TAXES AND EMPLOYMENT SUBSIDIES IN OPTIMAL REDISTRIBUTION PROGRAMS

Paul Beaudry
Charles Blackorby

Working Paper 6355

NBER WORKING PAPER SERIES

TAXES AND EMPLOYMENT SUBSIDIES IN OPTIMAL REDISTRIBUTION PROGRAMS

Paul Beaudry
Charles Blackorby

Working Paper 6355
http://www.nber.org/papers/w6355

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1998

We thank Tony Atkinson, Tim Besley, Richard Blundell, Craig Brett, Louis-André Gerard-Varet, Roger Guesnerie, Jon Kesselman, Thomas Piketty, Michel Poitevin, John Weymark and seminar participants at UBC, CIAR, Ecole Nationale des Ponts et Chausées, The Institute for Fiscal Studies, and Paris I for comments. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

Taxes and Employment Subsidies in Optimal Redistribution Programs
Paul Beaudry and Charles Blackorby
NBER Working Paper No. 6355
January 1998
JEL Nos. D82, H21, H23

ABSTRACT

This paper characterizes an optimal redistribution program when taxation authorities: (1) are uninformed about individuals' value of time in both market and non-market activities, (2) can observe both market-income and time allocated to market employment, and (3) are utilitarian. Formally, the problem is a special case of a multidimensional screening problem with two dimensions of unobserved attributes. In contrast to much of the optimal income taxation literature, we show that optimal redistribution in this environment involves distorting market employment upwards for low net-income individuals (through negative marginal income taxes or employment subsidies) and distorting employment downward for high net-income individuals (through positive marginal income taxes). It is also shown that workfare is only part of an optimal program if certain individuals have no access to market employment.

Paul Beaudry
Department of Economics
University of British Columbia
997-1873 East Mall
Vancouver, BC V6T 1Z1
CANADA
and NBER
beaudry@econ.ubc.ca

Charles Blackorby
Department of Economics
University of British Columbia
997-1873 East Mall
Vancouver, BC V6T 1Z1
CANADA

1. Introduction

In most countries income redistribution is achieved through a variety of programs: these include direct income taxation, employment programs, welfare and even possibly unemployment insurance. Viewed as a whole, these programs create intricate incentives and complex redistribution patterns. Reasoned economic policy should attempt to identify whether or not these programs are reasonably designed. That is, are they mutually coherent as well as being consistent with the goal of redistribution. For example, are high marginal tax rates on the poor a sign of an efficient or inefficient system? Is it justifiable to subsidize wages of low income earners? Are workfare programs desirable? The object of this paper is to provide further insights on these and related issues.

Our approach to the problem follows the optimal non-linear income taxation literature as pioneered by Mirrlees (1971),[1] that is, we approach redistribution as a welfare maximization problem constrained by informational asymmetries. However, we depart from Mirrlees formulation in order to focus on certain preoccupations and possibilities currently facing tax authorities. In particular, we address two issues. The first concerns the perceived need to target more effectively income transfers. For example, traditional welfare programs (or minimum revenue guarantees) are often criticized on the grounds that they transfer substantial income to individuals who value highly their non-market time, as opposed to transferring income only to the most needy. Although such a preoccupation is common, the literature is mostly mute on how to address this issue since the standard framework assumes that individuals value their non-market time identically. The second issue relates to the possibility of using work time requirements as a means of targeting transfers. Many social programs (such as most unemployment insurance programs) employ information on time worked in order to determine eligibility;[2] therefore it seems reasonable to allow for such a possibility when considering how best to redistribute income. Hence, the environment we examine is one where (1) taxation authorities are uninformed about individuals' potential value of time in market activities and about their potential value of time in non-market activities[3], and (2) income transfers can be contingent on both earned (market) income and on the allocation

[1] See also Mirlees 1997).

[2] In Canada, there is currently a large scale experiment aimed at encouraging welfare recipients to work, the self-sufficiency project (see Card and Robins (1996) for details). One particular aspect of this program is that it explicitly requires individuals to work 30 hours per week in order to be eligible for a transfer; recipients are required to mail in pay stubs showing their hours of work and earnings for the month.

[3] In our formulation, non-market activities can be interpreted as non-declared market activities.

of time to market employment.[4] Under the above assumptions, our redistribution problem formally becomes a multidimensional screening problem with two-dimensions of unobserved characteristics.[5]

Given the two-dimensional informational asymmetry, it is not surprising that the properties of the optimal redistribution program derived under our informational and observability assumptions are quite distinct from those found in the standard setup. More specifically, we show that optimal redistribution in our environment entails

- distorting upwards the employment level of low net-income earners through negative marginal tax rates, or equivalently, employment subsidies,;
- distorting downward the employment level of high net income earners through positive marginal tax rates;
- using public employment requirements (workfare) as a means to transfer income to individuals with very poor or non-existent market possibilities.

In addition to these general properties, we also indicate why certain counterintuitive outcomes can arise in optimum. For example, we discuss why, given two individuals earning the same gross income, it may be optimal to tax more heavily the individual who requires more time to earn this income, that is, the individual with the lower wage rate.

The above results provide a stark contrast with those of the non-linear taxation literature in large measure because there the informational asymmetry is restricted to the value of market time.[6] Recall that the main prescriptions derived by Mirrlees are that

- marginal tax rates be everywhere non-negative, and
- there be a zero marginal tax rate on the most productive individual(s).[7]

In order to facilitate comparison of our results with the standard ones we consider in Section 3 two special cases: perfect negative coorelation between the two sets of characteristics and perfect positive correlation between them. The former leads to an undistorted

[4] Dasgupta and Hammond (1980) and Maderner and Rochet (1995) also examine optimal redistribution in environments where taxation authorities can transfer income based on market-income and market allocation of time. However, in these papers there remains only one dimension of unobserved characteristics. See also Kesselman (1973) and Bloomquist (1981) for a related literature.

[5] Screening problems with two-dimension of unobserved characteristics are rather uncommon in the literature. See Rochet (1995) for an overview of this literature and a discussion of some of the difficulties associated with solving such problems.

[6] Preferences in standard treatments are restricted to the single-crossing property without which virtually nothing can be said. We use an assumption that can be viewed as a strong version of the single-crossing property.

[7] This assumes that such an individual exists (Mirlees (1997)). It has also been shown that, if the least productive individual is employed under the optimal scheme, then he or she also faces a zero marginal tax rate. However, this case is generally not considered to be very relevant since the less productive individual generally does not work under the optimal scheme.

solution that is similar to that obtained by Dasgupta and Hammond (1980); the latter leaves the highly skilled undistorted but highly taxed while providing distorting wage subsidies for the low end of distribution.

Since this seminal contribution, Mirrlees' analysis has been extended in several direction.[8] In particular, Besley & Coates (1995) have recently shown that it is not optimal to complement the optimal non-linear taxation schedule with workfare, that is, it is never efficient to make income transfers contingent on public employment.[9] One surprising aspect of these properties is how they appear to conflict with many of the current policies debates which, *de facto*, tend to favor active employment programs such as employment subsidies (negative marginal taxation) and workfare.[10] Hopefully, this paper shed new light on such policy debates.

The paper is structured as follows. In Section 2 we present the information constrained redistribution problem and derive properties of the associated optimal direct revelation mechanism. In Section 3 we allow the sets of characteristics to be correlated and in Section 4 we sketch the proof of the main proposition.[11] In Section 5 we derive the tax rate implications of the results derived in Section 2. We then discuss the implications of our results for workfare and other related issues.

2. The Environment and the Pattern of Second Best Distortions

The economy has two sectors—a formal market sector and an informal, non-market or household sector. An individual can work in the formal/market sector at a wage rate no greater than his or her intrinsic productivity. Income earned in the formal/market sector can be observed and hence taxed. The amount of time allocated to the formal/market sector can also be observed. Besides working in the formal/market sector, an individual can also allocate time to production in the informal/household sector.[12] Production in this sector

[8] Many of the extensions of Mirrlees original analysis involve giving more tools to the taxation authorities. For example, see Guesnerie and Roberts (1987) or Marceau and Boadway (1994).

[9] This is not the main emphasis of this Besley and Coates paper. In fact, their main result is to show that workfare can be part of an optimal program when income maintenance is the objective as opposed to welfare maximization. They view their result as mainly providing a positive theory of workfare as opposed to a normative prescription. See also Kanbur, Keen and Tuomola (1994) on this issue.

[10] It is precisely this observation which motives the work by Besley and Coates (1995) and Kanbur, Keen and Tuomola (1994).

[11] Otherwise, all proofs are relegated to the Appendix.

[12] In any case, a non tax-paying sector.

normalized to one; if individual i works for $h_i \geq 0$ hours in the formal sector, he or she has $1 - h_i$ hours available for producing goods in the informal/household sector.[14] Individuals have identical utility functions that are known and which depend upon the consumption of goods from both sectors of the economy. Individuals differ in their abilities and the ability level can vary across sectors. For example, one may be very productive in the formal/market sector but have low productivity in the informal/household sector or conversely.

Types are indexed by $i \in I = \{1, \ldots, NM\}$ where for each type there is a two-tuple (ω_i, θ_i); where $\omega_i \in \{\omega_1, \ldots, \omega_N\}$ is the productivity index of individual of type i in the formal/market sector and $\theta_i \in \{\theta_1, \ldots, \theta_M\}$ is the productivity index in the informal/household sector. We assume that $\theta_i > 0$ and, for simplicity, we assume that $\omega_i \neq \theta_i$ for all i. The percentage of the population that is type i is denoted p_i and we impose no restrictions on the distribution of types.[15] Consumption in the formal sector good is denoted c_i and is referred to as an individual's net (after tax) income. An individual's pre-tax income is denoted y_i, where pre-tax income is $y_i = h_i w_i$ with w_i being the wage rate received by type i in the formal/market sector. Individuals evaluate their well-being by means of a utility function, $U : \mathbf{R}_+ \mapsto \mathbf{R}$, which is defined on total consumption and assumed to be concave and differentiable. Individuals i's total consumption is given by $c_i + (1 - h_i)\theta_i$. Implicit in this formulation is the assumption that the goods from the two sectors are perfect substitutes and that the production technology is linear. Obviously these assumptions are restrictive, but appear to us as reasonable starting point for the analysis of a problem with two dimensions of heterogeneity.

An allocation in this economy is a mapping that associates with every type (ω_i, θ_i) a triplet composed of (1) a consumption level for the formal/market sector good, (2) the hours supplied in the formal/market sector and (3) the wage rate in formal sector employment (or alternatively the income in the formal sector). Therefore an allocation in this economy corresponds to a sequence of the form $\{c_i, h_i, w_i\}_{i=1}^{NM}$. A particular element of this sequence, say (c_j, h_j, w_j), is referred to as j's allocation.

---

13 Alternatively, production in the household sector can be viewed as leisure with individuals having different tastes for leisure.

14 It is assumed throughout the paper that the agent can choose how to allocate his or her time endowment in a continuous fashion. However, the results of this paper can be generalized to the case where h can only take on a discrete set of values (for example 0 and 1). The results generalize quite easily because the proofs do not exploit the continuous nature of the labor supply decision.

15 See Diamond for results obtained in the Mirlees' framework when the the distribution of types is constrained.

The government's objective is to maximize a utilitarian social welfare function[16] but is unable to implement a first-best optimum due to the asymmetry of information. In particular, it is assumed that the government cannot observe skill levels of individuals in either sector, that is, the government cannot observe either ω_i or θ_i. Under the above assumption, the government's maximization problem can be stated as follows.[17]

An optimal allocation is a $3NM$-tuple $\{\overset{*}{c}_i, \overset{*}{h}_i, \overset{*}{w}_i\}_{i=1}^{NM}$ that maximizes

$$\sum_{i=1}^{NM} p_i U\big(c_i + (1-h_i)\theta_i\big) \tag{2.1}$$

subject to

$$\sum_{i=1}^{NM} p_i c_i \leq \sum_{i=1}^{NM} p_i w_i h_i, \tag{2.2}$$

and for all i

$$U\big(c_i + (1-h_i)\theta_i\big) \geq U\big(c_j + (1-h_j)\theta_i\big), \quad \forall j \quad \text{s.t.} \quad w_j \leq \omega_i \tag{2.3}$$

$$U\big(c_i + (1-h_i)\theta_i\big) \geq U\big(\theta_i\big), \tag{2.4}$$

$$0 \leq h_i \leq 1 \tag{2.5}$$

In the above problem, (2.2) represents the materials balance constraint, (2.3) represents the incentive compatibility constraints and (2.4) represents the individual participation constraints. Since the incentive compatibility constraints in this problem are not standard, some clarification is in order. The implicit assumption in the above formulation is that an individual can costlessly mimic any other individual who has a lower market productivity, that is, individual i can choose to be employed in any job paying a wage $w \leq \omega_i$. In effect, the incentive compatibility constraint (2.3) insures that individual i finds his or her allocation at least as good as that of any agent employed at a wage no greater than his or her own market productivity ω_i. An individual's wage rate is observable but his or her inherent(potential) productivity is not. Such a distinction seems reasonable and has important consequences on the set of incentive compatible allocations.[18]

In fact, the government's maximization problem can be simplified by restricting attention to allocations where w_i is set equal to ω_i. This property is stated in the following Proposition along with a set of minimal restrictions on an optimal allocation.

[16] The results of this paper do not depend on a strict utilitarian perspective, but do depend on welfarism.

[17] Our formulation of the problem, which is standard in the taxation literature, restricts the government to a direct mechanism where individual allocations depend only on their own announcements.

[18] This formulation of incentive compatibility constraints has been examined by Dasgupta and Hammond (1980).

Proposition 1: If $\{\overset{*}{c}_i, \overset{*}{h}_i, \overset{*}{w}_i\}_{i=1}^{NM}$ is an optimal allocation, then

(a) $\overset{*}{w}_i = \omega_i$ for all i such that $\overset{*}{h}_i > 0$;

(b) If $\overset{*}{w}_i = \overset{*}{w}_j$, then $\overset{*}{h}_i > \overset{*}{h}_j \iff \overset{*}{c}_i > \overset{*}{c}_j$;

(c) If $\overset{*}{w}_i = \overset{*}{w}_j = \overset{*}{w}_k$ and $\overset{*}{h}_i > \overset{*}{h}_j > \overset{*}{h}_k$,

then

$\overset{*}{c}_j \geq \lambda \overset{*}{c}_i + (1-\lambda)\overset{*}{c}_k$ for $0 \leq \lambda \leq 1$ such that $\overset{*}{h}_j = \lambda \overset{*}{h}_i + (1-\lambda)\overset{*}{h}_k$;

(d) If $\overset{*}{h}_i \geq \overset{*}{h}_j$ and $\overset{*}{w}_i > \overset{*}{w}_j$, then $\overset{*}{c}_i \geq \overset{*}{c}_j$;

(e) If $\overset{*}{w}_i = \overset{*}{w}_j < \overset{*}{w}_k = \overset{*}{w}_l$ and, for $\lambda_1 \geq 1$ or ≤ 0 and $0 \leq \lambda_2 \leq 1$,

$\lambda_1 \overset{*}{h}_k + (1-\lambda_1)\overset{*}{h}_l = \lambda_2 \overset{*}{h}_i + (1-\lambda_2)\overset{*}{h}_j$

then

$\lambda_1 \overset{*}{c}_k + (1-\lambda_1)\overset{*}{c}_l \geq \lambda_2 \overset{*}{c}_i + (1-\lambda_2)\overset{*}{c}_j$.

Element (a) of Proposition 1 implies that we can find an optimal allocation by solving a simpler program than focuses only on the sequence $\{c_i, h_i\}_{i=1}^{NM}$. This simpler program is given by OP below. Elements (b) and (c) of Proposition 1 indicates that net income, c, must increase with h in a concave fashion, for a given level of w; and elements (d) and (e) indicates how individual allocations most compare across individuals paid different wages. The implications of Proposition 1 are easily seen on Figures 1A through 1C. In each of these figures we plot a set of individual allocations by projecting them on either the $c-h$ or $c-w$ plane. Figure 1A illustrates the content of elements (b) and (c) by plotting three individual allocations with the same wage rate. We have joined the three points by a line in order to help visualized the convexity property implied by element (c). We refer to the line that joins equal wage allocations in the $c-h$ space as a consumption-hours profile since it represents how an individual (of a given market ability) perceives his or her net return to supplying different amounts of labor. Figure 1B complements Figure 1A by illustrating that net income must be non-decreasing in the wage for a given level of hours, as implied by element (d). Finally, Figure 1C illustrates, as implied by elements (d) and (e), that consumption-hours profiles for a given wage level must essentially lie below the consumption-hours profiles for a higher wage (the statement of point (e) makes precise the reason for the qualifier "essentially").

