Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates: Comment

By MOMI DAHAN AND MICHEL STRAWCZYNSKI*

In a recent paper, Peter A. Diamond (1998) reopens the question of the optimal shape of marginal tax rates at high levels of income in the framework of the classical model of skills (James A. Mirrlees, 1971). As well known, this widely used model in the Public-Economics literature does not prescribe a clear-cut pattern for optimal marginal taxes, by calling for optimal marginal rates that lie between zero (for the top and bottom of the ability's scale)[1] and one. Consequently, it has been the practice to obtain the optimum tax shape by running simulations, which are based on different assumptions—as described below.

Until recently, a vast quantity of works seemed to generate a strong case supporting declining marginal tax rates at high levels of income[2] (Table 1). By contrast, Diamond's examples of a U-shaped optimal pattern imply rising marginal tax rates at high income levels, a finding that is in line with most actual income tax systems.

In this note we replace the assumption of linear utility of consumption made by Diamond (1998) to logarithmic utility of consumption. We present simulations based on these two functional forms that show that optimal income tax rates may decline or rise at high levels of income. The assumed form of utility of consumption is solely responsible for the shift from upward to downward sloping of the optimal tax structure at high levels of income.

This paper is organized as follows. Section I extends the first-order condition (FOC) for optimum marginal income tax rates for the case of concave utility of consumption. Section II presents simulations of declining and rising optimal tax rates at high levels of income. Section III concludes the paper.

I. The Optimum Shape with a Concave Utility of Consumption

Assume the following utility function:

$$u = U(C) + V(1 - L) \tag{1}$$

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 that V is concave, and we extend in this section Diamond's analysis to the case of concave utility of consumption that implies the presence of income effects.

The problem consists of maximizing a social utility function equal to the integral of $G(u)$, with skills distributed by $f(w)$, taking into account the budget constraints at the individual and macro level (the Lagrange multiplier of the latter being denoted by γ), the self-selection constraint, and the first-order condition at the

* Research Department, Bank of Israel, Kaplan Street, Kiriat Hamemshala, Jerusalem, 91007 Israel (e-mail: momid@bankisrael.gov.il; michels@bankisrael.gov.il). We thank Peter Diamond, David Frankel, Alejandro Gaviria, Giora Hanoch, Zvi Hercowitz, Laurence Kotlikoff, James Mirrlees, James Poterba, Efraim Sadka, Eytan Sheshinski, Shlomo Yitzhaki, and participants in seminars at the Bank of Israel, Boston University, the Hebrew University of Jerusalem, Tel Aviv University, and the Massachusetts Institute of Technology for very helpful comments. We benefited from useful suggestions by two anonymous referees. We thank Irina Blits for excellent research assistance.

[1] With a finite maximum for the skill distribution, the optimal marginal tax rate at the income level of the top skill is zero (Efraim Sadka, 1976; Jesus K. 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' pioneering simulations, optimal marginal taxes decline with income. Since the shape was close to linearity, this point was stressed neither by Mirrlees nor by other authors citing his work. A more recent paper that uses also a lognormal distribution and shows a declining pattern of marginal two-bracket linear tax rates is Joel Slemrod et al. (1994). Note that these results were obtained under the assumption that income is certain. With income uncertainty, simulations show that optimal taxes rise with income. Hal R. Varian (1980) and Strawczynski (1998) provide examples where differences in income are due to "luck." Matti Tuomala (1984b) provides an example that introduces labor income uncertainty to the classical model of skills.

TABLE 1—THE OPTIMUM SHAPE ACCORDING TO THE LITERATURE

	Optimum schedule (marginal taxes)			
$F(w)$	Mirrlees (1971)	Atkinson (1973)	Tuomala (1984a)	Kanbur and 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

Notes: All the simulations assume a lognormal distribution with a mean skill of 0.4 ($\mu = -1$). 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. The simulations shown in the table assume a utilitarian social planner except Anthony B. Atkinson (1973), who assumes an 'inequality aversion' coefficient of 2.

individual level. We obtain the following expression for the FOC of the optimum nonlinear marginal tax:[3]

$$(2)\quad \frac{\tau}{1-\tau} = \left[\frac{\varepsilon}{w}\right][U_C] \times \left[\frac{\int_w^{w_H}\left[\frac{\gamma}{U_C} - G_u\right]f(w)\,dw}{\gamma(1-F(w))}\right] \times \left[\frac{(1-F(w))}{f(w)}\right],$$

$$\varepsilon = 1 - \frac{V_{LL}L}{V_L}, \qquad f \equiv \frac{dF}{dw}$$

W_H—top skill.

