INCOME TAXATION WITH INTERGENERATIONAL MOBILITY: CAN HIGHER INEQUALITY LEAD TO LESS PROGRESSION

by

Tomer Blumkin and Efraim Sadka

Working Paper No.8-2003

April, 2003

The Foerder Institute for Economic Research
and
The Sackler Institute of Economic Studies

Income Taxation with Intergenerational Mobility: Can Higher Inequality Lead to Less Progression?

Tomer Blumkin* **Efraim Sadka****

April 2003

Abstract

We extend the optimal tax literature by assuming that earnings abilities are no longer purely innate abilities, but rather partly acquired through investments in human capital. This assumption adds endogenous mobility to the primarily static optimal tax models. The implications for the optimal redistribution policy are analyzed.

JEL Classification: H2, D6

Key Words: Mobility, Optimal Taxation, Re-distribution, Education, Inequality

* Department of Economics, Ben-Gurion University, Beer-Sheba 84105, Israel. E-mail address: tomerblu@bgumail.bgu.ac.il.

** The Eitan Berglas School of Economics, Tel Aviv University, Tel-Aviv 69978, Isreal. E-mail address: sadka@post.tau.ac.il

1. Introduction

The growth in earnings inequality over the last two decades is often attributed to the dramatic rise in the returns to education. For instance, according to the US bureau of the Census, for mid-career full-time year-round male workers, mean earnings of college graduate to high-school graduate ratio rose from 1.41 (in 1980) to 1.73 (in 1997). Family characteristics, such as parental wealth and educational attainment, are thought to bear a significant impact on the children's return to education [For some empirical evidence, see Rosenzweig and Wolpin (1994) and Altonji and Dunn (1996)]. Children coming from high-income families tend to obtain more education, and hence find themselves more often in the upper percentiles of the earnings distribution. Empirical studies [see, e.g., Solon (1992)] attest to the significant intergenerational correlation of incomes, which could well exceed 50 percent.

The rise in earnings inequality renders the debate on re-distribution policy, notably the one concerning the merits and pitfalls of progressive labor income taxation, ever so relevant. Strikingly though, in light of the evidence mentioned above, the theoretical literature on optimal income taxation, following the seminal contribution of Mirrlees (1971), has primarily focused on a 'static' model, where only intra-generational inequality matters.[1] Issues of social mobility, intergenerational inequality and changes in inequality across generations have not attracted due attention.

The purpose of this paper is to present a simple framework with just the key ingredients, so as to be able to shed some light on the linkage between intergenerational mobility and the design of the optimal labor income tax. Specifically, we allow for earning ability to be partly innate and partly acquired through education. We then address the issue of how a rise in mobility or in earnings inequality affects the optimal degree of progression of the labor income tax.

[1] Many dynamic models address the issue of optimal taxation, but maintain the single cohort assumption [see e.g., Brito et al (1991)].

The structure of the paper is as follows. The coming section presents the model. In section 3 we conduct comparative statics exercises. Section 4 concludes.

2. The Model

Consider the following two-period simple OLG framework.[2] The old cohort lives for one period and each old individual gives birth to one offspring that lives for two periods. The old cohort comprises of two types of individuals, which differ in innate earning ability. Agents with high ability earn an hourly wage rate of $\overline{w}$ and constitute a fraction γ of the population; whereas agents with low ability earn the wage rate of $\underline{w}$, $0<\underline{w}<\overline{w}$, and make up a fraction $1-\gamma$ of the population (which is normalized to unity for each cohort, with no loss in generality). For simplicity, we will set $\gamma=1/2$.

During the first period each parent takes two economic decisions. One concerns the allocation of time between leisure and labor. The other concerns the amount of education to acquire for his/her offspring. During the first period the offspring is passive and makes no economic decisions. In the second period the young cohort joins the labor market. We further simplify by assuming that labor supply of the two cohorts is perfectly inelastic.

We further assume that innate ability is (positively but imperfectly) correlated across time. That is, the ability of each descendant depends on family background, but also on the amount of education provided by the parent. Formally, denote by $p(e,w)$ the probability of upward mobility, that is the probability that a member of the young cohort will possess the high earning ability, $\overline{w}$, when he/she enters the labor market. This probability depends on the education level e and on the family background w (denoting the earning ability of the parent). We make the following natural assumptions:

$$\frac{\partial p(e,w)}{\partial e}>0; \frac{\partial^2 p(e,w)}{\partial e^2}<0; p(e,\overline{w})>p(e,\underline{w}) \text{ and } \frac{\partial p(e,\overline{w})}{\partial e}\geq\frac{\partial p(e,\underline{w})}{\partial e} \text{ for all } e.$$

Namely, the probability of upward mobility is rising with respect to education but at a diminishing rate and is increasing with respect to family background, captured by the parental earning ability. Moreover, the marginal contribution of education to upward mobility, namely $\partial p / \partial e$, is (weakly) rising with the family background w.

Acquiring education is costly and entails forgone consumption by the parents. The cost of education measured by forgone consumption is given by $d(e)=e$.

Parents are assumed to be altruistic (a la Barro).[3] The utility of each dynasty (parent cum offspring) of type w is given by:

$$(1) \qquad U(c_1, c_2, c_3, e, w) = u(c_1) + \beta\{p(e,w)c_3 + [1 - p(e,w)]c_2\} ,$$

where c_1, c_2 and c_3 denote, respectively, the consumption of the parent, the consumption of a low-ability and the consumption of a high ability descendant. The parameter $\beta > 0$ denotes the degree of altruism. We further assume that $u(\cdot)$ is increasing, concave and twice continuously differentiable.

Note that we choose a quasi-linear specification which significantly facilitates the analysis.[4] One particular implication of the special functional form is the fact that income effects are fully absorbed by the second generation level of consumption. This, however, does not limit the generality of the qualitative results we obtain, as the social planner is concerned with equity issues from a dynastical perspective.

There is a linear labor income tax, comprised of a flat tax rate, t, and uniform lump sum transfer, τ . We assume, as is usually the case in practice, that education costs

---

[2] A dynamic extension is presented in appendix C.

