

POLICY RESEARCH WORKING PAPER

2270

The Optimal Income Tax When Poverty Is a Public “Bad”

Waly Wane

Suppose that poverty is a negative externality affecting individuals' welfare. How does the introduction of concern about poverty affect the income tax policy of a social planner?

The World Bank
Development Research Group
Public Economics
January 2000



Summary findings

Wane considers poverty as an aggregate negative externality that affects people in different ways, depending on their aversion to poverty.

If society is on average averse to poverty, then the optimal income tax schedule displays negative marginal tax rates, at least for less skilled individuals. Negative

marginal tax rates play the role of a Pigouvian earnings subsidy, fostering the supply of poor individuals to provide labor. The result of no distortion at the endpoints, which is therefore violated, can be restored once the focus is shifted from individual to social distortions.

This paper — a product of Public Economics, Development Research Group — is part of a larger effort in the group to better understand public policy and poverty. Copies of the paper are available free from the World Bank, 1818 H Street, NW, Washington, DC 20433. Please contact Hedy Sladovich, room MC2-609, telephone 202-473-7698, fax 202-522-1154, email address hsladovich@worldbank.org. Policy Research Working Papers are also posted on the Web at www.worldbank.org/research/workingpapers. The author may be contacted at wwane@worldbank.org. January 2000. (34 pages)

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent.

The Optimal Income Tax
when
Poverty is a Public “Bad”

Waly Wane

Development Research Group, World Bank, 1818 H Street, NW, Washington D.C. 20433 USA.

Tel: +1 202 458 0155; Fax +1 202 522 1154; e-mail: wwane@worldbank.org.

1 Introduction

The optimal income taxation literature began with the seminal paper of Mirrlees (1971). In a society composed of individuals with different exogenous income earning possibilities, Mirrlees characterizes the best income tax function a social planner (or government) can implement. In his context, ‘best’ must be understood with reference to (i) the reaction of the individuals to the tax function, and (ii) the objective of the planner. The recognition of (i) is the main contribution of Mirrlees to this literature. Indeed, he introduced explicitly, in a tractable way and for the first time, the (dis)incentive effects associated with the taxation of (endogenous) income. In defining (ii), Mirrlees followed the tradition in welfare economics by postulating that the government maximizes a social welfare function, i.e. a function depending solely on the individuals’ welfare, social judgements and preferences being present through the weights assigned to different agents’ well-being. Mirrlees’ approach is in this sense “welfarist”.

Welfarism has been criticized on several grounds. In a stream of important papers gathered in a single book, Sen (1982) argues that welfarism is informationally restrictive. Welfarism rejects any information unrelated to personal individual welfare, i.e. non-utility information. Such information is given for instance by the possibility to exercise one’s ‘rights’, the freedom to express one’s views or ethical concepts such as ‘liberty’, ‘justice’, etc. From a policy standpoint also, welfarism is considered too narrow to provide sound policy prescriptions. Many important issues are missed or simply left out by the welfarist approach.

There are principles of social judgement that require essential use of non-utility information, and while such principles (e.g., liberty, non-exploitation, non-discrimination) are typically not much discussed in traditional welfare economics, they do relate closely to the subject matter of welfare economics. [Sen (1982), p. 338]

Alternative approaches to policy analysis, using non-utility information, have been developed following the work of Sen (1982). These approaches are usually ranked under the generic term of

'nonwelfarist'.

In setting the policy agenda, a social planner may then want to take into account, utility as well as non-utility information, departing from the usual welfarist approach. In the economics literature however, the welfarist and nonwelfarist viewpoints have been adopted each to the exclusion of the other. A recent exception is however provided by Atkinson (1995a) who explicitly recognizes that a government can pursue several objectives and has to combine them in a sensible way:

The objectives of securing individual freedom, avoiding dependency, or rewarding effort do not replace concern for the welfare of individuals. Rather we have now to recognize that we have a plurality of principles. [Atkinson (1995a) p. 63]

Atkinson therefore advocates a social objective function that trades off different, possibly conflicting, concerns. While this would take us away from the welfarist approach, in this paper we stick to the latter while accounting for the multidimensionality of welfare.

According to Sugden (1993), the industry standard in welfare economics is 'revealed preference welfarism' where the welfare of an individual is associated with the actions she undertakes. For instance, in the optimal taxation literature, a person's welfare is the utility derived from the bundle she chooses to consume. There are, however, many other dimensions to welfare.

This paper proposes a procedure that allows (some) a priori non-utility information to be taken into account within a welfarist approach. For welfarism to accommodate such information, the latter must be 'welfarized'. Let us consider aggregate poverty to illustrate. A person is poor if she is short of a given amount of income that would allow her to attain a given 'decent' standard of living. For each such person, a measure of poverty can be computed, using indexes that have been proposed in the literature, aggregate poverty being the sum of all these measures. It is broadly accepted that people are not indifferent with respect to self- or others' deprivation as suggested by the following quotes "[One might] feel depressed at the sight of misery" [Sen (1982), p. 8], and "If taxpayers care about reducing the visible signs of poverty [...]" [Besley and Coate

(1995), p. 189]. It is also generally thought that poverty can lead to criminality. If so, poverty can then be harmful for a society and it is considered here as an aggregate negative externality or public “bad” that reduces the utility of the individuals. It is then a kind of diffuse externality that Meade labeled “atmosphere” externalities. When the planner proposes a tax schedule, each individual chooses the bundle she prefers, revealing her preferences, but does not choose the level of externality, although it also affects her well-being.

Considering poverty, which is a non-utility information, as an externality is one way to welfarize it. People may still have different degrees of aversion to poverty. If one takes the criminality interpretation, people may be affected differently because of their geographic location. Sensitivity to poverty may simply differ among the individuals. Once the non-welfare information has been welfarized, one can accommodate it in the standard social welfare function and get back to welfarism. The social welfare function is then a function of the ‘comprehensive’ individuals’ welfares. In the following, when the planner considers comprehensive welfare, he is labeled a ‘poverty as a public bad’ (ppb) planner to distinguish it with the (standard) welfarist planner.

This paper is obviously not the first to consider poverty as an externality. The idea can be traced back at least to Zeckhauser (1971) who considers a representative citizen proposing income transfer mechanisms to poor people in order to reduce the externality the latter produce. In his setup only the rich are affected by the externality and he only considers linear taxation. The closest paper to ours is that of Kanbur, Keen and Tuomala (1994). Their objective is to find the shape of the nonlinear tax schedule that most alleviates poverty, measured by a given index. They then consider a Mirrleesian problem in which the objective function is replaced by an income-based poverty index. Their approach is non-welfarist since individuals’ utility appears only in the self selection constraints from which the government cannot escape. They obtain some interesting results such as the negativity of the optimal marginal tax rate at the lower end of the scale, even though this feature is not confirmed by their simulations. A more general, but completely overlooked, approach is Seade (1979) who discusses the properties of the optimal tax schedule a

non-utilitarian government, which can have any objective not directly linked to the individuals' welfare, has to implement. Besley and Coate (1992, 1995) also offer an interesting study for income maintenance programs. They focus on the effectiveness of workfare schemes in minimizing of the cost of these programs. Their papers do not, however, include a budget constraint; in this paper only self-financing policies are considered. Cremer et al. (1998) reconsider also the optimal design of the tax system in the presence of an aggregate externality. However, in their study, the externality is produced by a consumption good while in our case it comes from the paucity of income. Even though in both papers only aggregate externality matters, its nature is different. Indeed, in Cremer et al. there is no threshold effect and any additional consumption of the externality good increases the level of the aggregate externality, while in this paper only a proportion of the population creates an externality even though everybody enjoys income. They also consider both income and commodity taxation, whereas this paper is only concerned with income taxation.

The remainder of the chapter is organized as follows. The next section presents the general structure of the economy. Section 3 investigates the benchmark case of complete information. Section 4 is devoted to the asymmetric information economy and deals with nonlinear taxation. Section 5 derives an explicit solution of the nonlinear tax schedule for quasilinear preferences; and numerical simulations for specific poverty index, poverty line and distribution of skills are conducted. Section 6 concludes.

2 The Model

Following the traditional income taxation literature, it is assumed that individuals have the same preferences with regard to the only two goods present in the economy, namely consumption, x , and labor, y . Preferences are represented by a utility function $u(x, y)$ which is assumed to be strictly concave on (x, y) , at least twice continuously differentiable, increasing in x and decreasing in y . The

economy consists of a continuum of such individuals parameterized by a single variable n , which represents their income-generating possibilities or productivity. This parameter is distributed according to the density function $f(n)$ and the cumulative distribution function $F(n)$ over the support $[\underline{n}, \bar{n}]$.

For any agent of productivity n who works y units of time, her gross income is given by $z = n \cdot y$. Let us define $U(x, z, n) \equiv u(x, z/n)$ and impose the condition of Agent Monotonicity of Seade (1982)

$$(AM) \quad s_n \equiv \partial s(x, z, n) / \partial n < 0 \quad \forall (x, z)$$

where $s(x, z, n) \equiv -U_z(x, z, n) / U_x(x, z, n) > 0$ is the marginal rate of substitution between gross income and consumption. This condition states that at any point (x, z) in the consumption-gross income space, the indifference curves are flatter the higher the productivity of the agent. This is also assumption B of Mirrlees (1971) and the usual single crossing condition of the screening literature.

