Taxes and Employment Subsidies in Optimal Redistribution Programs[*]

Paul Beaudry, Charles Blackorby, and Dezsö Szalay [†]

August 23, 2007

[†]Paul Beaudry: University of British Columbia, and NBER. (Email: Beaudry@econ.ubc.ca)
Charles Blackorby: University of Warwick and GREQAM (Email: c.blackorby@warwick.ac.uk)
Dezsö Szalay: University of Warwick (Email: Dezso.Szalay@warwick.ac.uk)

Abstract

This paper explores how to optimally set taxes and transfers when taxation authorities: (1) are uninformed about individuals' value of time in both market and non-market activities and (2) can observe both market-income and time allocated to market employment. We show that optimal redistribution in this environment involves distorting market employment upwards for low wage individuals through decreasing wage-contingent employment subsidies, and distorting employment downwards for high wage individuals through positive and increasing marginal income tax rates. In particular, we show that whether a person is taxed or subsidized depends primarily on his wage, with the optimal program involving a cut-off wage whereby workers above the cutoff are taxed as they increase their income, while workers earning a wage below the cutoff receive an income supplement as they increase their income. Finally, we show that the optimal program transfers zero income to individuals who choose not to work.

1 Introduction

In most countries income redistribution is achieved through a variety of programs: these include direct income taxation, employment programs, welfare, unemployment insurance and pension schemes. Viewed as a whole, these programs create intricate incentives and complex redistribution patterns. Since the conditionality of these programs is quite varied, they generally result in a net tax-transfer system that depends not only on income but often depends on the extent of market participation as well. Reasoned economic policy should attempt to identify whether or not these programs are mutually consistent with the goal of redistribution.

The object of this paper is to explore the principles that should guide the evaluation of tax- transfer systems that depend on both market income and on quantity of time worked. In order to illustrate the types of issues we want to address, we start with an example of an individual who pays taxes or receives transfers from a government depending on his interaction with three different systems: an income tax system, a social assistance system (welfare) and an unemployment insurance system.[1] The example is inspired by the Canadian social system, however it has been purposely simplified to clarify issues and therefore the numerical values should be viewed as mainly illustrative.

Let y represent an individual's market income, let h represent the number of weeks (≤ 50) worked by an individual over a year and let T represent total taxes (net of transfers) paid by the individual over a year.

The income tax system:

If $y \leq \$6000$, there is no income tax; on income above $6000, a marginal income tax of 20% is applied (i.e., income tax equals max [.2(y-6000),0]).

The social assistance system (welfare):

If $y \leq \$6000$, the social assistance payment is $\$6000 - y$; if $y > \$6000$, there is no social assistance payment.

The unemployment insurance system:

Letting h be the number of weeks worked, if $h \leq 10$, the individual is not eligible for unemployment insurance; if $10 < h \leq 30$, then the individual is eligible for $h - 10$ weeks of unemployment insurance payments at 60% of weekly wages, up to a maximum payment of $400 per week; if $30 < h < 50$, the individual is eligible for $50 - h$ weeks of unemployment insurance payments at 60% of weekly wages, up to a maximum payment of $400 per week.

Consider the net tax implication of these three systems combined. The net amount of taxes paid (or transfer received) depends both on an individual's wage rate and on the number of weeks worked. Hence the pattern of tax rates faced by individuals varies with different market wage rates. In particular, consider the case where individual 1 earns $600 per week worked, and individual 2 earns $1000 per week. Then the net taxes-transfers, T, paid by individuals 1 and 2 as a function of annual income are given below. In calculating

---

[1]For simplicity, we have not included in the example the interaction with the pension system. However, the issues we address are also potentially relevant for pension systems since these programs have pay-outs that depend both on income earned and on amount worked.

these tax rates, we assume that an individual receives unemployment insurance payments for any eligible non-working weeks:[2]

Tax function of individual 1:

If $y \leq 6000$, $T = y - 6000$ (marginal rate of 100%);

If $6000 < y \leq 18000$, $T = -.4(y - 6000)$ (marginal rate of -40%);

If $18000 < y$, $T = -4800 + .8(y - 18000)$ (marginal rate of 80%);

Tax function of individual 2:

If $y \leq 6000$, $T = y - 6000$ (marginal rate of 100%);

If $6000 < y \leq 10000$, $T = .2(y - 6000)$ (marginal rate of 20%);

If $10000 < y \leq 30000$, $T = 800 - .2(y - 10000)$ (marginal rate of -20%);

If $30000 < y$, $T = -3200 + .6(y - 30000)$ (marginal rate of 60%).

There are three aspects to notice about this tax-transfer system. First, the tax rate depends not only on income but also depends on a worker's revealed market type, that is his wage rate. In particular, note that marginal tax rates are different at different income levels depending on a worker's wage rate. Second, the individuals face high marginal tax rates at both high and low income levels. Third, the individuals face negative marginal tax rates for intermediate income segments. Let us emphasize that all these features stand in stark contrast to the prescriptions one would derive from a Mirrlees' type optimal tax problem. However, given that the above example allows tax rates to be wage dependent, we immediately know that Mirrlees' analysis does not directly apply and hence an alternative framework is needed.