[3] Similar specifications of altruism are common in the recent literature stemming from the revived debate on inheritance and estate taxation [see e.g., Cremer and Pestieau (2001), Kopzcuk (2000) and Blumkin and Sadka (2002)].

[4] We would be able, for instance, to distinguish between the disincentive component and the re-distributive one in the comparative statics analysis conducted below.

are not tax deductible. The budget constraints faced by a typical dynasty of type *w* are given by:

(2) $c_1 \le (1-t)w - e - b + \tau$

(3) $c_2 \le (1-t)\underline{w} + b(1+r) + \tau$

(4) $c_3 \le (1-t)\overline{w} + b(1+r) + \tau$,

with *b* denoting bequest and *r* denoting the exogenously given interest rate.[5] We simplify the notation by letting *r*=0, with no loss in generality.

The first-order conditions (assuming an interior solution) are:

(5) $\dfrac{\partial u(c_1)}{\partial c_1} - \lambda_1 = 0$

(6) $\beta[1 - p(e,w)] - \lambda_2 = 0$

(7) $\beta p(e,w) - \lambda_3 = 0$

(8) $-\lambda_1 + \lambda_2 + \lambda_3 = 0$

(9) $\beta \dfrac{\partial p(e,w)}{\partial e}(1-t)(\overline{w} - \underline{w}) - \lambda_1 = 0$,

with λ_1, λ_2 and λ_3 denoting the *Lagrange* multipliers associated with constraints (2), (3) and (4), respectively.

Manipulating the first-order conditions yields the following two optimality conditions (with respect to consumption smoothing and investment in education):

(10) $\dfrac{\partial u(c_1)}{\partial c_1} = \beta$,

and

(11) $\dfrac{\partial p(e,w)}{\partial e}(1-t)(\overline{w} - \underline{w}) = 1.$

Equation (10) implies that the two types of parents have the same amount of consumption in the first period; namely, income differences are fully absorbed in the second period variation in consumption. Equation (11) implies that high-ability parents invest no less than low-ability parents in the education of their offspring.

Substituting the optimal solutions back into the direct utility function given in (1), one obtains the indirect utility function denoted by $V(w,t,\tau)$. We also denote by $e^*(w)$ the optimal investment in education of a parent with ability w.[6]

The government (social planner) determines the optimum tax rates so as to maximize some social welfare function, subject to a revenue constraint. We assume that the social preferences are represented by the following welfare measure:

(12) $$W = 1/2 \cdot W[V(\overline{w},\cdot)] + 1/2 \cdot W[V(\underline{w},\cdot)],$$

where $W'(\cdot) > 0$ and $W''(\cdot) < 0$. Namely, the welfare measure exhibits (strict) inequality-aversion.

The social planner is seeking to maximize the welfare measure in (12) by choosing the tax instruments, t and τ, subject to the following inter-temporal balanced budget constraint:[7]

(13)

$$t/2 \cdot \{[\overline{w} + p[e^*(\overline{w}), \overline{w}]\overline{w} + [1 - p[e^*(\overline{w}), \overline{w}]]\underline{w}\} + t/2 \cdot \{\underline{w} + p[e^*(\underline{w}), \underline{w}]\overline{w} + [1 - p[e^*(\underline{w}), \underline{w}]]\underline{w}\} \geq 2\tau$$

---

[5] This could be justified by the standard assumption of a small open economy.

[6] We suppress arguments to abbreviate notation throughout, whenever it causes no confusion.

[7] We assume that the tax parameters are fixed across the two periods. This seems a realistic assumption as it reflects the infrequent nature of tax reforms. Relaxing the assumption would not change our qualitative

This specification implicitly assumes that there are many parents of each type (but with an equal number), so that there is no macro risk.

We henceforth abbreviate notation and let all variables with an upper bar (lower bar, respectively) refer to high ability dynasties (low-ability dynasties, respectively). Particularly, we set $e^*(\overline{w}) \equiv \overline{e}$ and $e^*(\underline{w}) \equiv \underline{e}$.

Formulating the *Lagrangean*, one derives the first-order conditions for the two optimal tax instruments, t and τ, given, respectively, by:

(14)
$$-\overline{\theta}(\overline{\lambda_1}\overline{w} + \overline{\lambda_2}\underline{w} + \overline{\lambda_3}\overline{w})/2 - \underline{\theta}(\underline{\lambda_1}\underline{w} + \underline{\lambda_2}\underline{w} + \underline{\lambda_3}\overline{w})/2 + \mu[\overline{I}/2 + \underline{I}/2] + \mu/2 \cdot [t(\overline{w} - \underline{w})\frac{\partial p(\overline{e}, \overline{w})}{\partial e}\frac{\partial \overline{e}}{\partial t} + t(\overline{w} - \underline{w})\frac{\partial p(\underline{e}, \underline{w})}{\partial e}\frac{\partial \underline{e}}{\partial t}] = 0$$

(15)
$$\overline{\theta}(\overline{\lambda_1} + \overline{\lambda_2} + \overline{\lambda_3})/2 + \underline{\theta}(\underline{\lambda_1} + \underline{\lambda_2} + \underline{\lambda_3})/2 = 2\mu,$$

where, θ denotes the *lifetime* social marginal utility of income, I denotes *lifetime* expected income and is given by $I = w + p[e^*(w), w]\overline{w} + [[1 - p[e^*(w), w]]\underline{w}$, and μ denotes the *Lagrange* multiplier associated with the budget constraint in (13).

Substituting (15) into (14), some algebraic manipulations and re-arranging yields the following expression:

(16)
$$h(t, \overline{w}, \underline{w}) = -Cov(I, \theta) + \frac{t}{1-t} \cdot E(\theta)/2 \cdot [\frac{\partial \overline{e}}{\partial t} + \frac{\partial \underline{e}}{\partial t}] = 0,$$

with $E(\cdot)$ denoting the expectation operator and $Cov(\cdot)$ denoting the covariance operator.

Re-arranging the expression in (16) yields the following:

---

results. The tax rate for the first generation would be optimally set to unity by virtue of our assumptions (inelastic labor supply).

