THE OPTIMAL NON-LINEAR INCOME TAX

Momi Dahan and Michel Strawczynski

Discussion Paper No. 97.02

July 1997

Research Department, Bank of Israel, POB 780, 91007 Jerusalem, Israel

The optimal non-linear income tax

Momi Dahan and Michel Strawczynski*

July 1997

ABSTRACT

The dominant model for income taxation in the public finance literature is the classical model of skills (Mirrlees, 1971). Until recently, an influential number of works using this model seemed to support declining marginal tax rates at high income levels. In this paper we use Diamond's (1996) methodology in order to explore the critical assumptions that lead to increasing or decreasing marginal tax rates. We find that with a lognormal distribution of skills and zero income effects there is a case for increasing marginal tax rates at high income levels. By performing a Kernel estimation to Israeli data we find empirical support for the lognormal distribution of skills.

* Bank of Israel, Research Department, P.O.B. 780, Jerusalem, Israel. FAX: 972-2-6552657. Email: momid@bankisrael.gov.il and michels@bankisrael.gov.il. We are thankful to Irina Blits for excellent research assistance. We thank David Frankel, Giora Hanoch, Zvi Hercowitz, Efraim Sadka, Eytan Sheshinski, Shlomo Yitzhaki, participants at seminars at Tel Aviv University, the Hebrew University of Jerusalem and the Bank of Israel for helpful remarks, and Arie Beresteanu for his aid on the Kernel estimation.

The optimal non-linear income tax

1. Introduction

In actual income tax systems marginal rates rise with income. However, the widely accepted model for optimal income taxation in the public finance literature (Mirrlees, 1971), does not prescript a clear-cut pattern for optimal marginal taxes. The theoretical result of this model calls for optimal marginal rates that lie among zero (for the top and bottom of the ability's scale)[1] and one. Consequently, it is explicitly agreed in this literature that the optimum shape must be obtained by running simulations, which are based on different assumptions -- as described below. Until recently, a vast quantity of works seemed to generate a strong piece of evidence that supports declining marginal tax rates at high income levels; between others we find Mirrlees work itself[2], Tuomala (1984) and Slemrod et al. (1994). However, Diamond (1996) has reopened the question of the optimum shape by showing an example where optimal tax rates follow a U-shaped pattern; i.e., rising marginal tax rates at high income levels.

In this paper we use Diamond's methodology in order to calculate the optimum non-linear income tax. The departure from Diamond's paper is twofold: I) We aim

---

1 With a finite maximum for the skill distribution, the optimal marginal tax rate at the income level of the top skill is zero (Sadka, 1976a and Seade, 1977). If individuals choose to work, a zero optimal marginal tax rate is also obtained at the income level of the bottom skill. (Seade, 1977)

2 In Mirrlees' simulations optimal marginal taxes decline with income. Since the shape was close to linearity, this point was not stressed nor by Mirrlees nor by other authors citing his work.

at shedding light at the puzzling contradiction in the literature between rising and declining marginal taxes at high income levels and II) We concentrate on optimal taxes under a lognormal distribution -- which is in line with the benchmark assumption of a vast quantity of previous papers[3].

The paper is organized as follows. Section 2 presents the model and summarizes the results of the literature. Section 3 deals with the controversy in the literature between declining and rising marginal taxes for high income individuals. Section 4 studies the theoretical implications of our model for the case of a lognormal distribution and presents an empirical estimation of the distribution of skills using data from Israel. Section 5 concludes the paper and presents directions for further research.

2. The classical model and the optimum shape

A. 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. We assume all over the paper that V is concave, and we elaborate below on the properties of U. The budget constraint at the individual level is:

[3] In his benchmark case Mirrlees (1971) assumed a lognormal distribution with mean $n = 0.4$ (n represents skill) and $\sigma=0.39$. Other simulations using this distribution are found in Tuomala (1984),

(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:

$$(3)\quad \frac{V'}{U_C} = (1-\tau)w\,, \qquad\qquad \tau \equiv T'$$

where V' 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[4]:

$$(4)\quad \frac{du}{dw} = V'\frac{L}{w} > 0$$

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

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

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:

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

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

Kanbur and Tuomala (1994) and Slemrod et al. (1994).

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

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

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

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:

$$(8)\quad H_L = \left[\, \frac{\partial G}{\partial u}\frac{\partial u}{\partial L}\bigg|_{\bar{u}} + \gamma\left(w - \frac{dC}{dL}\bigg|_{\bar{u}}\right) \right]\frac{dF}{dw} + \frac{\lambda}{w}(V' - LV'') = 0$$

$$= \gamma\left(w - \frac{V'}{U_C}\right)\frac{dF}{dw} + \lambda\frac{V'}{w}(1+\varepsilon) = 0\,, \qquad \varepsilon \equiv -\frac{V''L}{V'}$$

$$(9)\quad -H_u = -\left(\frac{dG(u)}{du} - \gamma\frac{dC}{du}\right)\frac{dF}{dw} = \frac{d\lambda}{dw}$$

The transversality conditions are:

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

By integration of both sides in (9), and using the transversality condition (10), we obtain an expresion for λ:

$$(11)\quad \int_w^{w_H}\left(\frac{\gamma}{u_C} - \frac{dG}{du}\right)\frac{dF}{dw}dw = \int_w^{w_H}\frac{d\lambda}{dw}dw = \lambda(w_H) - \lambda(w) = -\lambda(w)$$

Using this expression and the F.O.C. of both the government and the individuals, we obtain the following expression for the optimum non-linear marginal tax:

$$(12)\qquad \frac{\tau}{1-\tau}=\left[\frac{(1+\varepsilon)}{w}\right]\left[U_C\right]\left[\frac{\int_{w}^{w_H}\left[\frac{\gamma}{U_C}-G'(u)\right]f\,dw}{\gamma(1-F)}\right]\left[\frac{(1-F)}{f}\right],\qquad f\equiv\frac{dF}{dw}$$