In this paper, we examine an optimal income tax problem in hope of providing guidance on how to design such a system. For example, we would like to know how to best set a tax and transfer system when the government can design the system to depend both on income and wage rates (or the number of weeks worked). Moreover, since we believe that one of the concerns of governments is to avoid transferring substantial income to individuals that simply do not want to engage in market employment, our analysis recognizes that individuals may have different valuations for their non-market time.

Our approach to the problem follows the optimal non-linear income taxation literature

---

[2]In this Canadian inspired example, we treat the unemployment insurance system as a transfer system that depends on whether an individual is employed or unemployed, and we treat the employment decision as under the worker's control. Obviously, this depiction departs from the aim of the Canadian Unemployment Insurance system which is directed at supporting involuntarily unemployed individuals. However, in practice, it is widely recognized, in both the press and by politicians, that a large fraction of unemployment insurance payments in Canada have traditionally been paid out to individuals that are more akin to voluntarily unemployed individuals than to involuntarily unemployed individuals. For example, there are many examples of individuals who have taken up employment for a duration that exactly corresponds to the minimum duration needed to obtain unemployment benefits, then, once they attain the duration needed to obtain benefits, they return to unemployment until benefits run out and then they start the cycle anew. For such individuals, the unemployment insurance system plays the role of a work contingent transfer scheme precisely as we have modelled it in the example. Part of the impetus for this paper derives from this observation, but instead of dismissing such transfers out of hand as undesirable since they did not reflect the original goal of the unemployment insurance system, our aim is to devise a framework in which we can discuss whether such employment contingent transfers may be desirable when viewed as part of a social transfer system.

as pioneered by Mirrlees (1971),[3] that is, we approach redistribution as a welfare maximization problem constrained by informational asymmetries. However, we depart in two directions from Mirrlees' formulation. 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 or pension programs – employ information on time worked (either in years, weeks or hours) in order to determine eligibility; therefore it seems reasonable to allow for such a possibility when considering how best to redistribute income. Hence, the environment we examine includes (1) taxation authorities which are uninformed about individuals' potential value of time in market activities and about their potential value of time in non-market activities,[4] and (2) income transfers that can be contingent on both earned (market) income and on the allocation of time to market employment and, as a result, also on the wage rate. 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
• A cutoff wage, where individuals with wage above the cutoff are taxed and individuals with wage below the cutoff are subsidized.
• For individuals below the cutoff wage, their employment level is distorted upwards as they face wage-contingent income subsidies that decrease as income increases.
• For individuals above the cutoff wage, their employment level is distorted downwards as they face positive and increasing marginal tax rates as they increase their income.
• Individuals that choose not to work receive no income transfer.

The above results provide a stark contrast with those of the standard non-linear taxation literature in large measure because in that literature the informational asymmetry is restricted to the value of market time. Since his seminal contribution, Mirrlees' analysis has been extended in several directions. Many of the extensions of Mirrlees' original analysis involve giving more tools to the taxation authorities. For example, see Guesnerie and

---

[3]See also Mirrlees (1997).

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

[5]Screening problems with two-dimensions of unobserved characteristics are becoming more common in the literature. See Armstrong (1996), Rochet and Choné (1998) for the state of the art in this literature and a discussion of some of the difficulties associated with solving such problems.

Roberts (1987) or Marceau and Boadway (1994).[6] In a different vein, Boone and Bovenberg (2004) extend the Mirrlees' model by introducing search costs and frictions. The model generates voluntarily unemployed individuals, involuntarily unemployed individuals and employed ones with heterogeneous levels of productivity. Search gives rise to bunching at the low end of the productivity distribution. One surprising aspect of much of the traditional optimal taxation literature is that it conflicts with current policy debates which, *de facto*, tend to favor active employment programs such as employment subsidies (negative marginal taxation). More recent works by Balestrino, Cigno, and Pettini (2003), Saez (2002), Choné & Laroque (2005) and Laroque (2005)[7] show that negative marginal tax rates can be optimal when one focuses on the extensive margin, that is, when labor supply is a zero-one decision. Moffitt (2006) examines the case where the government cares directly about the level of work of the poor, as opposed to having a welfarist objective. In this environment, Moffitt shows that negative marginal tax rates can be optimal. The current paper adds to the literature by highlighting why negative marginal tax rates can be optimal in a welfarist environment where individuals can adjust on both the intensive and the extensive margin. In particular, our approach prescribes a negative marginal tax rate on the margin where individuals choose their hours of work; an individual with a sufficiently low wage experiences an increase in net income in response to an increase in his hours worked that is greater than the associated increase in market income. This we believe captures the margin that is at the core of many policy discussions about negative marginal tax rates.[8] In Choné and Laroque (2005) and Laroque (2005), negative marginal tax rates arise when an increase in the wage – holding hours fixed – leads to a decrease in taxes. However, for an individual, the wage is not a choice variable, so individuals cannot try to improve their situation by taking advantage of the negative marginal rate.[9]

The paper is structured as follows. In Section 2 we present the constrained redistribution problem and discuss the laissez-faire and the first-best allocation. In Section 3 we discuss the case where individuals' market productivities are known but their non-market productivities

---

[6]See also Besley and Coate (1995), Brett (1998), Cuff (2000), and Kanbur, Keen and Tuomola (1994).

