<sup>10</sup> 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 glance. In particular, this result is at the opposite polar of the well known result of Sadka (1976) and Seade (1977). However, replacing the utility of leisure with a constant compensated elasticity of labor supply and using log-normal distribution instead of

---

<sup>10</sup> Moffitt and Wilhelm (2000) found that the elasticity of hours of work by the rich to the marginal tax rates reduction in the US during the eighties was close to zero. However, it is generally recognized that the reaction of the rich to tax changes is based on other behavioral channels, like incentives to work as self-employed, or working in jobs in which compensation is deferred or tax-sheltered.



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.

Although the result of the first proposition has no practical implications to public policy, our technique of driving the optimal asymptotic income tax rate helps to clarify the forces behind this unique result. The result of  $\tau=100\%$  appeared in Mirrlees (1971), but to the best of our knowledge it was never emphasized explicitly in the optimal income tax literature.

We now turn to characterize optimal tax rates at the top assuming that the relevant earnings distribution is Pareto. There is a growing consensus that the Pareto distribution fits reasonably well the empirical earnings distribution at high income levels (Poterba and Feenberg, 1993).

**Proposition 2:**

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

As before, we assume a Pareto distribution of the form:  $f=\alpha k^\alpha/w^{1+\alpha}$ . In addition, 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 he made that labor supply could not exceed one.

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. 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). Infinite labor supply is clearly unrealistic.

### **Proof**

Plugging these two assumptions into equation (15), the optimal asymptotic income tax rate converges to:

$$(17) \quad \lim_{w \rightarrow \infty} \frac{\tau}{1 - \tau} = \frac{b}{\alpha}$$

The asymptotic tax rate depends on the efficiency effect multiplied by the distribution effect which converges to a finite number ( $b/\alpha$ ) both with linear and non-linear utility of consumption. In this particular case (i.e., Pareto distribution and constant labor elasticity) this result holds for any form of non-linear utility of consumption, as long as the marginal utility of consumption goes to zero as wage goes to infinity. For a compensated elasticity of 0.25 and 0.5, the optimal asymptotic tax rate varies from 60% to 71.4%, respectively.<sup>11</sup> An asymptotic tax rate of 60 percent is more plausible given that the 0.5 estimate is closer to the elasticity found by Saez and Gruber (2002) for high income levels.

---

<sup>11</sup> Saez (2001) has used the same values. When the compensated elasticity is 0.25 and 0.5, Saez's parameter  $k$  equals, respectively, 2 and 4. Our parameter  $b$  equals  $(1+k)$ .

The optimal asymptotic tax rate is slightly different with a linear utility of consumption if we assume that the marginal social utility (of private utility),  $g$  converges to a positive number:

$$(18) \quad \lim_{w \rightarrow \infty} \frac{\tau}{1 - \tau} = \frac{b}{\alpha} \left[ 1 - \frac{g}{\gamma} \right]$$

It is easy to see that the asymptotic tax rate is identical for both the linear and non linear utility of consumption once we assume that  $g$  converges to zero as wage goes to infinity. Thus, the income effect is immaterial to the optimal asymptotic income tax rate. For example, a linear utility of consumption and log utility of consumption are clearly very different with regard to income effect but still the optimal asymptotic tax is identical (while the compensated elasticity is the same in this case).

This finding makes clear that the crucial factor behind Saez's high asymptotic tax rate is related to the assumption of a relatively low compensated elasticity of labor supply compared to other functional forms which are discussed below and in the appendix. It is also evident that the high asymptotic rate is not because of changes in the intensity of income effect. While the income effect does not play a role in shaping the asymptotic tax rate, it is still important for the optimal income tax schedule.

**Proposition 3:**

*For a Pareto distribution of skills and  $V(L) = -(1-L)^{-1}$ , 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

For this type of utility of leisure the optimal tax rate formula of equation (15) is the following:

$$(19) \quad \lim_{w \rightarrow \infty} \frac{\tau}{1 - \tau} = \lim_{w \rightarrow \infty} \left[ \frac{\varepsilon}{w} \right] \left[ \frac{w}{\alpha} \right] = \frac{\varepsilon}{\alpha},$$

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. In the case of log-utility of consumption the value of  $\varepsilon$  equals to 2.225 given that the optimal labor supply tends to 0.38 as wage goes to infinity. Using simple algebra, the computed optimal asymptotic income tax rate equals to 53% if  $\alpha=2$ . This case is important in light of the empirical literature. The compensated elasticity of labor supply in this case equals 0.8, which is in the plausible range of estimates.