In order to take advantage of the results of Proposition 1, we replace the original program by the program OP:

$$OP = \begin{cases} \{\overset{*}{c}_i, \overset{*}{h}_i\}_{i=1}^{NM} \quad \text{maximizes} \\ \sum_{i=1}^{NM} p_i U\big(c_i + (1-h_i)\theta_i\big) \quad \text{subject to} \\ \sum_{i=1}^{NM} p_i c_i \le \sum_{i=1}^{NM} p_i \omega_i h_i, \\ U\big(c_i + (1-h_i)\theta_i\big) \ge U\big(c_j + (1-h_j)\theta_i\big) \quad \forall j \quad \text{such that}, \omega_j \le \omega_i, \\ U\big(c_i + (1-h_i)\theta_i\big) \ge U\big(\theta_i\big), \quad \text{and} \\ 0 \le h_i \le 1 \quad \text{for all} \quad i. \end{cases} \tag{2.6}$$

Since a problem like OP is rather complex, there are many aspects that could be examined. For example, we could characterize which incentive compatibility constraints are binding at the optimum, or derive an algorithm for solving this program. However, this is not our goal. Our objective is to derive properties of OP that (1) emphasize the nature of the allocative distortions caused by the redistribution program, (2) highlight the pattern of taxes that support the solution, and (3) are easily comparable to those derived in the optimal taxation literature with one dimension of observable characteristics. To this end, we proceed in two steps. In this and the following section we highlight the type and pattern of distortions that arise in an optimum; then, following a sketch of the proof, we relate the propositions in this section to the structure of the implied tax rates.

First, let us define $\overset{*}{H}_i$ to be an individual's market labor supply in the absence of informational constraints, that is, $\overset{*}{H}_i = 1$ if $\omega_i > \theta_i$ and $\overset{*}{H}_i = 0$ if $\omega_i < \theta_i$; in addition, define $\underline{x} = \min_i\{x_i\}$, $\bar{x} = \max_i\{x_i\}$ for an arbitrary vector x of interest. With this notation, we state two propositions that emphasize the pattern of second-best distortions.

Proposition 2A: For any allocation $\{\overset{*}{c}_i, \overset{*}{h}_i, \omega_i\}$, where $\{\overset{*}{c}_i, \overset{*}{h}_i\}$ solves OP, there exists a non increasing function $g : [\underline{\omega}, \bar{\omega}] \mapsto [\underline{c}, \bar{c}]$ such that if

$$\overset{*}{c}_i > g(\omega_i) \longrightarrow \overset{*}{h}_i \le \overset{*}{H}_i, \tag{2.7}$$

$$\overset{*}{c}_i < g(\omega_i) \longrightarrow \overset{*}{h}_i \ge \overset{*}{H}_i, \tag{2.8}$$

furthermore

$$\bar{c} > g(\underline{\omega}) \quad \text{and} \quad \underline{c} = g(w') \quad \text{for some} \quad w' < \bar{\omega}. \tag{2.9}$$

Proposition 2B: For any allocation $\{\overset{*}{c}_i, \overset{*}{h}_i, \omega_i\}$, where $\{\overset{*}{c}_i, \overset{*}{h}_i\}$ solves OP, there exists a non decreasing function $f : [0,1] \mapsto [\underline{c}, \bar{c}]$ such that if

$$\overset{*}{c}_i > f(\overset{*}{h}_i) \longrightarrow \overset{*}{h}_i \leq \overset{*}{H}_i, \tag{2.10}$$

$$\overset{*}{c}_i < f(\overset{*}{h}_i) \longrightarrow \overset{*}{h}_i \geq \overset{*}{H}_i \tag{2.11}$$

furthermore

$$\underline{c} = f(0) \quad \text{and} \quad \bar{c} > f(1). \tag{2.12}$$

Propositions 2A and 2B correspond to restrictions imposed on the pattern of distortions associated with the projection of an allocation onto either the space $c-w$ or $c-h$. In order to illustrate the contents of the above propositions, we provide graphs and a simple example.

Figure 2A is a graphical illustration of the implications Proposition 2A. When projecting an optimal allocation on the space of wage and net income, c, the proposition indicates that this space can be divided into two areas by a non-increasing function. In the area to the northeast of the dividing line, the distortions on time allocated to the market can only be negative, that is, it can only involve reductions of market time relative to the first best allocation. In contrast, in the area to the southwest of the dividing line, any distortions on the allocation of market time must be towards overemployment. This pattern of predicted distortions is not vacuous since it is generally the case that an optimal allocation is characterized by employment distortions of both types. Loosely speaking, Proposition 2A indicates that an optimal allocation is characterized by low market performers (in terms of wages and net income) having their labor supply (weakly) distorted upwards, while strong market performers have their labor supply distorted downward.

The implications of Proposition 1A stand in stark contrast with those of its counterpart in the Mirrlees framework. In the Mirrlees framework, an allocation can also be represented in the $c-w$ space. The resulting locus of points (in the case of a discrete set of types) has the property that net income is increasing in the wage (ability). In this case, we know that types with low net income and low ability generally have their labor supply decision distorted downward since both the substitution effect (induced by a positive marginal tax rate) and the income effect (induce by a negative average tax rate) favor a reduction in labor supply. In contrast, the highest ability individual, who is also the highest net income earner, generally has his or her labor supply weakly distorted upwards (relative to the informationally unconstrained case) since he or she generally faces a zero marginal tax rate but a positive average tax rate, which, if there is an income effect, tends to favor more labor supply.[19] Note that this is pattern virtually opposite to that implied by Proposition 2A.

Figure 2B illustrates the implications of Proposition 2B. The idea is similar to that of Proposition 2A except now the relevant projection is on the space of net income and hours workers. The proposition indicates that this space can be divided by a non-decreasing function, with the north-west portion characterized by the possibility of underemployment only, while the south-east portion groups all individual allocations characterized by overemployment. To interpret these propositions, it is helpful to consider a numerical example.

Example 1

This example highlights why both upward and downward employment distortions can arise at an optimum, and how these distortions relate to Propositions 2A and 2B. In this example there are four equally likely types of individuals ($i = 1, 2, 3, 4$). Individuals 1 and 2 are both highly productive in the market, with their market productivity normalized to one, $\omega_1 = \omega_2 = 1$. However, individuals 1 and 2 differ with respect to their non-market value of time. In particular, individual 1 is strongly attached to the market with a non-market value of time, θ_1, equal to .2. In contrast, individual 2 is less attached to the labor market having a high non-market value of time, that is, θ_2 is equal to .9. Nonetheless, in the informationally unconstrained case, both these individuals would spend their full allocation of time in the market since $\omega > \theta$.

The two remaining individuals have low productivity in the market with $\omega_3 = \omega_4 = .1$. These individuals also differ with respect to their value of non-market time. Individual 3 has a non-market value of time, θ_3, equal to .2 and therefore in an informationally unconstrained economy does not allocate any time on the market. In contrast, individual 4 has a very low non-market value of time, $\theta = .05$, and therefore in an informationally unconstrained economy allocates all of his or her time to market even though market productivity of this individual is relatively low.

With the above specification of parameters and the additional assumption of log utility, the sequence $\{\overset{*}{c}_i, \overset{*}{h}_i\}_{i=1}^4$ that solves OP is $\{(.5, 1), (.39, .43), (.26, .30), (.41, 1)\}$. This solution is illustrated in Figure 3. At the optimum, individual 2's market allocation of time is downward distorted relative to the informationally unconstrained economy, whereas individual 3's allocation is distorted upwards. This property of "hours squashing" is a direct consequence of the redistribution program. It can also be seen in Figure 3 that the optimal allocation trivially satisfies Proposition 2B, since the surface can be divided by an upward sloping line

[19] For more details see Guesnerie and Seade (1982) and Weymark (1986a, 1986b, 1987).

with the individual allocations on both side satisfying the appropriate restrictions. A third element to note, although not transparent, is that the optimal redistribution program has considerably decreased the differences in income. In particular, in the informationally unconstrained economy, the difference in incomes was of the order of ten (for those actually earning income) whereas it is reduced to the order of two under the redistribution program.

To understand why the labor supply of individuals 2 and 3 are distorted in opposite directions, it is helpful to recognize that the government's goal is to take income away from individuals 1 and 2, and to redistribute it to individuals 3 and 4. However, the weak market attachment of individual 2 restricts the government's capacity to do so without distorting individual 2's work incentive. In particular, to extract more than .1 unit(each) of the market good from individuals 1 and 2, it is necessary to distort individual 2's employment level down in order to satisfy the participation constraint. The government does not like distorting down 2's employment level since it loses revenue by doing so ; nonetheless, the government chooses this outcome due to the fact that the lost revenue on individual 2 is more than compensated by the additional revenue extracted from 1. The reason for distorting upward individual 3's labor supply is quite different. In this case, the difficulty is to transfer income to individual 3 without inciting individual 2 to mimic. If individual 3 were to receive an income transfer without a work requirement (that is, without distorted labor supply), individual 2 would mimic individual 3. Individual 3's labour supply must be distorted upwards as a means of screening individual 3's low value of non-market time.

3. Correlated Skills

The results above are derived without placing any restrictions on the distributions of market skills or the non-market skills; nor are any assumptions made about the correlation of the two distributions. As an empirical matter, no particularly reasonable hypothesis about these distributions or their possible correlations comes to mind. Do individuals with high market skills also have a high value of their time in non market activities? Or, do those individuals with low market skills have valuable outside options? It seems to us that neither of these scenarios is compelling. Most likely, among those with high market valuations, some have good outside opportunities and some do not. The same seems probable for those with low market valuations. Nevertheless, from a theoretical point of view, an examination of the case of purely positive or purely negative correlation between the distributions is of some interest; such an assumptions renders our problem one-dimensional and permits further comparison with results in the extant literature. We pursue that line of argument in this section.

Suppose first that the distributions of characteristics are perfectly negatively correlated, that is,

$$\omega_i > \omega_j \longleftrightarrow \theta_i < \theta_j \quad \text{for all} \quad i, j, \tag{3.1}$$

and that not all market skills are greater than or less than all non market skills,[20] that is,

$$\min\{\omega_1, \ldots, \omega_M\} < \max\{\theta_1, \ldots, \theta_N\} \tag{3.2}$$

and

$$\max\{\omega_1, \ldots, \omega_M\} > \min\{\theta_1, \ldots, \theta_N\}. \tag{3.3}$$

Hence the higher the market value of one's time the lower is the value of one's outside option.

Proposition 3: If market and non market skills are perfectly negatively correlated, (3.1), and (3.2)—(3.3) hold, then

$$\overset{*}{h}_i = \overset{*}{H}_i \quad \text{for all} \quad i. \tag{3.4}$$

Proposition 3 states that if the two distribution are perfectly negatively correlated, then at the optimum no individual has his or her work effort distorted either up or down. This proposition is reminiscent of a result of Dasgupta and Hammond [1980] where, in a Mirrlees setup with observable hours, they showed that a first-best outcome could be achieved and utilities equalized.[21] Here, as in Dasgupta and Hammond, individuals with a high value of market time have a lower reservation price on generating market income that individuals with a low market value time because of the assumed negative correlation. In such a case, redistribution is made easy since high types are ready to accept high lump sum taxes in order to have the right to work at a high paying jobs. However, whenever this correlation is not perfect, such a scheme is not incentive compatible and hence a distortionary scheem is likely needed. In order to highlight this point, we now examine the case of perfect positive correlation.

Suppose that the distributions of characteristics are perfectly positively correlated, that is,

$$\omega_i > \omega_j \longleftrightarrow \theta_i > \theta_j, \quad \text{for all} \quad i, j \tag{3.5}$$

---

20 If not, the problem is trivial.

21 In our framework we cannot generally equalize utilities because of the participation constraints. That is, if there were no participation constraints, and the two distribution were perfectly negatively correlated, then, the Dasgupta-Hammond result would hold.

and that

$$\max\{\omega_1, \ldots, \omega_M\} > \max\{\theta_1, \ldots, \theta_N\} \tag{3.6}$$

and

$$\min\{\omega_1, \ldots, \omega_M\} < \min\{\theta_1, \ldots, \theta_N\}. \tag{3.7}$$

This again reduces our problem to one that has effectively one dimension of non observability. In this special case we can show that

Proposition 4: If market and non market skills are perfectly positively correlated, as defined by (3.5), and (3.6)–(3.7) hold, then

$$\overset{*}{h}_i = \overset{*}{H}_i \quad \text{for all} \quad i \quad \text{such that } \omega_i > \theta_i. \tag{3.8}$$

Proposition 4 indicates that, even in the special case of perfect negative correlation between characteristics, it is only the individuals whose market value is greater that his or her nonmarket value who have undistorted employment decisions at the optimum; they work full-time. In contrast, for an individual whose non market value exceeds his or her market value has, in general, their employment decisions are distorted upwards. In order to see this last possibility, it is useful to turn to a simple example where characteristics are positively correlated.

Example 2:

Consider a situation where there are only two types of equaly likely individuals. Individuals of type 1 are the high productivity individuals with market productivity of 1 and non-market productivity of .8. Individuals of type 2 are low productivity individuals with market productivity of .45 and a non-market productivity of .5. Assuming log utility, it can be verified that the optimal allocation in this case is for individual 1 to work full time and receive an after tax income equal to .87, while individual 2 should spend .17 of his time working and receive .21 in after tax (transfer) in income.[22] There are two aspects to note about this example. First, it satisfies the statment of Proposition 4 and the employment decision of individual 2 is upward distorted. This illustrates that upward distortions can arise. Individual 2 is upward distorted so that additional income can be transfered to him or her without inciting individual 1 to mimic. The implicit wage subsidy received by individual 2 is of the order of 200%, that is, by working 17% of the time this individual is receiving net income close to three times the market value of this time. However, this huge subsidy

[22] The exact numbers for individual 2 are .171429 and .207143.

is necessary limited to a set number of hours. If the subsidy were not phased out after .17 hours, both types of individuals would want to take advantage of it making the whole scheme infeasible. In effect, this example illustrates why phased-out wage subsidies – that is, wage subsidies which decreasse with the amount of time worked– are key to understanding how optimal redistribution is achieved in our environment.

4. A Sketch of the Proof[23]

The proofs of Propositions 2A and 2B, which are presented in the appendix, proceed by means of contradiction. In effect, the idea is to suppose that there exist an optimal allocation $\{\overset{*}{c}_i, \overset{*}{h}_i\}_{i=1}^{NM}$ where the proposition does not hold. Given this allocation, we construct a perturbation, satisfying the materials balance constraint, the incentive compatibility constraints, and the participation constraints, which is welfare improving ($dW > 0$). In this section we sketch the proof of the main claim of Proposition 2B, that is, the claim that there exist a non-decreasing function $f(\cdot)$ such that $\overset{*}{c}_i > f(\overset{*}{h}_i) \longrightarrow \overset{*}{h}_i \leq \overset{*}{H}_i$, and $\overset{*}{c}_i < f(\overset{*}{h}_i) \longrightarrow \overset{*}{h}_i \geq \overset{*}{H}_i$. We refer to this claim as **Claim 2B**. This sketch of the proof of Claim 2B highlights the main elements of our overall proof strategy and hopefully makes the the appendix more accessible.

4.1. Preliminary Results

A standard procedure for analyzing problems with incentive compatibility constraints begins by determining which constraints are binding at an optimum. For example, in many problems with one dimension of unobserved characteristics, only the adjacent incentive compatibility constraints are binding at the optimum.[24] However, when there two (or more) dimensions of unobserved characteristics, such simple characterizations are not available. Accordingly, our approach begins by characterizing the (necessary) properties of allocations that are linked together by incentive compatibility constraints at the optimum. We then use these properties to prove Claim 2B. In order to do so, we first define a set of concepts.[25]

23 The reader who is only interested in the implications of Propositions 2A and 2B should skip this section.

24 A standard procedures for solving problems with one dimensional problems first creates a relaxed problem by replacing the set of incentive constraints with only adjacent ones (downward or upward) and then demonstrates that the solution to the relaxed problem is in fact the solution to the original one. See, for example, Guesnerie and Seade (1982), Weymark (1986a, 1986b, 1987), Matthews and Moore (1987) and Besley and Coate (1995).

25 These concepts are very similar to those used by Guesnerie and Seade (1982). However, in Guesnerie and Seade, when two individuals are on an indifference curve this implies that an incentive constraint is binding. This is not the case in the present setup and hence we have resorted to a more general set of concepts. Note that the first half of our definition of a continuation is the same as being W-linked in the language of Guesnerie and Seade.

- **A Continuation:** Type j is a continuation of i if

$$U(\overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_i) = U(\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_i) \quad \text{and} \quad \omega_j \leq \omega_i. \tag{4.1}$$

A continuation links types i and j by an incentive compatibility constraint, where (1) it is type who can mimic type j and (2) it is type i who is indifferent between his or her own allocation and that associated with type j. Graphically, we depict the **continuation** relationship between a type i and a type j by an arrow which links the allocation of i with that of j (see Figures 4A and 4B). Notice that a continuation is not equivalent to an indifference relationship; it also involves the requirement that one be capable of mimicing, that is, $\omega_j \leq \omega_i$. The concept of an extended continuation simply expands this notion to the case where types are linked together by a series of incentive compatibility constraints.

- **An Extended Continuation:** Type j is an extended-continuation of i if there exists a sequence $k_1 \ldots, k_J$ such that type k_{l+1} is a continuation of k_l, $k_1 = i$ and $k_J = j$.

- **A Distinct-Continuation:** Type j is a distinct-continuation of i if j is a continuation and $(\overset{*}{c}_j, \overset{*}{h}_j) \neq (\overset{*}{c}_i, \overset{*}{h}_i)$.