[7]See also Diamond (1980) who first addressed the extensive margin in a Mirrlees' type model.

[8]The model of Balestrino, Cigno, and Pettini provides two dimensions of unknown ability, market and non-market. For some parameter values—positive skill correlation and assuming that the laissez-faire marginal utility of the low skilled is greater than that of the high skilled—they find results similar to ours, taxing high incomes and subsidizing low ones. There are however only two types and many of their results depend upon the curvature of the social welfare function and its interaction with the curvature of the utility function. The environment analyzed by Saez is substantially different from the one considered here and therefore a direct comparison is difficult. In particular, Saez (2004) examines an environment with costly job choice and moral hazard with respect to which jobs to choose. The framework does not allow agents to vary their market time, nor does it allow for differences across agents in the value of their market time, two aspects which are central to our work.

[9]Concurrently with this paper, Choné & Laroque (2006) have specified a model which can study both the intensive margin and the extensive margin as limiting cases. As our analysis considers intensive and extensive margins, our work is closely related to this paper. There are nevertheless several differences between the two papers. For example, when discussing the intensive margin Choné & Laroque (2006) have effectively only one dimension of heterogeneity and do not consider the intensive and extensive margins in the same model as we do here.

are not. In Section 4 we analyze the case where both the valuations of market and non-market time are unknown. Finally, in Section 5, we discuss how the optimal solution can be implemented by a simple social policy that depends on wage rates and market income, as was the case in our initial example. All proofs, with the exception of some in Section 4, are relegated to the appendix.

2 The Environment

Consider an economy that has two sectors—a formal market sector and an informal, non-market or household sector. Income earned in the formal sector can be observed and hence taxed. The amount of time allocated to the formal sector is also assumed to be observable. Since the wage rate earned in the formal sector can be deduced from market income and time spent working in the market, the wage rate earned can be treated as effectively observable. However, the individuals' intrinsic market productivity, that is the highest possible wage they can earn, is assumed unobservable. Besides working in the formal sector, an individual can also allocate time to production in the informal/household sector. Production in this sector is unobservable.[10] Each agent is endowed with a fixed number of hours which we have normalized to one; if an individual works for $h \geq 0$ hours in the formal sector, he has $1 - h$ hours available for producing goods in the informal sector. 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 sector or conversely.

Before describing this problem further, it is worth discussing the assumption about the observability of time worked, which could represent hours, weeks or years. This is particularly relevant since the more common assumption in the literature is that hours worked are not observable[11] and that only income is observable. In practice hours or weeks worked are used in many countries to determine eligibility for social programs. For example, in Canada, one of the biggest social programs is unemployment insurance. Eligibility and payments from the Canadian unemployment insurance system depend explicitly on income and the amount of time worked (both in terms of weeks and hours per week). This is a clear example of a large program that exploits information on time worked to determine transfers. Problems with measuring time worked do not appear to be very important.[12] Another example of

---

[10]Production in the household sector can be viewed as income that the government cannot see. This could reflect revenues from illegal activities or activities that allow tax avoidance, such as activities that are remunerated through non-taxed barter.

[11]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 is only one dimension of unobserved characteristics. See also Kesselman (1973) and Bloomquist (1981) for a related literature.

[12]There are obviously some groups in society for which it is very difficult to measure the amount of time worked, for example the self-employed. Accordingly, these groups are often excluded from programs such as unemployment insurance. Moreover, if transfers are made contingent on time worked, this may create an

work contingent transfers is the UK Working Tax Credit (WTC). This program requires a minimum number of hours of work per week to be eligible for a transfer, and the requirement varies with family situation. Finally, there are many social experiments that use or have used work time requirement as a condition for income transfer. Grogger and Karoly (2006) review results from several of these experiments; see also Moffitt (2004) for a discussion of how work requirements affect behavior. In Canada there is a large scale experiment aimed at encouraging welfare recipients to work; this program is called 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. Again, this illustrates that social programs currently use information on time-worked and therefore it seems relevant to allow for such a possibility in our analysis. Obviously, working time is not observable for everyone. Nonetheless, we believe that it is useful to examine the case where we assume it is observable, and later we discuss how our results would need to be modified if time worked is not observable for high wage individuals.