(17) $$\frac{t}{1-t} = \frac{Cov(I,\theta)}{E(\theta)/2\cdot[\frac{\partial \bar{e}}{\partial t}+\frac{\partial \underline{e}}{\partial t}]}$$

By virtue of the strict concavity of the social welfare measure, it follows that the numerator is negatively signed. It further follows from the concavity of $p(e,w)$ with respect to e, that the denominator is negatively signed by virtue of (11). Thus the optimal marginal tax rate, t, is positive, as expected.

3. The effects of mobility and initial inequality

We turn next to investigate the effects of mobility and initial inequality on the optimal tax-transfer redistribution policy. For this purpose, we simplify by specifying the functional form of the transition probability $p(e,w)$. We assume an additively separable form. Formally, we let the transition probability for an individual with high and low ability be given by $p(e,\bar{w}) = a + m + f(e)$ and $p(e,\underline{w}) = a - m + f(e)$, respectively, with $0 \le m \le a$ denoting the degree of mobility in the economy. A lower m reflects a higher degree of earning mobility in the economy. When m=0, family background bears no impact on the transition probability. This is the case of perfect mobility. Note crucially that we parameterize mobility in a wealth-neutral manner. Namely, any change in mobility (that is, in the parameter m) will neither directly nor indirectly affect the aggregate wealth in the economy.

3.1 Higher Mobility

Consider first the effect of a rise in the degree of earning mobility (a decrease in m) on the marginal tax rate, t. We examine the effect of a small decrease in the parameter m on the optimal t.

Formally we are seeking to sign the following expression:

(18) $$\left.\frac{\partial t}{\partial m}\right|_{h(t,m)=0} = -\frac{\partial h/\partial m}{\partial h/\partial t}$$

By virtue of the second order-condition of the social planner maximization, it follows that the denominator of the expression on the right-hand side of (18) is negatively signed. It follows that the sign of the expression on the left-hand side of (18) is determined by the sign of the numerator of the expression on the right-hand side of (18).

Differentiation of $h(\cdot)$ and some algebraic manipulations yield the following simplified expression:

(19) $$\frac{\partial h}{\partial m} = \frac{Cov(\theta, I) \cdot [\frac{\partial^2 \bar{e}}{\partial t \partial m} + \frac{\partial^2 \underline{e}}{\partial t \partial m}]}{[\frac{\partial \bar{e}}{\partial t} + \frac{\partial \underline{e}}{\partial t}]} + \frac{(\bar{I} - \underline{I}) \cdot [\bar{\theta} \cdot \frac{\partial \underline{\theta}}{\partial m} - \underline{\theta} \cdot \frac{\partial \bar{\theta}}{\partial m}]}{4 \cdot E(\theta)}$$

The first term on the right-hand side of (19) captures the disincentive component whereas the second term captures the re-distributive component. We turn to sign each one of the terms. Consider first the disincentive component.

By virtue of the separability of the transition probability, $p(e, w)$, with respect to *m*, it follows that the first term cancels out as its numerator is equal to zero.

Close inspection of the second term reveals that its sign is determined by the expression in brackets in the numerator. Dividing the expression by $\bar{\theta} \cdot \underline{\theta}$ (which does not change the sign) and using the fact that the *lifetime* social marginal utility of income is defined by $\theta = \frac{dW[V(w)]}{dV(w)}$, yields, after re-arrangement, the following expression:

(22) $$Sgn\left[\frac{\partial h}{\partial m}\right] = Sgn\left[\frac{-W''[V(\bar{w})]}{W'[V(\bar{w})]} \cdot \frac{\partial V(\bar{w})}{\partial m} - \frac{-W''[V(\underline{w})]}{W'[V(\underline{w})]} \cdot \frac{\partial V(\underline{w})}{\partial m}\right]$$

Employing the budget constraints (2)-(4) and (13), using the *envelope theorem*, it can be shown that (for details see appendix A):

$$(23) \qquad \frac{\partial V(\underline{w})}{\partial m} = -\beta \cdot (1-t) \cdot [\overline{w} - \underline{w}],$$

and

$$(24) \qquad \frac{\partial V(\overline{w})}{\partial m} = \beta \cdot (1-t) \cdot [\overline{w} - \underline{w}].$$

Hence, $Sgn\left[\frac{\partial h}{\partial m}\right]$>0.

In conclusion, a rise in the earning mobility (a decrease in *m*) yields unambiguously a lower marginal tax rate, hence a less progressive labor income tax, as one would expect.

3.2 Higher Earnings Inequality

Consider next the effect of a rise in earning inequality, captured by a mean preserving spread in the earning distribution, on the optimal marginal tax rate. Formally, let $\overline{w} = w + \Delta$ and $\underline{w} = w - \Delta$; where $\Delta > 0$. We examine the effect of a small rise in the parameter Δ on the optimal *t*.

Formally, we are seeking to sign the following expression:

$$(25) \qquad \left.\frac{\partial t}{\partial \Delta}\right|_{h(t,\Delta)=0} = -\frac{\partial h / \partial \Delta}{\partial h / \partial t}$$

Again, by virtue of the second order-condition of the social planner maximization, it follows that the sign of the expression on the left-hand side of (25) is determined by the sign of the numerator of the expression on the right-hand side of (25).

Differentiation of $h(\cdot)$ and some algebraic manipulations yield the following simplified expression:

(26) $$\frac{\partial h}{\partial \Delta}=\frac{Cov(\theta,I)\cdot[\frac{\partial^2 \bar{e}}{\partial t\partial\Delta}+\frac{\partial^2 \underline{e}}{\partial t\partial\Delta}]}{[\frac{\partial \bar{e}}{\partial t}+\frac{\partial \underline{e}}{\partial t}]}+\frac{(\bar{I}-\underline{I})\cdot[\bar{\theta}\cdot\frac{\partial\underline{\theta}}{\partial\Delta}-\underline{\theta}\cdot\frac{\partial\bar{\theta}}{\partial\Delta}]}{4\cdot E(\theta)}$$

The first term on the right-hand side of (26) captures the disincentive component whereas the second term captures the re-distributive component. We turn to sign each one of the terms. Consider first the disincentive component.