The notion of an extended continuations is useful when perturbing an allocation that reduces the utility level of some agent. For example, suppose that a perturbation reduces the consumption of type i, keeping h_i constant. In such a case, incentive compatibility requires that the consumption of all types in the extended continuation of i must also have their consumption reduced by the same amount. Hence, the set of extended continuations groups together types whose incentive compatibility constraints must be treated simultaneously. Similarly, the notion of a source is helpful when considering perturbations that increase the utility of some agent.

- **A Source :** Type j is a source of i if i is a continuation of j and

$$U(\overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_i) > U(\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_i) \quad \text{and,} \quad \overset{*}{h}_j > \overset{*}{h}_i. \tag{4.2}$$

- **A Fundamental Source:** Type j is a fundamental source of i if it is an extended source[26] of i and $h_j = 1$.

Graphically, we depict a source by a double arrow that links the allocation of type j to the allocation of type i (see for example Figure D). Finally, we define upward and downward

[26] An extended source is defined analogously to an extended continuation.

distorted types as types whose time allocation is distorted relative to the informationally unconstrained case.

- **Upward Distorted:** Type i is upward distorted if

$$\overset{*}{h}_i > \overset{*}{H}_i. \tag{4.3}$$

- **Downward Distorted:** Type i is downward distorted if

$$\overset{*}{h}_i < \overset{*}{H}_i. \tag{4.4}$$

With these definitions in mind, we are now in a position to state the following two simple Lemmas.

Lemma A: *If i is downward distorted, then, for all j that are distinct continuations of i,*

$$\overset{*}{h}_j < \overset{*}{h}_i \quad \textit{and} \quad \overset{*}{c}_j < \overset{*}{c}_i. \tag{4.5}$$

Lemma B: *If i is upward distorted, then, for all j that are distinct continuations of i,*

$$\overset{*}{h}_j > \overset{*}{h}_i \quad \textit{and} \quad \overset{*}{c}_j > \overset{*}{c}_i. \tag{4.6}$$

Lemmas A and B provide restrictions on the pattern of allocations which are consistent with types being linked by an incentive compatibility constraint. For example, Lemma A indicates that if a type i is downward distorted, then he or she can only mimic (in the incentive compatibility sense) a type whose allocation involves less consumption and fewer hours. Lemma B states the converse. The content of these two lemmas is depicted in Figures 4A and 4B. Intuitively, these results indicates that at an optimum, one must be indifferent only with respect to individual allocations that increase the distortion on hours worked.

Building on Lemmas A and B, Lemma C expresses a restriction on the pattern of allocations that are linked (directly or indirectly) by incentive compatibility constraints.

Lemma C: *If i is upward distorted, then, for all j that are extended continuations of i,*

$$\overset{*}{c}_j \geq \overset{*}{c}_i \quad \textit{and} \quad \overset{*}{h}_j \geq \overset{*}{h}_i. \tag{4.7}$$

Lemma C is depicted in Figure 4C; all types that are continuations of continuations of an upward distorted i must have allocations that lie to the north-east of i's allocation. This Lemma is useful when considering a perturbation that reduces the consumption level of an upward distorted i. Lemma C states that, in order to maintain incentive compatibility, we need only consider changing allocations that are to the north-east of i's allocation.

The counterpart to Lemma C, describes the restrictions imposed on sources emanating from a downward distorted type.

Lemma D: *If i is downward distorted, then the set of sources of i is non-empty and each member j of the set must be downward distorted or $\overset{*}{h}_j = 1$. Moreover, θ_i must be greater than θ_j.*

Lemma E: *If i is downward distorted and if f is a fundamental source of i ($\overset{*}{h}_f = 1$), then there is no k such that f is a distinct continuation of k.*

Graphically Lemma D implies that, starting from the allocation of a downward distorted type, one can connect sources (along indifference curves) so as to create an upward sloping line that starts at $(\overset{*}{h}_i, \overset{*}{c}_i)$ and finishes at a fundamental source. This pattern is illustrated in Figure 4D. This Lemma implies that an upward sloping line connecting a downward distorted type to its fundamental source always exists. Lemma E adds to Lemma D by indicating that a fundamental source has another very important property; no type wants to mimic (in an incentive compatibility sense) a fundamental source. This property is important because it implies that the consumption level of a fundamental source can be increased by a small amount without affecting incentive compatibility constraints.

4.2. Assembling the Pieces of the Argument

Given Lemmas A through E, we can now discuss the main elements in the proof of Claim 2B. As we have indicated, the proof is by contradiction. So let us suppose that there does not always exist a non-decreasing function $f(\cdot)$ such that $\overset{*}{c}_i > f(\overset{*}{h}_i) \longrightarrow \overset{*}{h}_i \leq \overset{*}{H}_i$, and $\overset{*}{c}_i < f(\overset{*}{h}_i) \longrightarrow \overset{*}{h}_i \geq \overset{*}{H}_i$. This assertion is equivalent to assuming that there exists an allocation $\{\overset{*}{c}_i, \overset{*}{h}_i\}_{i=1}^{NM}$ which includes a downward distorted type a, and an upward distorted type b, such that $\overset{*}{c}_b > \overset{*}{c}_a$, $\overset{*}{h}_b < \overset{*}{h}_a$. This configuration is illustrated in Figure 4E. The first aspect to note about such an allocation is that incentive compatibility necessarily requires that $\omega_b > \omega_a$, otherwise type a would mimic type b. Furthermore, since a is downward distorted and b is upward distorted, this implies that $\theta_b > \theta_a$ and hence that type b has a higher level of utility than type a. From an equity point of view, this is a situation where—if feasible–it would be desirable to transfer consumption from type b to type a. However, such a direct transfer is generally not feasible because of incentive compatibility constraints.

Hence, let us consider a more complex perturbation in order to show that the assumed configuration for the individual allocations of types a and b is impossible. The proposed perturbation reduces by ϵ the consumption of all types in the set of extended continuation of b (including b itself) and transfers the released resources to the fundamental source of a. For small enough ϵ, this perturbation satisfies the incentive compatibility constraints since, by Lemma E, no one wants to mimic a fundamental source and since, by construction, we are preserving incentive compatibility by simultaneously perturbing the individual allocations of the entire set of extended continuations of b. Furthermore, this perturbation is welfare improving if all the types in the extended continuation of b have utility levels that are greater than that of the fundamental source of a. Hence, the main issue becomes whether or not this last statement is true, that is, does there exist an extended continuation of b which initially has a utility level lower than that of the fundamental source of a?

In order to answer this last question, it is helpful to recall Lemmas C and D and consider the interplay between the extended continuations-lines emanating from b and extended source-lines emanating from a. We illustrated this interplay in Figure 4E where types b_1 and b_2 are the extended continuations of b, while types a_1, a_2 and a_f are the extended source set associated with a. If it is the case that one of the extended continuations of b, for example type b_2 in Figure 4E, has a lower level of utility than the fundamental source of a (denoted a_f in Figure 4E) it must be the case that the indifference curve of at least one extended continuation of b crosses from the left an extended source-line emanating from a's individual allocation. This possibility is illustrated in Figure 4E where the indifference curve of type b_2 crosses the source-line linking a_f and a_2. However, it is precisely this possibility that is ruled out by Lemma F.

Lemma F: *There do not exist i and j such that*

$$\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_i > \overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_i \tag{4.8}$$

and

$$\overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_j > \overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j. \tag{4.9}$$

Hence, Lemma F allows us to conclude that all extended continuations of b initially have a utility level that is higher than that of the fundamental source of a. This result is given in Lemma G. Hence, we know that the proposed perturbation would be welfare improving.

Lemma G: *If i is downward distorted, j is upward distorted, f is a fundamental source of i, k is an extended continuation of j, $\overset{*}{c}_j \geq \overset{*}{c}_i$, $\overset{*}{h}_j \neq \overset{*}{h}_i$, and $\omega_j \geq \omega_i$ then,*

$$U(\overset{*}{c}_k + (1 - \overset{*}{h}_k)\theta_k) > U(\overset{*}{c}_f) \tag{4.10}$$

It only remains to verify whether or not the proposed perturbation satisfies the participation constraints. This is done by Lemma H.

Lemma H: *If j is an extended continuation of an upward distorted i then either (1):there exists a k with $h_k = 0$,*

$$U(\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j) > U(\overset{*}{c}_k + \theta_j) \tag{4.11}$$

or (2):

$$U(\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j) > U(\theta_j) \tag{4.12}$$

In summary, Lemma E implies that the conjectured perturbation is incentive compatible (for small enough ϵ), Lemma G that it is welfare improving and Lemma H that it satisfies the participation constraints. Therefore, we arrive at a contradiction when we assume that an optimal allocation may have a downward distorted type a and an upward distorted type b with $c_b > c_a$ and $h_b < h_a$. Hence, it is necessarily the case that the desired function $f(\cdot)$ can always be constructed as implied by Claim 2B.

The remaining elements of Proposition 2B, which correspond to boundary conditions, show that there cannot exist an upward distorted type who has the highest level of consumption. The argument supporting this boundary condition is similar to the above and is therefore not repeated here.

5. The Implied Tax Rates

Propositions 2A and 2B restrict the pattern of distortions caused by the conjunction of incentive compatibility constraints, the participation constraints, and the desire for redistribution. However, we have not as yet made explicit the policy implications of these restrictions. Does a specific redistribution program satisfy these restrictions? In order to do so, we consider the implications of Propositions 2A and 2B for average and marginal tax rates. In principle, this provides implications of our model in terms observables.[27]

Because we have analyzed a case with a discrete set of types, marginal tax rates are not unambiguously defined. Therefore we first state what we mean by the marginal tax rate facing an individual.

In our framework, the relevant notion of a marginal tax rate corresponds to the change in taxes an individual experiences when his or her income changes due to a change in hours worked. In other words, if the tax function that implements the redistribution program is viewed as a function of income and hours worked, say $T(y, h)$, then, in the differentiable case a marginal tax rate would correspond to the total derivative of $T(y, h)$ with respect to h divided by w. We view this as the relevant notion since it reflects the dimension on which individuals actually have a choice. This notion can be operationalized in the case of a discrete set of types by looking at the slope of consumption-hours profiles implied by an allocation $\{\overset{*}{c}_i, \overset{*}{h}_i, \omega_i\}$. There are two difficulties with such an approach. First, marginal tax rates, evaluated at a particular individual allocation, are not uniquely defined; they depend on whether we consider an increase or a decrease in labor supply. Second, it may not be sufficient to know $\{\overset{*}{c}_i, \overset{*}{h}_i, \omega_i\}$ in order to calculate the marginal tax rates for all feasible directions of change in labor supply since consumption-hours profiles are not necessarily defined over the relevant range. Therefore, we begin by defining a complete set of left and right marginal alternatives facing each type in order to have well-defined marginal tax rates for all feasible changes in labor supply. Note that these definitions of marginal alternatives correspond to movements along consumption-hours profiles whenever possible. In the case where such a movements are not defined, we have extended the profiles by the use of a feasible option. Then, with the complete set of (direction specific) marginal tax rates, we defined an individual as facing a positive (negative) marginal rate if his or her marginal tax rate is positive (negative) in all feasible directions.

To this end, let us consider a situation where an optimal allocation is given by $\{\overset{*}{c}_i, \overset{*}{h}_i, \omega_i\}$ and let $J_i^l = \{j \mid \omega_j = \omega_i, \overset{*}{h}_j < \overset{*}{h}_i\}$, and $J_i^r = \{j \mid \omega_j = \omega_i, \overset{*}{h}_j > \overset{*}{h}_i\}$.

Definition: The *left marginal alternative of individual* i, denoted (c_i^l, h_i^l) is defined by (a) if $J_i^l \neq \emptyset$ and (b) otherwise.

a) $(\overset{*}{c}_{j'}, \overset{*}{h}_{j'})$, where j' is such that $\overset{*}{h}_{j'} \geq \overset{*}{h}_j \quad \forall j, j' \in J_i^l$,

b) $(\overset{*}{c}_i - \overset{*}{h}_i \theta_i, 0)$

[27] In practice, income redistribution is a achieved through a variety of programs. Although these programs are complex, it is nonetheless generally possible to aggregate all such programs in order to derive the implications in terms of the effective tax schedule facing any given individual.

Definition: The *right marginal alternative of individual* i, denoted (c_i^r, h_i^r) is defined by (a) if $J_i^r \neq \emptyset$ and (b) otherwise.

a) $(\overset{*}{c}_{j'}, \overset{*}{h}_{j'})$, where j' is such that $\overset{*}{h}_{j'} \leq \overset{*}{h}_j \quad \forall j, j' \in J_i^r$,

b) $(\overset{*}{c}_i, 1)$

The left marginal alternative of person i is the allocation received by an individual who has the same market skill level as i, works less than i but more than anyone else whom i could imitate; if there is no such individual, then the left marginal alternative of i is given by her or his outside option, θ_i. The right marginal alternative is defined analogously. We use these concepts to define the marginal tax rates.

Definition: The *marginal tax rate for individual* i is positive if either

(a) $\overset{*}{h}_i = 0$ and $w_i - \frac{\overset{*}{c}_i - c_i^r}{\overset{*}{h}_i - h_i^r} > 0$

(b) $0 < \overset{*}{h}_i < 1$, $w_i - \frac{\overset{*}{c}_i - c_i^r}{\overset{*}{h}_i - h_i^r} > 0$ and $w_i - \frac{\overset{*}{c}_i - c_i^l}{\overset{*}{h}_i - h_i^l} > 0$

(c) $\overset{*}{h}_i = 1$ and $w_i - \frac{\overset{*}{c}_i - c_i^l}{\overset{*}{h}_i - h_i^l} > 0$

Definition: The *marginal tax rate for individual* i is negative if either

(a) $\overset{*}{h}_i = 0$ and $w_i - \frac{\overset{*}{c}_i - c_i^r}{\overset{*}{h}_i - h_i^r} < 0$

(b) $0 < \overset{*}{h}_i < 0$, $w_i - \frac{\overset{*}{c}_i - c_i^r}{\overset{*}{h}_i - h_i^r} < 0$ and $w_i - \frac{\overset{*}{c}_i - c_i^l}{\overset{*}{h}_i - h_i^l} < 0$

(c) $\overset{*}{h}_i = 1$ and $w_i - \frac{\overset{*}{c}_i - c_i^l}{\overset{*}{h}_i - h_i^l} < 0$

These definitions of the marginal tax rate capture the sign of the marginal tax rate that individual i faces when labor supply is either increased or decreased. We say that the marginal tax rate of individual i is ambiguous if it is neither defined as positive or negative. Returning to example 1, the marginal tax rates can be calculated knowing the left and right marginal alternatives. The right marginal alternative of person 2 is person 1 and the left marginal alternative of person 2 is his or her outside option. The right marginal alternative of person 3 is person 4 and the outside option is the left marginal alternative. Person 2 is the left marginal alternative of person 1 and person 3 that of person 4. The marginal tax rates are given by

$$\begin{aligned} \text{MRT}_1 &= .81, \\ .1 \leq \text{MRT}_2 &\leq .81, \\ -7.6 \leq \text{MRT}_3 &\leq -1.14, \quad \text{and,} \\ \text{MRT}_4 &= -1.14. \end{aligned} \tag{5.1}$$

In the example, person 2 is downward distorted while person 3 is upward distorted relative to the informationally constrained solution. Hence, person 2 faces relatively high positive marginal tax rates whereas person 3 faces negative marginal (and average) tax rates. In fact, the two lower wage individuals have negative marginal tax rates and the two higher wage individuals have positive marginal tax rates. Another interesting observation is that the highest earner, who is individual 1, has a positive marginal tax rate. This contrasts with the results optimal taxation literature in which the highest earner faces a zero marginal tax rate when there is a discrete set of types.

From Proposition 1 (statement (c)), we know that there are likely at most N types (among NM) that have an ambiguous marginal tax rate since consumption-hours profiles are concave. Therefore, even if we focus only on types with unambiguous tax rates, we are actually focusing on the vast majority of types.[28] Let us denote the set of types with a non-ambiguous marginal tax rates as $\hat{I}$. With these definitions, and with the standard notion of an average tax rate, we can state the following proposition.

Proposition 5: For any optimal allocations $(\overset{*}{c}_i, \overset{*}{h}_i, \omega_i)$, there exists a non decreasing function $t : [0, 1] \mapsto [\underline{c}, \bar{c}]$ such that if $i \in \hat{I}$,

$$\overset{*}{c}_i > t(\overset{*}{h}_i) \longrightarrow \quad i\text{'s marginal tax rate is positive,} \tag{5.2}$$

$$\overset{*}{c}_i < t(\overset{*}{h}_i) \longrightarrow \quad i\text{'s marginal and average tax rate is negative} \tag{5.3}$$

In light of Proposition 2B, the interpretation of Proposition 5 is straightforward. It replaces the restrictions on the pattern of distortions emphasized in Proposition 2B with restrictions on the pattern of marginal tax rates. Proposition 5 also indicates that all types with negative marginal tax rates also have negative average tax rates, where average tax rates are defined by

$$\frac{w_i\overset{*}{h}_i - \overset{*}{c}_i}{w_i\overset{*}{h}_i}. \tag{5.4}$$

These restrictions in terms of taxes rates can be used to evaluate a redistribution program.[29] We view the main insight of Proposition 5 as prescribing the use of wage subsidies (negative marginal tax rates) as a means of redistributing income, which is in marked contrast to

[28] There may be more than N types with ambiguous tax rates when there is pooling.