Each individual chooses the amount of labor to supply depending on the tax schedule $x = \xi(z)$, relating gross income to consumption, the social planner proposes. By picking up a specific bundle (x, z) from the proposed tax schedule, the individual reveals that it is her preferred bundle which is the one that maximizes the component of her welfare she has a control upon, namely $u(x, y)$ that is labeled here ‘intrinsic’ utility. However, her “true” welfare depends also on other variables that are out of her control.

It is assumed that the only relevant feature of the social state that affects the welfare of individuals is the level of poverty that exists in the society. Poverty can affect welfare if, as it is often argued, it is among the main social causes for the emergence of crime. Another simple argument is that people just dislike to live in or cohabit with misery. Let us denote by \mathcal{P} the aggregate level of poverty that will be made explicit later in the analysis. The overall welfare level

of an n -agent choosing the bundle $(x(n), z(n))$ is therefore

$$V(x(n), z(n), n, \mathcal{P}) = U(x(n), z(n), n) - \beta(n)\mathcal{P} \quad (1)$$

where $\beta(n)$, which is common knowledge, measures the aversion to poverty of the n -agents. The agents are poverty-lovers if it is negative, indifferent if it equals zero and poverty-averse if it is strictly positive. The (loss of) welfare of the agents derived from the sight of poverty is labeled ‘extrinsic’ welfare.¹

The government wants to choose the tax schedule $x = \xi(z)$ that maximizes social welfare, which is the sum of the individuals’ true welfare. Its objective is therefore:

$$Max_{x=\xi(z)} \int_{\underline{n}}^{\bar{n}} V(x(n), z(n), n, \mathcal{P})f(n)dn \quad (2)$$

The government also must consider the reaction of the individuals with respect to this budget constraint and the resource constraint. This latter is given by

$$\int_{\underline{n}}^{\bar{n}} (z(n) - x(n))f(n)dn = \bar{R} \quad (3)$$

where \bar{R} is the revenue the government wants to levy for the financing of (say) a public good. When it is zero, the planner is only concerned by redistribution. When faced with the tax schedule, the agents choose the amount of labor to supply in order to maximize their welfare. Since the extrinsic welfare is treated parametrically, the individuals in fact only maximize the intrinsic part of their welfare. The incentive problem is left unaffected by the introduction of the aggregate externality and is therefore the same as that under the traditional income taxation literature. The first order condition is

$$U_x \cdot x'(n) + U_z \cdot z'(n) = 0 \quad \text{or} \quad \frac{du(n)}{dn} = U_n(x(n), z(n), n) \quad (4)$$

which is the second constraint the government has to consider. Where $u(n)$ is the indirect intrinsic utility, that is the maximum utility the individual can get from the tax schedule. Note that we

¹ The different components of the true welfare are taken to be separable. This assumption seems quite natural since the agents have no control over the second component and behave parametrically with respect to it.

are taking the first-order approach to the problem, i.e., only first order constraints are included. To ensure that the tax schedule derived from this problem is implementable and the optimal one, it must be verified that the resulting gross income is a nondecreasing function of ability, which means that the second-order constraint is fulfilled.²

It remains now to describe the precise form of the aggregate externality. Poverty is considered in this paper in its narrowest sense. Indeed, deprivation is defined solely in terms of income. An individual is poor if her disposable income falls short of a prespecified level x^* called the poverty line in the poverty literature. The poverty line helps to identify poor people in a society. The second step then consists of measuring the extent of poverty by constructing a poverty index. The measurement of poverty has generated a huge literature which survey is beyond the scope of this paper, the interested reader is referred to Ravallion (1994). Several measures of poverty exist; this paper focuses on measures that take into account not only its existence, but also its severity.³ The poverty measure of an agent whose consumption is x is then given by the function $P(x, x^*)$, satisfying the conditions $P(x, x^*) \geq 0$, $P_x(x, x^*) < 0$ and $P_{xx}(x, x^*) > 0$ for all $x \in [0, x^*)$. Finally $P(x, x^*) = 0$ for $x \geq x^*$ and $P(x^*, x^*) = P_x(x^*, x^*) = 0$. The poverty index used here, which aggregates all individual poverty measures, is supposed to belong to the class of additively decomposable indices. For a given poverty line, poverty in the economy is measured by:⁴

$$\mathcal{P}(x^*) = \int_{\underline{n}}^{\bar{n}} P(x(n), x^*) f(n) dn$$

where $\{x(n), n \in [\underline{n}, \bar{n}]\}$ is the consumption profile in the economy.

The objective of the government is therefore to choose the tax schedule to

$$\text{Max} \int_{\underline{n}}^{\bar{n}} \{u(n) - \beta P(x(n), x^*)\} f(n) dn \quad (5)$$

² Introducing explicitly the second order constraint allows to deal with bunching problems which are not the primary focus of the paper. On this see Ebert (1992), Mirrlees (1986), Brito and Oakland (1977) among others.

³ The Headcount ratio, for instance, indicates only what proportion of the population is poor and thus deals only with the existence of poverty. More sophisticated measures have been proposed in the literature to deal with the intensity of poverty, for instance the P_α or FGT measures introduced by Foster, Greer and Thorbecke (1984). For a good and comprehensive survey, the reader is referred to Atkinson (1987).

⁴ The index chosen here is very general and has many indices as special cases. However, given its assumed properties, the Sen (1982, Essay 17) types of measures, which are based on a Gini index, are precluded.

under the differential equation (4) which represents the incentive compatibility constraints and the resource constraint (3). The parameter $\beta = \int_{\underline{n}}^{\bar{n}} \beta(n) dF$ is the average or social aversion to poverty and can be computed by the government since individual aversions are known.⁵ It is (reasonably) assumed to be strictly positive, i.e., the society is, on average, averse to misery.

3 The First-Best

Let us first consider a first-best world in which the government knows who is of what ability. The agents do not have the possibility to mimick other lower ability agents and the government can directly impose consumption-labor bundles that are contingent on the individuals' productivity. The incentive problem having vanished, only the resource constraint (3) needs then to be considered. The maximization program is therefore as follows:

$$\begin{cases} \text{Max } \int_{\underline{n}}^{\bar{n}} \{u(x(n), y(n)) - \beta P(x(n), x^*)\} f(n) dn. \\ \text{subject to } \int_{\underline{n}}^{\bar{n}} (ny(n) - x(n) - \bar{R}) f(n) dn \geq 0 \end{cases}$$

The first order conditions are easily derived and one obtains $u_x - \beta P_x = \gamma$, $u_y = -n\gamma$ and $s = -u_y/nu_x = \gamma/(\gamma + \beta P_x)$. The marginal social valuation of an additional unit of consumption is equated among all the individuals. This condition boils down to the equalization of marginal utilities in a utilitarian framework or if the society is indifferent to poverty ($\beta = 0$). In this setting however, income has an additional positive impact through the reduction of poverty which is valued by the society at the rate β . One of the most prominent features of income transfer mechanisms is the marginal rate of tax the agents face since it captures the (dis)incentive effects

⁵ This objective is obtained after the following transformations:

$$\begin{aligned} \int_{\underline{n}}^{\bar{n}} V(x(n), z(n), n, \mathcal{P}(x^*)) dF &= \int_{\underline{n}}^{\bar{n}} U(x(n), z(n), n) dF - \mathcal{P}(x^*) \int_{\underline{n}}^{\bar{n}} \beta(n) dF \\ &= \int_{\underline{n}}^{\bar{n}} u(n) dF - \int_{\underline{n}}^{\bar{n}} \beta P(x(n), x^*) dF = \int_{\underline{n}}^{\bar{n}} \{u(n) - \beta P(x(n), x^*)\} dF. \end{aligned}$$

Thanks to the separability between the intrinsic and extrinsic welfares and to the decomposability properties of the poverty index, the social welfare which originally takes into account the individual aversions, $\beta(n)$, to social poverty, can be expressed equivalently by using the social aversion, β , to individual poverty.

and hence distortions associated with a given mechanism. In our first best world, the marginal tax rate for the n -agents is given by:

$$1 - s(x(n), y(n)) = \frac{\beta P_x(x(n), x^*) s(x(n), y(n))}{\gamma} = \frac{\beta P_x(x(n), x^*)}{\gamma + \beta P_x(x(n), x^*)} \quad (6)$$

It is immediately apparent that those whose consumption is above the poverty line face no distortion and receive lump-sum transfers as is usual in a complete information setup. Interestingly however, the poor face a negative marginal tax rate, which basically amounts to earnings subsidies.⁶ By subsidizing their earnings at the margin, the government induces the poor to work harder, earn even more income and thereby reduce the negative externality they create. It is in fact a kind of corrective Pigouvian subsidy which make agents internalize the externality. Indeed, the subsidy on the last earned unit of income is equal to its marginal social benefit measured in terms of (this individual) poverty reduction. The Pigouvian subsidy is however nonlinear, i.e., it is not the same for everybody and depends on the level of consumption. Poorer individuals are offered higher subsidies. The subsidy is also higher in societies that are more averse to poverty.

It is also of interest to determine the optimal distribution of consumption and labor in the economy. This problem has received much attention in the standard welfarist framework, see for instance Sadka (1976). It is easy to show that the optimal allocation requires the supply of labor to be an increasing function of productivity. For efficiency reasons, the more productive an agent is the more she works. The redistribution of the income so produced depends mainly on the preferences. For instance, if they are separable between consumption and labor, the optimal consumption allocation is the egalitarian one. The consumption pattern depends on the Edgeworth-complementarity of consumption and leisure, i.e., on the sign of u_{xy} . The following proposition (proof in the appendix) summarizes the results of this section.