The first and second terms in brackets represent the standard efficiency and income effects. The third term is the inequality aversion effect. If we assume a decreasing social marginal utility ($G_{uu} < 0$), 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 the ratio of individuals above a particular income level, $1 - F$, to the individuals in that particular income level itself, f.[4] A high marginal tax rate at a particular income level distorts the decision for that income level, but this new higher marginal tax acts as a lump-sum tax on higher income levels. This is so since the decision at the margin is not affected by marginal tax rates in previous brackets. The higher $(1 - F)$ relative to f, the higher is the quantity of individuals that are paying lump-sum taxes, and consequently the higher is the optimum marginal tax rate.

Figure 1 shows the "distribution effect" for several well-known income distributions (uniform, exponential, Pareto, and lognormal). The aim of that figure is to show the shape (declining or rising) of income tax rates implied by the "distribution effect" only. The chosen parameters of the distributions are based on previous studies.[5]

Equation (2) provides the basis for simulating all different optimal shapes shown in the literature—varying the assumptions on the different components of the model. For

[3] A formal analysis is presented in an Appendix available on request from the authors. We assume that government intervention is purely redistributive. As explained by Joseph E. Stiglitz (1987 p. 1008), for the case of a separable utility function ($u_{C(1-L)} = 0$) and linear utility of consumption, ε is the compensated elasticity of labor supply.

[4] Note that this term is equal to 1 over the hazard rate.

[5] The parameters for Pareto and lognormal distributions were taken from David R. Feenberg and James M. Poterba (1993) and Ravi Kanbur and Tuomala (1994), respectively.

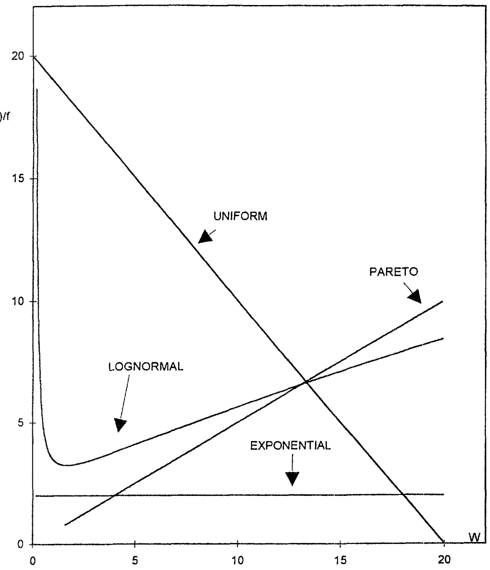

FIGURE 1. DISTRIBUTION EFFECT

Notes: Parameters: Uniform—from 0 to 20.
Lognormal—$\mu = 1 \quad \sigma = 1$.
Exponential—$\lambda = \frac{1}{2}$.
Pareto—$x_0 = 1.5 \quad \theta = 2$.

the sake of comparison with the previous literature, we will concentrate on the case of logarithmic utility of leisure, which constitutes one of Diamond's (1998 p. 90) examples with increasing optimal tax rates, and it was also the benchmark assumption of Mirrlees, who obtained declining tax rates at high levels of income.

II. Rising or Declining Optimal Rates at High Levels of Income?

We assume logarithmic utility of leisure, a Pareto distribution of skills, and an inequality-averse social planner. The efficiency effect [the first term in equation (2)] in this case equals $(1 - \tau)U_C$. The inexistence of income effects implies that $U_C = 1$, which means that the structure of marginal income tax rates depends on the inequality aversion and distribution effects in equation (2). An inequality-averse social planner means that the third term in equation (2) increases with income at high levels of income since G_{uu} is negative. Using a Pareto distribution implies that the distribution effect increases all over the range.[6] Therefore optimal marginal income tax rates unequivocally rise with income at high levels of income. This result holds also with a lognormal distribution of earnings[7] which is frequently used in the optimal income tax literature and has empirical support.[8]

However, the optimal structure of marginal income tax rates at high levels of income is unclear once we assume concave utility of consumption. As can be shown in equation (2), the presence of U_C drives the optimal marginal income tax rates down due to the standard income effects. A concave utility of consumption implies that income effects are weaker for rich individuals, which calls for lower taxes at high levels of income. But at the same time it works to raise tax rates because of its impact on the inequality aversion effect. Therefore we are bound to use simulations to determine the shape of marginal tax rates in that case.