By virtue of (11) it follows that the denominator is negatively signed. It further follows that the $Cov(\cdot)$ is negative, by virtue of the strict concavity of the welfare function. We need to determine the sign of the expression in brackets in the numerator in order to sign the disincentive component.

Differentiating the expression in (11) with respect to *t* and re-arranging yields the following:

(27) $$\frac{\partial e}{\partial t}=\frac{\partial \mathrm{p(e,w)}/\partial e}{(1-t)\cdot\partial^2 p(e,w)/\partial e^2},$$

It then follows that,

(28) $$Sgn\left[\frac{\partial^2 e}{\partial t\partial\Delta}\right]=Sgn\left[\frac{\partial[-\frac{\partial^2 p(e,w)/\partial e^2}{\partial \mathrm{p(e,w)}/\partial e}]}{\partial\Delta}\right]$$

By virtue of (11) and the construction of the mean-preserving spread it follows that $\partial e/\partial\Delta>0$. Noting that the term in the numerator of the expression on the right-hand side of (28) is the familiar measure of absolute risk aversion, it follows that the sign of the expression is negative (positive) if the measure is decreasing (increasing) correspondingly. If *f*(*e*) is iso-elastic, for instance, the measure is decreasing.[8]

We turn next to determine the sign of the re-distributive term.

Re-iterating our derivation above yields, after re-arrangement, the following expression:

$$(29)\qquad Sgn\left[\frac{(\bar{I}-\underline{I})\cdot[\bar{\theta}\cdot\frac{\partial\underline{\theta}}{\partial\Delta}-\underline{\theta}\cdot\frac{\partial\bar{\theta}}{\partial\Delta}]}{4\cdot E(\theta)}\right]=Sgn\left[\frac{-W''[V(\bar{w})]}{W'[V(\bar{w})]}\cdot\frac{\partial V(\bar{w})}{\partial\Delta}-\frac{-W''[V(\underline{w})]}{W'[V(\underline{w})]}\cdot\frac{\partial V(\underline{w})}{\partial\Delta}\right]$$

Employing the budget constraints (2)-(4) and (13), using the *envelope theorem*, it can be shown that (for details see appendix A):

$$(30)\qquad \frac{\partial V(\underline{w})}{\partial\Delta}=\beta(1-t)\cdot[2\underline{p}-2]+2\beta\cdot t\left\{[\underline{p}-(1-\bar{p})]+\Delta\frac{\partial\bar{p}}{\partial\Delta}+\Delta\frac{\partial\underline{p}}{\partial\Delta}\right\},$$

and

$$(31)\qquad \frac{\partial V(\bar{w})}{\partial\Delta}=\beta(1-t)\cdot 2\bar{p}+2\beta\cdot t\left\{[\underline{p}-(1-\bar{p})]+\Delta\frac{\partial\bar{p}}{\partial\Delta}+\Delta\frac{\partial\underline{p}}{\partial\Delta}\right\}$$

Reformulating the expression on the right-hand side of (29) substituting (30) and (31) it follows:

$$(32)\qquad \begin{aligned}&\left[\frac{-W''[V(\bar{w})]}{W'[V(\bar{w})]}\cdot\frac{\partial V(\bar{w})}{\partial\Delta}-\frac{-W''[V(\underline{w})]}{W'[V(\underline{w})]}\cdot\frac{\partial V(\underline{w})}{\partial\Delta}\right]=\\&\left[\frac{-W''[V(\bar{w})]}{W'[V(\bar{w})]}+\frac{-W''[V(\underline{w})]}{W'[V(\underline{w})]}\right]\cdot\beta(1-t)+\left\{\left[\frac{-W''[V(\bar{w})]}{W'[V(\bar{w})]}\right]\cdot\bar{\Omega}-\left[\frac{-W''[V(\underline{w})]}{W'[V(\underline{w})]}\right]\cdot\underline{\Omega}\right\}\end{aligned}$$

Where,

[8] We assume this functional form in the numerical example below.

$$\underline{\Omega} = \beta(1-t)\cdot[2\underline{p}-1] + 2\beta\cdot t\left\{[\underline{p}-(1-\overline{p})] + \Delta\frac{\partial\overline{p}}{\partial\Delta} + \Delta\frac{\partial\underline{p}}{\partial\Delta}\right\}$$

$$\overline{\Omega} = \beta(1-t)\cdot[2\overline{p}-1] + 2\beta\cdot t\left\{[\underline{p}-(1-\overline{p})] + \Delta\frac{\partial\overline{p}}{\partial\Delta} + \Delta\frac{\partial\underline{p}}{\partial\Delta}\right\}$$

The first term on the right-hand side of (32) captures the short run effect; namely, the rise in earnings inequality deriving from the change in the skill distribution of the first generation (background effect). This term is positively signed and implies an upward adjustment in the marginal tax rate to accommodate the rise in inequality. Consider next the second term. This term captures the long run effect (the one associated with the change in the skill distribution of the second generation in our framework which affects investment decision of the parent generation). Suppose that investment in education is sufficiently intensive, such that $\underline{\Omega} > 0$ (A sufficient condition is: $\underline{p} > 1/2$). Suppose the limit case of perfect mobility (*m*=0), which implies that $\overline{\Omega} = \underline{\Omega}$.[9] It is natural to assume that *W*(·) exhibits decreasing absolute risk aversion. Thus the second term on the right hand side of (32) is negatively signed. For sufficiently high degree of mobility, by virtue of continuity, the second term is negatively signed as well. This term captures the mobility factor, which mitigates the 'background effect' and suggests a downward adjustment in the marginal tax rate.