[29] For the policy oriented reader, the restrictions embedded in Proposition 5 may appear rather weak or at least incomplete. This is a criticism often made of the optimal taxation literature. Nevertheless, it is our belief that many redistribution systems may not satisfy this condition and therefore it may offer a way to identify direction in which a system can be improved.

most of the optimal tax literature which prescribes marginal tax rates to be everywhere non negative. Moreover, a key a aspect to note about wage subsidies in our environment is that they are generally phased-out as individuals supply more work (as implied by Proposition 1.(c)). This constrasts with many commonly proposed wage-subsidy programs which do not include a phase-out, and hence, are often considered too expensive to implement. It should be noted that wage-subsidy phase-outs are optimal in our environment both because they stop some high market-value individuals from taking advantage of such programs as well as allowing some low-market value individuals to take advantage of the relatively higher value of their non-market time.

6. The Taxation of Equal Earners

Proposition 5 can be viewed as expressing restrictions on the type of tax functions that implement an optimal allocation. Does optimal redistribution impose any additional restrictions on tax functions? Focusing on tax functions of the form $T(y, h)$, does the solution to OP entail that it be monotone in either of its components (at least for the segments chosen in equilibrium)? In particular, is the tax function non-increasing in h (for a given y)? That this is not necessarily the case, is shown in the following example[30] which also nicely illustrates the nature of constraints and tradeoffs that a fiscal authority may face. Moreover, it illustrates how certain properties that at first glance appear opposed to the principle of redistribution, may actually be consistent with it.[31]

Example 3:

Consider a four type example with log utility and where $\{\omega_i, \theta_i\}_{i=1}^4$ is equal to $\{(1, 0.001), (1, .9), (.8, .5), (.1, .05)\}$[32]. In an informationally unconstrained situation, all four individuals would spend their full labor allocation in the market and respectively receive incomes $(1, 1, .8, .1)$. Therefore the main objective of the government in this situation is to transfer income to individual 4. To achieve this objective, it is easy to verify that the government chooses to distort down the labor supply of individual 2 to a point where he or she earns the same market income as individual 3 does working full-time. Moreover, even if these

30 Maderner and Rochet (1995) also find the possibility that an optimal tax function of the form $T(y, h)$ may be increasing in h. They interpret such a phenomena as reflecting incentive towards worksharing.

31 It can also be shown that $T(y, h)$ is not necessarily increasing in y for h fixed. This, however, depends on a highly special patterns of covariance between ω and θ.

32 For the calculation to be exact, set $\theta_2 = .8966$

two individuals earn the same income, individual 3 is asked to pay more taxes. In fact the sequence $\{\overset{*}{c}_i, \overset{*}{h}_i\}$ that solve OP in this case is given by $\{(.72,1),(.72,.8),(.63,1),(.63,1)\}$ and is illustrated in Figure 5.[33] As can be seen, both individuals 2 and 3 earn .8 units of gross income, individuals 2 gets .72 units of net income while individual 3 get .63 units of net income even though working more. This examples demonstrates that the optimal tax function may not be decreasing in hours (In many cases it is decreasing; see Example 1.) The reason the government taxes individual 3 so heavily is that this agent can be taxed with out creating distortions; in balancing efficiency versus equity, the inequitable treatment of 3 versus 2 is dominated by efficiency considerations. Intuitively, individual 3 is highly taxed because he or she belongs to a group that is highly attached to the labor market and therefore highly taxable, while individuals with the higher wages are viewed as having a more elastic labor supply.

7. Workfare

The analysis to this point demonstrates that an optimal redistribution program (in our environment) generally involves subsidizing employment of low market performers, by means of negative marginal tax rates, and by taxing the employment income of high market performers. As noted above, these results contrast markedly with much of the optimal taxation literature in which it is never optimal to transfer income through employment subsidies (negative marginal tax rates). From a policy point of view, workfare—a public work requirements—is a related and frequently discussed means of achieving redistribution. However, it has been shown by Besley and Coates (1995), that it is generally not welfare improving within the Mirrlees framework to complement the optimal non-linear income tax with workfare.[34]

To pose questions about workfare, we first define what is meant by workfare and then indicate how it can be integrated into our analysis. To be distinct from the subsidization of private sector employment, we consider workfare to be a requirement to work in a public sector employment program; in addition, we assume that individuals of all types have the same productivity in such a program and that the government does not learn an individual's

33 The actual consumption of individual one is .72002; he or she is the fundamental source of individual two.

34 In Brett (1997) public work is treated as a separate input. In this case, workfare may be employed as a last resort in an optimal redistribution program.

type by having time allocated to the workfare program. This latter assumption appears reasonable given the likely institutional structure of most workfare programs.

We denote marginal productivity in the workfare program by ω^f (which can be negative) and the time requirement in workfare as h_i^f. With the addition of a workfare option, the government's problem is to choose an allocation of the form $\{\overset{*}{c}_i, \overset{*}{h}_i, \overset{*}{\omega}_i, \overset{*}{h}{}_i^f\}_{i=1}^{NM}$ which solves the following problem.

$$\max \sum_{i=1}^{NM} p_i U\big(c_i + (1 - h_i - h_i^f)\theta_i\big) \tag{7.1}$$

subject to

$$\sum_{i=1}^{NM} p_i c_i \leq \sum_{i=1}^{NM} p_i(w_i h_i + \omega^f h_i^f), \tag{7.2}$$

and for all i

$$U\big(c_i + (1 - h_i - h_i^f)\theta_i\big) \geq U\big(c_j + (1 - h_j - h_j^f)\theta_i\big), \quad \forall j \quad \text{s.t.} \quad w_j \leq \omega_i, \tag{7.3}$$

$$U\big(c_i + (1 - h_i - h_i^f)\theta_i\big) \geq U\big(\theta_i\big), \tag{7.4}$$

$$0 \leq h_i + h_i^f \leq 1, \quad h_i \geq 0, \quad h_i^f \geq 0 \tag{7.5}$$

Individuals can concurrently work in the workfare program and in private sector employment as long as the total time in these activities is no greater that the individual's endowment.

The following proposition relates the redistribution-workfare program to our previous analysis. In particular, starting from a given distribution of types, the proposition exploits the construction of a modified economy in with all types j with $\omega_j < \omega^f$ are relabeled so that $\omega_j = \omega^f$.

Proposition 6: If $\{\tilde{c}_i, \tilde{h}_i\}_{i=1}^{NM}$ solves OP for the modified economy (where types j with $\omega_j < \omega_f$ are relabeled so that $\omega_j = \omega^f$), then $\{\hat{c}_i, \hat{h}_i, \hat{w}_i, \hat{h}_i^f\}_{i=1}^{NM}$ solves the redistribution-workfare problem if for all i,

(1) $\hat{c}_i = \tilde{c}_i$,

(2) $\hat{w}_i = \omega_i$ and

(3) If $\omega_i > w^f$, $\hat{h}_i = \tilde{h}_i$ and $\hat{h}_i^f = 0$

(4) If $\omega_i = w^f$, $\hat{h}_i^f = \tilde{h}_i$ and $\hat{h}_i = 0$

Proposition 6 implies that workfare is to be used only if an individual's social productivity is greater in the workfare program than in private sector employment; otherwise, it is preferable to use employment subsidies instead of workfare requirements as a means of redistributing income. In part, this statement supports Besley and Coates (1995) finding against the use of workfare (when the objective is welfare maximization) since it is generally assumed that public works programs are unlikely to have greater social product than that associated with private employment. However, to be realistic, we must consider the possibility of individuals without private sector employment possibilities. Formally, in our notation this case corresponds to types with $\omega = -\infty$. It is exactly in such a case that it can be optimal to use workfare in our setup even with $w^f < 0$, while it is not welfare improving to do so in the Mirrlees setup (which is the situation analyzed by Besley and Coates, Section VII). The following example illustrates why a workfare requirement can be used to improve welfare even when $w^f < 0$.

Example 3 There are two individuals with log utility. Individual of type one has a market productivity normalized to one and values his or her non-market time at .8, that is, $\omega_1 = 1$ and $\theta_1 = .8$. Individual 2 is unemployed and values his or her non market time at .1, that is, $\omega_2 = -\infty$ and $\theta_2 = .1$. In the absence of the possibility of workfare, the social optimal would be characterized by transferring .1 unit from individual 1 to individual 2, thereby providing a utility level for individual 2 of log(.2). In this case, individual 2 can be considered to be in a welfare program without any work requirements. Now let us assume that the government can also impose work requirements, but that the work itself is socially unproductive with $w^f = -.2$. In this case, the social optimum is characterized by requiring individual 2 to spend .2 of his or her time in the workfare program in order to receive .16 units of the good. The utility of the type 2 individual is increased to log(.24). In effect, by tying the income transfer to workfare it has become possible to transfer more income to the unemployed individual since the resulting package remains unattractive to the employed individual. Hence, this example demonstrates why a socially costly workfare program may potentially may be a desirable way to tie income transfers for individuals with low or non-existant private sector employment possibilities.

8. Conclusion

The object of this paper is to explore the principles that govern the design of an optimal redistribution program in which taxation authorities have both reasons and tools to favor a programs that target transfers more effectively than simple negative income tax schemes.[35]

[35] Avenues of future research include examining the value of rendering some informal activities observable through monitoring, and rendering the acquisition of skill endogenous.

To this end we have analyzed a variant of the optimal taxation problem pioneered by Mirrlees. Our departure consists of allowing for a greater scope of unobserved heterogeneity in the population and allowing the government to transfer income based on both market income and market labor supply. Our main finding is that, in contrast to much of the optimal taxation literature, optimal redistribution in this environment is achieved using employment subsidies on low market performers, positive marginal tax rates on high market performers and, in last recourse, workfare as a mean of transferring income to individuals with very poor or non-existant market opportunities.

How should these results interpreted? In our view, these results are not a call for redesigning income tax system to include a dependence on annual hours worked. Instead we view these results as supporting the potential relevance of certain active labor market programs as a complement to income tax as a means of redistributing income.[36] For example, these results provide potential support for programs, such as the Canadian Self-Sufficiency Project, which supplements income to low wage earners who choose to work. More generally, we view our results as suggesting the use of phased-out wage subsidies as a means of redistributing income to low earners, that is, wage subsidies that decrease as an individual chooses to supply more labor. Such phased-out subsidy programs, in effect, allow substantial transfers to the most needy in society without inciting either high market-value individuals or high non-market value individuals to take advantage of it.

[36] We also view these results as providing minimal guidelines of how such programs should interact with the income tax system in terms of the implied pattern of effective marginal tax rates.

APPENDIX

Proof of Proposition 1:

Point (a). By contradiction. Suppose that for some j, $\overset{*}{h}_j > 0$ and $\overset{*}{w}_j \neq \omega_j$. Then consider setting $\overset{*}{w}_i = \omega_i$ for all i. In this case, all the constraints of the problem remain satisfied and the material balance constraint is relaxed. It is then possible to take the additional resources and share them equally, thereby creating an Pareto improvement.

Point (b). From incentive compatibility we know that $\overset{*}{c}_i - \overset{*}{c}_j \geq (\overset{*}{h}_i - \overset{*}{h}_j)\theta_i$ and $\overset{*}{c}_j - \overset{*}{c}_i \geq (\overset{*}{h}_j - \overset{*}{h}_i)\theta_j$. Since $\theta > 0$, if $\overset{*}{h}_i > \overset{*}{h}_j$ the first inequality implies $\overset{*}{c}_i > \overset{*}{c}_j$; and if $\overset{*}{c}_i > \overset{*}{c}_j$ then the second inequality implies $\overset{*}{h}_i > \overset{*}{h}_j$.

Point (c). The two incentive compatibility constraints associated with type j not mimicing either type i or k imply that, for any $0 < \lambda < 1$,

$$\overset{*}{c}_j - \overset{*}{h}_j\theta_j \geq \lambda\overset{*}{c}_i + (1-\lambda)\overset{*}{c}_k - \lambda\overset{*}{h}_i\theta_j - (1-\lambda)\overset{*}{h}_k\theta_j$$

.

If λ is further chosen such that $\overset{*}{h}_j = \lambda\overset{*}{h}_i + (1-\lambda)\overset{*}{h}_j$, the above inequality implies $\overset{*}{c}_j \geq \lambda\overset{*}{c}_i + (1-\lambda)\overset{*}{c}_k$.

Point (d). Incentive compatibility implies that $\overset{*}{c}_i + (1-\overset{*}{h}_i)\theta_i \geq \overset{*}{c}_j + (1-\overset{*}{h}_j)\theta_i$, and since, $\overset{*}{h}_i \geq \overset{*}{h}_j$ by assumption, this implies $\overset{*}{c}_i \geq \overset{*}{c}_j$.

Point (e). Proof by contradiction. Suppose there exist a λ_1 and a λ_2 such that, λ_1 is either ≤ 0 or ≥ 1, $0 \leq \lambda_2 \leq 1$,

$$\lambda_1\overset{*}{h}_k + (1-\lambda_1)\overset{*}{h}_l = \lambda_2\overset{*}{h}_i + (1-\lambda_2)\overset{*}{h}_j \tag{A.1}$$

and

$$\lambda_1\overset{*}{c}_k + (1-\lambda_1)\overset{*}{c}_l < \lambda_2\overset{*}{c}_i + (1-\lambda_2)\overset{*}{c}_j \tag{A.2}$$

- Two cases. First case. If $\lambda_1 \geq 0$, then (A.2) minus θ_k times (A.1) implies:

$$\lambda_1(\overset{*}{c}_k+(1-\overset{*}{h}_k)\theta_k)+(1-\lambda_1)(\overset{*}{c}_l+(1-\overset{*}{h}_l)\theta_k) < \lambda_2(\overset{*}{c}_i+(1-\overset{*}{h}_i)\theta_k)+(1-\lambda_2)(\overset{*}{c}_j+(1- \ hs_j)\theta_k) \tag{A.3}$$

Furthermore, incentive compatibility associated with type k not mimicing either i or j implies:

$$\lambda_2(\overset{*}{c}_i + (1-\overset{*}{h}_i)\theta_k) + (1-\lambda_2)(\overset{*}{c}_j + (1- \ hs_j)\theta_k) \leq \overset{*}{c}_k + (1-\overset{*}{h}_k)\theta_k \tag{A.4}$$

Combing both (A.3)and (A.4)implies

$$\overset{*}{c}_l + (1 - \overset{*}{h}_l)\theta_k > \overset{*}{c}_k + (1 - \overset{*}{h}_k)\theta_k$$

which violates the incentive compatibility constraint associate with type k not mimicing type l.

Second case. If $\lambda_2 \leq 0$, a similar contradiction can be shown by subtracting θ_l times (A.1) from (A.2) instead of subtracting θ_k times (A.1) from (A.2). ■

Lemma A1: *If i is distorted, there must exist a k such that i is a continuation of k.*

Proof: Without loss of generality, suppose that i is downward distorted and that there is no k of which it is a continuation. Now pick $dc_i > 0$ and $dh_i > 0$ such that

$$dc_i - \theta_i dh_i = 0; \tag{A.5}$$

this deviation is physically feasible because i is downward distorted, $\overset{*}{h}_i < \overset{*}{H}_i$. Moreover, this deviation is such that the resulting allocation is incentive compatible, since there is no k that, by assumption, wants to mimic i, and this deviation satisfies the participation constraint since it leaves i indifferent. The deviation also releases resources since $w_i > \theta_i$. The resulting resources can be divided equally among all types (lump sum) in order to create a Pareto improvement. This contradicts the optimality of the solution. ■

Lemma A: *If i is downward distorted, then, for all j that are distinct continuations of i,*

$$\overset{*}{h}_j < \overset{*}{h}_i \quad \textit{and} \quad \overset{*}{c}_j < \overset{*}{c}_i. \tag{A.6}$$

Proof: First we established by Lemma A1 that because i is downward distorted there must exist a k such that i is a continuation of k

Now suppose (A.6) is false, that is,

$$\overset{*}{h}_j \geq \overset{*}{h}_i \quad \text{or} \quad \overset{*}{c}_j \geq \overset{*}{c}_i. \tag{A.7}$$

That j is a distinct continuation of i implies that

$$\overset{*}{c}_i + \theta_i(1 - \overset{*}{h}_i) = \overset{*}{c}_j + \theta_i(1 - \overset{*}{h}_j) \quad \text{and} \quad w_i \geq w_j. \tag{A.8}$$

Therefore, (A.8) and (A.7) imply that

$$\overset{*}{h}_j > \overset{*}{h}_i \quad \text{and} \quad \overset{*}{c}_j > \overset{*}{c}_i. \tag{A.9}$$

That i is a continuation of k implies

$$\overset{*}{c}_k + \theta_k(1 - \overset{*}{h}_k) = \overset{*}{c}_i + \theta_k(1 - \overset{*}{h}_i) \quad \text{and} \quad w_k \geq w_i. \tag{A.10}$$

Case 1: Suppose $\theta_i > \theta_k$; in this case k prefers the allocation of j to that of i and can do so yielding a contradiction.