In the simulations below we use logarithmic and linear utility of consumption, logarithmic utility of leisure, an inequality-averse social planner, and Pareto and lognormal distributions of skills. Figures 2A and B show the simulated shape of marginal income tax rates for four different cases.[9] Figure 2A compares two cases that correspond to rows 1 and 2 in Table 2,

[6] A Pareto distribution defined as $f = \alpha k^{\alpha} w^{-(1+\alpha)}$ implies that $(1 - F) = (k/w)^{\alpha}$ and $(1 - F)/f = w/\alpha$. Thus the "distribution effect" increases as w increases.

[7] As implied by the analysis presented in Tony Lancaster (1990 p. 47), in this case the "distribution effect" is U-shaped.

[8] There is evidence supporting a Pareto distribution at the upper tail of the distribution (Feenberg and Poterba, 1993) which is the main focus of this Note. However, for the whole distribution, J. Aitchison and J. A. C. Brown (1957), H. F. Lydall (1968), and Yoram Weiss (1972) argued that the lognormal distribution fits fairly well the distribution of earnings in homogeneous sectors of the labor market. A detailed discussion may be found in A. Frank Cowell (1995).

[9] The simulation is based on an approximation that assumes that $G_u \cong 0$ at high levels of income, i.e., a concave social utility function (a similar approximation for $U_C \cong 0$ and a utilitarian social planner is presented by Tuomala, 1984 p. 364). We normalize optimal taxes to the level in Mirrlees: 19 percent for $F(w) = 0.9$. Note that the shapes of the tax structure at high levels of income in cases that use linear utility of consumption are known without simulations. We provide simulations for those cases just for comparison.

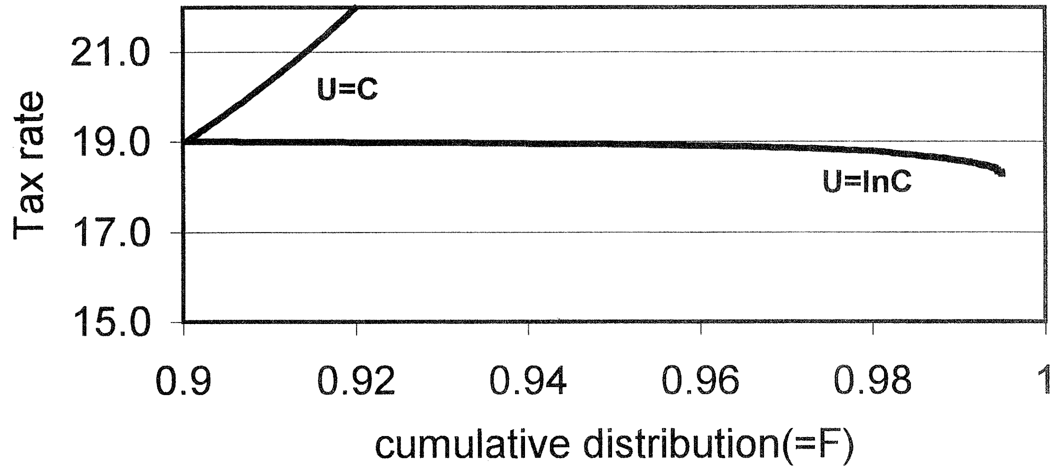

FIGURE 2A. OPTIMAL TAXES AT HIGH LEVELS OF SKILL
THE CASE OF PARETO DISTRIBUTION
(ALFA = 1.5 AS IN DIAMOND, 1998)