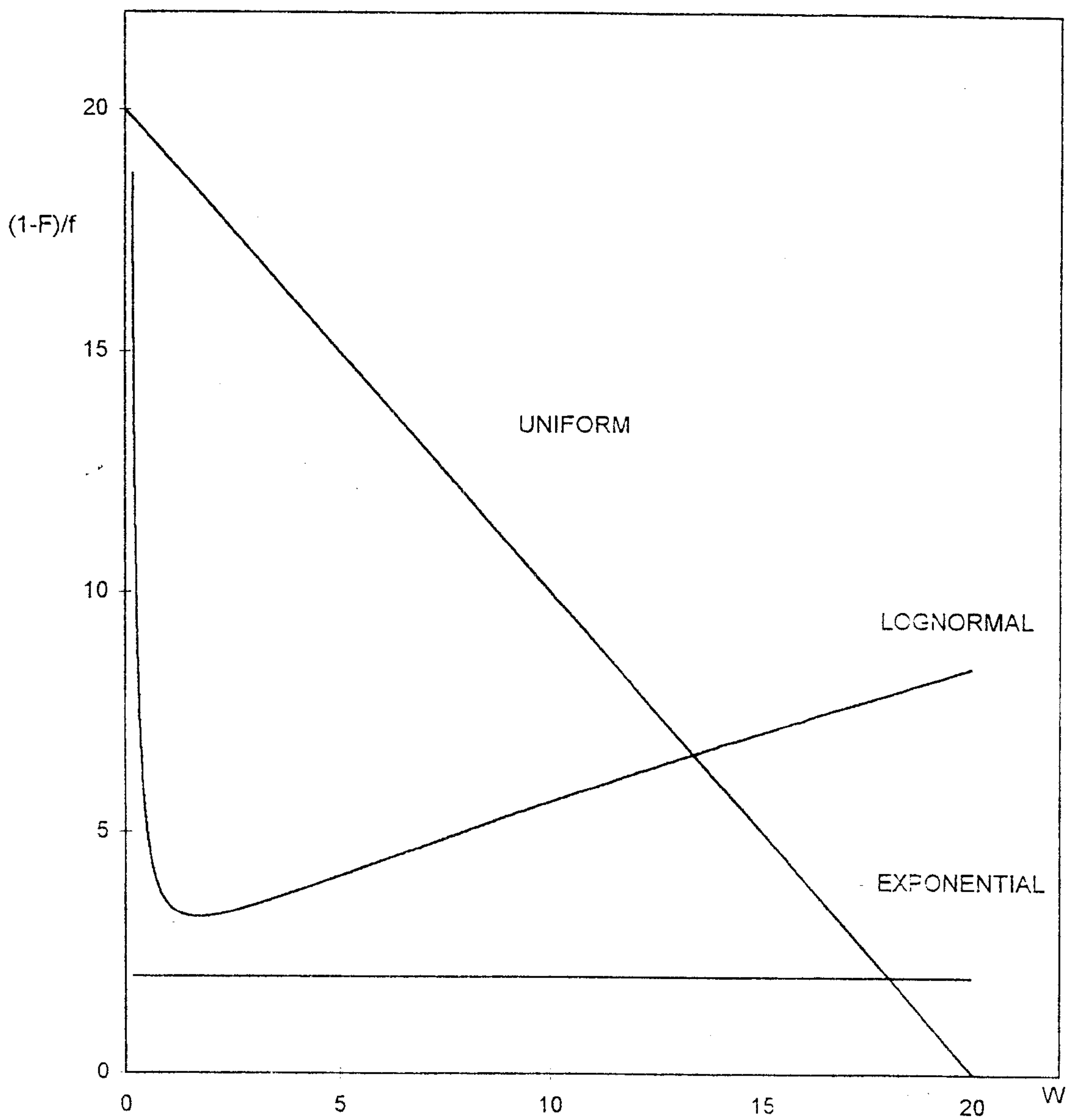
The first term in brackets represents the *"efficiency effect"* : the lower the labor supply elasticity and the lower the wage (skill), the higher the marginal tax as a consequence of efficiency considerations. The second term (U_C) represents *"income effects"*. Income effects play a role, since when the utility of consumption is convex ($U_{CC}<0$), changes in the level of consumption affect the chosen amount of leisure at the optimum; with a decreasing U_C, this effect causes optimal taxes to decrease with the level of income. The intuition is as follows: the imposition of taxes imply a negative income effect; since leisure is a normal good, individuals choose to work more and thus mitigate the negative income effect. This 'mitigation' effect is lower for rich individuals, since they have a lower marginal utility of consumption. As a consequence of that, labor supply for the rich is more elastic – which calls for lower taxes. The third term is the *"inequality aversion"* effect; if we assume a decreasing social marginal utility ($G''<0$), the higher w, the higher the optimum marginal tax rate as a consequence of this effect. Note also that when U_C decreases with w, the impact of the inequality aversion effect increases, since transferring one dollar from the rich to the poor increases social utility. The last term is the *"distribution effect"*. This term is composed by the ratio of individuals above the income level, 1-F, to the individuals in the income level itself, f. To understand this effect, we must note that when we rise the marginal tax rate at a low income level, on one hand we distort the decision for this income level, but on the other hand this new higher marginal tax acts as a lump-sum tax on higher income levels; this is so since at high income levels the decision at the margin is not affected by marginal tax rates in previous brackets. The higher (1-F), the higher the quantity of individuals that are paying higher lump-sum taxes, and consequently the higher is the optimum marginal tax. On the other hand, the higher f the lower the optimum tax, since the higher is the quantity of individuals affected by the distortion. Figure 1 shows the "distribution effect" for different income distributions (uniform, exponential and log-normal).

[INSERT FIGURE 1 HERE]

The intuition becomes clear by looking at the uniform distribution case: since f is equal in all income levels, the marginal tax declines all over the range - reflecting the fact that as we advance in the income axis, marginal taxes act as lump-sum for fewer individuals.

By using equation 12 we may obtain all different optimal shapes shown in the literature - according to the assumptions on the different components of the model. Since there is a vast number of papers dealing with the different assumptions, we summarize below the results of the literature on these components.

Figure 1 - The distribution effect



Parameters: Uniform – from 0 to 20,
Lognormal – $\mu = 1 \quad \sigma = 1$,
Exponential – $\lambda = \frac{1}{2}$

B. Results in the literature

Leaving aside papers that consider income uncertainty[5], a number of influential papers using the classical model of income taxation seem to generate a strong case for declining marginal tax rates at high income levels (table 1).

Table 1 - The optimum shape according to the literature

	optimum schedule (marginal taxes)			
	Mirrlees	Atkinson	Tuomala	Kanbur and
F (n)	(1971)	(1973)	(1984a)	Tuomala (1994)
0.1	24	33	68	70
0.5	22	30	62	80
0.9	19	26	50	78
0.99	17	21	35	48

Note: All the simulations assume a lognormal distribution with a mean skill of 0.4 (μ=-1). The first three papers assume that the variance of the logarithm of skills is 0.39, while the last paper assumes a variance of 1. All papers assume an elasticity of substitution (e) of 0.5, except Mirrlees (1971) where e=1. All papers assume the existence of income effects: in the first two papers U_C equals 1/C while in the last two it equals $1/C^2$. The revenue requirement is 7, 2, 10 and 10 percent of total income, respectively. All papers assume a utilitarian social planner except Atkinson (1973), who assumes an 'inequality aversion' coefficient of 2.

[5] With income uncertainty, optimal taxes rise with income. Varian (1980) provides an example where differences in income are due to 'luck'. Tuomala (1984b) provides an example that introduces labor income uncertainty to the classical model of skills.

Although results of these works differ in the level of optimal taxes and in the degree of linearity, it is a common feature that marginal taxes fall at high income levels[6]. However, a recent paper by Diamond (1996) that uses the same model showed rising marginal taxes at high income levels. This fact raises the following question: which is the critical assumption that turns the results from rising to declining marginal rates?

3. Rising or declining rates at high income levels?

In order to answer this question we list all critical assumptions one by one. Since the existence of income effects is a critical assumption, we analyze it separately in sub-section A, and then in sub-section B we summarize all other assumptions.

A. Rising or declining rates at high income levels? (or: What is the assumption that drives Mirrlees' results of declining marginal tax rates?)

Mirrlees' cases were obtained for $V=\ln(1-L)$. Under this assumption the *"efficiency effect"* in equation 12 is equal to $(1-\tau)U_C$, and we may write equation 12 as follows:

[6] Marginal taxes for the highest percentile (table 1) suggest that the result of a zero marginal tax at the top of the distribution is local. For further analysis on this point see Tuomala (1984a, p. 364).

$$(12)' \quad \frac{\tau}{(1-\tau)^2} = \left[(U_C)^2\right] \left[\frac{\int_{w}^{w_H} [\frac{\gamma}{U_C} - G'(U)]\, f dw}{\gamma(1-F)} \right] \left[\frac{(1-F)}{f}\right], \qquad f \equiv \frac{dF}{dw}$$

Equation 12' shows that the term denominated as *"income effects"* is crucial for the result of declining marginal tax rates.[7] To see this point we must note that at high income levels the last two terms *increase* with income: i) with G''<0 the distribution effect implies increasing marginal taxes all over the range, and ii) we show in the next section that with a lognormal distribution the distribution effect implies increasing marginal taxes at high income levels. Since income effects is the single term that decreases at high income levels, it becomes clear that this is the term that drives the result of declining marginal taxes. Figure 2 shows the results of a simulation that compares optimal taxes with and without income effects.[8] In this figure U(C)=lnC corresponds to the case analyzed by Mirrlees, while U(C)=C implies the inexistence of income effects, as assumed by Diamond (1996). Thus, figure 2 shows that income effects constitute the critical assumption that explains the transition at high income levels from Mirrlees' declining optimal rates to Diamonds' rising rates.