Case 2: Suppose that $\theta_k > \theta_i$ and consider the following change in allocation for type i. Set $c_i = c_j$ and $h_i = h_j$. The resulting allocation remains incentive compatible and individually rational since the utility level of i has not changed and there can't exist at type l that would now want to mimic i since l would have already mimicked j. Furthermore, such a change in allocation frees up resources since i is downward distorted. Therefore, this again leads to a contradiction since these resources could be equally divided among all types in order to induce a Pareto improvement. ■

Lemma B: *If i is upward distorted, then, for all j that are distinct continuations of i,*

$$\overset{*}{h}_j > \overset{*}{h}_i \quad \textit{and} \quad \overset{*}{c}_j > \overset{*}{c}_i. \tag{A.11}$$

Proof: The argument proceeds just as in the proof of Lemma A. ■

Lemma C1: *If i is upward distorted, then, for all j that are continuations of i, such that*

$$\overset{*}{c}_i = \overset{*}{c}_j \quad \textit{and} \quad \overset{*}{h}_i = \overset{*}{h}_j, \tag{A.12}$$

either j is upward distorted or the set of distinct continuations of j is empty.

Proof: Suppose, to the contrary, that j is both downward distorted[37] and has a distinct continuation, say k. Because j is downward distorted, $w_j > \theta_j$ and as it is a continuation of i, $w_i \geq w_j$. By assumption i is upward distorted so that $\theta_i > w_i$ and hence $\theta_i > \theta_j$. Given that j is downward distorted and has a distinct continuation, k, we know from Lemma A that

$$\overset{*}{h}_k < \overset{*}{h}_j \quad \text{and} \quad \overset{*}{c}_k < \overset{*}{c}_j. \tag{A.13}$$

Because k is a distinct continuation of j, $w_j > w_k$ and hence, $w_i > w_k$. Using this, (A.12), (A.13), and the fact that $\theta_i > \theta_j$, shows that i prefers $(\overset{*}{c}_k, \overset{*}{h}_k)$ to $(\overset{*}{c}_i, \overset{*}{h}_i)$ and that i could imitate k. This cannot be optimal and the above supposition must be false. ■

Lemma C2: *If i is upward distorted, then, for all j that are distinct continuations of i, j is either upward distorted or the set of distinct continuations of j is empty.*

[37] Given (A.12), $\overset{*}{h}_j = 1$ is not possible because i is downward distorted.

Proof: Suppose, to the contrary, that j is both downward distorted and the set of distinct continuations of j is not empty. Because i upward distorted, $\theta_i > w_i$; j is a continuation of i implies that $w_i \geq w_j$; and j being downward distorted implies that $w_j > \theta_j$. Hence, it follows that $\theta_i > \theta_j$. If k is a distinct continuation of j, $w_j \geq w_k$, then, by Lemma A, $\overset{*}{h}_k < \overset{*}{h}_j$ and $\overset{*}{c}_k < \overset{*}{c}_j$. Because $w_i \geq w_j$, this means that i prefers the allocation of k and can imitate k yielding a contradiction. ■

Lemma C: *If i is upward distorted, then, for all j that are extended continuations of i,*

$$\overset{*}{c}_j \geq \overset{*}{c}_i \quad \textit{and} \quad \overset{*}{h}_j \geq \overset{*}{h}_i. \tag{A.14}$$

Proof: Consider an extended continuation $(j_1, j_2, \ldots, j_K)$. If j_1 is is a distinct continuation, then by Lemma B, (A.14) is satisfies; if non distinct, (A.14) is trivially satisfied. If j_1 is downward distorted, then from Lemma C1 and Lemma C2 we know that $j_1 = j_K$ and the proof is complete. If upward distorted, repeat the above argument and the proof is completed by induction. ■

Lemma D: *If i is downward distorted, then the set of sources of i is non-empty and each member j of the set must be downward distorted or $\overset{*}{h}_j = 1$. Moreover, θ_i must be greater than θ_j.*

Proof: First of all, suppose that the set of sources is empty. This means that either there is no j such that i is a continuation of j or that if i is a continuation of j, then either

$$\overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_i \leq \overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_i \tag{A.15}$$

or

$$\overset{*}{h}_j \leq \overset{*}{h}_i \tag{A.16}$$

or both. If there is no j such that i is a continuation of j, then it cannot be optimal for i to be downward distorted. Thus if i is downward distorted there exists some j such that i is a continuation of j, that is

$$\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j = \overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_j \quad \text{and} \quad w_j \geq w_i. \tag{A.17}$$

Among all j such that i is a continuation of j, pick that one with the smallest θ_j.

Next suppose that $\overset{*}{h}_j \leq \overset{*}{h}_i$. We know that $w_j \geq w_i > \theta_i$. If j is downward distorted and i is a distinct continuation of j, then, Lemma A yields a contradiction. Thus, if j is downward distorted, it must be that

$$\overset{*}{h}_i = \overset{*}{h}_j \quad \text{and} \quad \overset{*}{c}_i = \overset{*}{c}_j. \tag{A.18}$$

If $\theta_i > \theta_j$, because θ_j is the smallest and because $\omega_j > \theta_j$, we can increase j's consumption and hours worked with indifference and satisfy the incentive compatibility constraints. This generates a surplus and i being a continuation of j cannot be part of the optimal solution. Hence, we have that $\theta_j > \theta_i$. But this in turn means that there is no reason for j to be downward distorted and because i is downward distorted $\overset{*}{h}_j = 1$ is not possible. Thus, j is not downward distorted and we have that $\theta_j > w_j$ and hence that $\theta_j > \theta_i$. If i is a distinct continuation of j Lemma B yields a contradiction and (A.18) holds. But this implies that i being a continuation of j cannot be part of the optimal solution. Thus, $\overset{*}{h}_j > \overset{*}{h}_i$.

Suppose that (A.15) is true. Substituting (A.17) into (A.15) shows that $\theta_j > \theta_i$ and hence there is no reason for i to be downward distorted and a continuation of j. Thus, θ_j cannot have been the smallest among the set of those j such that i is a continuation of j and the set of sources of i is non empty.

Now applying (A.17) and the definition of a source, (4.2), shows that $\theta_i > \theta_j$ and hence that j is not upward distorted; j is either downward distorted or $\overset{*}{h}_j = 1$. ■

Lemma E: *If i is downward distorted and if j is a fundamental source of i ($\overset{*}{h}_j = 1$), then there is no k such that j is a distinct continuation of k.*

Proof: Suppose that such a k exists. Then we have

$$\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j = \overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_j \quad \text{and} \quad w_j \geq w_i, \tag{A.19}$$

$$\overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_i > \overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_i \quad \text{and} \quad \overset{*}{h}_i < 1, \tag{A.20}$$

and

$$\overset{*}{c}_k + (1 - \overset{*}{h}_k)\theta_k = \overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_k \quad \text{and} \quad w_k \geq w_j. \tag{A.21}$$

From (A.19)—(A.21) and the downward distortion of i we obtain

$$w_k \geq w_j \geq w_i > \theta_i > \theta_j \tag{A.22}$$

where the last inequality follows from Lemma D. If $\theta_k < \theta_j$ then k is downward distorted and Lemma A yields $\overset{*}{h}_j < \overset{*}{h}_k < 1$, a contradiction. If $\theta_k > \theta_j$, then from (A.19) and (A.21) we have

$$\overset{*}{c}_k + (1 - \overset{*}{h}_k)\theta_k < \overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_k; \tag{A.23}$$

the latter means that k prefers the allocation of i and, because $w_k \geq w_i$, could obtain it. This cannot be part of an optimal solution. ■

Lemma F: *There do not exist i and j such that*

$$\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_i > \overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_i \tag{A.24}$$

and

$$\overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_j > \overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j. \tag{A.25}$$

Proof: Suppose there exist i and j such that (A.24) and (A.25) hold; it follows that

$$(\overset{*}{h}_j - \overset{*}{h}_i)(\theta_j - \theta_i) > 0. \tag{A.26}$$

If $\theta_i > \theta_j$ and hence that $\overset{*}{h}_i > \overset{*}{h}_j$, it must be that j cannot imitate i and hence $w_j < w_i$. In this case however i prefers the bundle of j and can imitate j; this is not incentive compatible. ■

Lemma G: *If i is downward distorted, j is upward distorted, f is a fundamental source of i, k is an extended continuation of j, $\overset{*}{c}_j \geq \overset{*}{c}_i$, $\overset{*}{h}_j \neq \overset{*}{h}_i$, and $\omega_j \geq \omega_i$ then,*

$$U(\overset{*}{c}_k + (1 - \overset{*}{h}_k)\theta_k) > U(\overset{*}{c}_f) \tag{A.27}$$

Proof: In order to prove Lemma G it is helpful to define the following concepts. For any upward distorted j, let

$$\mathcal{J} = (j_0, j_1, \dots, j_M) \tag{A.28}$$

be a set of types that are distinct extended continuations of j where $j_0 = j$, and for all j_m and $j_{m+1} \in \mathcal{J}$, $\overset{*}{h}_{j_m} < \overset{*}{h}_{j_{m+1}}$ and $\overset{*}{c}_{j_m} + (1 - \overset{*}{h}_{j_m})\theta_{j_m} = \overset{*}{c}_{j_{m+1}} + (1 - \overset{*}{h}_{j_{m+1}})\theta_{j_m}$.

Now define the function $J : [\overset{*}{h}_j, 1] \mapsto [\overset{*}{c}_j, \bar{c}]$ such that

$$J(h) = \overset{*}{c}_{j_m} + (h - \overset{*}{h}_{j_m})\theta_{j_m} \quad \text{for} \quad h \in [\overset{*}{h}_{j_m}, \overset{*}{h}_{j_{m+1}}] \tag{A.29}$$

and

$$J(h) = \overset{*}{c}_{j_M} + (h - \overset{*}{h}_{j_M})\theta_{j_M} \quad \text{for} \quad h \in [\overset{*}{h}_{j_M}, 1] \tag{A.30}$$

Graphed in the $c-h$ space, the function $J(h)$ corresponds to a continuous and positively slopped line that connects by linear segments a set of distinct continuations of j and, if $\overset{*}{h}_{j_M} < 1$, extends this line along j_M indifference curve up to the point where $h = 1$.

For any downward distorted i, let

$$\mathcal{I} = (i_0, i_1, \ldots, i_N) \tag{A.31}$$

be a set of types that are distinct extended sources of j where $i_0 = i, i_N = f$ (f being a fundamental source of i), and for all i_n and $i_{n+1} \in \mathcal{I}$, $\overset{*}{h}_{i_n} < \overset{*}{h}_{i_{n+1}}$ and $\overset{*}{c}_{i_n} + (1 - \overset{*}{h}_{i_n})\theta_{i_{n+1}} = \overset{*}{c}_{i_{n+1}} + (1 - \overset{*}{h}_{i_{n+1}})\theta_{i_{n+1}}$.

Now define the function $I : [\overset{*}{h}_i, 1] \mapsto [\overset{*}{c}_i, \bar{c}]$ such that

$$I(h) = \overset{*}{c}_{i_{n+1}} + (h - \overset{*}{h}_{i_{n+1}})\theta_{i_{n+1}} \quad \text{for} \quad h \in [\overset{*}{h}_{i_n}, \overset{*}{h}_{i_{n+1}}] \tag{A.32}$$

Graphed in the $c-h$ space, the function $I(h)$ corresponds to a continuous and positively slopped line that connects by linear segments a set of distinct sources of i that extend all the way to on of its fundamental sources.

The functions $J(h)$ and $I(h)$ are not uniquely defined since the sets $\mathcal{J}$ and $\mathcal{I}$ are not generally unique. All statements invoking these functions refer to any function satisfying the above definition.

Given these definitions, we now show that if i and j satisfy the statement of the Lemma, then

$$J(h) > I(h) \quad \text{for} \quad h \in [\max(\overset{*}{h}_i, \overset{*}{h}_j), 1]. \tag{A.33}$$

Once it is shown that the above strict inequality must hold, it is then trivial to show that (A.27) must hold since (A.27) is just a special case when $h = 1$.

Let us first show that $J(\max(\overset{*}{h}_i, \overset{*}{h}_j)) \leq I(\max(\overset{*}{h}_i, \overset{*}{h}_j))$ leads to a contradiction. If $\overset{*}{h}_i = \max(\overset{*}{h}_i, \overset{*}{h}_j)$, then we know (from the definition of $I(h)$) that $I(\overset{*}{h}_i) = \overset{*}{c}_i$ and by assumption that $\overset{*}{c}_j \geq \overset{*}{c}_i$; however, as $\theta_j > \theta_i$ this leads to a contradiction. If $\overset{*}{h}_j = \max(\overset{*}{h}_i, \overset{*}{h}_j)$, then we know (from the definition of $J(h)$) that $J(\overset{*}{h}_j) = \overset{*}{c}_j$. Moreover, from the last statement in Lemma D, we know that $I(h) < \overset{*}{c}_i + (h - \overset{*}{h}_i)\theta_i$ for $h > \overset{*}{h}_i$ and hence $I(\overset{*}{h}_j) < \overset{*}{c}_i + (\overset{*}{h}_j - \overset{*}{h}_i)\theta_i$. By the fact that i is downward distorted, j is upward distorted, and $w_j \geq w_i$, we know that $\theta_j > \theta_i$. Combining these elements we obtained that $c_j = J(\overset{*}{h}_j) \leq I(\overset{*}{h}_j) < \overset{*}{c}_i + (\overset{*}{h}_j - \overset{*}{h}_i)\theta_j$. Rewriting these inequalities we obtain that $\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j < \overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_j$, which is inconsistent with the incentive compatibility constraint associated with type j not mimicing type i. Hence if $I(h) \geq J(h)$ for some $h \in [\max(\overset{*}{h}_i, \overset{*}{h}_j), 1]$, it must be that $J(h)$ and $I(h)$ cross (or touch).

So let us denote by $\bar{h}$ be the first point where $J(\bar{h}) = I(\bar{h})$ and let j'_m be the type with the largest value of h in $\mathcal{J}$ that is strictly less then $\bar{h}$ and let $i_{n'}$ be the point in $\mathcal{I}$ with the smallest value of h that is larger or equal to $\bar{h}$. From the definitions it follows that

$$\overset{*}{c}_{j_{m'}} + (1 - \overset{*}{h}_{j_{m'}})\theta_{j_{m'}} \leq \overset{*}{c}_{i_{n'}} + (1 - \overset{*}{h}_{i_{n'}})\theta_{j_{m'}}, \tag{A.34}$$

$$\overset{*}{c}_{i_{n'}} + (1 - \overset{*}{h}_{i_{n'}})\theta_{i_{n'}} = \overset{*}{c}_{i_{n'-1}} + (1 - \overset{*}{h}_{i_{n'-1}})\theta_{i_{n'}}, \tag{A.35}$$

$$\overset{*}{c}_{i_{n'}} + (1 - \overset{*}{h}_{i_{n'}})\theta_{i_{n'}} < \overset{*}{c}_{j_{m'}} + (1 - \overset{*}{h}_{j_{m'}})\theta_{i_{n'}}, \tag{A.36}$$

and

$$\overset{*}{c}_{j_{m'}} + (1 - \overset{*}{h}_{j_{m'}})\theta_{j_{m'}} > \overset{*}{c}_{i_{n'-1}} + (1 - \overset{*}{h}_{i_{n'-1}})\theta_{j_{m'}}. \tag{A.37}$$

Combining (A.34) and (A.37) yields

$$\overset{*}{c}_{i_{n'-1}} + (1 - \overset{*}{h}_{i_{n'-1}})\theta_{j_{m'}} < \overset{*}{c}_{i_{n'}} + (1 - \overset{*}{h}_{i_{n'}})\theta_{j_{m'}} \tag{A.38}$$

which in conjunction with (A.35) implies that

$$\theta_{i_{n'}} > \theta_{j_{m'}}. \tag{A.39}$$

From (A.36) it is then clear that

$$\omega_{j_{m'}} > \omega_{i_{n'}}, \tag{A.40}$$

which implies in (A.34) holds with equality.

Since (A.34) holds with equality and $\omega_{j_{m'}} > \omega_{i_{n'}}$, Lemma E implies that $i_{n'}$ is downward distorted (since it is not a fundamental source) and Lemma A implies that $j_{m'}$ is upward distorted (since $0 < \overset{*}{h}_{j_{m'}} < 1$). Hence we know that $\theta_{j_{m'}} > \omega_{j_{m'}}$ and $\omega_{i_{n'}} > \theta_{i_{n'}}$. In conjunction with (A.39) we find that $\omega_{j_{m'}} < \omega_{i_{n'}}$, which contradicts (A.40). This completes the demonstration that $J(h) > I(h)$ (on their shared domain) and therefore implies that (A.27) must hold by the very fact that $\mathcal{J}$ and $\mathcal{I}$ can be chosen such that $J(1) = \overset{*}{c}_k + (1 - \overset{*}{h}_k)\theta_k$ and $I(1) = \overset{*}{c}_f$.