⁶ Zeckhauser (1971) presents a similar result even though in his setting only the representative rich citizen suffers from poverty externality.

Proposition 1 *The optimal allocation $(\bar{x}(n), \bar{y}(n))$ is such that:*

- (i) *all nonpoor receive lump-sum transfers while poor individuals face a corrective nonlinear (unless preferences are separable) Pigowian subsidy,*
- (ii) *the supply of labor is an increasing function of ability, i.e., $d\bar{y}(n)/dn > 0$,*
- (iii) *the consumption function pattern depends on the Edgeworth-complementarity between consumption and leisure, i.e., $d\bar{x}(n)/dn \cdot u_{xy} \geq 0$.*

4 Unobservable Abilities

This section considers the case where the productivity of an agent is private information which is by no means available to the government. Let us, for the sake of clarity, restate the maximization program of the government:

$$\text{Max}_{x(n), z(n)} \int_{\underline{n}}^{\bar{n}} \{u(n) - \beta P(x(n), x^*)\} f(n) dn$$

subject to

$$\frac{du(n)}{dn} = U_n(x(n), z(n), n) \quad \text{and} \quad \int_{\underline{n}}^{\bar{n}} (z(n) - x(n)) f(n) dn = \bar{R}.$$

This is a usual optimal control problem with the indirect intrinsic utility $u(n)$ as the state variable and earned income $z(n)$ as the control variable. The multipliers associated to the constraints are, $\mu(n)$ and γ , respectively and the formal resolution of the problem is postponed to the appendix. Assuming that it is optimal for everybody to work and that there is no bunching, the marginal tax rate faced by an n -agent is given by:

$$t(z(n)) = (1 - s(x(n), z(n), n)) = \frac{\beta P_x s}{\gamma} + \frac{\mu(n) U_x s_n}{\gamma f} \quad (7)$$

where

$$\mu(n) = \int_n^{\bar{n}} (U_x - \beta P_x - \gamma) (1/U_x) \exp\left(\int_n^p (U_{nx}/U_x) dm\right) f(p) dp \quad (8)$$

with the endpoint conditions

$$\mu(\underline{n}) = \mu(\bar{n}) = 0. \quad (9)$$

It is easy to show that $\mu(n) < 0$ for $n \in (\underline{n}, \bar{n})$. Indeed, because of the transversality conditions (9), the integrand cannot be always positive or negative, it must change sign at least once. Since

P is a convex function of x which is increasing with n , then $-P_x$ is a decreasing function; so is U_x . Hence the integrand $U_x - (\beta P_x + \gamma)$, which is the marginal social benefit of giving an additional unit of income to an agent in a lump sum manner, is monotonically decreasing and must change sign only once. It must be positive for low values of n and negative for high values, hence the negativity of μ . From equation (8) we can also deduce that $\gamma > 0$. Contrary to the welfarist case, it is no longer possible to assess the sign of the marginal tax rate solely from the sign of μ , because there is a second effect at play here, which is that of poverty reduction.

To better understand the optimal tax schedule chosen by the government, let us rewrite equation (7) in the following more telling way:

$$\left[(1 - s(x(n), z(n), n)) - \frac{\beta P_x s}{\gamma} \right] f(n) = -\frac{\mu(n)}{\gamma} \frac{\partial}{\partial z} \left[\frac{du(n)}{dn} \right] \quad (10)$$

where $U_x s_n = -\partial [du(n)/dn] / \partial z$. Departures from the first-best and hence distortions are captured by the right hand side of (10) which has the same form than in the standard problem. While in the familiar income tax literature the relevant distortions are individual distortions, in our setting one must consider *social* distortions. These latter answer to the following question: how does the incentive problem oblige the society (or its representant, the planner) to move from its optimal allocation? Social distortions are a mere generalization of individual distortion. Indeed, in the standard welfarist problem, individual and social distortions coincide since individual decisions do not have any impact on the society as a whole. On the contrary, in this setup, when an agent chooses a consumption level that falls below the poverty line, she takes a decision that affects the society through her contribution to the aggregate negative externality. Individual distortions exist when the marginal tax rate is different from zero i.e., people receive non lump-sum transfers, while social distortions are present whenever the left hand side of (10) is not zero. Along the same line than Brito and Oakland (1977) it is possible to provide a clear economic interpretation to equation (10). The left hand side of this equation,

$$\left[(1 - s(x(n), z(n), n)) - \frac{\beta P_x s}{\gamma} \right] f(n)$$

is simply the marginal social benefit from the n -agents on the last unit of income they are required to earn. The marginal tax rate $1 - s(x(n), z(n), n)$ is the amount the government retains on the additional unit of income earned by the n -agent which is to be redistributed directly or indirectly if it is positive. If, on the contrary, it is negative, it is the cost the government has to bear to induce the agent to earn this last unit. Therefore $s(x(n), z(n), n)$ is the increase in consumption that must be given to the agent for her to voluntarily earn this last unit of income. Once poverty is taken into consideration, this increase in consumption entails a marginal reduction in the contribution to poverty of these individuals which value to society is $-\beta P_x s / \gamma$ in terms of the government's revenue.⁷ The marginal benefit is greater the higher the number of these individuals $f(n)$.

The right hand side of (10) can similarly be interpreted as the marginal social cost imposed by the presence of the incentive compatibility constraints. It is clear that $-\mu(n)$ is the cost the planner has to bear if it wants to increase the utility of the type- n individual by one unit. This cost is implied by the existence of all the potential mimickers of n -agents, i.e. all p -agents with $p > n$. $U_x - \beta P_x - \gamma$ is the marginal social benefit of allocating one unit of consumption to the p -individual. It is the difference between the marginal social valuation ($U_x - \beta P_x$) — where both the effects on welfare and poverty are counted for — and the cost of this unit in terms of the government's revenue (γ). The number of units of consumption x needed to increase the utility of the p -consumer by one unit is $(1/U_x)$. Thus $-(U_x - \beta P_x - \gamma) \times (1/U_x)$ is the cost to bear to increase the utility of the p -agent by one unit followed by an increase of one unit of the utility of the n -agent.⁸

The total cost of this operation is obtained when the rate of change of the utility of the p -agent

⁷ To reason in terms of an increase of one unit of labor is straightforward and only requires rewriting equation (10) the following way $[(1 - s(x(n), z(n), n)) - \frac{\beta P_x s}{\gamma}] n f(n) = -\frac{\mu(n)}{\gamma} \frac{\partial}{\partial y} \left[\frac{du(n)}{dn} \right]$.

⁸ Note that the cost (or benefit) of an additional unit of income is measured taking into account both intrinsic welfare and the impact on poverty, while the number of units of consumption needed to increase the utility of the agent is measured solely with respect to the intrinsic welfare. Indeed, while the former captures the societal valuation of income, the latter shows that individual incentives depend only on intrinsic welfare.

is introduced. This is given by

$$T_{np} = \exp\left(\int_n^p (U_{nx}/U_x)dm\right)$$

with $U_{nx} = -U_{xz}z/n^2$. When preferences are separable between consumption and leisure, $U_{nx} \equiv 0$ implying $T_{np} = 1$ increasing the utility of the n -agent by one unit requires to also increase the utility of all p -agents (with $p \geq n$), by one unit to respect the self selection constraints. If consumption and leisure are Edgeworth-substitutes ($U_{xz} \geq 0$) then $U_{nx} \leq 0$, implying that $T_{np} < 1$ and for an increase of one unit of the n -agent, the necessary increase in the utility of those who are more productive is less than one unit. The social cost, for an increase of one unit in the utility of the n -agent, imposed by the p consumers is thus $-(U_x - \beta P_x - \gamma)(1/U_x)T_{np}$. Integrating over $p > n$ gives the aggregate cost the planner has to pay, i.e. $-\mu(n)$. The only instrument the government has at its disposal to effectuate such an increase in utility is to alter the tax schedule to produce this effect on the n^{th} consumer; $\frac{\partial}{\partial z} \left[\frac{du(n)}{dn} \right]$ gives the change in the utility of the n -agent that is needed for the incentive compatibility constraints still to hold when she is required to increase her pre-tax income by one unit.

The rhs of (10) is thus the marginal cost of inducing the n -agent to earn one more unit of income in terms of the distortions that arise due to the presence of more skilled workers. This cost is given in terms of the government's revenue.

Equations (7) or (10) tell thus nothing else than the usual economic rule which states that at the optimum the marginal social costs of extracting one more unit of labor, from any agent of given productivity, must be equal to the marginal social benefits.

The formula characterizing the marginal tax rate (7) is rather interesting. The distortion any individual must face has two components. The second term of equation (7) has the same form than the one encountered in the familiar income taxation literature and stems only from the existence of asymmetric information which forces the planner to introduce distortions to satisfy the self selection constraints. On the contrary, the first term captures the 'first-best distortion' i.e. the

distortion the agent must face if productivities were observable, it is the term that is obtained in equation (6) and arises because of the need to induce agents to internalize the externality they produce.⁹ Rewriting (7) as

$$1 - s(x(n), z(n), n) = \frac{\mu(n)U_x s_n}{\gamma f} - \left[-\frac{\beta P_x s}{\gamma} \right]$$

it is clear that marginal tax rate is the difference between the marginal cost induced by private information and the marginal benefit from the reduction of the externality. The above equation shows that all the non-poor people face a positive marginal tax rate. Indeed, for the non-poor, the last term in the rhs is zero. It has been shown that the first term is strictly positive except at the top and bottom of the distribution, thanks to the transversality conditions. Therefore, the marginal tax rate $1 - s$ is strictly positive for the non poor. It is however zero for the most able since consumption is increasing $x'(n) > 0$ implying $x(\bar{n}) > x^*$ and there exist non poor. Thus the no distortion at the top result applies as long as not everybody is poor at the optimum. Moreover, this rate is of precisely the same structure than that proposed by the welfarist planner. Using continuity arguments, some poor people also must face a positive marginal tax rate at least just below the poverty line. The lower segment of the distribution must face a negative marginal tax rate. Let us for instance consider the least able agents. From the transversality condition, the first term in the above equation is zero and their consumption is below the poverty line so the marginal tax they face is negative, their earnings are subsidized at the margin, which induces them to work hard. The planner is ready to subsidize the marginal unit of earned income whenever the marginal benefit associated with it is greater than the marginal cost.