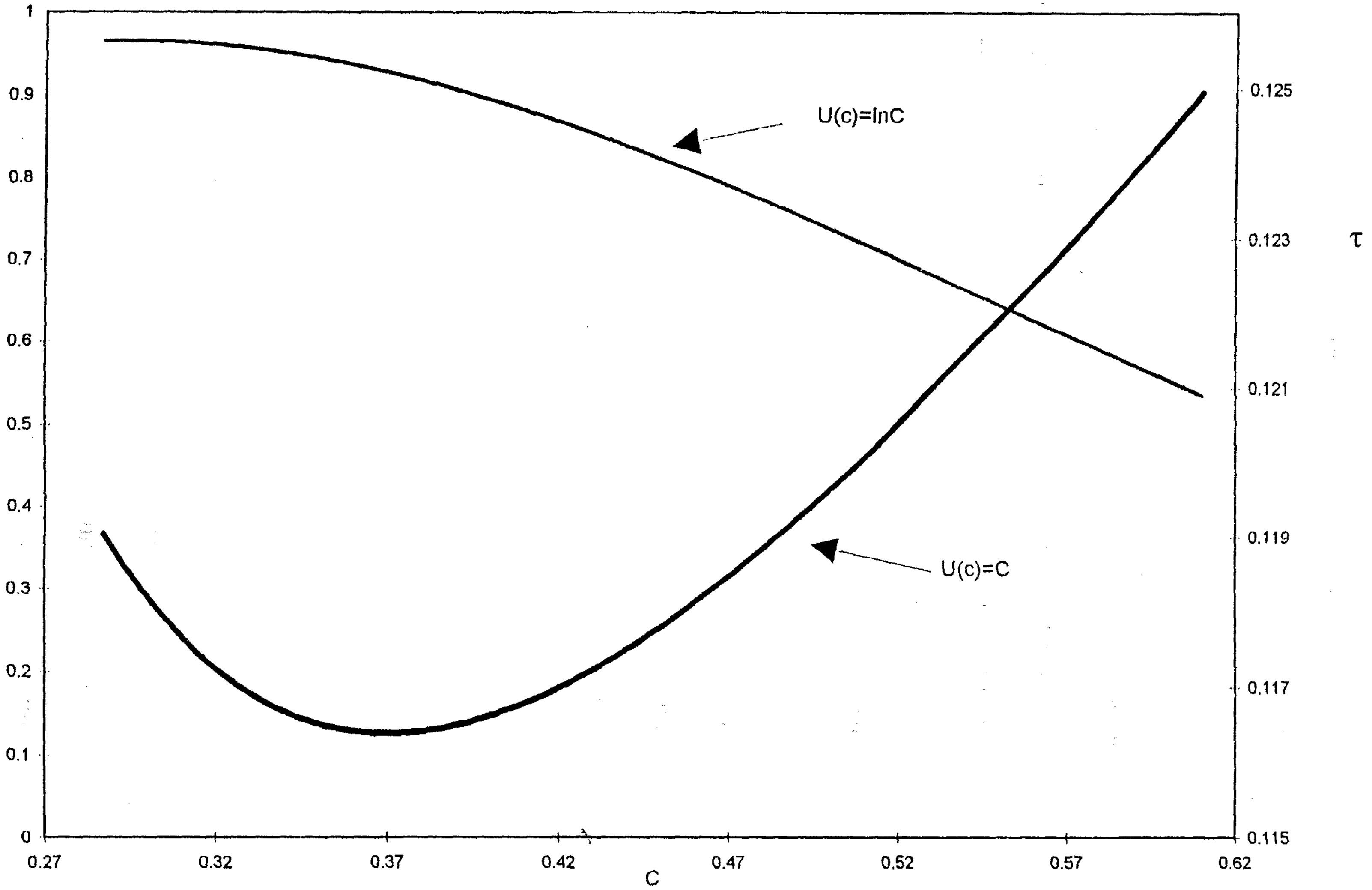
[INSERT FIGURE 2 HERE]

---

[7] Note that the elasticity of substitution is the same (=1) for both the linear and logarithmic case: e=(∂A/∂B)*(B/A), where A=(1-L)/C and B=(∂u/∂C)/[∂u/∂(1-L)]. For U(C)=C, A=(1-L)/C and B=1-L, i.e., e=1. For U(C)=ln C, A=B=1,i.e., e=1.

[8] The details of the simulation are explained in appendix A. In general terms, the assumption we need in order to avoid income effects is u=f[C-V(L)], where u and V are convex functions.

Figure 2 - OPTIMAL TAXES AT HIGH INCOME LEVELS



It is important to stress that all the other results shown in table 1 (Atkinson, 1973, Tuomala, 1984a and Kanbur and Tuomala, 1994) assume the existence of income effects. Thus, by using equation 12 we know that the result of declining marginal tax rates at high income levels as obtained in these studies is affected by this assumption.

B. Other assumptions

a. The elasticity of substitution

Stern (1976) found that assuming a more realistic assumption on the elasticity of substitution between consumption and leisure - according to available empirical evidence - leads to higher (linear) marginal taxes than those found by Mirrlees. Tuomala (1984a) had shown that the elasticity of substitution affects not only the *level* of taxes, but also the *shape* of optimum taxation[9]. According to his simulations different values of a constant elasticity of substitution can explain the change from an almost linear shape as found by Mirrlees to a significantly non-linear shape, with optimum marginal taxes falling at high income levels. However, it is worth to note that his example implies both the reduction of the elasticity of substitution and an increase in the impact of income effects. In appendix A.2 we show that a similar result is obtained by changing only the degree of income effects.

[9] The range of values for the elasticity of substitution is chosen between 0.5 and 1. Sadka (1976b) shows that with higher values of the elasticity of substitution, it is possible to get an optimal pattern with decreasing average tax rates - even when efficiency effects do not play a role (i.e., with lump-sum taxes/transfers) both for a utilitarian and a Rawlsian social planner.

It would be possible to obtain rising marginal tax rates by assuming that the elasticity of substitution goes down while utility goes up. Nevertheless, to the best of our knowledge there is no empirical support for such an assumption. Moreover, all the papers cited above assumed a constant elasticity of substitution[10].

b. Labor supply elasticity

It is possible to have rising marginal tax rates if the compensated labor supply elasticities go down as income goes up. We are not aware of any serious attempt to assume explicit differences in elasticities of labor supply as a function of income. It seems that the reason is twofold: I) most evidence on labor supply elasticities supports a low value, and II) there is no clear-cut evidence on the variability of elasticities across income levels[11].

c. The social planner

The assumption on the type of social planner may affect both the level and shape of the optimum tax system. Atkinson (1973) found that if the social planner maximizes the utility of the poorest individual ("Rawlsian"), marginal tax rates are higher than the ones obtained with a utilitarian social planner, as in Mirrlees (1971). Moreover, the shape of the optimum system is substantially non-linear. However, it is remarkable that even in the Rawlsian case marginal tax rates decline at high income levels. Slemrod et *al.* (1994) found that choosing a more "inequality averse" social

---

[10] This is true except for the example presented by Diamond (1996), where $V=a(1-L^k)$. In this case the elasticity of substitution is $(1-k)L/(1-L)$; i.e., it changes with labor supply.