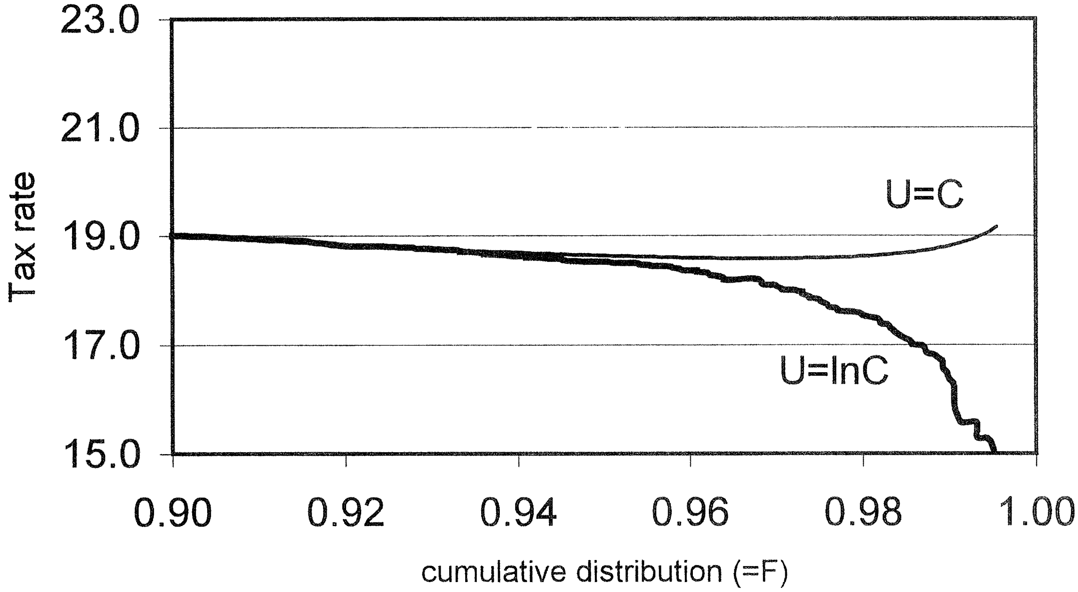

FIGURE 2B. OPTIMAL TAXES AT HIGH LEVELS OF SKILL
THE CASE OF LOGNORMAL DISTRIBUTION
(MEAN = −1 AND STANDARD DEVIATION = 0.39 AS IN MIRRLEES, 1971)

respectively. The solid line corresponds to Diamond's results while the dash line represents our first example, that replaces linear utility by logarithmic utility of consumption. The shift from linear utility of consumption (that implies no income effects) to logarithmic utility of consumption (that implies the presence of income effects) reverses the result from rising to declining marginal income tax rates.[10] When income effects are not present, the optimal tax schedule

[10] The logarithmic utility of consumption implies weak income effects relative to other common functions used in

TABLE 2—SUMMARY OF RESULTS[a]

	Utility of leisure	Utility of consumption	Social planer	Distribution of skills	Optimal tax schedule at high levels of income
Diamond (1998)	Logarithmic	Linear	Inequality averse	Pareto	Rising
Our example—Figure 2A	Logarithmic	Logarithmic	Inequality averse	Pareto	Declining
Our example—Figure 2B	Logarithmic	Linear	Inequality averse	Lognormal	Rising
Mirrlees (1971 p. 203)	Logarithmic	Logarithmic	Inequality averse	Lognormal	Declining

[a] The results shown in column 5 are the same with exponential distribution of skills keeping the same all other assumptions.

hits the ceiling of Figure 2A relatively quickly because the asymptotic tax rate goes to one.[11]

Figure 2B also makes clear that the assumed utility of consumption is critical by comparing Mirrlees' example to our second example (these two examples correspond to rows 3 and 4 in Table 2, respectively). Again the only difference between these two examples is the assumed utility of consumption. While Mirrlees' example shows declining optimal marginal income tax rates, our example with no income effects produces rising income tax rates at high levels of income.

The message that arises from these simulations under the assumptions summarized in Table 2 is that the structure of the optimal marginal income tax rates at high levels of income is sensitive to the assumed form of the utility of consumption. It also helps to put a bridge between the declining shape of Mirrlees' example and the rising shape of Diamond's example.

The assumed utility of consumption is important also for the calculation of asymptotic marginal income tax rates. The asymptotic tax rate converges to a constant under the following assumptions: constant elasticity of labor, a Pareto distribution, linear utility of consumption, and an inequality-averse social planner. In contrast, the asymptotic tax rate is unclear once we replace linear by concave utility of consumption, since the second term in equation (2) goes to zero and the fourth term goes to infinity.

the literature. For example, Tuomala (1984a) and Kanbur and Tuomala (1994) use $U = -1/C$.

[11] In this case the asymptotic tax rate equals to one since the left-hand side of equation (2) goes to infinity; the relevant equation in this case is: $\tau/(1 - \tau)^2 = aw$, where a is a constant.

III. Conclusions

This Note shows that income effects play an important role in determining the optimal shape of income tax structure. First, it shows that the result of rising marginal income tax rates presented by Diamond (1998) is sensitive to the assumed utility of consumption. Replacing linear by logarithmic utility of consumption that implies the presence of income effects, produces an opposite result of declining marginal tax rates at high levels of income in the simulations.

Second, it shows that the assumed utility of consumption plays a critical role in Mirrlees' example (1971). The income tax structure is upward (rather than downward) sloping at high levels of income using linear utility of consumption that implies no income effects.

REFERENCES

Aitchison, J. and Brown, J. A. C. *The lognormal distribution*. London: Cambridge University Press, 1957.