The sign of the expression in (25) thus depends on the sign of the three terms and cannot be unambiguously signed. However, when the mobility component (second term in the re-distributive component) turns out to be sufficiently dominant, it may be the case that a mean-preserving spread would imply a lower marginal tax rate. Thus, a higher initial inequality leads to a less progressive tax system! A numerical example in appendix B illustrates this result. The result is embedded in the effect of taxation on the ratio $\overline{p}/\underline{p}$, which, in a sense, constitutes some measure for the relative success prospect of the offspring of a typical wealthy family to remain wealthy, vis-à-vis the success prospect of

[9] A high degree of mobility implies a lower marginal tax rate in the optimum, which follows from our previous comparative static exercise. This further boosts investment in education.

the offspring of a poor family to move up along the socio-economic scale. Formally the ratio is increasing with respect to the tax rate, *t*, which implies that the share of individuals coming from a poor background in the wealthy population of the second generation is diminishing.

It should be emphasized that we choose a two period model, which is biased towards the 'background effect'. Extending the model to a longer duration will shift the weight to the long-run component and amplify the impact of upward mobility. Indeed, in the dynamic model we present in appendix C, in the limit case, the 'background effect' disappears and the optimal tax rate is set to zero, regardless of the static level of inequality. This reflects the stark difference between the notions of static inequality and dynamic inequality. Education thus serves to mitigate the initial increase in earning inequality, by increasing the transition probability. In the case of perfect mobility this increase operates like a lump-sum transfer that reduces overall inequality.

4. Conclusions

The optimal tax literature focuses primarily on static setups with exogenously given distributions of earnings abilities. In the words of Mirrlees (1971), these abilities are purely innate. In this paper we consider a somewhat more general framework in which these abilities are partly innate but partly also acquired through education. In other words, we allow for economic (or social) mobility whose extent is endogenously determined through investment in human capital decisions. We find that mobility may plausibly reduce the optimal degree of progression of the tax-transfer system.

References

Blumkin, T. and E. Sadka (2002) "Estate Taxation", forthcoming in Journal of Public Economics

Brito, D, Hamilton, J., Slutsky, S. and J. Stiglitz (1991) "Dynamic Optimal Income Taxation with Government Commitment", Journal of Public Economics 44, 15-35

Cremer, H. and P. Pestieau (2001) "Non Linear Taxation of Bequests, Equal Sharing Rules and the Tradeoff between Intra- and Inter-family Inequalities", Journal of Public Economics, 79, 35-53

Kopzcuk, W. (2000) "On Re-distribution Using Estate Taxation", University of Michigan, Department of Economics, mimeo

Mirrlees, J. (1971) “An Exploration in the Theory of Optimum Income Taxation”, Review of Economic Studies, 38, 175-208

Rosenzweig, M. and K. Wolpin, (1994) “Are There Increasing Returns to the Intergenerational Production of Human Capital?”, Journal of Human Resources 29, 670-693

Altonji, J. and T. Dunn (1996) “The Effects of Family Characteristics on the Return to Education”, Review of Economics and Statistics 78, 692-704

Solon, G. (1992) “Intergenerational Income Mobility in the United States”, American Economic Review 82, 393-408

Appendix A: Derivations

Derivation of equations (23) and (24)

Using the notation in the text and employing the individual budget constraints [(2)-(4)], it follows that:

(A1)

$$V(w,t,\tau) = \max_{c_1,c_2,c_3,e} \begin{bmatrix} u(c_1) + \beta\{p(e,w)c_3 + [1 - p(e,w)]c_2\} + \lambda_1[(1-t)w - e - b + \tau - c_1] \\ + \lambda_2[(1-t)\underline{w} + b + \tau - c_2] + \lambda_3[(1-t)\overline{w} + b + \tau - c_3] \end{bmatrix}$$

By assumption the transition probabilities are given by $p(e,\overline{w}) = a + m + f(e)$ and $p(e,\underline{w}) = a - m + f(e)$. It follows that:

(A2) $\partial p(e,\overline{w}) / \partial m = -\partial p(e,\underline{w}) / \partial m = 1$

Moreover, by subtracting the budget constraint in (3) from the budget constraint in (4) one obtains:

(A3) $c_3 - c_2 = (1-t) \cdot [\overline{w} - \underline{w}]$

Using the envelope theorem, it follows that:

(A4) $\partial V(w) / \partial m = \partial p(e,w) / \partial m \cdot \beta \cdot [c_3 - c_2]$

The result follows by substituting (A2) and (A3) in (A4).

Derivation of equations (30) and (31)

We first turn to derive equation (30). By construction of the mean-preserving spread, it follows that $\partial \overline{w} / \partial \Delta = -\partial \underline{w} / \partial \Delta = 1$. Differentiating (A1) with respect to Δ, using the envelope theorem, one obtains the following expression:

$$\text{(A5)} \qquad \frac{\partial V(\underline{w})}{\partial \Delta} = -\lambda_1(1-t) - \lambda_2(1-t) + \lambda_3(1-t) + [\lambda_1 + \lambda_2 + \lambda_3] \cdot \partial \tau / \partial \Delta$$

Substituting for λ_1, λ_2 and λ_3 from (6)-(8) into (A1) and differentiating the budget constraint faced by the planner given in (13) with respect to Δ (to obtain an explicit expression for $\partial \tau / \partial \Delta$) yields the result.

Turning next to derive equation (31), differentiate (A1) to obtain:

$$\text{(A6)} \qquad \frac{\partial V(\underline{w})}{\partial \Delta} = \lambda_1(1-t) - \lambda_2(1-t) + \lambda_3(1-t) + [\lambda_1 + \lambda_2 + \lambda_3] \cdot \partial \tau / \partial \Delta$$

Following the same steps of the previous derivation yields the expression in the text.

Appendix B: A Mean-preserving Spread in Earnings Inequality

We illustrate that a rise in inequality may lead, counter-intuitively, to a less progressive tax system by means of a simple numerical example. The parametric assumptions are consistent with the model's assumptions and given by:

$$U(c_1) = \ln(c_1)$$
$$\gamma = 0.5$$
$$\beta = 0.8$$
$$p(\overline{w}, e) = 0.11 + \sqrt{e}$$
$$p(\underline{w}, e) = 0.09 + \sqrt{e}$$
$$W(V) = \sqrt{V}$$
$$\overline{w} = 2.4$$
$$\underline{w} = 1.6$$

The following table represents the calculated optimal marginal tax rate, *t*, for different values of Δ, the mean preserving spread:

Δ	0.1	0.2	0.3	0.4	0.5
t	0.083	0.082	0.081	0.080	0.078

We can observe that the marginal tax rate is decreasing with respect to the size of the spread.