Let types be indexed by $i, j \in I \times J$, where $I = \{1, \ldots, n\}$ and $J = \{1, \ldots, m\}$. $i \in I$ is the productivity of an individual with type i in the formal/market sector and $j \in J$ is the productivity of an individual with type j in the informal/household sector.[13] p_{ij} denotes the joint probability that an agent's productivities take values i and j, respectively. For the time being we impose no restrictions on the probability distribution of i and j. Assumptions are introduced below when needed. To ease notation, we shall assume that $m \geq n$, which implies that in each group of individuals with the same market productivity there is a type who is equally productive in both market and non-market activities, so $j = i$ for that individual.[14]

Individuals evaluate their well being according to the utility function

$$U\left(h \cdot w + (1 - h) \cdot j - T\right) \tag{1}$$

where $w \leq i$ is the wage rate earned in the market sector and T is the amount of taxes paid to or subsidies received, respectively, from the government. We assume that $U\left(\cdot\right)$ is increasing, differentiable, and concave. Note that the argument of the utility function is the net-income of the individual thus assuming that the individual consumes two goods that are perfect substitutes. The first good is bought from net market income; $c_i = h \cdot w - T$ is the amount consumed of this good. In addition, the individual consumes $c_j = (1 - h) \cdot j$ units of the good he produces in the informal sector.

---

incentive for firms and workers to collude to exploit the redistribution system. Although this is a possibility that should be kept in mind, we abstract from it in the current analysis since it does not appear to be a widespread concern in the actual implementation of programs which do depend on work time information.

    [13]Normalizing the type space in this way is appealing because it allows to extend our results to the continuous case by simply replacing sums by integrals.

    [14]This is not crucial; it just avoids a case distinction. For the case where $m < n$, one has to define an object $f\left(i\right)$, which is the largest j for individuals with market productivity i such that $j \leq i$. If $m \geq n$ then $f\left(i\right) \equiv i$; if $m < n$, then $f\left(i\right) < i$ for some i.

The government's objective is to maximize a utilitarian social welfare function.[15] But the government is unable to implement a first-best optimum due to the asymmetry of information. In particular, the government cannot observe skill levels of individuals in either sector, that is, the government cannot observe either i or j. By the revelation principle, we can restrict attention to direct, incentive-compatible mechanisms, where individuals are asked to announce a type $(\hat{i}, \hat{j})$ and the government chooses an allocation of work-time between the sectors, $h_{\hat{i}\hat{j}}$, a tax to be paid by the individual, $T_{\hat{i}\hat{j}}$, and a job allocation, $w_{\hat{i}\hat{j}}$, such that the individual is presumably able to do this job, that is $w_{\hat{i}\hat{j}} \leq \hat{i}$. It is immediate that the job allocation decision is trivial. At any solution to the government's problem, every individual must work in his most productive job. Otherwise a Pareto improvement can be created. We prove this statement in section 4. For $w_{\hat{i}\hat{j}} = \hat{i}$, the government's problem can be written as follows

$$\max_{\substack{\{h_{ij}, T_{ij}\} \, i=1,\ldots,n \\ j=1,\ldots,m}} \left\{ \sum_i \sum_j p_{ij} \cdot U\left(j + h_{ij} \cdot (i - j) - T_{ij}\right) \right\} \text{ s.t., for all } (i,j): \qquad (2)$$

$$U(j + h_{ij} \cdot (i - j) - T_{ij}) \geq U(j + h_{\hat{i}\hat{j}} \cdot (\hat{i} - \hat{j}) - T_{\hat{i}\hat{j}}) \ \forall \hat{j}, \forall \hat{i} \leq i, \qquad (3)$$

$$U(j + h_{ij} \cdot (i - j) - T_{ij}) \geq U(j), \qquad (4)$$

$$\sum_i \sum_j p_{ij} \cdot T_{ij} = 0, \quad \text{and} \quad 0 \leq h_{ij} \leq 1. \qquad (5)$$

In the above problem, (3) represents the incentive compatibility constraints, (4) represents the participation constraints and constraint (5) represents the materials balance constraint. Since the incentive compatibility constraints in this problem are not standard, some clarification is in order. An individual can costlessly mimic any other individual who has a lower market productivity; that is, individual (i, j) can choose to be employed in any job paying a wage $w \leq i$. In effect, the incentive compatibility constraint (3) ensures that individual (i, j) finds his allocation at least as good as that of any agent employed at a wage no greater than his own market productivity i. The participation constraints, (4), reflect our assumption that the government cannot impose a positive tax on an individual with no market income, that is, the fruits of non-market activity are not transferable to the government. Under this assumption, any individual can guarantee a minimum level of utility by simply not working.

This is a problem of multi-dimensional screening, and thus potentially complex to solve. However, the incentive compatibility constraints (3) reveal a crucial difference to the general problem of multi-dimensional screening. To clarify this difference we have rewritten the individual's income as the sum of the value of his time out of the market, j, and the gain from market participation, which we define as the after-tax excess income $h_{\hat{i}\hat{j}} \cdot (\hat{i} - j) - T_{\hat{i}\hat{j}}$. Note that both (3) and (4) depend effectively only on the after-tax excess income[16]. Moreover,

---

[15]Most of the results of this paper can be derived under the more general assumption that the government maximizes a quasi-concave Paretian welfare function, as opposed to being a strict utilitarian.