[11] A significant difference is found only between men and women (Atkinson and Stiglitz, p. 51).

planner implies higher marginal taxes, but does not affect the result that in a two-bracket system the first marginal tax rate is higher than the second.

d. The revenue requirement

A higher revenue requirement raises the level of marginal taxes. If marginal utility is convex, it may also affect the optimum shape of the tax through the income effect. Mirrlees assumed a low revenue requirement, up to 9 percent of resources. In Kanbur and Tuomala (1994) changes in the revenue requirement still imply a remarked decline of marginal taxes at high income levels. Slemrod et *al.* (1994) found that raising the revenue requirement raises optimal taxes, but does not affect their finding in the context of a two-bracket system of a higher first marginal tax rate. As explained before, a rise of the revenue requirement implies higher income effects.

e. The distribution of income

According to the model shown above, clearly the distribution of income is an important factor: since different distributions imply a different ratio for (1-F)/f, it is crucial to learn about the impact of the income distribution on the optimum shape of taxes. Since Diamond (1996) did not elaborate explicitly on the lognormal distribution case - which is the benchmark distribution on empirical grounds - we turn in the following section to this case.

4. Optimal taxes with a lognormal distribution of skills

In this section we characterize the optimal shape for the case of a lognormal distribution of skills, which is the benchmark distribution in the income tax literature. Figure 3 characterizes the distribution effect for two pairs of values of the mean and variance of the lognormal distribution, according to the scenarios in Mirrlees (1971).

[INSERT FIGURE 3 HERE]

Figure 3 shows that the distribution effect implies a U-shaped pattern. In sub-section A we generalize this empirical observation by using two propositions.

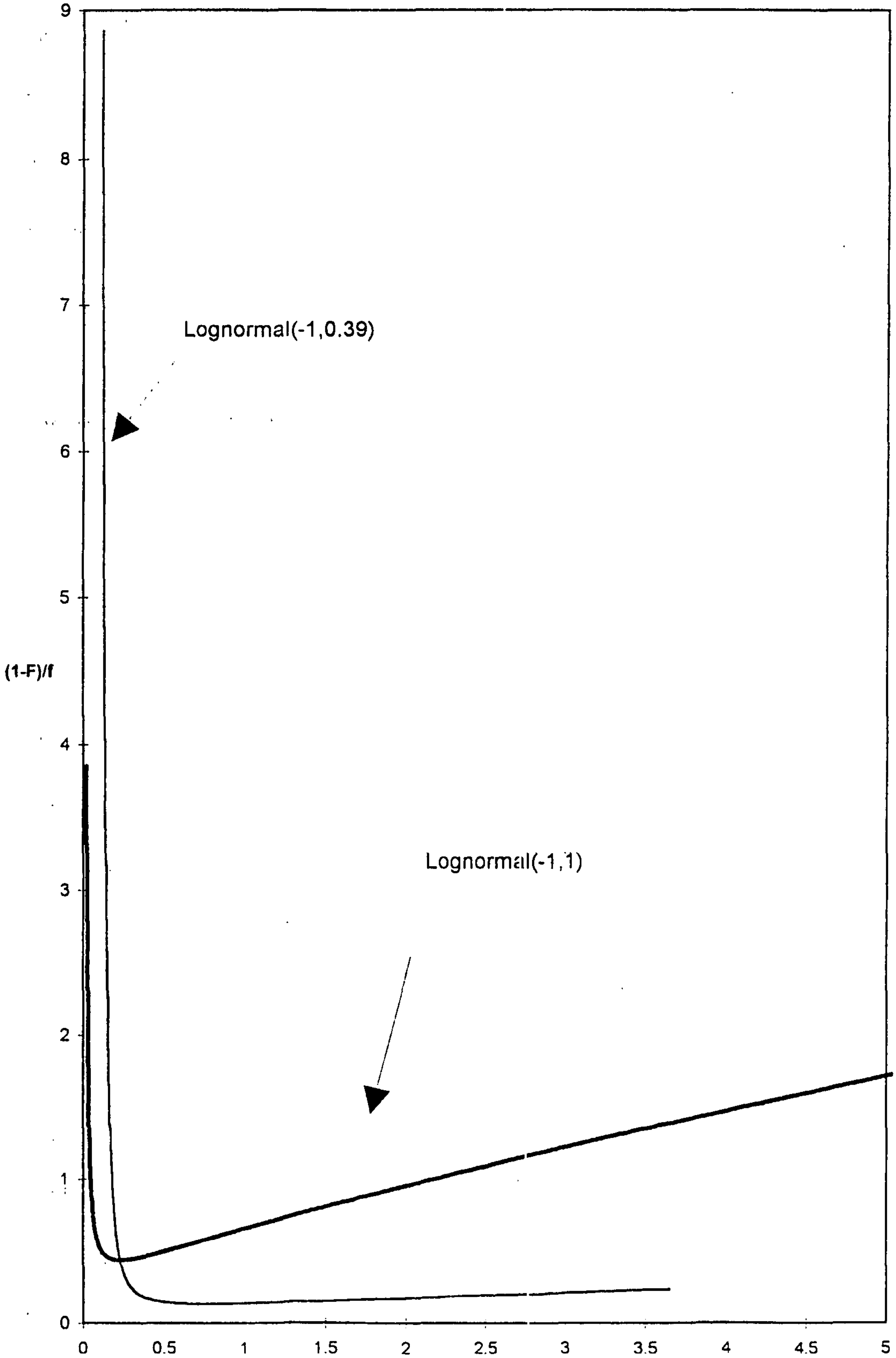
A. Optimal taxes with a lognormal distribution of skills

In order to characterize optimal taxes we use the following two propositions.

Proposition 1 - With a lognormal distribution, the distribution effect is U-shaped.

Proof (adapted from Lancaster, 1990, p. 47) - We may write the ratio (1-F)/f for the lognormal case as $D(w)=\sigma w/h(x)$, where log w is distributed normal with mean $\mu(x)$ and variance σ^2 , $x=(\ln w-\mu)/\sigma$ and h(x) is the standard normal hazard rate. The properties of h(x) are well-known: it increases monotonically from zero as x increases from $-\infty$, and as $x\to\infty$, h(x) approaches x. Thus, for large x, $D(w)\sim w\sigma^2 / (\ln w-\mu)$;i.e., D(w) approaches infinity. Furthermore, since the lognormal density starts from zero, $D(0)= \infty$. Thus, the function D(w) first decreases from infinity and then, ultimately, increases to infinity. The single minimum occurs at the value of x satisfying $h(x)= \sigma + x$.

Figure 3 - The distribution effect for the lognormal case



In appendix B we characterize this minimum for key values of the distribution-mode, median and mean.