Why does the introduction of poverty concerns lead to the apparition of negative MTR while this never happens in the usual welfarist context? The following answer is offered. The welfarist planner aims, *in fine*, at redistributing utility. This can be achieved by redistributing either consumption or leisure. There is no need to introduce negative marginal tax rate inducing the

⁹ The marginal tax rate formula found in Kanbur, Keen and Tuomala (1994) is of exactly the same form, however, the authors do not offer a clear interpretation of the result.

agents to work harder and subsidizing them at the margin if the same effect on welfare can be achieved by substituting leisure for consumption. This reasoning no longer holds once care is taken of poverty. Even if leisure still increases welfare, it now has an opportunity cost in terms of poverty. Consumption is more important for this planner since it has positive effects on both. The planner thus subsidizes at the margin each additional earned unit of income, for a certain segment of the poor population, to reduce the amount of the public bad. This kind of scheme resembles a great deal to the Earned Income Tax Credit (EITC) policy which refunds paid taxes to the deserving working poor and fosters labor supply, see Eissa and Liebman (1996). In fact the EITC implies negative marginal tax rates. Once the welfarist planner takes into account poverty, it switches from a system of “subsidized leisure” to a system of “subsidized labor”.

Since there has been argued that the relevant distortion to be taken into account are the social ones, let us just write them for the extreme points of the distribution using the transversality conditions:

$$1 - s(\underline{n}) - \beta P_x(x(\underline{n}), x^*)s/\gamma = 1 - s(\bar{n}) - \beta P_x(x(\bar{n}), x^*)s/\gamma = 0.$$

This is exactly the condition obtained in the first best equation (6), there is thus no social distortion at the endpoints. This is exactly the counterpart of the classical result of no (individual) distortion at the endpoints as given for instance by Seade (1977). The discussion of this section is summarized in the following:

Proposition 2 *If there is no bunching and everybody provides some labor, then the optimal tax schedule is such that*

- (i) all the non poor, except the most able, face a strictly positive marginal tax rate,*
- (ii) poor individuals can face both negative or positive marginal tax rates,*
- (iii) although there (might) exist **individual** distortions, there are no **social** distortions at the endpoints.*

It would also be interesting to analyze the case where with the welfarist optimal tax schedule, some agents provide no work and then check the effect of the introduction of poverty on the result of the optimality of induced unemployment.

5 Explicit Solution and Numerical Simulation

5.1 A closed form solution

The high complexity of the first order conditions makes the derivation of the general solution of the optimum income tax problem impossible. With further simplifications, there are a few papers that provide closed form solutions of the optimal non-linear tax schedule. Among them are Lollivier and Rochet (1983), Guesnerie and Laffont (1984) who apply the solution of the former authors to a principal-agent model, Diamond (1998) who gives analytical expressions of the marginal tax rate for special distributions or utility functions and finally Ebert (1992). The trick to obtain closed form solutions of the optimal tax schedule is to consider preferences that are quasi linear w.r.t earned income (and hence labor). In this section, we use Ebert's example. Unlike him, however, we are not interested in dealing with the cumbersome issue of bunching, which demonstrates the relevance of the second order approach; instead we rather want to focus on and emphasize the role of poverty when it becomes an important dimension of the social policy maker's objective.

Suppose that all the individuals have the same quasilinear utility function over consumption and labor given by $u(x, y) = v(x) - y$, the true welfare of an n -agent, which also includes her aversion to generalized poverty, expressed in terms of pre- and post-tax income is therefore:

$$V(x, z, n, \mathcal{P}(x^*)) = v(x) - \frac{z}{n} - \beta(n)\mathcal{P}(x^*), \quad \text{for } n \in [\underline{n}, \bar{n}] \quad (11)$$

with v strictly increasing, concave and twice continuously differentiable. Defining $G(n) = \int_{\underline{n}}^n \frac{f(m)}{m} dm$, the solution of the problem PPB (bunching excluded) is the following:

$$\gamma = G(\bar{n}) \quad (12.1)$$

$$\mu(n) = n[G(\bar{n})F(n) - G(n)] \quad (12.2)$$

$$\frac{\beta P_x(x(n), x^*) + G(\bar{n})}{v'(x(n))} = \frac{nf(n)G(\bar{n}) - G(n) + G(\bar{n})}{f(n)} \quad (12.3)$$

$$t(n) = \frac{\beta P_x(x(n), x^*)}{\beta P_x(x(n), x^*) + G(\bar{n})} + \frac{G(n) - G(\bar{n})F(n)}{nf(n)[\beta P_x(x(n), x^*) + G(\bar{n})]} \quad (12.4)$$

$$V(n) = u(n) - \beta(n)\mathcal{P}(x^*) = \frac{1}{n} \left(K + \int_{\underline{n}}^n v(x(m)) dm \right) - \beta(n) \int_{\underline{n}}^{\bar{n}} P(x(n), x^*) f(n) dn \quad (12.5)$$

$$z(n) = n(v(x(n)) - u(n)) \quad (12.6)$$

$$K = \int_{\underline{n}}^{\bar{n}} \{ [nf(n) - 1 + F(n)] v(x(n)) - x(n)f(n) \} dn \quad (12.7)$$

The first remark is that when poverty is not a matter for concern or in a society that is indifferent to poverty on average, this solution coincides with that of the welfarist planner (see Ebert (1992)). It is also impossible to directly give an explicit solution of the after-tax income since this latter depends on both the poverty line and index. However, once they are specified, consumption levels can be computed. The intrinsic utility level and labor supply function are of precisely the same form than under the welfarist solution but one must keep in mind that these functions depend on consumption and hence the solutions are not necessarily equal.

Suppose the ppb planner inherits the optimal welfarist planner's tax schedule $(\bar{x}(n), \bar{z}(n))$. How will this schedule be amended to take account of poverty and obtain the optimal ppb tax schedule $(\hat{x}(n), \hat{z}(n))$? Let us also denote the marginal tax rates by $\bar{t}(n)$ and $\hat{t}(n)$ with obvious notations. The discussion is given for "reasonable" poverty lines, i.e., poverty lines such that there exist both poor and non-poor people under the optimal welfarist schedule. First of all, one observes that

the marginal cost of public funds, γ , and $\mu(n)$ depend only on the productivity distribution, this property comes from the quasi linearity of individuals' preferences. The following, easily derived, result also comes from that property:

Lemma 3 *Define n^* such that $\bar{x}(n^*) = x^*$. For quasi linear utility functions as defined by (11), the optimal tax schedule considering poverty as a public bad is such that:*

- (i) $\hat{x}(n) = \bar{x}(n) \geq x^*$ and $\hat{t}(n) = \bar{t}(n) \forall n \geq n^*$ and
- (ii) $x^* > \hat{x}(n) > \bar{x}(n)$ and $\hat{t}(n) < \bar{t}(n) \forall n < n^*$.

The nonpoor individuals under the welfarist scheme face the same marginal tax rate and enjoy the same consumption under the new tax schedule. On the contrary, all the poor face lower marginal tax rates which induce them to work harder and they also all enjoy higher consumptions. The aggregate amount of poverty is therefore lowered. Note however that the poor population is still the same, therefore the Headcount ratio has not changed. In terms of distortions, using (12.4), it is easy to show that the social distortion associated to an n -agent is equal to the marginal tax rate the agent would face were the planner a welfarist one i.e. the individual distortion. This result also is a direct consequence of the assumed preferences. Indeed, the incentives problem is in fact completely independent of the level of consumption and labor in this case since we have $\partial[du(n)/dn]/\partial z = 1/n^2$.

There is an important property worth noting for poverty indexes that satisfy $\lim_{x^* \rightarrow \infty} P_x(x, x^*) = 0 \forall x$, as the FGT indexes. For the ppb planner who uses such indexes, when the poverty line is very high (converging to infinity), its optimal tax rate converges to the welfarist optimal rate $\bar{t}(n) \forall n$. This result is quite intuitive, indeed for high poverty lines almost everybody in the society will be poor; hence altering the tax rates have slight effects on the aggregate poverty. The planner itself then feels powerless in front of poverty and thus prefers to concentrate on the intrinsic welfare side of the economy accepting the high level of the "public bad".

5.2 Numerical simulations

In the preceding section, a closed form solution to the optimal tax problem, when poverty is taken into consideration, has been derived. It corroborates the results found in the theoretical model.

However, some results are still stated in a general form, such as the negativity of some tax rates at the lower tail of the distribution. To have a more precise idea about the solution, we need to specialize even more the example and take explicit functions.