[16]To see this formally, apply the monotonic transformation U^{-1} (the inverse of U) to both sides of (3)

the after-tax excess income depends only on the message sent about market productivity, $\hat{i}$, but not on the market productivity itself. The dependence on the market productivity is only implicit in the sense that to each $\hat{i}$ there is an upper bound which is equal to i. These elements of the problem contribute to making it tractable.

To help understand the constraints imposed by the informational asymmetries, we begin by characterizing the laissez-faire and the first-best outcomes when both i and j are assumed to be observable.

2.1 Laissez Faire and First Best

In a laissez-faire world all types whose market productivity is greater than or equal to their non-market productivity, $i \geq j$, work full time, $h_{ij} = 1$, and all types whose non-market productivity exceeds their respective market productivity, $i < j$, do not work, $h_{ij} = 0$. This allocation of labour across the formal and informal markets is efficient. The individual utilities at this allocation are given by $U(\max\{i,j\})$ and utility levels range from a high of $U(m)$ to a low of $U(1)$. The social planner's objective is to reduce this range by means of taxes and subsidies.

In the first-best situation, the government is assumed to know the productivities of each individual both in the market and in non-market employment. The problem is to find the optimal redistribution of income among individuals under the constraint that the redistribution is feasible and that individuals are willing to participate. Formally, the government's problem can be stated as

$$
\max_{\substack{\{T_{ij},h_{ij}\}_{i=1,\ldots,n} \\ j=1,\ldots,m}} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \cdot U\left(j + h_{ij} \cdot (i-j) - T_{ij}\right) \right\} \quad s.t.
$$

$$
\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \cdot T_{ij} = 0,
$$

$$
h_{ij} \cdot (i-j) - T_{ij} \geq 0, \quad \text{and} \quad 0 \leq h_{ij} \leq 1 \quad \text{for all } i,j.
$$

The problem is strictly concave in the choice variables. The optimal allocation of working times is

$$
h_{ij}^* = \begin{cases} 1 & \text{if } i \geq j \\ 0 & \text{otherwise.} \end{cases}
$$

The first-order condition for T_{ij}^* is

$$
U'\left(j + h_{ij}^* \cdot (i-j) - T_{ij}^*\right) - \lambda \leq 0; \qquad h_{ij}^* \cdot (i-j) - T_{ij}^* \geq 0
$$

$$
\left(U'\left(j + h_{ij}^* \cdot (i-j) - T_{ij}^*\right) - \lambda\right) \cdot \left(h_{ij}^* \cdot (i-j) - T_{ij}^*\right) = 0
$$

and (4) and subtract j from both sides.

where λ is the multiplier on the budget constraint. Thus, either the participation constraint is strictly binding and $h_{ij}^* \cdot (i-j) - T_{ij}^* = 0$, or the participation constraint is not binding, $h_{ij}^* \cdot (i-j) - T_{ij}^* > 0$, and the individual's marginal utility is set equal to the marginal utility of everyone who receives a strictly positive net excess income. Hence, utility for these individuals must be equalized, that is $j + h_{ij}^* \cdot (i-j) - T_{ij}^* = c$ for all (i,j) such that $h_{ij}^* \cdot (i-j) - T_{ij}^* > 0$. It follows that the after-tax excess incomes at the optimum, equal to $c-j$, depend only on the non-market productivity j but not on market productivity i. An individual therefore receives a strictly positive after-tax excess income of $c-j = h_{ij}^* \cdot (i-j) - T_{ij}^*$ if $j < c$ and the individual receives a zero after-tax excess income if $j \geq c$.

Using these definitions and the optimal allocation of working time we can restate the government's budget constraint as

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \cdot \max\{c - j, 0\} = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \cdot \max\{i - j, 0\}$$

which determines the optimal level of c^*.

We illustrate in Figure 1 the properties of these two allocations where, for ease of exposition, we depict the case in which there are just two distinct levels of market productivity $i'' > i'$. In the laissez-faire world, there are three types of income heterogeneity. Individuals

Figure 1: The laissez-faire and the first-best allocations are depicted for individuals with market productivity i' (left panel) and i'' (right panel). The dots are total income $((ih + (1 - h)j)$ in the case of laissez faire, and the squares are total income $(ih + (1 - h)j - T)$ for the case of the first best. The triangles are hours worked in market activities for the case of laissez-faire and the circles are hours worked in market activities in the first-best.
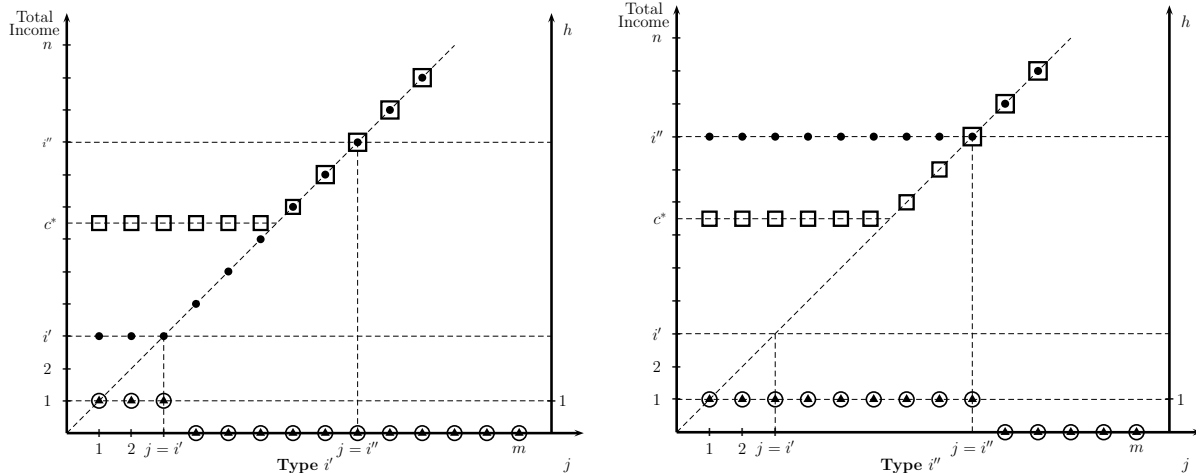
with a value of non-market time $j \leq i'$ work in both market productivity groups, but those with market productivity i'' have a higher total income. Individuals with $i' < j \leq i''$ work