Proposition 2 - Assume zero income effects and all the commonly accepted assumptions as in Mirrlees (1971): a social planner with decreasing social marginal utility (G''<0), a logarithmic utility of leisure, and a lognormal distribution. Then, optimal tax rates ***rise*** *for high income individuals (excluding the top itself).*

<u>Proof -</u> We apply equation 12 for V=ln(1-L) and U=C. Since in this case $1+\varepsilon=1/(1\text{-}L)$ and by the F.O.C. of the individual $1/(1\text{-}L)=w(1\text{-}\tau)$, we obtain the following equation:

$$(14) \qquad \frac{\tau}{(1-\tau)^2} = \left[\frac{\int_w^\infty [\gamma - G'(U)]\, f dm}{\gamma(1-F)} \right] \left[\frac{(1-F)}{f} \right], \qquad f \equiv \frac{dF}{dm}$$

i.e., in this case the efficiency effect is not at work. The first term in the right-hand side is the distribution effect; Since G''<0 by assumption, this term implies rising marginal rates all over the range. Finally, we have shown in proposition 1 that with a log-normal distribution the distribution effect implies rising marginal taxes at high income levels.

Proposition 2 is rather suggestive. It says that under Mirrlees' assumptions, if income effects are not present at high income levels, we can obtain rising marginal taxes at high income levels, a result which is in line with most systems worldwide, but is opposite to the simulations in the literature as shown in section 2.

We explore now the empirical viability of a lognormal distribution of skills.

B. The distribution of skills: an empirical test

In order to test the empirical viability of the model we test the empirical distribution of skills using Israeli data[12]. The Income Survey of the Central Bureau of Statistics in Israel includes both total income per household and number of working hours per-week. Using this data we obtain the wage per-working hour, which is our proxy for skills (marginal product). The next step is to test for the best fit in terms of the hole distribution of income, for the years 1993 and 1994. As explained before, empirical findings in the labor market show that it is not possible to find statistically significant differences in the elasticity of labor among income groups. However, there is a substantial difference in elasticities between men and women. Since we aim at looking at the distribution effect (i.e., efficiency considerations are neutral), we test the distribution of abilities only for *men* that are also head of households. Figures 4 and 5 show the results for the data as extracted straightforward from the survey.

[INSERT FIGURES 4 AND 5 HERE]

[12] In accordance to our theoretical framework we assume that 'skills' are exogenous. In reality, skills represent innate abilities and acquired human capital (through education and 'on the job training'). For a further discussion on the empirical distribution of skills see Sahota (1978).

Figure 4 - The distribution of skills in 1993

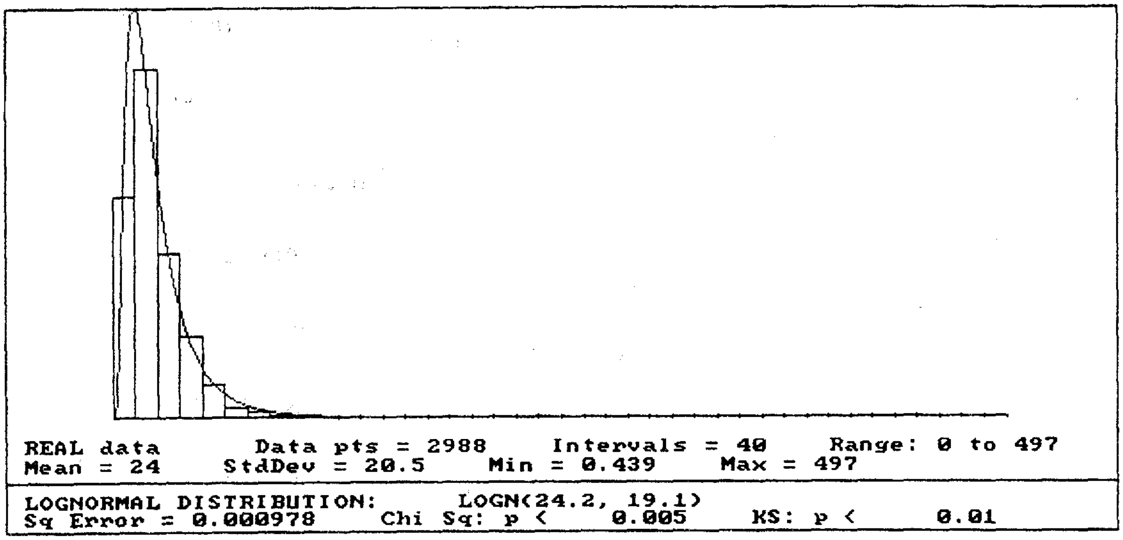
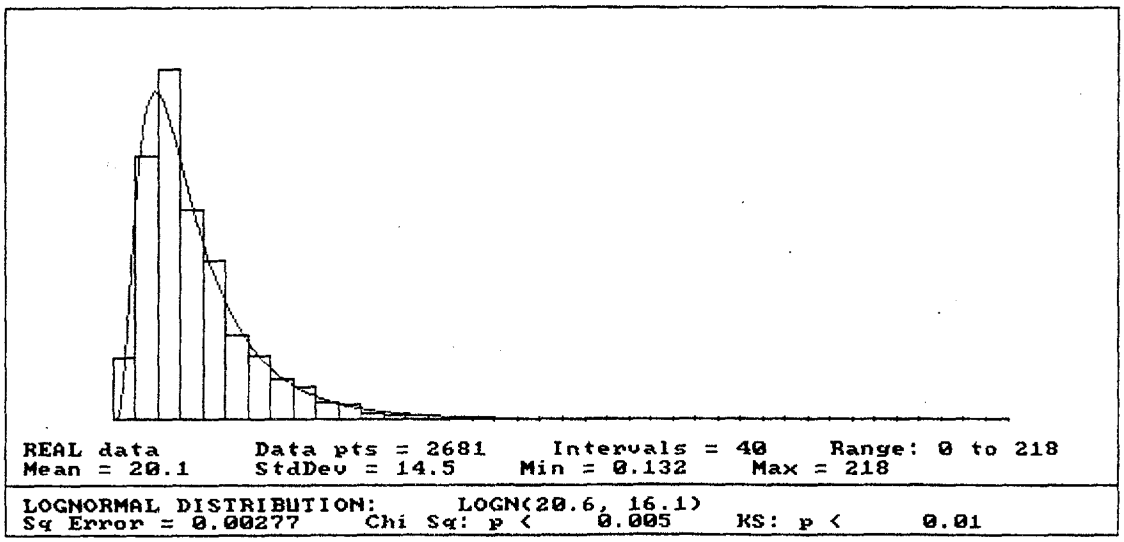


Figure 5 - The distribution of skills in 1994



It is highly remarkable that in both years the best approximation for the distribution of skills is the lognormal distribution. However, the results are not significant according to both Kolmogorov-Smirnov test and Chi-square test.