Appendix C: Dynamic Extension

In this part, we extend the two-period model into a fully dynamic framework.

Parents are assumed to be altruistic (a la Barro) with respect to the sequence of future generations. To simplify, we assume that preferences are represented by a *quasi-linear* utility function, and defined recursively for a typical dynasty of type *w* by:

(C1) $U(c,e,w) = c + \beta\{p(e,w)V(\overline{w}) + [1 - p(e,w)]V(\underline{w})\}$

With *c* denoting consumption, $V(\cdot)$ denoting the maximal utility derived by the offspring (taking into consideration all future generations) and $\beta > 0$ denoting the degree of altruism. Note that by virtue of the stationarity of the transition probabilities (we will confine attention to the steady-state) the value function, $V(\cdot)$, is independent of time.

We denote by ρ the (implicit) discount rate associated with β, the degree of altruism. That is, $\beta = 1/(1+\rho)$. We henceforth focus on the case in which the equilibrium interest rate, *r*, satisfies $\rho = r$. It follows that consumption smoothing becomes irrelevant, and hence, we can solve for the optimum of each generation separately.

The cost of education, measured by forgone consumption, is given by *e*. We assume that education costs are not tax deductible.

In the presence of a linear labor income tax, comprised of a flat tax rate, *t*, and a uniform lump-sum transfer, τ, each generation is faced with the following budget constraint:

(C2) $w(1-t) + \tau = c + e$

One can write the following asset condition, which determines the continuation value of a typical dynasty of type w (henceforth, suppressing the tax parameters for notational convenience):

(C3) $$V(w) = \max_e \{w(1-t) + \tau - e + \beta p(e,w)[V(\overline{w}) - V(\underline{w})] + \beta V(\underline{w})\}$$

Solving for the optimal investment in education yields the following first-order-condition:

(C4) $$1 = \beta p'(e,w)[V(\overline{w}) - V(\underline{w})]$$

Denote by $\overline{e}$ and $\underline{e}$ the optimal levels of education determined by individuals of types $\overline{w}$ and $\underline{w}$, correspondingly. Let the induced transition probability measures be denoted by $\overline{p} = p(\overline{e}, \overline{w})$ and $\underline{p} = p(\underline{e}, \underline{w})$, correspondingly.

Normalizing the size of population at each time s to unity, we denote by $\overline{q}_s$ and $\underline{q}_s$ the size (fraction) of population with earning ability $\overline{w}$ and $\underline{w}$ at any time s, respectively. It follows by construction that $\overline{q}_s + \underline{q}_s = 1$ for all s.

The evolution of the economy is determined by the following condition:

(C5) $$\overline{q}_s = \overline{q}_{s-1}[\overline{p} - \underline{p}] + \underline{p}$$

By virtue of (C4) and using the properties of the transition probability it follows that $\overline{p} > \underline{p}$ and that for any initial condition, $\overline{q}_0$, the economy converges to a steady-state distribution of earning abilities given by:

(C6) $\bar{q} = \frac{\underline{p}}{1-(\bar{p}-\underline{p})}$; $\underline{q} = \frac{1-\bar{p}}{1-(\bar{p}-\underline{p})}$

With the intention of addressing policy issues, we assume that the preferences of the social planner are represented by some *Bergson-Samuelson* welfare measure of the form:

(C7) $W = \sum_{s=0}^{\infty} \delta^s \{\bar{q}_s W[V(\bar{w})] + \underline{q}_s W[V(\underline{w})]\}$

Where $W''(\cdot) < 0$ and $W'(\cdot) > 0$. Namely, the welfare measure exhibits (strict) inequality-aversion.

Note, that the welfare of the offspring generations is already incorporated into the social welfare function, *W*, through the parent's utility, $V(\cdot)$, at time *s*=0. However, the social planner may also assign a positive weight to the welfare of the offspring generations per se. This is captured by the parameter $\delta \geq 0$. When $\delta = 0$, the offspring's welfare is fully 'laundered out'. We will henceforth focus on the limiting case, $\delta \to 1$, where the social planner is seeking to maximize the welfare flow at the steady state.

We turn next to characterize the optimal linear labor income tax. We first take some preliminary steps before formulating the social planner's program. By manipulation of the asset condition in (C3), one can show that the continuation value functions for high ability and low ability dynasties are given respectively by equations (C8) and (C9):

(C8) $V(\bar{w}) = \frac{1}{(1-\beta)} \cdot [\bar{\gamma}\bar{c} + (1-\bar{\gamma})\underline{c}]$,

(C9) $V(\underline{w}) = \frac{1}{(1-\beta)} \cdot [\underline{\gamma}\bar{c} + (1-\underline{\gamma})\underline{c}]$.

Where $\bar{\gamma} = \frac{1-\beta(1-\underline{p})}{1-\beta(\bar{p}-\underline{p})}$, $\underline{\gamma} = \frac{\beta \underline{p}}{1-\beta(\bar{p}-\underline{p})}$, $\bar{c} = \bar{w}(1-t) + \tau - \bar{e}$ and $\underline{c} = \underline{w}(1-t) + \tau - \underline{e}$

Differentiating (C8) and (C9) with respect to the labor income tax rate, *t*, using the envelope theorem, it follows that:

$$\text{(C10)} \quad (1-\beta)\frac{\partial V(\overline{w})}{\partial t} = -\overline{\gamma}\overline{w} - (1-\overline{\gamma})\underline{w}$$

$$\text{(C11)} \quad (1-\beta)\frac{\partial V(\underline{w})}{\partial t} = -\underline{\gamma}\overline{w} - (1-\underline{\gamma})\underline{w}$$

Fully differentiating (C4) with respect to *t*, noting that $\overline{\gamma} > \underline{\gamma}$, one can show that $\frac{\partial \overline{p}}{\partial t} < 0$; $\frac{\partial \underline{p}}{\partial t} < 0$ and, by virtue of (C6), $\frac{\partial \overline{q}}{\partial t} < 0$. In words, more progressive taxation; namely, a rise in the marginal tax rate, *t*, reduces the gains for education and results in lower upward mobility, hence a lower fraction of high-ability dynasties in the steady state.