only if their market productivity is i'' but not if it is i'. So for these individuals income heterogeneity is determined by the difference $i'' - j$. Finally, types with $j > i''$ do not work, and their incomes are higher than the incomes of types that are active in the formal sector.

The first-best allocation eliminates income heterogeneity between all individuals (i', j) and (i'', j) for all $j \leq i''$. Individuals with a high market productivity and a value of non-market time below c^* all pay the same tax $T_{i''j} = i'' - c^*$; those with a high market productivity and $c^* < j \leq i''$ pay taxes that just make them indifferent between engaging in market activities and informal activities, $T_{i''j} = i'' - j$. Likewise, individuals with a low market productivity and a value of non-market time below c^* all receive the same subsidy that elevates their incomes to the level c^*, $T_{i'j} = -(c^* - i')$. Individuals with $i' < j \leq c^*$ receive a type dependent subsidy, that is the lower the larger is their value of non-market time, $T_{i'j} = -(c^* - j)$. Finally, individuals with value of non-market time higher than i'' do not pay taxes nor do they receive subsidies. Notice that the participation constraints constrain the set of feasible redistribution schemes since individuals can always choose to engage in non-taxable non-market activities. This is key to understanding our problem.

Obviously, the first-best allocation creates incentive problems. If an individual faces the above transfer scheme, and could lie about his type, he would want to claim that he has a low value of non-market time, and an even lower value of market time. In Figure 1 he would claim to be type $(i', i' + 1)$. This way the individual would receive the biggest transfer and enjoy the fruits of his non-market activity. It is also of interest to examine which incentive constraints would bind if an individual could only lie about his non-market type j. In this case, an individual with high market productivity i'' (which is above c^*) would claim to have value of non-market time $j = i''$, so as to pay no taxes. That is, these individuals would claim to be type (i'', i''). Individuals with low market productivity i' (which is below c^*) would want to claim to have the lowest possible value of j subject to $j > i'$. That is they claim they are of type $(i', i' + 1)$, which implies that they can enjoy the fruits of their non-market activity and receive the largest possible transfer. What is important to notice in this case is that the direction in which the incentive constraints bind depends upon an individual's market type. Individuals with a high market productivity who are taxed have incentives to exaggerate their value of non-market time. Overstating one's non-market type implies that the outside option is more tempting and hence that the individual is willing to pay only a smaller amount of taxes in exchange for the right to participate in market activities. Individuals with a low market productivity who are subsidized have incentives to understate their value of non-market time. For an individual who spends some time outside of the market this amounts to understating total income, which implies that the individual should receive a larger transfer. The property that the direction into which an individual is tempted to deviate depends on the level of market productivity contributes to making the solution to this screening problem non-standard.

3   Observable Market Characteristics

Let us begin with the case where market characteristics are observable, while the value of non-market time is not. In this case, it is helpful to further break down the problem into a few steps by treating individuals with the same market productivity as a group. For any such group, we first examine how best to proceed if we want to extract a total revenue of T from the group, and then we examine how best to proceed if we want to transfer to a group a total subsidy of S. Once this is known, we examine the problem of which groups to tax and which groups to subsidize, and by how much.

In order to provide a complete characterization of the optimal redistribution problem, we now introduce restrictions on the distribution of i and j. Let p_i denote the marginal probability density function of i. For any i such that $p_i > 0$ let $p_j(i)$ denote the probability density function of j conditional on i and let $P_j(i) = \sum_{k=1}^{j} p_k(i)$ denote the associated distribution function. We assume throughout the paper that the conditional distribution of j conditional on i has full support for any group that has positive frequency; formally, whenever $p_i > 0$ then $p_j(i) > 0$ for all j[17]. Moreover, we impose the following regularity condition.

Assumption 1: For any i, $\frac{P_j(i)}{p_j(i)} \geq \frac{1}{2}\frac{P_{j-1}(i)}{p_{j-1}(i)} + \frac{1}{2}\frac{P_{j+1}(i)}{p_{j+1}(i)}$ for all j such that $1 < j < m$.