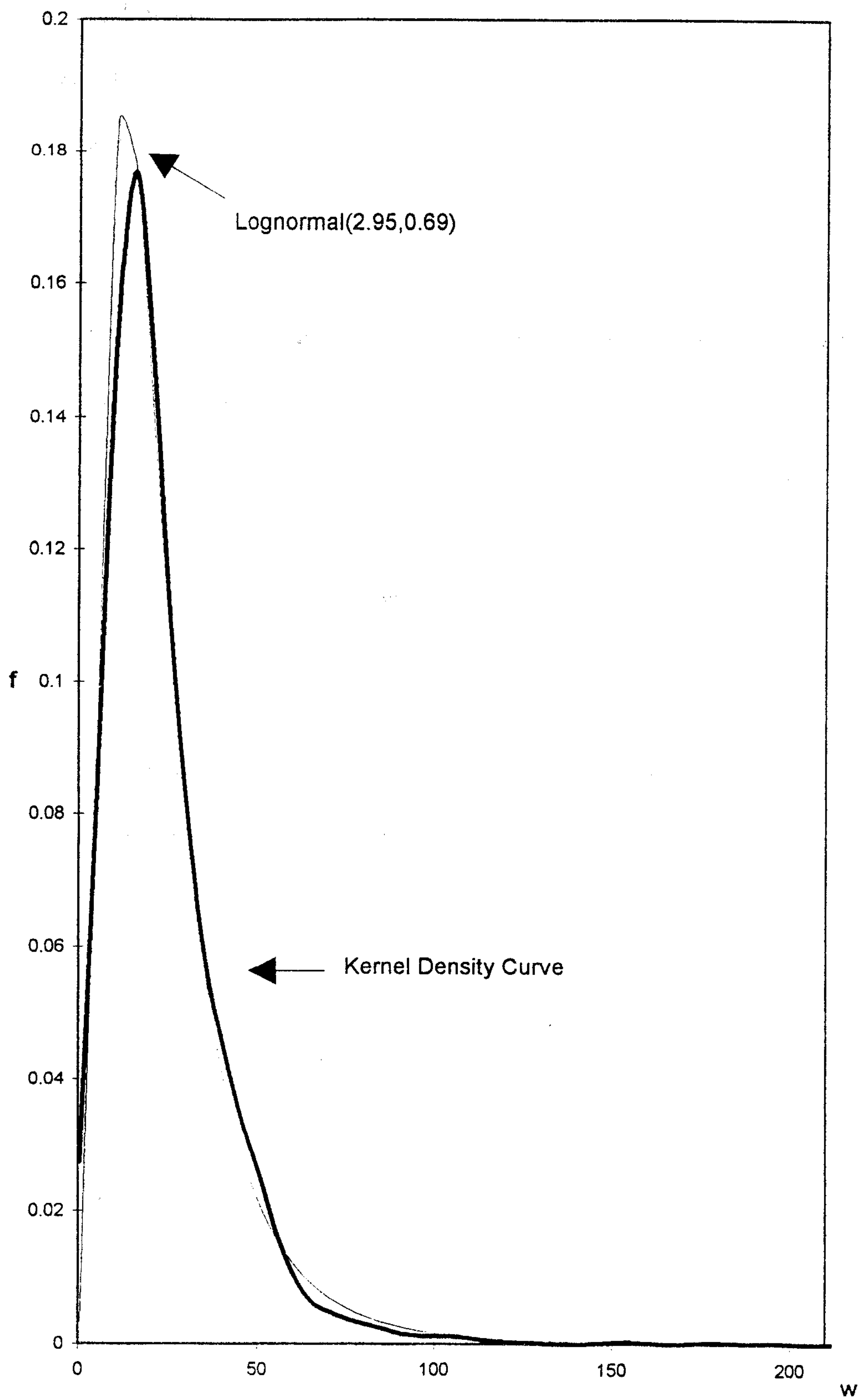
As well known, it is very difficult to get significant results by using raw data as taken straightforward from the sample, since it is based on discrete observations. In order to get a further inside on the significance of the results we decided to perform a Kernel estimation of the distribution function. The Kernel estimation is based on smoothing the raw data, and in its simple form it is a smoothed histogram. The properties of this estimation are well documented in the literature (see f.e. Burkhouser, Crews, Daly and Jenkins, 1996). The main advantage of the Kernel estimation is that it shows simultaneously the distribution's level, the modality and the spread of the distribution. For our purposes the main advantage of the Kernel estimation is that it eliminates unnecessary noise created by the use of discrete observations, as is the case of the raw data.

Figure 6 shows the results of the Kernel estimation for 1994, compared to the theoretical lognormal distribution by using the logarithm of the mean and variance as obtained from the sample.

[INSERT FIGURE 6 HERE]

In order to test whether the Kernel estimation is a good approximation to the theoretical distribution we perform the Kolmogorov-Smirnov test. This test consists on calculating the maximal difference (in absolute value) between the empirical

Figure 6 - Kernel estimation



distribution and the theoretical one, D_n , as compared to critical values as obtained by applying the Kolmogorov theorem. The null hypothesis is $F_0=F$, where F_0 is the empirical cumulative distribution and F is the theoretical distribution - which serves as a benchmark. The lower the value of D_n , the more likely we do not reject the null hypothesis.

Table 1 - A Kolmogorov-Smirnov test to the distribution of skills

	Critical values for different significance levels	
D_n	1 %	5 %
0.1124	0.1615	0.1347

Note: n(intervals)=99. Critical values are calculated according to Bickel and Doksum (1977), p. 483.

Results show that D_n is lower than the critical values for both 1 and 5 percent significance levels[13]. We conclude then that we cannot reject that the distribution of skills is lognormal.

[13] Note that the higher the level of significance, the more likely we do not reject the null hypothesis. As well-known, there is a trade off between type 1 error (rejecting the null hypothesis when it is true) and type 2 error (not rejecting the null hypothesis when it is false). The critical values provided in the table are consistent with the accepted level of significance as they appear in the standard application of the Kolmogorov-Smirnov test (Bickel and Doksum, 1977).

6. Conclusions and directions for further research

This paper simulates optimal income taxes by using the classical model of skills (Mirrlees, 1971). We find that under Mirrlees' assumptions, the existence of income effects explains the switch from declining optimal marginal tax rates to rising optimal rates at high income levels. The result of rising marginal tax rates at high income levels is in line with most systems worldwide, but contrasts with an important stream of the literature which shows declining optimal rates.

Our analysis suggests two main directions for further research. First, it is important to explore in empirical grounds the significance of income effects at high income levels. Second, our analysis calls for further research on how important is the assumption of income effects for a more general specification of the utility function. In particular, there is a need to solve the problem for the case where income effects and the elasticity of substitution can be changed once at a time, and not together as done until now in the frame of the simulations of the optimal non-linear income tax.[14]

[14] The C.E.S. function $u=[\alpha C^{-\rho} + (1-\alpha)(1-L)^{-\rho}]^{-\beta/\rho}$ allows for changes in income effects by performing changes in the parameter β. This utility function for the case of $\beta=1$ was assumed in Stern (1976) and Slemrod et. al (1994), both in the context of linear schedules.

Appendix A- A simulation of optimal taxes at high income levels

In order to simulate optimal taxes at high income levels we use the following approximation for marginal taxes - assuming a social planner with G'' < 0:[15]

$$(13)\quad \tau = \frac{-s_w U_c}{f(w)} \int_w^{w_H} \frac{1}{U_c} f(m)dm$$

where s_w is the marginal rate of substitution between labor and consumption[16] and w is the relevant level of skill for the calculation of the optimal marginal tax.

A.1) From declining to rising optimal marginal tax rates

Using the calculation of s_w for the different cases allows us to write the following approximations:

$$(13)'\quad \text{For } U(C)=\ln C,\ \frac{\tau}{(1-\tau)^2} = \frac{1}{f(w)C^2} \int_w^{w_H} Cf(m)dm$$

$$\text{For } U(C)=C,\ \frac{\tau}{(1-\tau)^2} = \frac{1}{f(w)} \int_w^{w_H} f(m)dm$$

The results of the simulation were shown in figure 2.

A.2) Increasing the degree of income effects

We use the following examples: V=ln(1-L), and U=lnC and -1/C respectively.