We turn next to formulate the social planner's program. We focus on the steady state (see our discussion above). We assume that the planner is seeking to maximize the welfare flow in the steady state subject to an inter-temporal balanced budget constraint, which in a steady state implies period-by-period balanced constraint. Formally,

$$\text{(C12)} \quad \begin{aligned} &\max_{t,\tau}\{\overline{q}W[V(\overline{w})] + \underline{q}W[V(\underline{w})]\} \\ &s.t. \\ &t[\overline{q}\overline{w} + \underline{q}\underline{w}] - \tau = 0 \end{aligned}$$

Formulating the *Lagrangean*, one derives the following first-order conditions (with respect to *t* and τ):

(C13) $$\frac{\partial \bar{q}}{\partial t}\{W[V(\bar{w})]-W[V(\underline{w})]\}+\frac{\bar{q}\bar{\mu}}{(1-\beta)}[-\bar{\gamma}\bar{w}-(1-\bar{\gamma})\underline{w}]+$$
$$\frac{\underline{q}\underline{\mu}}{(1-\beta)}[-\underline{\gamma}\bar{w}-(1-\underline{\gamma})\underline{w}]+\lambda[\bar{q}\bar{w}+\underline{q}\underline{w}+t(\bar{w}-\underline{w})\cdot\frac{\partial \bar{q}}{\partial t}]=0$$

(C14) $$\frac{\bar{q}\bar{\mu}+\underline{q}\underline{\mu}}{(1-\beta)}=\lambda .$$

With $\bar{\mu}$ and $\underline{\mu}$ denoting the social marginal utility of income for a high-ability and a low-ability dynasty, respectively and λ denoting the *Lagrange* multiplier associated with the budget constraint in (C12).

Substituting (C14) into (C13), some algebraic manipulations and re-arranging yield the following expression:

(C15) $$t=\frac{-Cov[w,\theta]+(1-\beta)\frac{\partial \bar{q}}{\partial t}\{W[V(\bar{w})]-W[V(\underline{w})]\}}{-[(\bar{w}-\underline{w})\cdot\frac{\partial \bar{q}}{\partial t}]\cdot E(\theta)}$$

Where θ denotes the *lifetime* social marginal utility of income, given by $\bar{\theta}=\bar{\gamma}\bar{\mu}+(1-\bar{\gamma})\underline{\mu}$ and $\underline{\theta}=\underline{\gamma}\bar{\mu}+(1-\underline{\gamma})\underline{\mu}$ for a low-ability and a high ability dynasty, respectively.

We turn next to sign the expression in (C15). Since $\frac{\partial \bar{q}}{\partial t}<0$ the denominator is positively signed. The sign of the numerator depends on two elements. By the strict concavity of the welfare function, it follows that $\underline{\mu}>\bar{\mu}$. As $\bar{\gamma}>\underline{\gamma}$ the covariance term is negatively signed, hence the first element is positively signed. By virtue of our earlier derivations, it follows that the second element is negatively signed. Thus determining whether the optimal labor income tax is progressive or regressive; namely, whether the labor income tax rate, *t*, is positive or negative, depends on the relative magnitude of two competing forces. The first captured by the covariance element in the numerator is the re-distributive component, which suggests a progressive labor income tax. The second force at play is a *Pigouvian* element, negatively signed, that is aimed at internalizing the discrepancy between the welfare as perceived by the social planner and welfare from the parents' generation perspective. When aversion for inequality as captured by the welfare function is sufficiently strong the optimal tax is progressive and vice versa.

Close inspection of the optimal condition in (C15) above reveals the distinction from the traditional 'static' framework. Two differences are worth noting.

First the social marginal utility from income is a weighted-average defined over the range of earning levels. This reflects the mobility consideration in optimal tax design. Thus, although, at any point in time, 'static' earnings inequality persists, when the impact family background bears on mobility tends to be negligible; namely, as $\underline{\gamma} \to \overline{\gamma}$, the re-distributive component tends to disappear. When $\beta \to 1$, namely, for sufficiently high degree of altruism, the optimal tax rate approaches zero, as family background effect tends to disappear and the Pigouvian term vanishes as well.[10]

Second, the non-zero marginal labor income tax rate affects the amount of education hence the level of mobility. In fact, as we assume for simplicity that labor supply is perfectly inelastic, this is the only source of distortion in the economy.

[10] To verify the latter, note that by employing (C8) and (C9), $\lim_{\beta\to 1}[V(\overline{w})-V(\underline{w})]<\infty$. Concavity of the welfare measure, *W*, implies that $\lim_{\beta\to 1}\{W[V(\overline{w})]-W[V(\underline{w})]\}<[V(\overline{w})-V(\underline{w})]\cdot W'[V(\underline{w})]<\infty$. By fully differentiating (C4) and (C6) with respect to *t*, and employing (C10) and (C11), it follows that $\lim_{\beta\to 1}\left[\frac{\partial \overline{q}}{\partial t}\right]<\infty$.