Let us assume that $v(x) = Ln(x)$, $P(x, x^*) = \left(\frac{x^* - x}{x^*}\right)^2$ and $f(n) = 5/6 - n/5$ for $n \in [1, 4]$. The example also assumes that all the agents have the same aversion to poverty $\beta(n) \equiv 1$ for all n . The poverty line has been set at $x^* = 2.5$. Let now see which tax schedules would be implemented by the welfarist and ppb planners in such a society. First of all, the aggregate poverty levels are of 0.279 and 0.063 respectively. Poverty has been sharply reduced, in fact it has been cut to a quarter of its initial level, under the ppb tax schedule. In the following graphics, the functions relative to the welfarist planner are in bold, while they are in dashed form for the ppb one. Figure 1 shows the consumption enjoyed by the agents under the different regimes while figure 2 displays the patterns of the marginal tax rates. As demonstrated in lemma 3, all the nonpoor under the welfarist regime have the same consumption and face the same marginal tax rate. Turning to the poor, the switch from the welfarist to the ppb regime translates into lower marginal tax rates at all levels which induces them to work harder and enjoy higher consumptions. In the lowest interval, they all face negative tax rates. Note however that the poor population is still the same even though aggregate poverty has been reduced. Finally, the curve of welfarist marginal tax rates represents also the curve of social distortions introduced by the ppb planner.

[See figure 1]

[See figure 2]

How are the additional resources needed to increase the consumption level of the poor generated? The answer to this question is provided by figure 3 which displays the labor supply functions. As predicted, the poor work much harder in the ppb regime compared to the welfarist one. More strikingly, the nonpoor whose consumption is maintained at the original level now

enjoy more leisure. This comes from the need to respect the incentive compatibility constraints. For the high ability persons not to claim the earnings subsidy package, the government has to increase their leisure at the same consumption level thereby increasing their intrinsic welfare.

[See figure 3]

Switching from the welfarist to the ppb schedule necessarily implies a tradeoff between aggregate poverty and intrinsic social welfare. Since poverty has decreased benefiting to all, one can expect the utility derived solely from the consumed bundle to decrease on average. Figure 5 shows that with respect to the welfarist schedule, the high skilled workers are those who gain in terms of intrinsic utility from the changing of policy while the lower part of the distribution loses in the process. Once true welfare is taken into account by introducing the negative effect of poverty, the gain from the welfarist to the ppb schedule profits to a larger part of the society and at a greater extent, see figure 6. However, the poorest of the poor are again the losers, which gives the flavor of a “repugnant conclusion”. In fact it is easy in this example to make the poorest benefit more in terms of true welfare by assuming that aversion to poverty is a decreasing function of ability instead of assuming it constant. Marginal tax rates, consumptions, supplies of labor are still the same if social aversion to poverty is maintained at $\beta = 1$. For instance $\beta(n)$ is decreasing if one thinks that poor areas are more prone to criminality and therefore the reduction in poverty will profit more to those areas in terms of reduced criminality.

[See figure 4]

[See figure 5]

Turning to the curves representing net individual transfers (figures 6 and 7), the most striking result in the simulations is that the individuals at the extremes of the distribution (the very poor and the very rich) pay taxes while those in the middle range receive positive net transfers. The

tax liability curve decreases first, confirming the negative marginal tax rates at the bottom and the inducement to work hard, before starting to increase. The tax burden of the rich has been much decreased and shifted on the poor who have to self-support. The welfarist tax function is the strict S predicted by Seade (1982).

[See figure 6]

[See figure 7]

As in Mirrlees (1971), the tax schedules are close to be linear. There is a whole bunch of bundles that were available under the welfarist schedule but are not under the ppb one. These bundles display low pretax and posttax incomes. They correspond to a high leisure, high poverty type of economy. On the contrary, the ppb planner allows only some minimum income to be earned. It is worth noting that the poor population is the same because of the class of poverty indexes considered which, as shown by Bourguignon and Fields (1997), implies transfers to the poorest of the poor. The welfarist tax schedule has thus been amended in such a way that the increase in consumption has been greater the poorer the agent. It cannot be optimal to push some poor out of the poverty trap unless this is done for all. Indeed, it is always, in terms of poverty reduction, more efficient to increase the consumption of a poorer individual.

[See figure 8]

6 Conclusion

This article has tried to tackle an optimal income tax problem in which the planner considers the comprehensive welfare of the individuals. This comprehensive welfare is intended to capture the multidimensional nature of welfare which is determined by variables the individuals have control upon and others that go beyond their control. The only relevant source of (dis)utility that falls into this second category is taken here to be aggregate poverty and its implied consequences.

Income poverty is assumed to be a negative aggregate externality that lowers, possibly differently, the utility of the agents.

In the complete information case, all non-poor people receive as usually a lump-sum income transfer while poor face a Pigouvian nonlinear earnings subsidy. Once ability becomes private information, the planner in assessing the nonlinear optimal tax schedule, follows the usual rule of equating the marginal social benefit to the marginal social cost of extracting a unit of earned income from any agent. The novelty is that poverty reduction now enters the marginal social benefit of transferring income to the poor. This can lead to the apparition of negative marginal tax rate at the lower end of the distribution of ability. The tax schedule then provides strong incentives to the less skilled people to work harder and help themselves to get out of poverty, and lessen by the same token the burden on the whole population. This tax scheme resembles a great deal to the Earned Income Tax Credit initiated in the United States in 1987 as an extension to the Tax Reform Act of 1986. As the EITC, the schedule that is proposed here creates complex labor supply responses. Numerical simulations corroborate the theoretical results that have been derived. While the emergence of negative marginal tax rates overturns the famous result of no distortion at the endpoints, it is shown that this result can be restored by focusing on social rather than individual distortions.

More generally, we expect the analysis to carry over to situations in which the comprehensive welfare includes many other sources of utility in addition to poverty. The major problem will be to find the way to welfarize each item and express true welfare in a tractable way.

Appendix A: Solution of the First Best

In this appendix, only the first-best problem of the ppb planner is considered. That of the welfarist is readily derived by setting $\beta = 0$.

Let us write the maximization program:

$$\begin{cases} \text{Max} \int_{\underline{n}}^{\bar{n}} \{u(x(n), y(n)) - \beta P(x(n), x^*)\} dF. \\ \text{subject to} \int_{\underline{n}}^{\bar{n}} (z(n) - x(n) - \bar{R}) dF \geq 0 \end{cases}$$

The following first order conditions are easily computed:

$$[1] u_x(x(n), y(n)) - \beta P_x(x(n), x^*) = \gamma, \quad [2] u_y(x(n), y(n)) = -n\gamma.$$

Finally, [3] $1 - s - \beta P_x(x(n), x^*)/\gamma = 0$, this last equation gives the size of the Pigouvian subsidy. We note $SW(x(n), y(n))$ the social welfare associated to a given allocation.

Proof of Proposition 1 (ii): At any optimal allocation $(x(n), y(n))$, the more productive individuals provide a higher supply of labor, i.e. $d\bar{y}(n)/dn \geq 0$.

It is assumed that the consumption and labor supply functions are differentiable. Let us assume that there exist an interval $[n_0, n_1]$ such that $\bar{y}'(n_0) = \bar{y}'(n_1) = 0$ and $\bar{y}'(n) < 0 \forall n \in]n_0, n_1[$, the less able work harder on this interval. Let $0 < \epsilon < \text{Min}\{f(n), n \in [n_0, n_1]\}$ be a real positive number that exists since $f(n) > 0 \forall n$. Consider now the allocation $(\hat{x}(n), \hat{y}(n))$, such that

1. $\forall n \in [\underline{n}, n_0 \cup]n_1, \bar{n}]$, $(\hat{x}(n), \hat{y}(n)) = (\bar{x}(n), \bar{y}(n))$,
2. for a proportion ϵ of $n \in [n_0, n_1]$, $(\hat{x}(n), \hat{y}(n)) = (\bar{x}(n), \bar{y}(n))$,
3. for a proportion ϵ of $n \in [n_0, n_1]$, $(\hat{x}(n), \hat{y}(n)) = (\bar{x}(n_0 + n_1 - n), \bar{y}(n_0 + n_1 - n))$.

2 and 3 are possible because of the non anonymity of the first best. The planner can then assign different allocations to otherwise identical agents.