■

Lemma H: *If j is an extended continuation of an upward distorted i then*

(1) If there exists a k with $h_k = 0$ and $\omega_k \leq \omega_i$, then

$$U(\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j) > U(\overset{*}{c}_k + \theta_j) \tag{A.41}$$

(2) Otherwise

$$U(\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j) > U(\theta_j) \tag{A.42}$$

Proof: Incentive compatibility implies that the relationships (A.41) and (A.42) hold with a weak inequality, therefore all that must be shown is that these relationships cannot hold with strict equality. Suppose they do hold with equality then, by the same argument as that used in Lemma C2, we know that j must be upward distorted. But if j is upward distorted and one of the two relationships holds with equality, then change the individual allocation of j to $\{\overset{*}{c}_j = \overset{*}{c}_k, \overset{*}{h}_j = 0\}$ if there is equality in case (1) and change it to $\{\overset{*}{c}_j = 0, \overset{*}{h}_j = 0\}$ if there is equality in case (2). The above modification of the allocation is incentive compatible (The argument here is similar to that of Lemma A.), and generates a surplus. Since the surplus can always be divided equally among individuals and thereby improve welfare, this leads to a contradiction. ■

Proof of Proposition 2A A function $g(\cdot)$ satisfying (2.7) and (2.8) always exist if there cannot exits a downward distorted i and an an upward distorted j such that $w_j \geq w_i$, $\overset{*}{c}_j \geq \overset{*}{c}_i$, and $(w_j, \overset{*}{c}_j) \neq (w_i, \overset{*}{c}_i)$. Therefore, let us assume that such i and j exist, and show that it leads to a contradiction. From the fact that i is downward distorted and j is upward distorted we can immediately infer that

$$\theta_j > w_j \geq w_i > \theta_i. \tag{A.43}$$

Because $w_j \geq w_i$, it must be the case that (A.25) is false for otherwise j would imitate i. Hence, either

$$\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j = \overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_j \tag{A.44}$$

or

$$\overset{*}{c}_j + (1 - \overset{*}{h}_j)\theta_j > \overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_j. \tag{A.45}$$

is true. If, (A.44) holds, then, as $w_j \geq w_i$, i is a continuation of j. If $\overset{*}{c}_j > \overset{*}{c}_i$, then i is a distinct continuation of j and by Lemma B, $\overset{*}{c}_j < \overset{*}{c}_i$, a contradiction. Therefore, $\overset{*}{c}_j = \overset{*}{c}_i$ and $\overset{*}{h}_j = \overset{*}{h}_i$. Because i is downward distorted, Lemma C1 implies that the set of distinct continuations of i is empty. That is, there does not exist k such that

$$\overset{*}{c}_i + (1 - \overset{*}{h}_i)\theta_i = \overset{*}{c}_k + (1 - \overset{*}{h}_k)\theta_i \quad \text{and} \quad w_i \geq w_k. \tag{A.46}$$

In this case, consider decreasing $\overset{*}{c}_i$ by ϵ ($dc_i < 0$) and transferring the resulting savings to a fundamental source of i denoted f. From Lemma D, such a fundamental source always exits. By Lemma H and the fact that $\theta_i < \theta_j$, type i's participation constraint cannot be strictly binding nor can can i be indifferent between his or her allocation and that of some individual k with $\overset{*}{h}_k = 0$. Therefore, by Lemma E and the fact that i has no distinct continuations, for ϵ small enough, such a transfer does not interfere with any of the incentive compatibility constraints or participation constraints. Moreover, from the definition of a source, $U(\overset{*}{c}_i+(1-\overset{*}{h}_i)\theta_i) > U(\overset{*}{c}_f)$ and therefore such a deviation would be welfare improving, hence a contradiction.

Thus suppose that (A.45) is true and let f be a fundamental source of i. From Lemma G we know that for every member of the extended continuations of j, their utility level must be strictly greater than that of f. Reduce by ϵ the consumption of all types that are extended continuations of j (including j itself), and transfer the resulting savings to a fundamental source of i. This deviation is welfare improving, by construction it is incentive compatible, and by Lemma H it satisfies the participation constraints; therefore, we have a contradiction. Hence, if $w_j \geq w_i$, $\overset{*}{c}_j \geq \overset{*}{c}_i$, and $(w_j, \overset{*}{c}_j) \neq (w_i, \overset{*}{c}_i)$, then j is downward distorted.

We now turn to proving the boundary conditions given by (2.9). The first boundary condition holds if there cannot exist an upward distorted type i with $\overset{*}{c}_i = \bar{c}$. Therefore, let us assume that such an i exists and show that it leads to a contradiction. By Lemma B, i cannot have a distinct continuation and all nondistinct continuations are upward distorted. There cannot exist a j that is downward distorted, since it would be welfare improving to transfer income from i (and all non-distinct continuations of i) to a fundamental source of j. However, since $\overset{*}{c}_i \geq \omega_i$, $\overset{*}{h}_i = \bar{c}$ (by individual rationality and the fact that i is upward distorted), there must exist a j such that $\overset{*}{h}_j = 1 = \overset{*}{H}_j$ and $\overset{*}{c}_j < \omega_j$, otherwise the budget would not balance. If there is no k such that j is a distinct continuation of k, then again a welfare improvement can be obtained by transferring income from i (and all non-distinct continuations of i) to j. Assume that there is such a k. If so, since k cannot be downward distorted, it must be the case that $\theta_k > \omega_k \geq \omega_j > \omega_i$. But then, by the participation constraint of k, it must be that $U(\overset{*}{c}_k + (1-\overset{*}{h}_k)\theta_k) > U(\overset{*}{c}_j)$, which contradicts the existence of such a k. Hence $g(\cdot)$ can always be chosen to satisfy the first boundary condition.

The second boundary condition holds if there cannot exist an upward distorted type i with $w_i = \bar{w}$. Therefore, let us assume that such an i exists. If the are many of such types, let i be the one with the greatest θ. If there does not exist a j such that j is a continuation of i and $\theta_j > \theta_i$, then a perturbation of the kind $dc_i - dh_i\theta_i = 0$, with $dc_i < 0$, is feasible and releases resources, therefore allowing for a Pareto improvement. However, such a type j cannot exist by the definition of type i (which stipulates that there does not exist a type j with $\omega_j \geq \omega_i$ and $\theta_j > \theta_i$). ■

Proof of Proposition 2B: A function $f(\cdot)$ satisfying (2.10) and (2.11) can always be constructed if there cannot not exit a downward distorted i and an an upward distorted j such that $\overset{*}{h}_j \leq \overset{*}{h}_i$, $\overset{*}{c}_j \geq \overset{*}{c}_i$, and $(\overset{*}{c}_j, \overset{*}{h}_j) \neq (\overset{*}{c}_i, \overset{*}{h}_i)$. Therefore let us assume that such an i and j exits and show that it leads to a contradiction. The incentive compatibility constraint associated with i not mimicing j implies that $w_j > w_i$. But, by the proof of Proposition 2A, we know that there cannot not exist a downward distorted i and an an upward distorted j such that $w_j \geq w_i$, $\overset{*}{c}_j \geq \overset{*}{c}_i$, and $(\overset{*}{c}_j, w_j) \neq (\overset{*}{c}_i, w_i)$, hence a contradiction.

The first boundary condition given in the proposition is trivially satisfied since all types with $\overset{*}{h} = 0$ have $\overset{*}{c} = \underline{c}$. The second boundary condition is satisfied for exactly the same reason as provided in the proof of the first boundary condition of Proposition 2A. ■

Proof of Proposition 3: First suppose that there exists i such that $\omega_i < \theta_i$ and that $\overset{*}{h}_i > 0$. By Lemma A1, there exists k such that i is a continuation of k. If $\theta_k \leq \theta_i$ then there is, as in the reasoning of Lemma A1 no reason for i to be upward distorted; hence it must be the case that $\theta_k > \theta_i$. But by the hypothesis of negative correlation this implies that $\omega_k < \omega_i$ in which case i cannot be a continuation of k yielding a contradiction. If $\omega_i < \theta_i$ then $\overset{*}{h}_i = 0$.

Now suppose that there exists i such that $\omega_i > \theta_i$ and the $\overset{*}{h}_i < 1$. By Lemma D these exists a source of I, say, f, which by Lemma F can be moved upwards. If there are no continuations of i we can move the source chain from i to f upward raising the utilities of everyone in the chain and hence this could not have been part of an optimal solution. Suppose that there is a continuation of i, say k. Then, $\omega_k \leq \omega_i$ which implies by the hypothesis of negative correlation that $\theta_k \geq \theta_i$ and hence that the utility of k is greater than or equal to that of i and hence of f. This is true for all extended continuations of i and hence the entire source chain can be moved upwards and this cannot have been part of an optimal solution. ■

Proof of Proposition 4: Consider first the case where $\omega_i > \theta_i$ and suppose that $\overset{*}{h}_i < 1$. By Lemma D, i has a source, say, f, where $\omega_f > \omega_i$ and by the hypothesis of positive correlation that $\theta_f > \theta_i$. This contradicts the claim that f is a source of i and hence $\overset{*}{h}_i = 1$.

Now consider the case where $\omega_i < \theta_i$. Example 2 in the text is a case where $\overset{*}{h}_i > 0$ and hence proves that complete efficiency cannot be attained in this case.■

Proof of Proposition 5: A function $t(\cdot)$ satisfying the statement of Proposition 3 can always be constructed if there cannot not exit a i with a positive marginal tax rate and a j with a negative marginal and average tax rate such that $\overset{*}{h}_j \leq \overset{*}{h}_i$, $\overset{*}{c}_j \geq \overset{*}{c}_i$, and $(\overset{*}{c}_j, \overset{*}{h}_j) \neq (\overset{*}{c}_i, \overset{*}{h}_i)$. Once that it recognized that (1) if i faces a positive marginal rate then $\overset{*}{h}_i \leq \overset{*}{H}_i$, and (2) if j faces a negative marginal tax rate then $\overset{*}{h}_j \geq \overset{*}{H}_j$ and the average tax rate is negative; then it is obvious from Proposition 2B that the above statement is true. Therefore, all that needs to be verified is statements (1) and (2) above. If i faces a positive tax rate and $\overset{*}{h}_i > 0 = \overset{*}{H}_i$, then i would prefer his or her left marginal alternative to his or her own allocation, which contradicts incentive compatibility. Conversely, if j faces a negative marginal tax rate and $\overset{*}{h}_j < 1 = \overset{*}{H}_j$, then j would prefer his or her right marginal alternative to his or her own allocation, which is again a contradiction. Moreover, $\overset{*}{h}_j \geq \overset{*}{H}_j$ only if the average tax rate is negative. ■

Proof of Proposition 6: By contradiction. Suppose there is a $\{c'_i, h'_i, w'_i, h'^f_i\}_{i=1}^{NM}$ which is superior to the one defined in the proposition. Then it must be the case that there exist at least one j such that $h_j \neq 0$ and $h^f_j \neq 0$; otherwise $\{\tilde{c}_i, \tilde{h}_i\}_{i=1}^{NM}$ would not solve OP for the modified economy. However, if there exist such a j, we can free up resources by considering the following perturbation

(1) If $\omega_j > \omega^f$, set $h_j = h'_j + h'^f_j$ and $h^f_j = 0$

(1) If $\omega_j = \omega^f$, set $h^f_j = h'_j + h'^f_j$ and $h_j = 0$

Since such a perturbation allows all incentive compatibility constraints to remain satisfied and allows an equal redistribution of positive resources to every type, it allows for a Pareto improvement and thereby leads to a contradiction. ■

REFERENCES

Besley, T. and S. Coates, 1995, The Design of Income Maintenance Programmes, *The Review of Economic Studies*, **62**, 187-221.

Bloomquist, N.S., 1981, A comparison of Tax Bases for a Personal Tax, *The Scandinavian Journal of Economics*, **83**, 390-407.

Brett, C., 1997, Who Should Be on Workfare? The Use of Work Requiremnt as Part of an Optimal Tax Mix, *Oxford Economic Papers*, forthcoming.

Card, D. and P.K. Robins, 1996, Do Financial Incentive Encourage Welfare Recipients to Work?, Social Research and Demonstration Corporation, Canada.

Dasgupta, . and P. Hammond, 1980, Fully Progressive Taxation, *Journal of Public Economics*, **13**, 141-154.

Diamond, P., 1997, Optimal Income Taxation: an Example with a U-Shaped Pattern of Optimal Marginal Tax Rates, *American Economic Review*, forthcoming.

Guesnerie, R., and K. Roberts, 1987, Minimum Wage Legislation as a Second Best Policy, *European Economic Review*, **31**, 491-8

Kanbur, R., M. Keen, and M. Tuomola, 1994, Optimal Nonlinear Income Taxation and the Alleviation of Income-Poverty, *European Economic Review*, **38**, 1613-1633.

Kesselman, J., 1973, A Comprehensive Approach to Income Maintenance: Swift, *Journal of Public Economics*, **2**, 59-88.

Kesselman, J., 1977, Tax Instruments and Economic Equality, *Canadian Journal of Economics*, **10**, 289-299.

Maderner, N., and J.-C. Rochet, 1995, Is it Legitimate to Encourage Work Sharing?, *Scandinavian Journal of Economics*, 621-633.

Marceau, N. and R. Boadway, 1994, Minimum Wage Legislation and Unemployment Insurance as Instruments for Redistribution, *Scandinavian Journal of Economics*, **96**, 67-81.

Mattews, S. and Moore, J. H., 1997, "Optimal Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening", *Econometrica*, **55**, 441-467.

Mirrlees, J.A., 1971, An exploration of the Theory of Optimal Taxation, *The Review of Economic Studies*, **38**, 175-208.

Mirlees, J.A., 1997, Information and Incentives: The Economics of Carrots and Sticks, *Economic Journal*, **107**, 1311-1329.

Rochet, J.-C., 1995, Ironing, Sweeping and Multidimensional Screening, Cahier no 95.11.374, GREMAQ.

Weymark, J., A Reduced-Form Optimal Nonlinear Income Tax Problem, *Journal of Public Economics*, **30**, 199-217.

Weymark, J., Bunching Properties of Optimal Nonlinear Income Taxes, *Social Choice and Welfare*, **3**, 213-232.

Weymark, J., Comparative Static Properties of Optimal Nonlinear Income Taxes, *Econometrica*, **55**, 1165-1186.

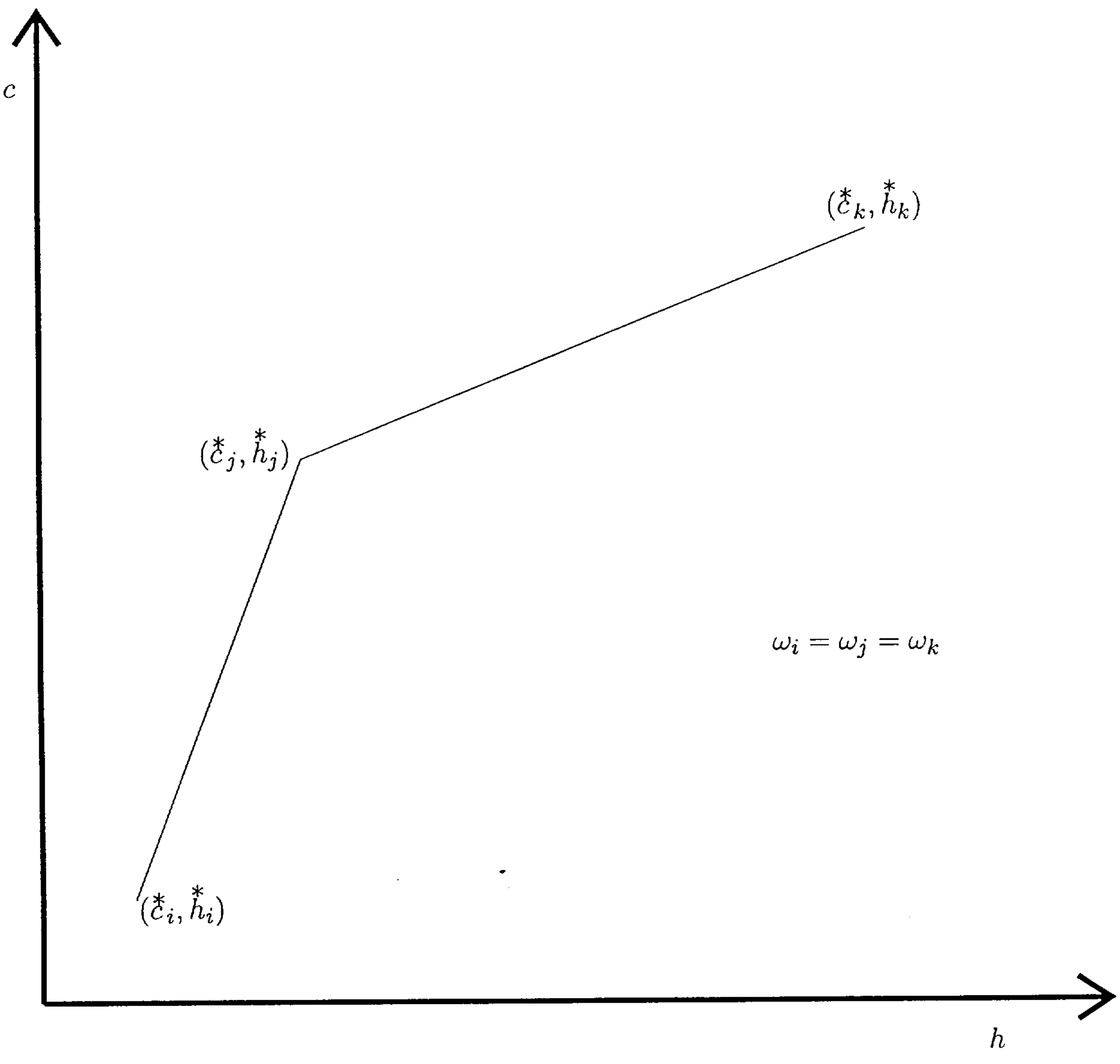

c
$(\overset{*}{c}_k, \overset{*}{h}_k)$
$(\overset{*}{c}_j, \overset{*}{h}_j)$
$\omega_i = \omega_j = \omega_k$
$(\overset{*}{c}_i, \overset{*}{h}_i)$
h