THE FOERDER INSTITUTE FOR ECONOMIC RESEARCH and
THE SACKLER INSTITUTE FOR ECONOMIC STUDIES
The Eitan Berglas School of Economics
Tel-Aviv University

Recent List of Working Papers

19-2001	Itzhak Gilboa David Schmeidler	Inductive Inference: An Axiomatic Approach
20-2001	Itzhak Gilboa Eva Gilboa-Schechtman	Mental Accounting and the Absentminded Driver
21-2001	Antoine Billot Alain Chateauneuf Itzhak Gilboa Jean-Marc Tallon	Bargaining Over an Uncertain Outcome: The Role of Beliefs
22-2001	Itzhak Gilboa David Schmeidler	Subjective Distributions
23-2001	Hans K. Hvide Aviad Heifetz	Free-Entry Equilibrium in a Market for Certifiers
24-2001	Robert J. Aumann Aviad Heifetz	Incomplete Information
25-2001	Aviad Heifetz Ella Segev	The Evolutionary Role of Toughness in Bargaining
26-2001	Tomer Blumkin Efraim Sadka	Estate Taxation
27-2001	Nathan Rosenberg Manuel Trajtenberg	A General Purpose Technology at Work: The Corliss Steam Engine in the Late 19th Century US
28-2001	Ran Spiegler	Testing Threats in Repeated Games
29-2001	Bronwyn Hall Adam Jaffe Manuel Trajtenberg	The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools
30-2001	Ran Spiegler	Argumentation and Decision Biases in Multi-Issue Debates
31-2001	Shabtai Donnenfeld Alfred A. Haug	Currency Invoicing of U.S. Imports
32-2001	Itay Goldstein Ady Pauzner	Demand Deposit Contracts and the Probability of Bank Runs
33-2001	Itay Goldstein Ady Pauzner	Contagion of Self-Fulfilling Financial Crises Due to Diversification of Investment Portfolios

----

Copies of the working papers of the two Institutes can be obtained from: Mrs.Stella Padeh, The Foerder Institute for Economic Research, Tel-Aviv University, Tel-Aviv, 69978 Israel, Tel: 972-3-640-9255; fax: 972-3-640-5815; e-mail: foerder@post.tau.ac.il **Papers from 2000 onwards (and some from 1998 and1999) can be downloaded from our website as follows: http://econ.tau.ac.il/research/search_workingPapers.asp.**

34-2001	Yoram Weiss Robert M. Sauer Menachem Gotlibovski	Immigration, Search and Loss of Skill
1-2002	Gene M. Grossman Elhanan Helpman	Outsourcing in a Global Economy
2-2002	Hedva Ber Asher Blass Oved Yosha	Monetary Policy in an Open Economy: The Differential Impact on Exporting and Non-Exporting Firms
3-2002	Ran Spiegler	Justifiability and Delay Tactics in a Concession Game
4-2002	Chaim Fershtman Uri Gneezy Frank Verboven	Discrimination and Nepotism: The Efficiency of the Anonymity Rule
5-2002	Pierre Andre Chiappori Yoram Weiss	Marriage Contracts and Divorce : An Equilibrium Analysis
6-2002	Ariel Rubinstein	Comments on the Risk and Time Preferences in Economics
7-2002	Jean-Marie Viaene Itzhak Zilcha	Human Capital and Cross-Country Comparison of Inequality
8-2002	Chaim Fershtman Aviad Heifetz	Read My Lips, Watch for Leaps: A Theory of Endogenous Instability
9-2002	Sarit Cohen Zvi Eckstein	Labor Mobility of Immigrants: Training, Experience, Language and Opportunities
10-2002	Michele Piccione Ariel Rubinstein	Modeling the Economic Interaction of Agents with Diverse Abilities to Recognize Equilibrium Patterns
11-2002	Dmitri Byzalov	A Dynamic Model of OPEC: Collusive Behavior and Demand-Reducing Investment
12-2002	Aviad Heifetz Ella Segev	Escalation and Delay in Long International Conflicts
13-2002	Gene M. Grossman Elhanan Helpman	Outsourcing versus FDI in Industry Equilibrium
14-2002	Gilat Levy	A Model of Political Parties
15-2002	Assaf Razin Efraim Sadka	FDI Flows, Portfolio Flows and Domestic Investment: The Role of Information
16-2002	Itay Goldstein Ady Pauzner	Demand Deposit Contracts and the Probability of Bank Runs
17-2002	Aviad Heifetz Ady Pauzner	Backward Induction with Players who Doubt Others' Faultlessness
18-2002	Allan Drazen	Fiscal Rules from a Political Economy Perspective
19-2002	Allan Drazen	Conditionality and Ownership in IMF Lending: A Political Economy Approach

20-2002	Chaim Fershtman Hans K. Hvide Yoram Weiss	A Behavioral Explanation of the Relative Performance Evaluation Puzzle
21-2002	Neil Gandal	The Effect of Native Language on Internet Usage
22-2002	Martin J. Osborne Ariel Rubinstein	Sampling Equilibrium with an Application to Strategic Voting
23-2002	Kfir Eliaz Ran Spiegler	Anticipatory Feelings and Attitudes to Information
24-2002	Burkhard C. Schipper	Imitators and Optimizers in Cournot Oligopoly
25-2002	Alex Cukierman Itay Goldstein Yossi Spiegel	The Choice of Exchange Rate Regime and Speculative Attacks
26-2002	Alex Cukierman Stefan Gerlach	The Inflation Bias Revisited: Theory and Some International Evidence
27-2002	Aviad Heifetz Chris Shannon Yossi Spiegel	What to Maximize if you Must
28-2002	Maristella Botticini Zvi Eckstein	From Farmers to Merchants: A Human Capital Interpretation of Jewish Economic History
29-2002	Zvi Hercowitz	Concentration of Capital Ownership and Investment Fluctuations
30-2002	Jeffrey Church Neil Gandal David Krause	Indirect Network Effects and Adoption Externalities
1-2003	Zvi Eckstein Yoram Weiss	On the Wage Growth of Immigrants: Israel, 1990-2002
2-2003	Bernhard Eckwert Itzhak Zilcha	The Effect of Better Information on Income Inequality
3-2003	Chaim Fershtman Oren Bar-Gill	Law and Preferences
4-2003	Chaim Fershtman Hans K. Hvide Yoram Weiss	Cultural Diversity, Status Concerns and the Organization of Work
5-2003	Assaf Razin Efraim Sadka	Privatizing Social Security Under Balanced-Budget Constraints: A Political-Economy Approach
6-2003	Assaf Razin Efraim Sadka	A Brazilian-Type Debt Crisis: Simple Analytics
7-2003	Yossi Hadar David Pines	Population Growth and its Distribution between Cities: Positive and Normative Aspects
8-2003	Tomer Blumkin Efraim Sadka	Income Taxation with Intergenerational Mobility: Can Higher Inequality Lead to Less Progression