What is the social welfare level? After straightforward but tedious computations one can show that

$$\begin{aligned} SW(\hat{x}(n), \hat{y}(n)) &= SW(\bar{x}(n), \bar{y}(n)) \\ &+ \epsilon \int_{n_0}^{n_1} [u(\bar{x}(n_0 + n_1 - n), \bar{y}(n_0 + n_1 - n)) - u(\bar{x}(n), \bar{y}(n))] dn \\ &- \beta \epsilon \int_{n_0}^{n_1} [P(\bar{x}(n_0 + n_1 - n), x^*) - P(\bar{x}(n), x^*)] dn \end{aligned}$$

Let us operate a change of variable by posing $m = n_0 + n_1 - n$. It is then easy to show that the last two terms vanish. The social welfare is thus the same with both allocations. One could have anticipated this since the social welfare function is symmetric in the individuals true welfare. The aggregate consumption also is the same thanks to the symmetry. What about aggregate produced income?

$$\begin{aligned} \int_{\underline{n}}^{\bar{n}} n \hat{y}(n) f(n) dn &= \int_{\underline{n}}^{n_0} n \bar{y}(n) f(n) dn + \int_{n_1}^{\bar{n}} n \bar{y}(n) f(n) dn \\ &\quad + \int_{n_0}^{n_1} n \bar{y}(n) [f(n) - \epsilon] dn + \epsilon \int_{n_0}^{n_1} n \bar{y}(n_0 + n_1 - n) dn \\ &= \int_{\underline{n}}^{\bar{n}} n \bar{y}(n) f(n) dn + \epsilon \int_{n_0}^{n_1} n [\bar{y}(n_0 + n_1 - n) - \bar{y}(n)] dn \end{aligned}$$

It is now sufficient to consider the last term above, we denote $(*)$. Note also $\bar{n} = (n_0 + n_1)/2$.

$$\begin{aligned} (*)/\epsilon &= \int_{n_0}^{\bar{n}} n [\bar{y}(n_0 + n_1 - n) - \bar{y}(n)] dn + \int_{\bar{n}}^{n_1} n [\bar{y}(n_0 + n_1 - n) - \bar{y}(n)] dn \\ &= \int_{n_0}^{\bar{n}} n [\bar{y}(n_0 + n_1 - n) - \bar{y}(n)] dn + \int_{\bar{n}}^{n_0} (n_0 + n_1 - n) [\bar{y}(n) - \bar{y}(n_0 + n_1 - n)] (-dn) \\ &= 2 \int_{n_0}^{\bar{n}} (\bar{n} - n) [\bar{y}(n) - \bar{y}(n_0 + n_1 - n)] dn \end{aligned}$$

$\bar{y}(n) - \bar{y}(n_0 + n_1 - n) > 0$ since $\bar{y}'(n) < 0$ on $]n_0, n_1[$. This implies that $(*) > 0$ and thus $(\bar{x}(n), \bar{y}(n))$ cannot be optimal. ■

Proof of Proposition 1 (iii): For efficiency reasons, whatever the objective of the planner, the more able always work harder. How is then the produced revenue redistributed among the consumers? From the first order condition [1]:

$$u_x(\bar{x}(n), \bar{y}(n)) - \beta P_x(\bar{x}(n), x^*) = u_x(\bar{x}(n'), \bar{y}(n')) - \beta P_x(\bar{x}(n'), x^*)$$

After some straightforward computations, this equation becomes

$$\int_{\bar{y}(n')}^{\bar{y}(n)} u_{xy}(\bar{x}(n'), Y) dY = \int_{\bar{x}(n')}^{\bar{x}(n)} [\beta P_{xx}(X, x^*) - u_{xx}(X, \bar{y}(n))] dX$$

Take without loss of generality $n > n'$. Since $P_{xx} \geq 0$ and $u_{xx} < 0$ the integrand at the right hand side is positive. Therefore

$$dx(n)/dn \cdot u_{xy} \geq 0,$$

i.e. consumption is decreasing if it is Edgeworth-complement with leisure, constant if preferences are separable and increasing if it is Edgeworth-substitute with leisure.

Appendix B: Solution of the non-linear income taxation

The resolution of program (PPB) is achieved by using the maximum principle. The state and control variables must then be defined. We choose the gross income $z(n)$ as the control variable. The indirect utility $u(n)$ will serve as the state variable. The resource constraint is directly introduced in the objective function. Since the utility function is monotonic, it can be inverted to get the consumption $x(n)$ as a function of $u(n)$ and $z(n)$, i.e. there exists a function Ψ such that $x = \Psi(u, z)$. It is easy to show that

$\Psi_z = -U_x/U_x = s$ and $\Psi_u = 1/U_x$ and $x(n)$ is implicitly taken into account in the computations. The program of the government is therefore:

$$\text{Max} \int_{\underline{n}}^{\bar{n}} [u(n) - \beta P(x(n), x^*) + \gamma(z(n) - x(n))] dF$$

subject to

$$\frac{du}{dn} = U_n(x(n), z(n), n)$$

The Hamiltonian is then:

$$H(x, z, w, n) = [u(n) - \beta P(x(n), x^*) + \gamma(z(n) - x(n))]f(n) + \mu(n)U_n(x(n), z(n), n) \quad (12)$$

where μ and γ are the multipliers associated to the different constraints, the first order conditions are:

$$\begin{aligned} \frac{\partial H}{\partial u} &= [1 - \beta P_x(\cdot, x^*)\Psi_u + \gamma \frac{\partial}{\partial u} \{z(n) - x(n)\}]f(n) + \mu \frac{\partial U_n}{\partial u} = -\mu' \\ \frac{\partial H}{\partial z} &= [-\beta P_z(\cdot, x^*)\Psi_z + \gamma \frac{\partial}{\partial z} \{z(n) - x(n)\}]f(n) + \mu \frac{\partial U_n}{\partial z} = 0 \\ \mu(\underline{n}) &= \mu(\bar{n}) = 0. \end{aligned}$$

The last conditions are the transversality conditions. Rearranging the first order conditions, one ends up with the, slightly modified, traditional equations:

$$-\mu(n)U_x s_n + (\gamma(1 - s(n)) - \beta P_x(\cdot, x^*)s(n))f(n) - \phi'(n) = 0, \quad (B.1)$$

$$\mu'(n) + \mu(n) \frac{U_{nz}}{U_x} + (1 - (\beta P_x(\cdot, x^*) + \gamma)/U_x)f(n) = 0 \quad (B.2)$$

with the endpoint conditions $\mu(\underline{n}) = \mu(\bar{n}) = 0$,

where $(1 - s(n))$ represents the marginal tax rate which expression can be obtained from the first condition. The second condition is a differential equation from which we can derive the expression of the multiplier. Let us write the expression of the marginal tax rate:

$$t(z(n)) = (1 - s(x(n), z(n), n)) = \frac{\beta P_x s}{\gamma} + \frac{\mu(n)U_x s_n}{\gamma f} \quad (B.3)$$

Transforming the first condition, we obtain another differential equation:

$$\mu'(n) + \mu(n) \frac{U_{nz}}{U_x} + (1 + \gamma/U_x)f(n) = 0 \quad (B.4)$$

solving these differential equations directly gives us the following two expressions for μ :

$$\mu(n) = \int_n^{\bar{n}} \left(1 - \frac{(\beta P_x + \gamma)}{U_x}\right) \exp\left(\int_n^p (U_{nx}/U_x) dm\right) f(p) dp \quad (B.5)$$

$$\mu(n) = \int_n^{\bar{n}} \left(1 + \frac{\gamma}{U_x}\right) \exp\left(\int_n^p (U_{nz}/U_x) dm\right) f(p) dp \quad (B.6)$$

in the first equation the poverty measure appears since it concerns consumption, but one can immediately see that this expression vanishes in the equation concerning z the earned income which is "independent" of poverty.

Appendix C: Solution for the quasi linear case

A closed form solution to the optimal tax problem is provided in this appendix. The consumers' preferences are taken to be quasi linear of the following form:

$$u(x, y) = v(x) - y$$

where for simplicity everyone is assumed to be equally averse to poverty. The utility function is indexed by the ability and rewritten:

$$V(x, z, n, \mathcal{P}(x^*)) = U(x, z, n) - \beta \mathcal{P}(x^*) = v(x) - \frac{z}{n} - \beta \mathcal{P}(x^*)$$

for later reference we have:

$$U_x = v'(x), U_z = -\frac{1}{n}, s = \frac{1}{nv'(x)}, s_n = -\frac{1}{n^2v'(x)}, U_{nx} = 0, U_{nz} = \frac{\partial}{\partial z} \left[\frac{du(n)}{dn} \right] = \frac{1}{n^2}$$

The planner's objective function is:

$$\text{Max} \int_{\underline{n}}^{\bar{n}} [u(n) - \beta \mathcal{P}(x(n), x^*)] dF(n)$$

under the incentive compatibility and revenue constraints. Using equation (B.6), neglecting the term relative to bunching, and the transversality conditions, it is easy to compute:

$$\gamma = \int_{\underline{n}}^{\bar{n}} \frac{f(n)}{n} dn = G(\bar{n})$$

using again (B.6), one obtains

$$\mu(n) = n[G(\bar{n})F(n) - G(n)]$$

(12.3) is easily obtained from (B.1), replacing all the terms by their explicit expressions. Hence:

$$\frac{\beta P_x(x(n), x^*) + G(\bar{n})}{v'(x(n))} = \frac{nf(n)G(\bar{n}) - G(n) + G(\bar{n})F(n)}{f(n)}$$

The consumption of the workers is derived from this equation. However, due to the poverty index which contains the consumption level, one cannot get the consumption as a function of n , f and v only. The level of consumption for the different individuals must be computed once f and v are specified. Afterwards, from (B.3), replacing s by $1 - t$, we derive the marginal tax rate

$$t(n) = \frac{\beta P_x(x(n), x^*)}{\beta P_x(x(n), x^*) + G(n)} + \frac{G(n) - G(\bar{n})F(n)}{nf(n)[\beta P_x(x(n), x^*) + G(\bar{n})]}$$