Figure 1A

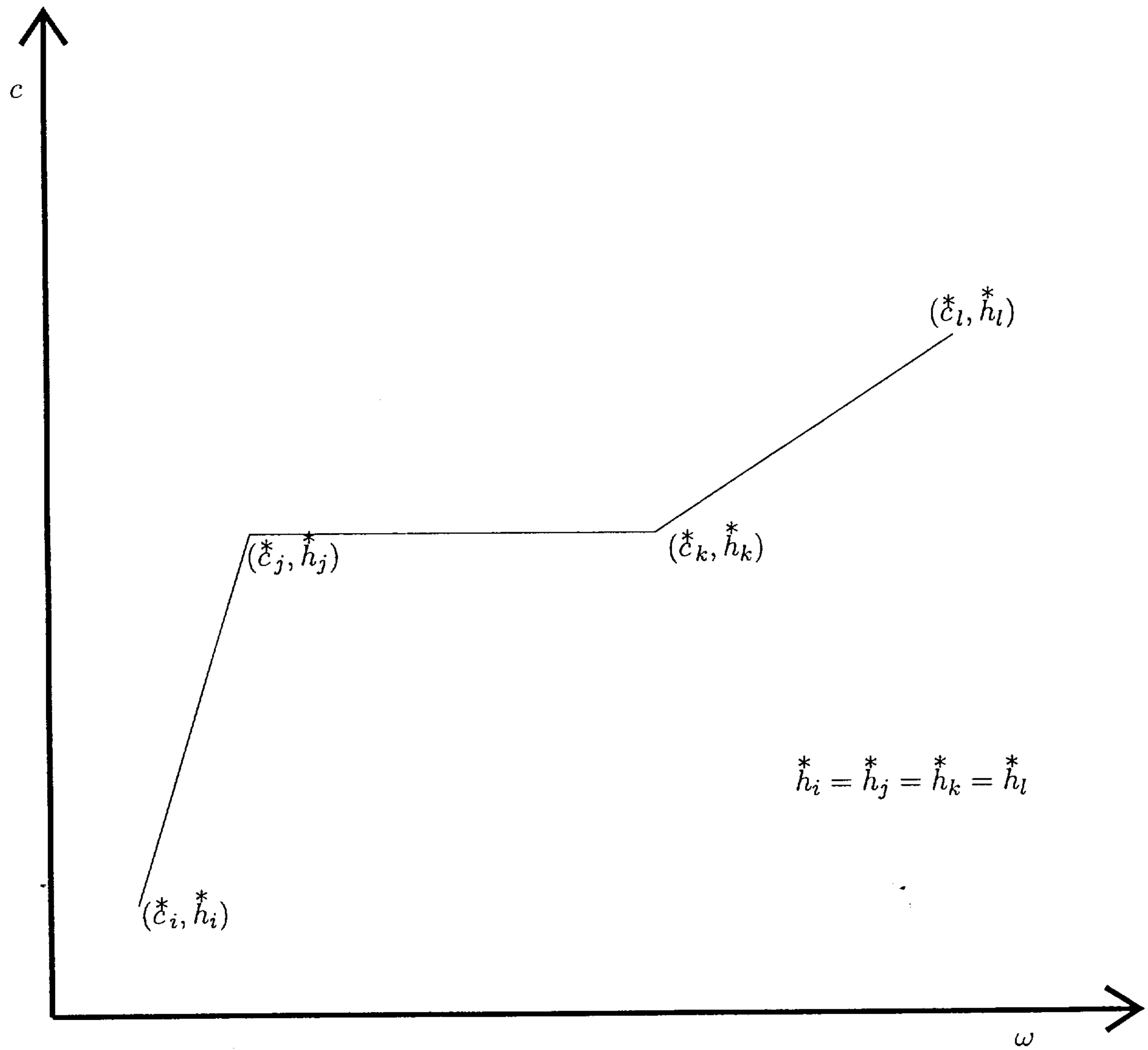

c
$(\overset{*}{c}_l, \overset{*}{h}_l)$
$(\overset{*}{c}_j, \overset{*}{h}_j)$
$(\overset{*}{c}_k, \overset{*}{h}_k)$
$\overset{*}{h}_i = \overset{*}{h}_j = \overset{*}{h}_k = \overset{*}{h}_l$
$(\overset{*}{c}_i, \overset{*}{h}_i)$
ω

Figure 1B

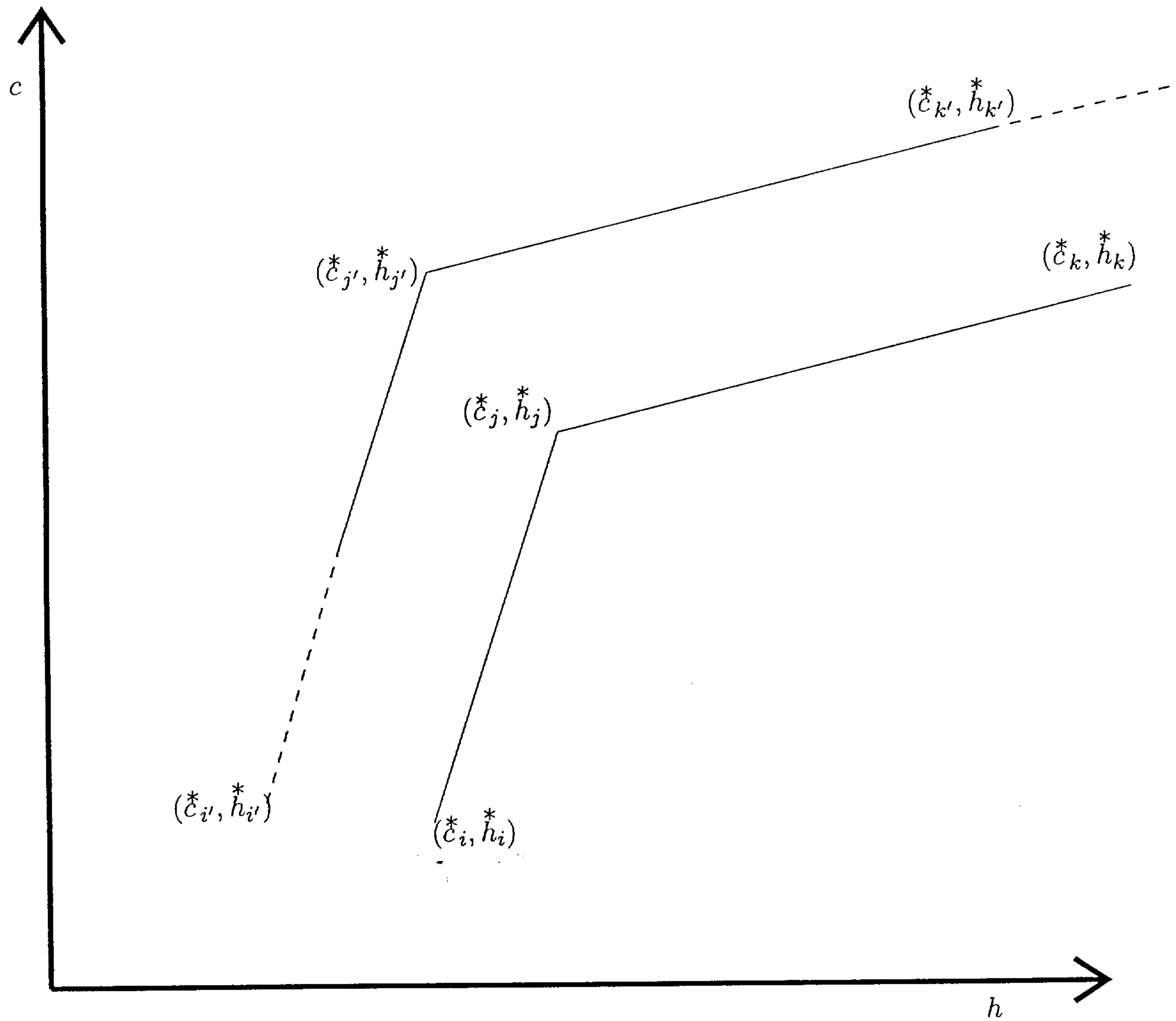
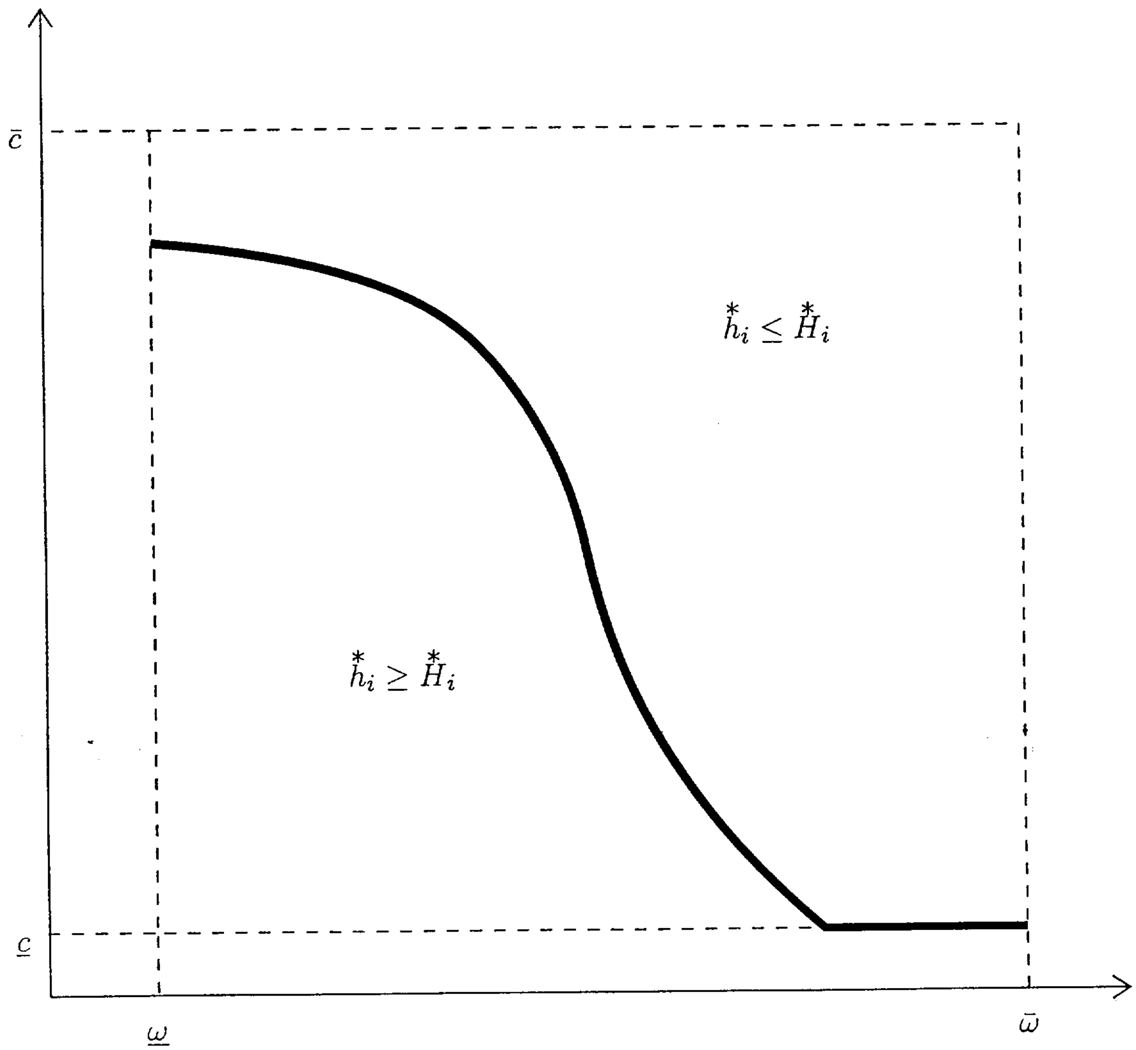

c
h
$(\overset{*}{c}_{k'}, \overset{*}{h}_{k'})$
$(\overset{*}{c}_{j'}, \overset{*}{h}_{j'})$
$(\overset{*}{c}_{k}, \overset{*}{h}_{k})$
$(\overset{*}{c}_{j}, \overset{*}{h}_{j})$
$(\overset{*}{c}_{i'}, \overset{*}{h}_{i'})$
$(\overset{*}{c}_{i}, \overset{*}{h}_{i})$

Figure 1C



$\bar{c}$
$\overset{*}{h}_i \leq \overset{*}{H}_i$
$\overset{*}{h}_i \geq \overset{*}{H}_i$
$\underline{c}$
$\underline{\omega}$
$\bar{\omega}$

Figure 2A

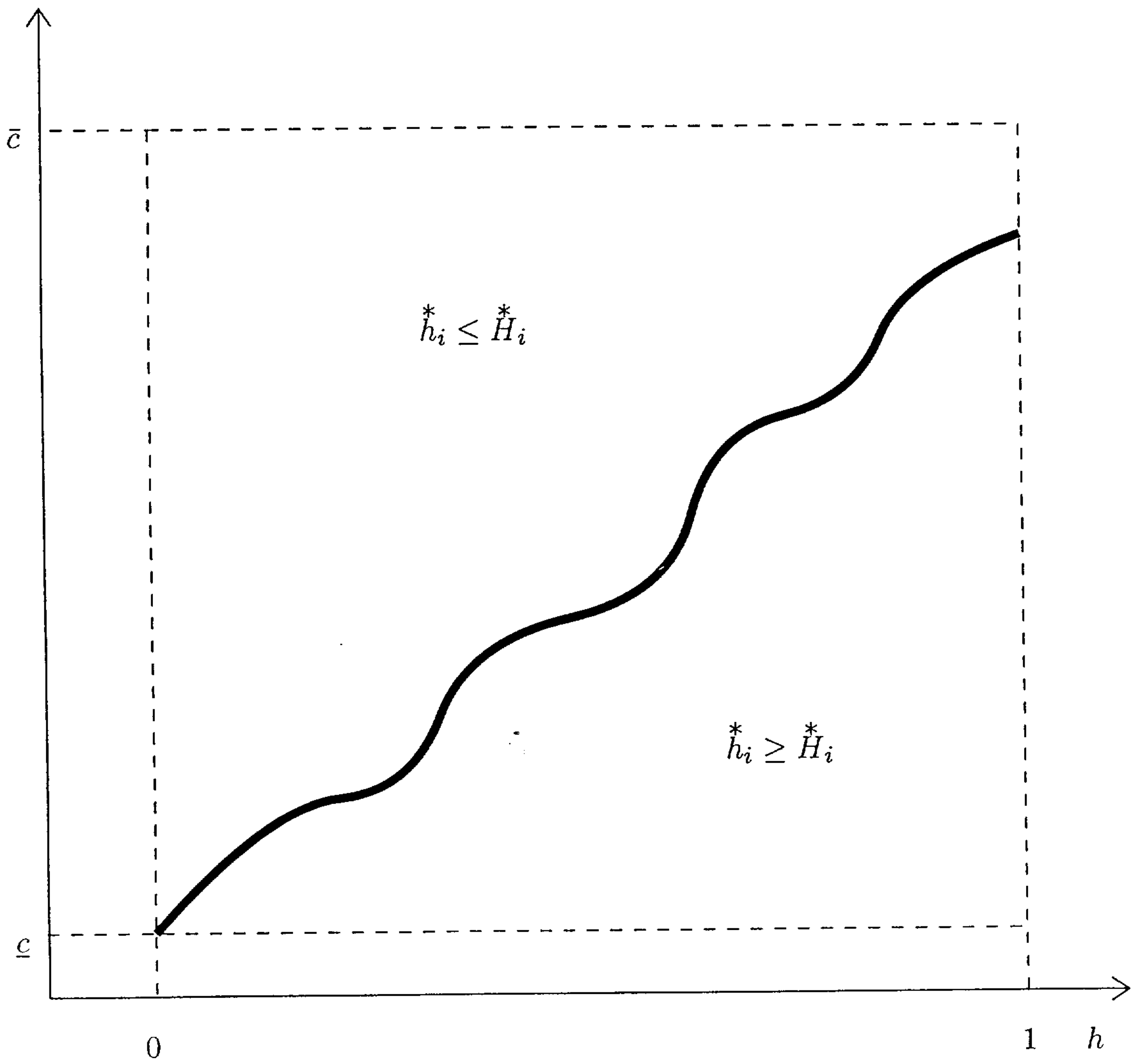

$\bar{c}$
$\overset{*}{h}_i \leq \overset{*}{H}_i$
$\overset{*}{h}_i \geq \overset{*}{H}_i$
$\underline{c}$
0
1
h