Since U_C equals respectively 1/C and $1/C^2$, the impact of income effects is

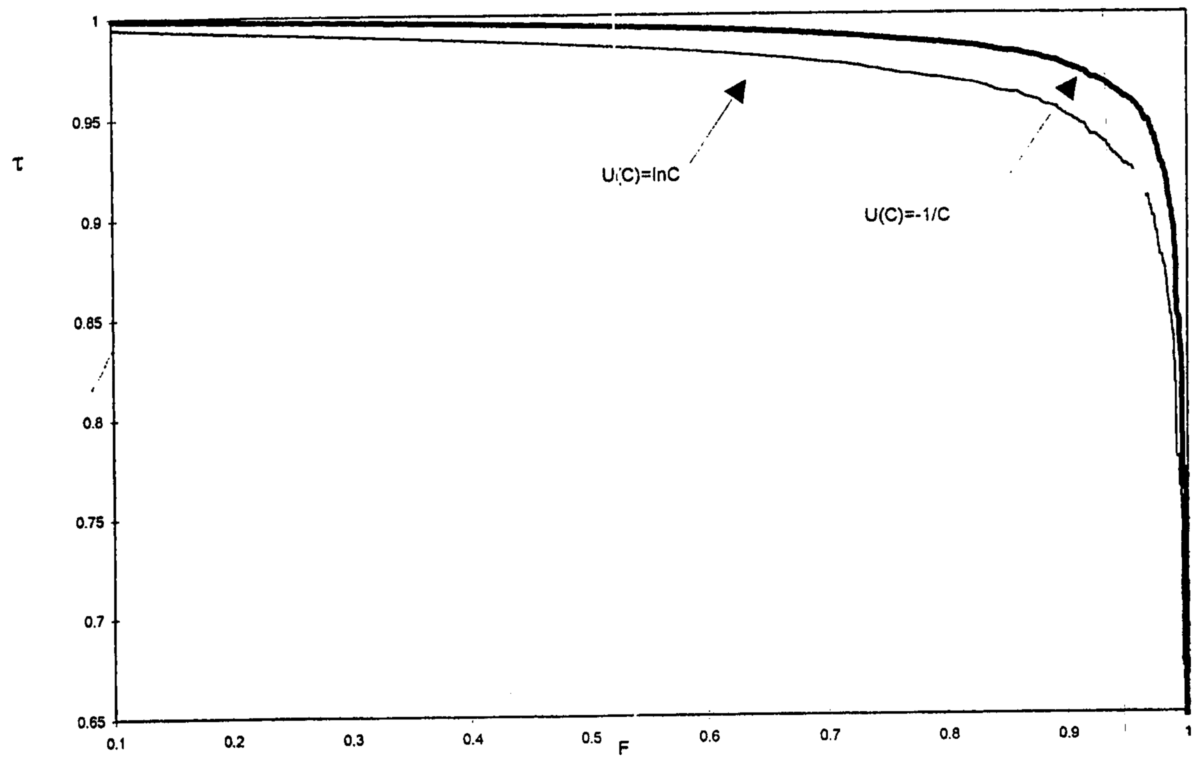
[15] The use of the approximation must be restricted to high income levels, since it is true only for low levels of social marginal utility. Tuomala (1984a, p.364) shows the same approximation for the case of a utilitarian social planner and a convex utility of consumption.

different in these two cases, while leaving constant the elasticity of substitution (=1).

Figure 7 shows the results of increasing income effects. The results are similar to those of Tuomala (1984a, p. 362), but here the reason that changes the shape of optimal taxes is the existence of income effects. Since in Tuomala's example both income effects and the elasticity of substitution play a role, this figure calls for further research on the reasons that are driving the result.

[16] s equals V'/wU_C and s_w equals $\partial s/\partial w$.

Figure 7 - Changing the degree of income effects



Appendix B - The optimal shape and the lognormal distribution

In this appendix we characterize the optimal shape using key values of the lognormal distribution: the mode, the median and the mean. For simplicity, we adopt 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.

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 $\sigma \approx 0.8$ marginal taxes decline until the median and since then they rise (i.e., the minimum marginal tax is at the median). If $\sigma > (<) 0.8$ 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 $\sigma \approx 2$ marginal taxes decline until the mean and since then they rise (i.e., the minimum marginal tax is at the mean). If $\sigma > (<) 2$ 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.

Bibliography

[1] Atkinson A. and J. Stiglitz, *Public Economics.*

[2] Atkinson A. (1973), "How progressive should income tax be?" in: M. Parkin and A.R. Nobay, eds., *Essays in modern economics* (Logman, London).

[3] Bickel P.J. (1977), *Mathematical Statistics - Basic Ideas and Selected Topics*, editor: E.L. Lehmann.

[4] Burkhauser R., A. Crews, M. Daly and S. Jenkins (1996), "Where in the world is the middle class? A cross-national comparison of the vanishing middle class using Kernel density estimates", presented at the N.B.E.R. Summer Institute, july 1996.

[5] Diamond P. (1996), "Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates", presented at the Conference Twenty-five Years of the New Public Economics, Jerusalem, january.

[6] Kanbur R. and M. Tuomala (1994), "Inherent inequality and the optimal graduation of marginal tax rates", *Scandinavian Journal of Economics*, 96 (2), 275-282.

[7] Lancaster T. (1990), *The econometric analysis of transition data*, Econometric Society Monographs no 17, Cambridge University Press.

[8] Mirrlees J. (1971), "An exploration in the theory of optimal income taxation", *Review of Economic Studies*, 38, 135-208.

[9] Myles G. (1996), *Public Economics*, Cambridge University Press.

[10] Sahota G. (1978), "Theories of personal income distribution: a survey", *Journal of Economic Literature*, 16, 1-55.

[11] Sadka E. (1976a), "On income distribution, incentive effects and optimal income taxation", *Review of Economic Studies*, 43, 261-268.

[12] Sadka E. (1976b), "On progressive income taxation", *American Economic Review*, 66, 931-935.

[13] Seade J. (1977), "On the shape of optimal tax schedules", *Journal of Public Economics*, 7, 203-236.

[14] Slemrod J., S. Yitzhaki, J. Mayshar and M. Lundholm (1994), "The optimal two-bracket linear income tax", *Journal of Public Economics*, 53 (2): 269-290.

[15] Stern N. (1976), "On the specification of models of optimum income taxation", *Journal of Public Economics*, 6, 123-162.

[16] Tuomala M. (1984a), "On the optimal income taxation - some further numerical results", *Journal of Public Economics* 23, 351-366.

[17] Tuomala M. (1984b), "Optimal degree of progressivity under income uncertainty', *Scandinavian Journal of Economics* 80 (2), 184-193.

[18] Varian H. (1980), "Redistributive taxation as social insurance", *Journal of Public Economics* 14, 49-68.

רשימת המאמרים בסדרה

91.01	R. Melnick and Y. Golan - Measurement of Business Fluctuations in Israel.
91.02	י. ארטשטיין, צ. זוסמן - דינאמיקה של עליות שכר בישראל: כוחות שוק והשוואות בינענפיות.
91.03	M. Sokoler - Seigniorage and Real Rates of Return in a Banking Economy.
91.04	E.K. Offenbacher - Tax Smoothing and Tests of Ricardian Equivalence Israel 1961-1988.
91.05	ג. עופר, ק. פלוג, נ. (קלינר) קסיר, - קליטה בתעסוקה של עולי בריה"מ בשנת 1990 והלאה: היבטים של שמירה והחלפת משלחי יד.
91.06	צ. זוסמן, ד. זכאי, - פערים בין בכירים וזוטרים ומשברים במערכת ציבורית: שכר הרופאים בשנים 1974 עד 1990.
91.07	M. Beenstock, Y. Lavi and S. Ribon The Supply and Demand for Exports in Israel.
91.08	R. Ablin - The Current Recession and Steps Required for Sustained Sustained Recovery and Growth.
91.09	צ. הרקוביץ, ל. (רובין) מרידור - ההשלכות המקרו-כלכליות של עלייה המונית לישראל.
91.10	M. Beenstock - Business Sector Production in the Short and Long Run in Israel: A Cointegrated Analysis.
91.11	א. ארנון, ר. עמיחי, - ההפרטה וגבולותיה.
91.12	ק. פלוג, נ. קסיר (קלינר) - עלות העבודה בתעשייה הישראלית.
91.13	A. Marom - The Black-Market Dollar Premium: The Case of Israel.
91.14	A. Bar-Ilan and A. Levy - Endogenous and Exoqenous Restrictions on Search for Employment.
91.15	M. Beentstock and S. Ribon- The Market for Labor in Israel.