To derive the labor supply function, it is useful to remind the first order condition

$$U_x x' + U_z z' = 0 \implies nv'(x(n))x'(n) = z'(n) \implies z(n) = \int z'(m)dm - K_1$$

assuming that the government taxes for redistribution, $R = 0$, the revenue constraint is rewritten:

$$\int_{\underline{n}}^{\bar{n}} z(n)f(n)dn = \int_{\underline{n}}^{\bar{n}} x(n)f(n)dn$$

that is

$$\int_{\underline{n}}^{\bar{n}} \left[\int_{\underline{n}}^n mv'(x(m))x'(m)dm - K_1 \right] f(n)dn = \int_{\underline{n}}^{\bar{n}} x(n)f(n)dn$$

$$K_1 = \int_{\underline{n}}^{\bar{n}} \left[\int_{\underline{n}}^n mv'(x(m))x'(m)dm - x(n) \right] f(n)dn$$

$$\begin{aligned} K_1 &= \int_{\underline{n}}^{\bar{n}} \left[\int_{\underline{n}}^n mv'(x(m))x'(m)dm - x(n) \right] f(n)dn \\ &= \int_{\underline{n}}^{\bar{n}} \left[nv(x(n)) - \int_{\underline{n}}^n v(x(m))dm - x(n) \right] f(n)dn - \underline{n}v(x(\underline{n})) \\ &= \int_{\underline{n}}^{\bar{n}} [(nf(n) - 1 + F(n))v(x(n)) - x(n)f(n)] dn - \underline{n}v(x(\underline{n})) = K - \underline{n}v(x(\underline{n})) \end{aligned}$$

Finally, it is straightforward to compute

$$z(n) = nv(x(n)) - \int_{\underline{n}}^n v(x(m))dm - K,$$

since $u(n) = v(x(n)) - \frac{z(n)}{n}$ we get

$$u(n) = \frac{1}{n} \left(K + \int_{\underline{n}}^n v(x(m))dm \right)$$

and

$$z(n) = n(v(x(n)) - u(n))$$

Proof of Lemma 3: To obtain the welfarist planner tax schedule, it is sufficient to take $\beta = 0$ in the above solution. Let us denote the optimal ppb planner solution with a 'hat' and that of the welfarist with a 'bar'. Let us first prove lemma 3(i). Denoting the rhs of (12.3) by $\theta(n)$, we have

$$\frac{\beta P_x(\hat{x}(n), x^*) + G(\bar{n})}{v'(\hat{x}(n))} = \theta(n) = \frac{G(\bar{n})}{v'(\bar{x}(n))},$$

all the functions in the rhs of (12.3) are well defined and therefore $\theta(n)$ is single-valued. From the above equation we deduce:

$$\frac{v'(\hat{x}(n))}{v'(\bar{x}(n))} = \frac{\beta P_x(\hat{x}(n), x^*) + G(\bar{n})}{G(\bar{n})} \quad \text{for all } n$$

The argument n is omitted in the following without any risk of confusion. Suppose $\bar{x} > x^* > \hat{x}$ for a given agent, this implies $v'(\bar{x}) < v'(\hat{x})$ from the concavity of v . Therefore, the lhs of the above equation is greater than 1. Coming to the rhs, $P_x(\hat{x}, x^*) < 0$ from our assumption, hence the rhs is less than 1. Contradiction. Therefore whenever an agent is non-poor under the welfarist schedule, she is also non-poor under the ppb schedule. Moreover since v is monotonic, her consumption does not change, because the rhs is equal to 1. Suppose now $\bar{x} < x^* < \hat{x}$, implying the lhs is less than 1, we have the same contradiction since the rhs is equal to 1. Therefore, poor people under the welfarist schedule remain poor under the ppb one. Since $\hat{x} < x^*$, the rhs is less than 1, again using the concavity of v , this means that $\hat{x} > \bar{x}$, i.e. even though poor remain poor, their consumption is increased. The same type of argument is used to prove lemma 3 (ii).

Acknowledgements

I am highly indebted to Helmuth Cremer for his many insightful comments and suggestions. For their helpful comments, I thank Jonathan Eaton, Efraim Sadka, the seminar participants at Aix-Marseille, Toulouse and the 1998 IIPF congress (Cordoba). I would especially like to express my gratitude to Georges Casamatta and Francesca Barrigozzi for the many discussions we had. The usual disclaimer applies.

References

- Atkinson, A. B., 1987. On the Measurement of Poverty. *Econometrica* 55, 749–764.
- Atkinson, A. B., 1995a. *Public Economics in Action*. Clarendon Press Oxford.
- Atkinson, A. B., 1995b. On Targeting Social Security: Theory and Western Experience with Family Benefits. *Public Spending and the Poor: Theory and Evidence*. in van de Walle, D., Nead, K. (Ed), John Hopkins University Press.

- Besley, T., Coate, S., 1992. Workfare vs. welfare: incentive arguments for work requirements in poverty alleviation programs. *American Economic Review* 82, 249–261.
- Besley, T., Coate, S., 1995. The Design of Income Maintenance Programmes. *Review of Economic Studies* 62, 187–221.
- Bourguignon, F., Fields, G., 1997. Discontinuous losses from poverty, generalized P_α measures, and optimal transfers to the poor. *Journal of Public Economics* 63, 155–175.
- Brito, D. L., Oakland, W.H., 1977. Some properties of the optimal income-tax. *International Economic Review*, 18:407–423.
- Cremer, H., Gahvari, F., Ladoux, N., 1998. Externalities and optimal taxation. *Journal of Public Economics* 70, 343–364.
- Diamond, P., 1998. Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review* 88, 83–95.
- Ebert, U., 1992. A reexamination of the optimal nonlinear income tax. *Journal of Public Economics* 49, 47–73.
- Eissa, N., Liebman, J.B., 1996. Labor supply responses to the earned income tax credit. *Quarterly Journal of Economics* CXI, 605–637.
- Foster, J.E., Greer, A.F., Thorbecke, E., 1984. A Class of Decomposable Poverty Measures. *Econometrica* 52, 761–766.
- Guesnerie, R., Laffont, J.J., 1984. A complete solution to the principal-agent problem with an application to the control of a self-managed firm. *Journal of Public Economics* 25, 329–369.
- Kanbur, R., Keen, M., Tuomala, M., 1994. Optimal nonlinear income taxation for the alleviation of income poverty. *European Economic Review* 38, 1613–1632.
- Lollivier, S., Rochet, J.C., 1983. Bunching and second order conditions. A note on optimal tax theory. *Journal of Economic Theory* 31, 392–400.
- Mirrlees, J.A., 1971. An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38, 175–208.
- Mirrlees, J.A., 1986. The theory of optimal taxation. *Handbook of Mathematical Economics*, vol. III, Arrow, K.J., Intriligator, M., (Ed).
- Ravallion, M., 1994. *Poverty Comparisons*. Chur, Switzerland: Harwood Academic Press *Fundamentals of Pure and Applied Economics* Volume 56.
- Sadka, E., 1976. Social welfare and income distribution. *Econometrica* 44, 1239–51.
- Seade, J.K., 1977. On the shape of optimal tax schedules. *Journal of Public Economics* 7, 203–235.
- Seade, J.K., 1979. Optimal nonlinear policies for non-utilitarian motives. University of Warwick, Working paper 143.
- Seade, J.K., 1982. On the sign of the optimum marginal income tax. *Review of Economic Studies* 49, 637–643.
- Sen, A.K., 1982. *Choice, welfare and measurement*. John Hopkins University Press.

Sugden, R., 1993. Welfare, resources and capabilities. *Journal of Economic Literature*, XXXI, 1947–1962.

Zeckhauser, R.J., 1971. Optimal mechanisms for income transfer. *American Economic Review* 61, 324–34.