Figure 2B

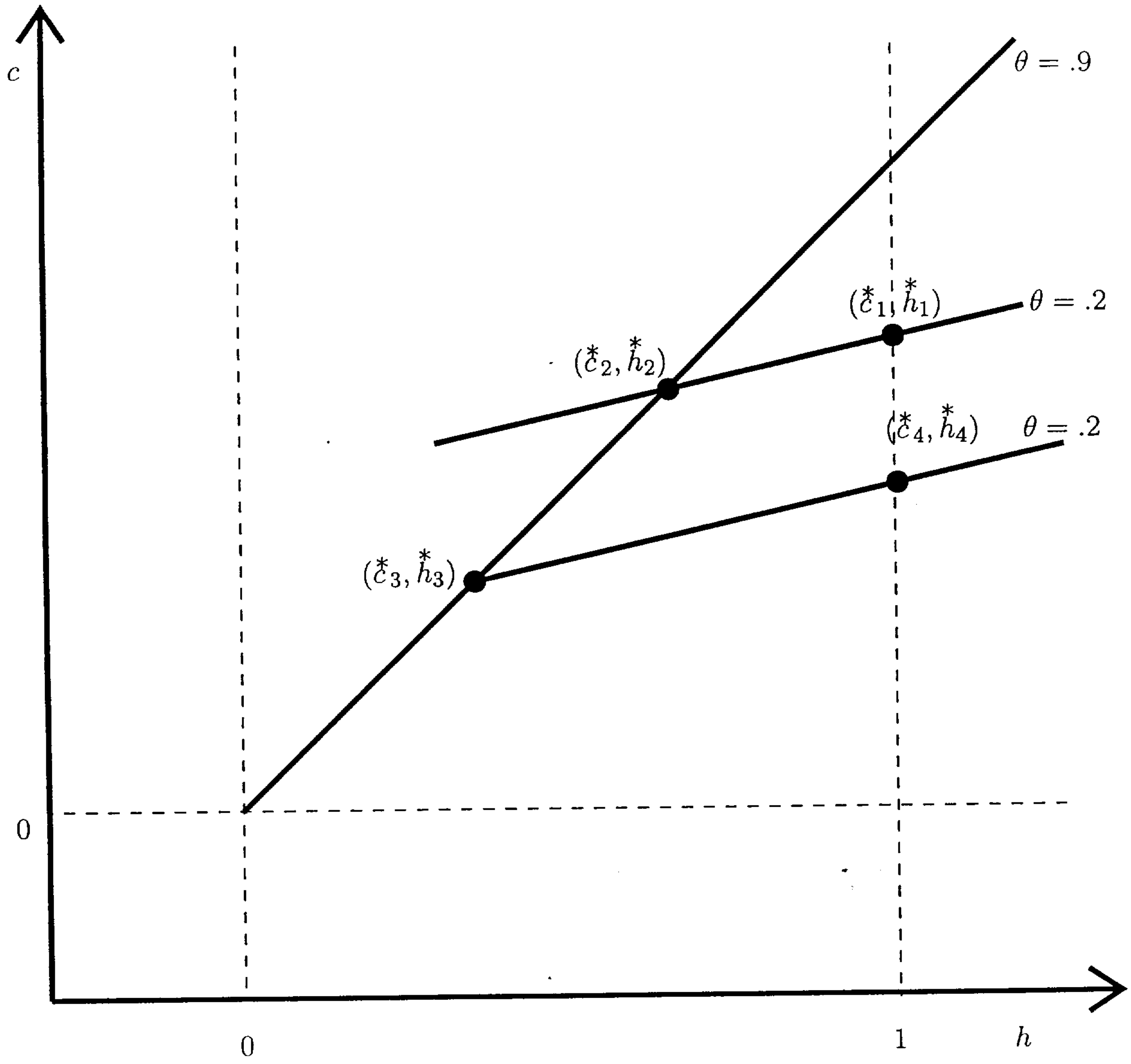


Figure 3

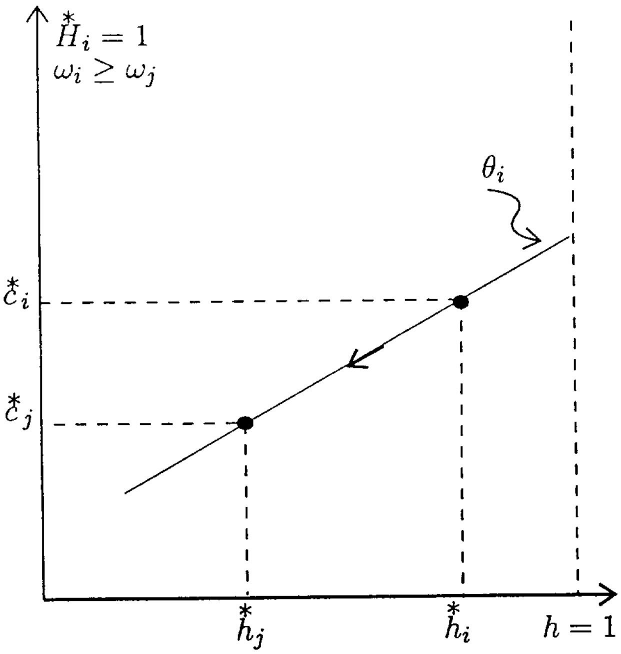


Figure A

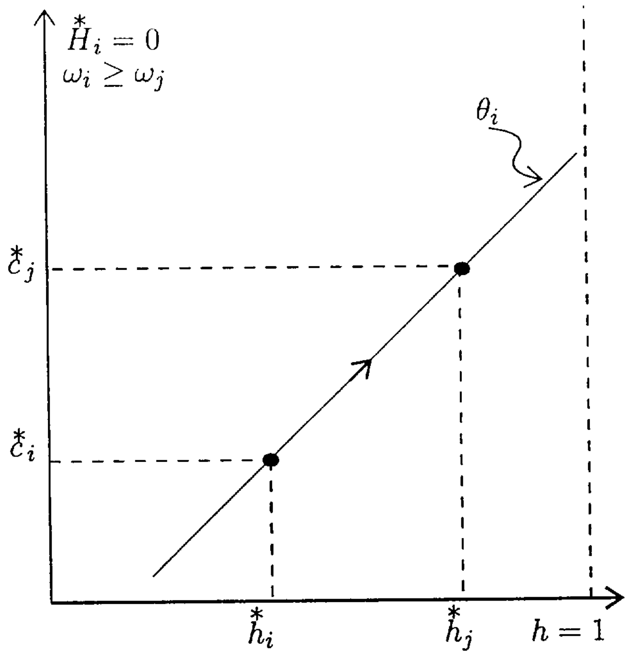


Figure B

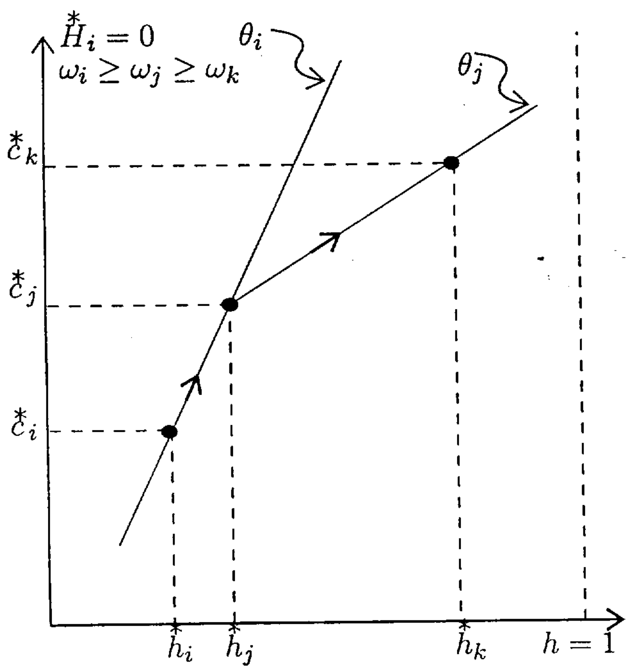


Figure C

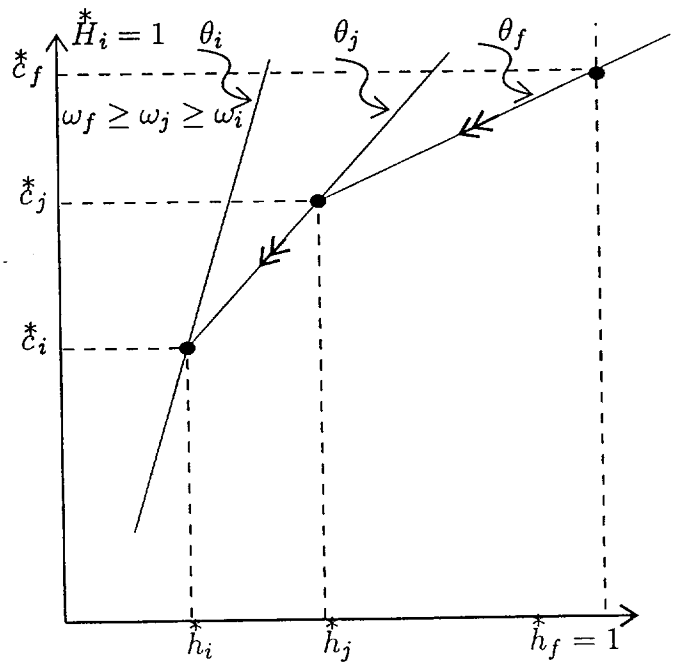
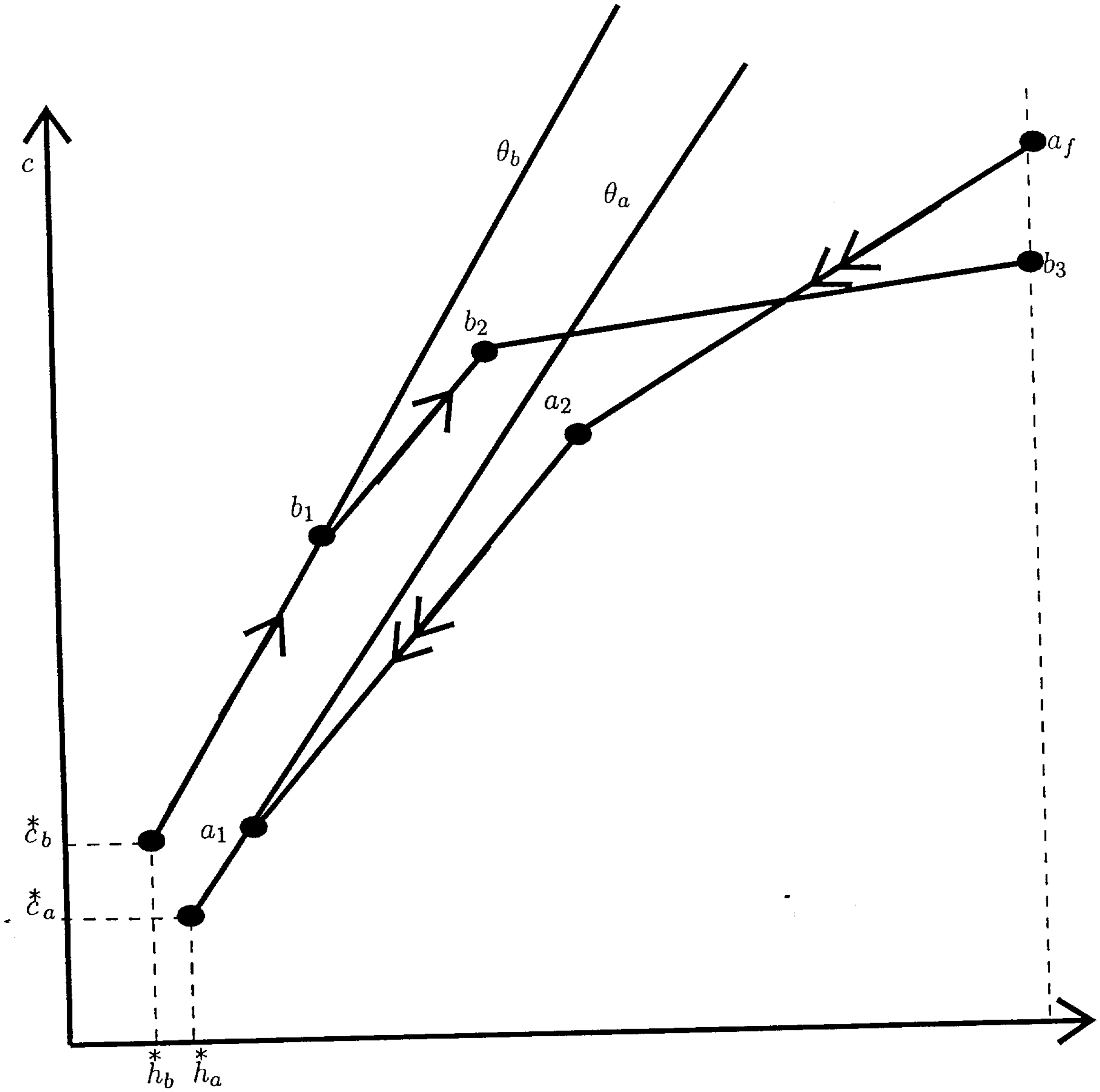


Figure D



c
θ_b
θ_a
a_f
b_3
b_2
a_2
b_1
a_1
$\overset{*}{c}_b$
$\overset{*}{c}_a$
$\overset{*}{h}_b$
$\overset{*}{h}_a$

Figure 4E

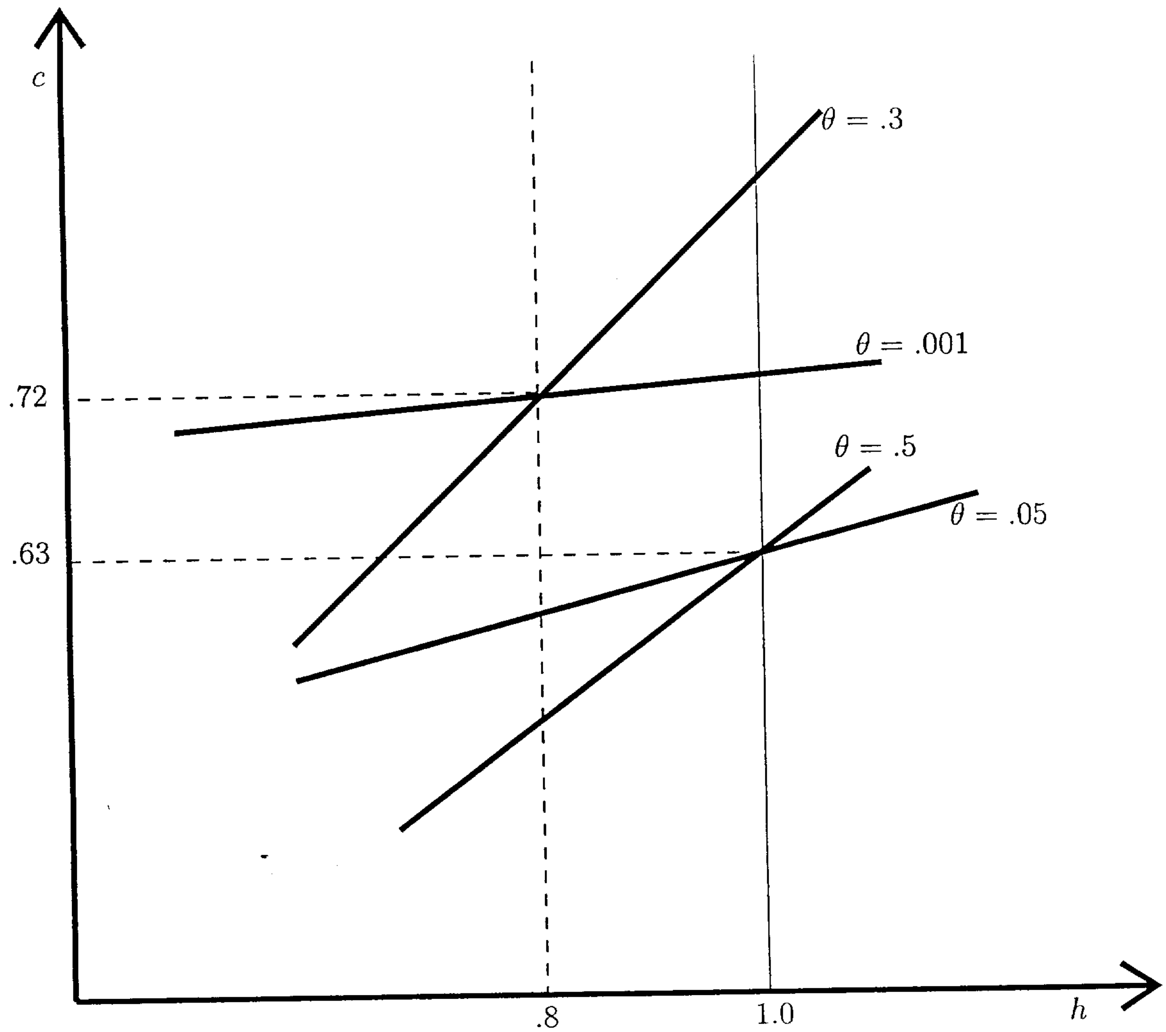


Figure 5