In the case of  $-1/c$  utility of consumption, labor supply goes to zero and consequently  $\varepsilon$  goes to one and the compensated elasticity of labor supply goes to infinity. When  $\alpha$  equals 2, the optimal asymptotic tax rate is 33%.<sup>12</sup> The combination of infinite compensated elasticity of labor supply and optimal tax rate greater than zero is surprising. Though, the fact that the compensated elasticity goes to infinity does not imply that  $\varepsilon$  goes to zero. In fact,  $\varepsilon$  goes to one as  $L$  goes to zero. The optimal asymptotic tax rate reflects two conflicting forces: the efficiency effect component goes down as wage increases but the distribution effect pulls the tax rate up. It turns out that those two conflicting forces cancels each other and  $\tau/(1-\tau)$  approaches a finite number as wage goes to infinity.

---

<sup>12</sup> Using log-utility of leisure, instead of  $V(L)=-\ln(L)$ , would yield the same asymptotic tax rate (see in the appendix).

As mentioned above, the joint assumptions on utility of consumption and leisure implies a 0.5 elasticity of substitution between leisure and consumption. Following the empirical evidence, this particular functional form became a benchmark case in the literature.

Summarizing the results for Pareto distribution, we have found that for plausible assumptions either on compensated elasticity of labor supply or the elasticity of substitution between consumption and leisure, the top marginal tax rate is finite, in a range between 33% and 60%.

**Table 1 – Optimal asymptotic tax rates: Pareto Distribution <sup>a</sup>**

Utility of Consumption	Utility of Leisure		
	Constant Labor Elasticity <sup>b</sup>	-1/(1-L)	Logarithmic <sup>c</sup>
Linear	60%	100%	100%
Logarithmic	60%	53%	50%
-1/c	60%	33%	33%

- a. Assuming that  $\alpha = 2$ .
- b. Assuming a constant compensated elasticity of 0.5.
- c. Discussed in the appendix.

#### **4. Optimal asymptotic tax rates with a lognormal distribution**

While Pareto distribution fits the right tail, the whole distribution is best characterized by lognormal distribution (Aitchison and Brown, 1957). Moreover, almost all simulations in the old literature (until the nineties) were based on a lognormal distribution.<sup>13</sup> Deriving the optimal asymptotic tax rate for a lognormal distribution

<sup>13</sup> Mirrlees (1971), Tuomala (1984), Kanbur and Tuomala (1994).

helps to link the more recent literature with previous results. Proposition 4 characterizes the optimal asymptotic tax rate using a lognormal distribution.

**Proposition 4:**

*With a lognormal distribution of skills and a non-linear utility of consumption the optimal asymptotic tax rate converges to zero both for a constant compensated elasticity of labor supply and utility of leisure of the form  $V(L)=-\ln(1-L)$ .*

**Proof:**

In this case equation (15) is as follows:

$$(20) \quad \lim_{w \rightarrow \infty} \frac{\tau}{1 - \tau} = \lim_{w \rightarrow \infty} \frac{\varepsilon S}{\log w - \varpi},$$

This proposition emphasizes once more the importance of the assumption on the distribution of skills. 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. It is well known for normal distribution that 1 over the hazard rate of normal distribution goes to zero when  $w$  goes to infinity. Thus, a lognormal distribution pulls the optimal tax rate down at high levels of income compared to Pareto distribution.

When using a constant compensated elasticity of labor, the optimal asymptotic tax rate is zero regardless of the type of utility of consumption that is employed. This result shows that the switch from a low to a relatively high asymptotic tax rate could not be achieved by a shift from log-utility of leisure to a constant compensated

elasticity of labor supply. A Pareto distribution (instead of log-normal) is essential to avoid a zero income tax rate at the top.

The optimal asymptotic tax rate is zero with a log utility of leisure and a non-linear utility of consumption (discussed in the appendix). 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 result shows that Mirrlees simulation of the top marginal tax rate is in fact a bad approximation.

**Table 2 – Optimal asymptotic tax rates: Lognormal Distribution**

Utility of Consumption	Utility of Leisure		
	Constant Labor Elasticity	-1/(1-L)	Logarithmic <sup>a</sup>
Linear	0	100%	100%
Logarithmic	0	0	0
-1/c	0	0	0

a Discussed in the appendix.

## 5. Summary and Conclusions

Our paper provides a simple analytical expression for the optimal asymptotic tax rate. While all four effects are essential for the shape of income tax schedule only two effects play a role in determining the optimal asymptotic tax rate; under certain

conditions both income and inequality aversion effects are immaterial for the optimal income tax rate at the top.

As long as the marginal utility of consumption converges to zero, a policy maker needs only two types of information to set the optimal top income tax rate: the compensated elasticity of labor supply and the shape of skills 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, which 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%.

We found that the optimal asymptotic tax rate converges to a finite value using plausible assumptions about the efficiency effect, and a Pareto distribution of skills, which is considered today as the benchmark distribution for high income levels. For empirically accepted estimations of the compensated labor supply elasticity and the elasticity of substitution, we found that the optimal tax rates at the top are between 33% and 60%, a range that covers marginal tax rates imposed in developed countries.