Figure 1: Consumption Functions under the Different Regimes

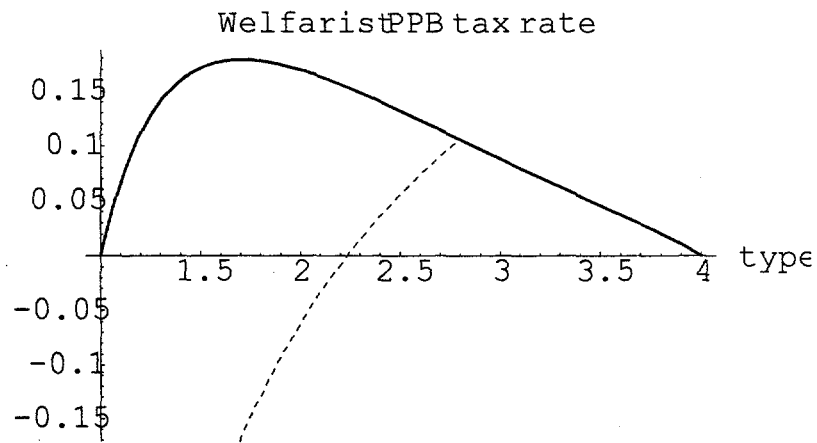


Figure 2: Marginal Tax Rates under the Different Schedules

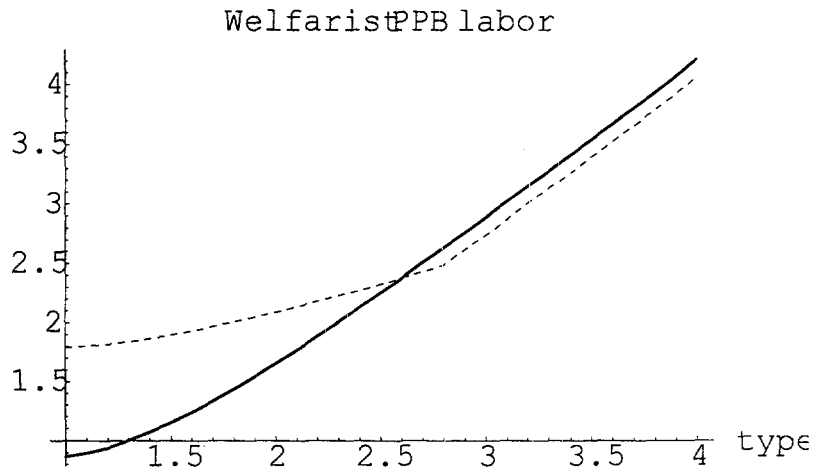


Figure 3: Labor Supply Functions

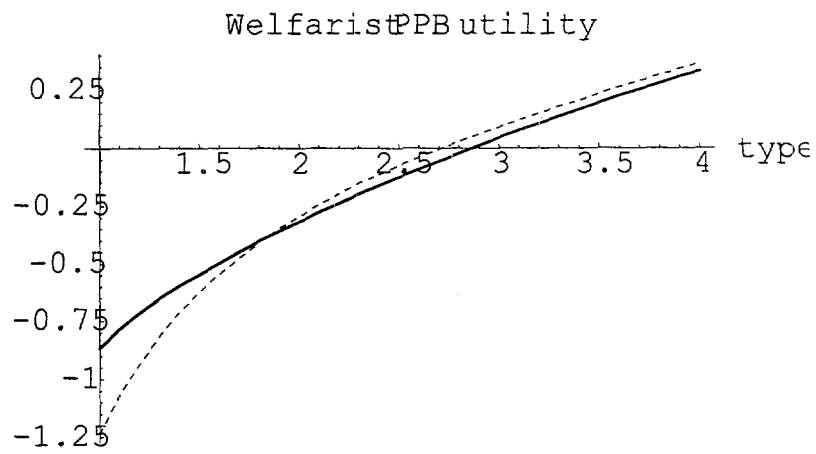


Figure 4: The Intrinsic Welfare Levels

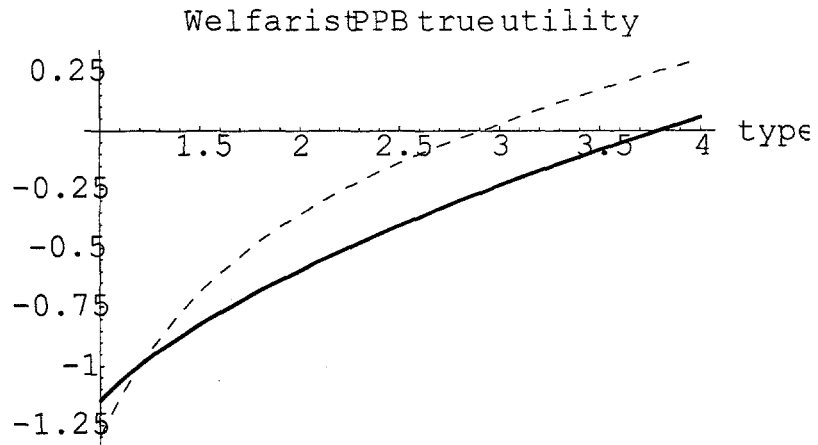


Figure 5: The True Welfare Levels

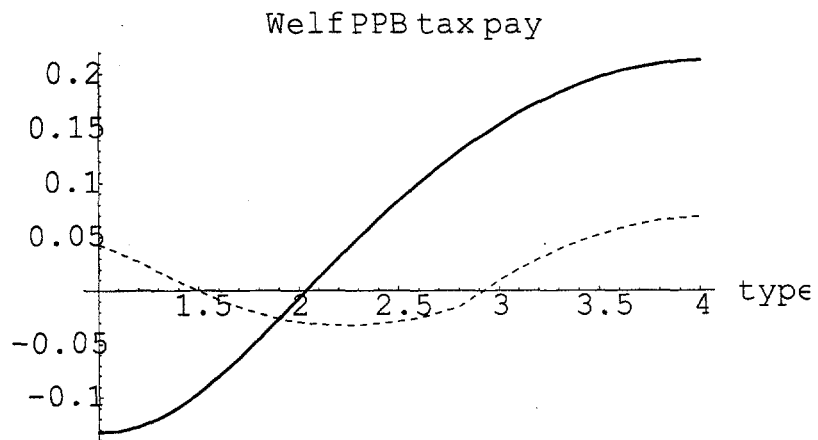


Figure 6: Welfarist and ppb Tax Liabilities by Type

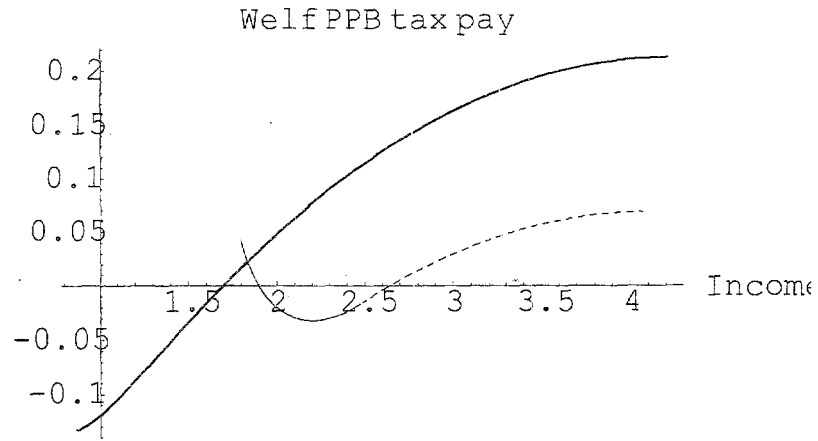


Figure 7: Welfarist and ppb Tax Liabilities by Earned Income

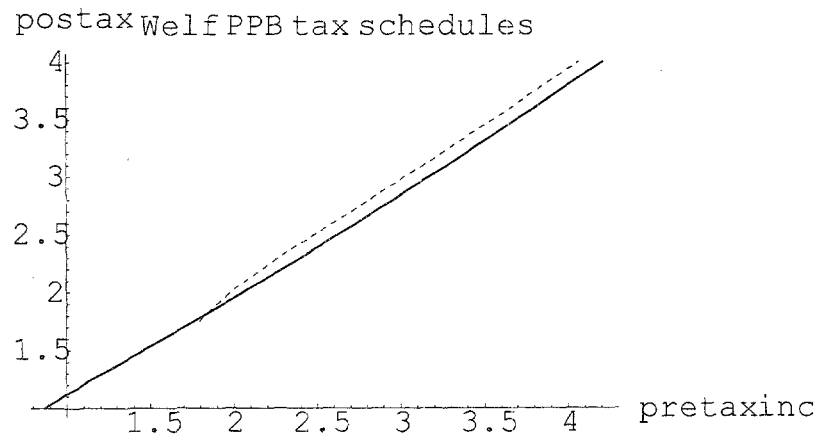


Figure 8: The Welfarist and ppb Tax Schedules

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2252	Productivity Growth, Capital Accumulation, and the Banking Sector: Some Lessons from Malaysia	Ejaz Ghani Vivek Suri	December 1999	N. Mensah 80546
WPS2253	Revenue Recycling and the Welfare Effects of Road Pricing	Ian W. H. Parry Antonio Miguel R. Bento	December 1999	R. Yazigi 37176
WPS2254	Does "Grease Money" Speed Up the Wheels of Commerce?	Daniel Kaufmann Shang-Jin Wei	December 1999	H. Sladovich 37698
WPS2255	Risk and Efficiency in East Asian Banks	Luc Laeven	December 1999	R. Vo 33722
WPS2256	Geographical Disadvantage: A Heckscher-Ohlin-von Thunen Model of International Specialization	Anthony J. Venables Nuno Limão	December 1999	L. Tabada 36896
WPS2257	Infrastructure, Geographical Disadvantage, and Transport Costs	Nuno Limão Anthony J. Venables	December 1999	L. Tabada 36896
WPS2258	Market Access Bargaining in the Uruguay Round: Rigid or Relaxed Reciprocity?	J. Michael Finger Ulrich Reincke Adriana Castro	December 1999	L. Tabada 36896
WPS2259	Predicting Currency Fluctuations and Crises: Do Resident Firms Have an Informational Advantage?	Daniel Kaufmann Gil Mehrez Sergio Schmukler	December 1999	E. Khine 37471
WPS2260	Regional Integration Agreements: A Force for Convergence or Divergence?	Anthony J. Venables	December 1999	L. Tabada 36896
WPS2261	Is Knowledge Shared within Households?	Kaushik Basu Ambar Narayan Martin Ravallion	December 1999	M. Mason 30809
WPS2262	How Inadequate Provision of Public Infrastructure and Services Affects Private Investment	Ritva Reinikka Jakob Svensson	December 1999	H. Sladovich 37698
WPS2263	When is Growth Pro-Poor? Evidence from the Diverse Experiences of India's States	Martin Ravallion Gaurav Datt	December 1999	J. Israel 85117
WPS2264	Do More Unequal Countries Redistribute More? Does the Median Voter Hypothesis Hold?	Branko Milanovic	December 1999	P. Sader 33902

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2265	The Political Economy of Distress In East Asian Financial Institutions	Paola Bongini Stijn Claessens Giovanni Ferri	January 2000	R. Vo 33722
WPS2266	The Impact of Adult Deaths on Children's Health in Northwestern Tanzania	Martha Ainsworth Innocent Semali	January 2000	S. Fallon 38009
WPS2267	Do High Interest Rates Defend Currencies during Speculative Attacks?	Aart Kraay	January 2000	R. Bonfield 31248
WPS2268	The Structure of Social Disparities In Education: Gender and Wealth	Deon Filmer	January 2000	S. Fallon 38009
WPS2269	Context Is Everything: Measuring Institutional Change in Transition Economies	Nauro F. Campos	January 2000	J. Victor 36549

1-19