91.16 ד. אלקיים, - השפעת המדיניות המוניטרית על פער הריביות במגזר השקלי הלא צמוד 1986 עד 1990.

92.01 מ. דהון, - בחינת מדד הדחף הפיסקאלי של ה-IMF עבור המשק הישראלי לשנים 1964 עד 1990.

92.02 O. Bar Efrat - Interest Rate Determination and Liberalization of International Capital Movement: Israel 1973 - 1990.

92.03 Z. Sussman and D.Zakai - Wage Gaps between Senior and Junior Physicians and Crises in Public Health in Israel, 1974-1990.

92.04 צ. ויס, ע. לויתן - התפתחות תשלומי העברה בישראל, 1965 עד 1989.

92.05 O. L iviatan - The Impact of Real Shocks on Fiscal Redistribution and Their Long-Term Aftermath.

92.06 A. Bregman, M. Fuss and H. Regev - The Production and Cost Structure of the Israeli Industry: Evidence from Individual Firm Data.

92.07 M. Beenstock, Y. Lavi and A. Offenbacher - A Macroeconometric-Model for Israel 1962 -1990:A Market Equilibrium Approach to ggregate Demand and Supply.

92.08 ס. ריבון, - מודל חודשי לשוק הכסף.

92.09 R. Melnick - Financial Services, Cointegration and the Demand for Money in Israel.

92.10 מ. ברון, - העליות לארץ והשפעתן על הפסיפס הדמוגרפי של האוכלוסייה והן ההון האנושי.

92.11 ד. זינגר, - גורמים הקובעים את ההסתברות של פירמות להיסגר.

92.12 R. Melnick Forecasting Short-Run Business Fluctuations in Israel

92.13 K. Flug, N. Kasir and G. Ofer - The Absorption of Soviet Immigrants in to the Labor Market from 1990 Onwards: Aspects of Occupational Substitution and Retention.

92.14 א. ארנון, ת. פרשטמן, - הפרטת מונופולים טבעיים: הריצה אחר הבלתי מוכח.

B. Eden - How to Subsidize Education and Achieve Voluntary Integration: An Analysis of Voucher Systems. 93.01

א. ברגמן, א. מרום, - גורמי צמיחה בסקטור העסקי בישראל(1958 עד 1988). 93.02

מ. דהן, - צמיחה כלכלית תחת איום ביטחוני. 93.03

ק. פלוג, נ. (קלינר)קסיר - קליטה בתעסוקה של עולי חבר המדינות - הטווח הקצר. 93.04

מ. דהן, - האם קיימת יריבות בין שיוון בחלוקת ההכנסות להתפתחות כלכלית? המקרה של ישראל. 93.05

צ. הרקוביץ, ל. מרידור - ההשלכות המקרו-כלכליות של עלייה המונית לישראל: עדכון ובחינה מחודשת. 93.06

A.Arnon, D. Gottlieb - An Economic Analysis of the Palestinian Economy: The West Bank and Gaza, 1968-1991. 93.07

צ. הרקוביץ, ל. מרידור, נ. קנטור - הגירה וצמיחה בתנאים של ניידות הון בלתי משוכללת: גל העלייה לישראל בראשית שנות התשעים. 93.08

K. Flug, N. Kasir - The Absorption in the Labor Market of Immigrants from the CIS - the Short Run. 93.09

R. Ablin - Exchange Rate Systems, Incomes Policy and Stabilization Some Short and Long-Run Considerations. 94.01

B.Eden - The Adjustment of Prices to Monetary Shocks When Trade is Uncertain and Sequential. 94.02

מ. ברון, - התחזית הדמוגרפית ולקחיה. 94.03

K. Flug, Z. Hercowitz and A. Levi - A Small -Open-Economy Analysis of Migration. 94.04

R. Melnick and E. Yashiv - The Macroeconomic Effects of Financial Innovation: The Case of Israel. 94.05

צ. הרקוביץ, מ. סטרבצ'ינסקי, - מדיניות חוב ציבורי בישראל 94.06

א. בלס, - חוזים כחסמי כניסה בשיווק דלק לתחנות תילדוק: בחינת החלטת הממונה על הגבלים עיסקיים לפיה מערכת ההסדרים הקיימת היא בגדר הסדר כובל. 94.07

מ. דהן, - צמיחה כלכלית, פעילות בלתי חוקית והתחלקות הכנסות. 94.08

95.09	ז. שיפר - על סעיף הדיור בבעלות הדיירים במדד המחירים לצרכן.
95.10	א. בן בסט - משטר שער החליפין, המדיניות המוניטרית ויעד האינפלציה.
95.11	A. Arnon and A. Spivak - Monetary Integration Between the Israeli, Jordanian and Palestinian Economies.
95.12	נ. לויתן - המשבר במכסיקו - פירושים ולקחים.
95.13	ר. כהן, ב. סורני - איתור יעד ביניים למדיניות מוניטרית בישראל 1988-1994. א. בלס - ביצועי קופות הגמל ומבנה שוק: 1987-1994.
96.02	ח. בר, ע. יושע, י. יפה - פיקוח על בנקים, בורסה ובעלי עניין: עדויות על פירמות בישראל.
96.03	A. Blass and R. S. Grossman - A Harmful Guarantee? The 1983 Israel Bank Shares Crisis Revisited
96.04	Z. Sussman and D. Zakai - The Decentralization of Collective Bargaining and Changes in the Compensation Structure in Israeli`s Public Sector.
96.05	M. Strawczynski - Precautionary Savings and The Demand for Annuities.
96.06	מ. דהן, צ. הרקוביץ - מדיניות פיסקאלית והחיסכון הלאומי במשק פתוח.
96.07	י. גיברה, ד. צידון - יעילותם של המכרז המוניטרי וההלוואה במכסה ככלים למדיניות מוניטרית.
96.08	מ. דהן א. בן-פורת - גל העלייה ואי-השוויון בהתחלקות ההכנסות.
96.09	Y. Lavi and A. Spivak - The Impact of Pension Schemes on Saving in Israel: Empirical Analysis.
96.10	M. Strawczynski - Social Insurance and The Optimum Piecewise Linear Income Tax.
96.11	א. בלס, י. גיברה - הודעות מוניטריות ותגובת שוק ההון: ההשפעה הראשונית על המק"ם והאג"ח וההשלכות למנגנון התמסורת.
96.12	ז. שיפר - פיצול (Stripping) של אגרות חוב ממשלתיות צמודות מדד לפי תוכן ויצירת אגרות חוב צמודות למחירי הדיור

ד. עיני - בחינה מחודשת של הרכב סל המטבעות האופטימלי לישראל. 97.01

M. Dahan and M. Strawczynski - The Optimal Non-Linear Income Tax. 97.02