A non-linear utility of consumption combined with a lognormal distribution of skills, which was very common in the old income taxation literature, leads to a polar result: a zero optimal asymptotic tax rate. This result is obtained assuming that the utility of leisure is  $-(1-L)^{-1}$  or a log-utility of leisure. In contrast, with a linear utility of consumption we get the other polar result: an optimal asymptotic tax rate of 100%. An exception of this result is obtained in the case of a constant compensated elasticity of labor supply, where the optimal asymptotic tax rate converges to zero even with a linear utility of consumption.

## **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 taxation: 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.

Gruber, Jon and Saez, Emmanuel, "The elasticity of taxable income: evidence and implications", *Journal of Public Economics*, 2002, 84, pp. 1-32.

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.

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

Moffitt, Robert A. and Mark O. Wilhelm, "Taxation and the labor supply decisions of the affluent", in Joel Slemrod, Ed., *Does Atlas Shrug?*, New York: Russell Sage Foundation, 2000, pp. 193-234.

Myles D. Gareth, *Public Economics*, Cambridge University 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.

Stern, Nicholas. "On the specification of models of optimum income taxation", *Journal of Public Economics*, 1976, 6, pp. 123-162.

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.

## **Appendix: The Log-utility of leisure case**

The log utility of leisure was very common in the old literature on optimal income tax. Analyzing this type of utility of leisure in addition to the cases that are covered in the text helps to link the results in the more recent to the old literature. It allows us to trace the changes in the assumptions that are responsible of the changes in the results.

### **a) The optimal labor supply**

In many respects, log-utility of leisure, that was typical in the old literature on optimal income tax (for example, Mirrlees 1971), is similar to utility of leisure of the type  $V(L)=-\ln(1-L)$  discussed in the text. The optimal labor supply is the same for both types of utility of leisure in the case of  $\mu < 1$  ( $L=1$ ) and  $\mu > 1$  ( $L=0$ ). In the case where  $\mu=1$ , the optimal labor supply has an internal solution but it is higher for log-utility of leisure ( $L=0.5$ ). In this case, the compensated elasticity of labor supply goes to one as wage goes to infinity. Thus, for  $\mu=1$  the labor supply is clearly within a plausible range but the compensated elasticity of labor supply is at the upper range of the empirical literature.

### **b) The optimal asymptotic tax rate**

#### **Proposition A1:**

*With a log-utility-of-leisure, the asymptotic tax rate converges to one both for a Pareto and Lognormal distribution of skills, when the utility of consumption is linear.*

Using equation (15) with a Pareto distribution,  $\tau/(1-\tau)$  goes to  $\varepsilon/\alpha$ . As shown in the previous sub-section, with a linear utility of consumption  $L$  approaches one and  $\varepsilon$

goes to infinity as wage goes to infinity and therefore the optimal asymptotic tax rate goes to 100%.

In the Lognormal case, the distribution effect goes to  $w\sigma/(\log w - \bar{\omega})$  and  $\tau/(1-\tau)$  goes to  $\varepsilon/(\log w - \bar{\omega})$ . In this case  $\varepsilon = -V_L = U_C w(1-\tau)$  and therefore  $\tau/(1-\tau)^2$  goes to  $w\sigma/(\log w - \bar{\omega})$  that goes to infinity as wage goes to infinity and the tax rate goes to 100%.

**Proposition A2:**

*For a Pareto distribution of skills and a 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  $-1/c$ .*

In the case of log-utility of consumption the value of  $\varepsilon$  equals to 2 given that the optimal labor supply tends to 0.5 as wage goes to infinity. Using simple algebra, the computed optimal asymptotic income tax rate equals to 50% if  $\alpha=2$ . This case is important in light of the empirical literature. The compensated elasticity of labor supply in this case equals 1, which is in the plausible range of estimates.

In the case of  $-1/c$  utility of consumption, labor supply goes to zero and consequently  $\varepsilon$  goes to one and the compensated elasticity of labor supply goes to infinity. When  $\alpha$  equals 2, the optimal asymptotic tax rate is 33%.

**Proposition A3:**

*With a lognormal distribution of skills, a log-utility of leisure and a non-linear utility of consumption, the optimal asymptotic tax rate converges to zero.*

In the Lognormal case, the distribution effect goes to  $ws/(\log w - \bar{\omega})$  and  $\tau/(1-\tau)$  goes to  $\varepsilon s/(\log w - \bar{\omega})$ . In this case, as wage goes to infinity  $\varepsilon$  goes to one ( $L$  goes to zero) and  $\tau/(1-\tau)$  goes to  $s/(\log w - \bar{\omega})$ , that goes to zero. Thus, the optimal marginal tax rate goes to zero.

As discussed in the text 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.