Atkinson, Anthony B. "How Progressive Should Income Tax Be?" in M. Parkin and A. R. Nobay, eds., *Essays in modern economics*. London: Longman, 1973, pp. 90–109.

Cowell, A. Frank. *Measuring inequality*, LSE Handbooks in Economics. London: Prentice Hall, 1995.

Diamond, Peter A. "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 Tuomala, Matti. "Inherent Inequality and the Optimal Graduation of Marginal Tax Rates." *Scandinavian Journal of Economics*, June 1994, *96*(2), pp. 275–82.

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

Lydall, H. F. *The structure of earnings*. Oxford: Oxford University Press, 1968.

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

Sadka, Efraim. "On Income Distribution, Incentive Effects and Optimal Income Taxation." *Review of Economic Studies*, June 1976, *43*(2), pp. 261–67.

Seade, Jesus K. "On the Shape of Optimal Tax Schedules." *Journal of Public Economics*, April 1977, *7*(2), pp. 203–35.

Slemrod, Joel; Yitzhaki, Shlomo; Mayshar, Joram and Lundholm, Michael. "The Optimal Two-Bracket Linear Income Tax." *Journal of Public Economics*, February 1994, *53*(2), pp. 269–90.

Stiglitz, Joseph E. "Pareto Efficient and Optimal Taxation and the New New Welfare Economics," in Alan J. Auerbach and Martin Feldstein, eds., *Handbook of public economics*. New York: Elsevier Science, 1987, pp. 991–1042.

Strawczynski, Michel. "Social Insurance and the Optimum Piecewise Linear Income Tax." *Journal of Public Economics*, September 1998, *69*(3), pp. 371–88.

Tuomala, Matti. "On the Optimal Income Taxation: Some Further Numerical Results." *Journal of Public Economics*, April 1984a, *23*(3), pp. 351–66.

______. "Optimal Degree of Progressivity Under Income Uncertainty." *Scandinavian Journal of Economics*, June 1984b, *86*(2), pp. 184–93.

Varian, Hal R. "Redistributive Taxation as Social Insurance." *Journal of Public Economics*, August 1980, *14*(1), pp. 49–68.

Weiss, Yoram. "The Risk Element in Occupational and Educational Choices." *Journal of Political Economy*, November–December 1972, *80*(6), pp. 1203–13.

LINKED CITATIONS

You have printed the following article:

Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates: Comment
Momi Dahan; Michel Strawczynski
The American Economic Review, Vol. 90, No. 3. (Jun., 2000), pp. 681-686.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28200006%2990%3A3%3C681%3AOITAEW%3E2.0.CO%3B2-D

[Footnotes]

[1] **On Income Distribution, Incentive Effects and Optimal Income Taxation**
Efraim Sadka
The Review of Economic Studies, Vol. 43, No. 2. (Jun., 1976), pp. 261-267.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28197606%2943%3A2%3C261%3AOIDIEA%3E2.0.CO%3B2-D

[8] **The Risk Element in Occupational and Educational Choices**
Yoram Weiss
The Journal of Political Economy, Vol. 80, No. 6. (Nov. - Dec., 1972), pp. 1203-1213.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28197211%2F12%2980%3A6%3C1203%3ATREIOA%3E2.0.CO%3B2-C

References

Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates
Peter A. Diamond
The American Economic Review, Vol. 88, No. 1. (Mar., 1998), pp. 83-95.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28199803%2988%3A1%3C83%3AOITAEW%3E2.0.CO%3B2-8

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

An Exploration in the Theory of Optimum Income Taxation
J. A. Mirrlees
The Review of Economic Studies, Vol. 38, No. 2. (Apr., 1971), pp. 175-208.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28197104%2938%3A2%3C175%3AAEITTO%3E2.0.CO%3B2-V

On Income Distribution, Incentive Effects and Optimal Income Taxation
Efraim Sadka
The Review of Economic Studies, Vol. 43, No. 2. (Jun., 1976), pp. 261-267.
Stable URL:
http://links.jstor.org/sici?sici=0034-6527%28197606%2943%3A2%3C261%3AOIDIEA%3E2.0.CO%3B2-D

The Risk Element in Occupational and Educational Choices
Yoram Weiss
The Journal of Political Economy, Vol. 80, No. 6. (Nov. - Dec., 1972), pp. 1203-1213.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28197211%2F12%2980%3A6%3C1203%3ATREIOA%3E2.0.CO%3B2-C

NOTE: *The reference numbering from the original has been maintained in this citation list.*