Monotonicity of inverse hazard rates is a standard restriction in much of the screening literature. For our purposes, it is more useful to require that the inverse hazard rate has non-increasing increments.[18] A simple distribution that satisfies this assumption is the uniform distribution. This is an attractive feature given that the uniform represents a diffuse prior, which we want to permit since we know very little about the actual distribution of j. Assumption 1 allows us to obtain simple and explicit solutions using standard methods of proof. It can be shown[19] that many of our qualitative results can be derived without this assumption at the cost of substantially more analytical complications and less transparency. Note that we place no restriction on the marginal distribution of i throughout the paper.

3.1 The Problem of Optimally Collecting Taxes

Consider the problem of how best to collect a total tax revenue T from a group of individuals with market productivity i. Since we are treating i as observable, and j as unobservable, this problem is almost a dual of Mirrlees' original problem.[20] As in many screening problems, it is helpful to reduce the dimension of the optimization problem by exploiting the binding

---

[17]The qualifier "whenever $p_i > 0$" applies to all statements we make about the conditional distribution of j conditional on i and is suppressed in what follows.

[18]Since we do not impose monotonicity of the inverse hazard rate, Assumption 1 is neither stronger nor weaker than monotonicity of the inverse hazard rate.

[19]See the working paper under Beaudry & Blackorby (2004).

[20]In contrast to Mirrlees' analysis, which assumes that only income is observable, we assume that both the time worked and income are observable. In Mirrlees' model, the first-best would be implementable under our assumptions. However, in addition, we also have the individual's valuation of non-market time unobservable. For a model with observable hours worked, but unobservable income and no further source of asymmetric information, see Maskin and Riley (1985).

incentive compatibility constraints. To this end, let us first define v_{ij} as the optimal after-tax excess income of type (i, j) by

$$v_{ij} \equiv \max_{\hat{j}} \left\{ h_{i\hat{j}} \cdot (i - j) - T_{i\hat{j}} \right\}. \tag{6}$$

Using this notation, Lemma 1 provides an equivalent formulation of the government's problem.

Lemma 1 *The government's problem of levying a tax T from individuals with observable market productivity i is equivalent to the following problem:*

$$W_i(T) \equiv \max_{\{h_{ij}\}_{j=1,\ldots,m}, v_{im}} \sum_{j=1}^{m} p_j(i) \cdot U \left(j + v_{im} + \sum_{k=j+1}^{m} h_{ik} \right) s.t.$$

$$\sum_{j=1}^{m} p_j(i) \cdot \left(h_{ij} \cdot (i - j) - v_{im} - \sum_{k=j+1}^{m} h_{ik} \right) = T,$$

$$h_{ij} \geq h_{i,j+1} \ for \ j \leq m - 1, \quad v_{im} \geq 0, \quad and \quad 0 \leq h_{ij} \leq 1.$$

From the discussion above, we know that at the first-best allocation some individuals have incentives to overstate their non-market type, j. Hence, the constraints ensuring that no individual *exaggerates* his value of non-market time must be binding at the solution to the constrained program; otherwise this solution would coincide with the first-best allocation, which is not incentive compatible. In fact, the incentive constraint for any type (i, j) not to mimic his right-wards adjacent type $(i, j + 1)$ must be binding; if that were not the case, then the allocation could be improved upon by simply changing the taxes T_{ij}, leaving the allocation of working times h_{ij} unchanged, to reduce the inequality without affecting incentive compatibility. Imposing all the right-wards adjacent incentive constraints with equality and solving recursively, we observe that an allocation can be optimal only if excess incomes satisfy the condition

$$v_{ij} = v_{im} + \sum_{k=j+1}^{m} h_{ik}. \tag{7}$$

Working times satisfy the incentive compatibility condition (3) for fixed i only if they are monotonic. Vice versa, the monotonicity condition and condition (7) are jointly sufficient for conditions (3) and (4) for fixed i. Hence, substituting condition (7) into the original objective function and the resource constraint gives rise to an equivalent representation of the problem.

We now solve the reduced problem in Lemma 1. Assumption 1 implies that the solution to the reduced problem satisfies the monotonicity constraint $h_{ij} \geq h_{i,j+1}$. Hence, we can neglect this constraint. Taking derivative with respect to h_{ij}, we find that the effect of a

marginal change in h_{ij} is

$$\sum_{j'=1}^{j-1} p_{j'}(i) \cdot U'\left(j' + v_{im} + \sum_{k=j'+1}^{m} h_{ik}\right) + \lambda_i \cdot (p_j(i) \cdot (i-j) - P_{j-1}(i)) \tag{8}$$

where λ_i is the Lagrangian multiplier attached to the resource constraint. An increase in h_{ij} increases the incomes of all individuals with $j' < j$, because, by the chain of binding incentive compatibility constraints (and thus condition (7)), an increase in h_{ij} increases the rents that have to be left to these types. This has two implications. First, it increases the utilities of individuals with $j' < j$ in the government's objective and second, it decreases the resources available for redistribution. Finally, an increase in h_{ij} also directly affects the resources available for redistribution by an amount $p_j(i) \cdot (i-j)$, the number of individuals with value of non-market time j times their productivity gap between market and non-market employment.

The structure of the solution to this problem becomes apparent when we study how this trade-off changes with j. It is useful to divide (8) by $P_{j-1}(i)$ to obtain

$$\frac{\sum_{j'=1}^{j-1} p_{j'}(i) \cdot U'\left(j' + v_{im} + \sum_{k=j'+1}^{m} h_{ik}\right)}{P_{j-1}(i)} + \lambda_i \cdot \left(\frac{p_j(i)}{P_{j-1}(i)} \cdot (i-j) - 1\right). \tag{9}$$

The first of these terms is the average marginal utility of individuals with value of non-market time smaller than j; this term is non-increasing in j, since individuals with a higher value of time are weakly better off than individuals with a lower value of time. The second term is the productive gain from allocating individual (i,j) to market activities relative to the rent cost of doing so. Assumption 1 implies that this term is decreasing in j. Thus the higher is j the less attractive it is from a pure resource point of view to allocate the individual to market activities. Indeed, the monotonicity properties of these trade-offs imply that the optimal allocation takes a very simple form. It is optimal to let individuals with a small value of non-market time work full time in the formal sector. Then, there is a marginal type $j = t$, such that the optimal working time of type (i,t) exactly balances the marginal effects in (9). All types with value of non-market time higher than t do not work at all.

To complete the characterization of the optimum, we need to determine the identity of the marginal type, t, and the hours this type spends in the formal sector, $h_{it}^* \equiv h_t$. Given the structure of the solution, this is straightforward. Since $h_{ij}^* = 1$ for $j < t$, the total incomes of types (i,j) with $j < t$ are independent of j, and we can express the first-order condition for h_{it}^* as[21]

$$U'(t - 1 + h_t) = \lambda_i \cdot \left(1 - \frac{p_t(i)}{P_{t-1}(i)} \cdot (i-t)\right). \tag{10}$$

---

[21]Although not in general true, the assumption that $m \geq n$ implies that $v_{im} = 0$.

From the incentive constraint of type $(i, t-1)$, who must not have any incentive to mimic the marginal type, (i, t), we can recover the taxes paid by the inframarginal types. From the participation constraint of the marginal type, we can recover the taxes the marginal type pays. Substituting these amounts into the government's budget constraint, we obtain

$$T = P_{t-1}(i) \cdot [i - (t-1) - h_t] + p_t(i) \cdot [h_t \cdot (i-t)]. \tag{11}$$

t is then simply the largest $j \leq i$ such that equation (11) has a solution for some $h_t \in [0,1]$. The solution exists only if T is at most equal to $T^{\max} \equiv \max_{t \leq i}(i - (t-1)) P_{t-1}(i)$. If $T > T^{\max}$, then there is no solution to the problem since it is not feasible to collect such a high level of taxes. We summarize this discussion in Proposition 1.

Proposition 1 *If the government wants to levy a tax revenue T from a group of individuals with market productivity i, where $T \leq T^{\max}$, then the optimal allocation partitions the individuals into three groups. There is a marginal type $j = t$ who works part-time, $h_{it}^* = h_t$, pays taxes $T_{it}^* = h_t \cdot (i - t)$, and has total income t. All individuals with $j < t$, work full time, $h_{ij}^* = 1$, pay taxes $T_{ij}^* = i - (t-1) - h_t$, and have total incomes equal to $t - 1 + h_t$. Individuals with $j > t$ do not work nor do they pay taxes.*

To simplify notation, we have suppressed the fact that t and h_t depend on i and T, but this dependence is made explicit later on when needed. It is instructive to consider how the solution depends on the amount of taxes to be collected, T. Suppose first T is small, in particular smaller than $P_{i-1}(i)$. Such a small amount of taxes can be collected without distorting the allocation away from first-best. The identity of the marginal type is $t = i$ and the amount of taxes paid by the inframarginal types is $\frac{T}{P_{i-1}(i)}$. The marginal type pays zero taxes, but since he is indifferent as to where to work, we can have him spend time h_t in the formal sector. By the incentive constraint of the inframarginal types, h_t determines the net after-tax incomes of the inframarginal types. The higher is T, the smaller must be h_t so that we can extract more taxes from the inframarginal types. The amount $T = P_{i-1}(i)$ can only be collected if type (i,i) does not work at all in the formal sector, $h_t = 0$. If T is still higher, then the identity of the marginal type changes to $t = i - 1$. Now the time the marginal type spends in the market determines not only the incomes left to the inframarginal types, but also - through the participation constraint of the marginal type - the taxes the marginal type has to pay. For $T < T^{\max}$, the total amount of taxes collected is decreasing in h_t; hence, the higher is T, the less time the marginal type spends in the market. The maximum amount of taxes that can be collected when the marginal type is type $(i, i-1)$ is equal to $2P_{i-2}(i)$. Still higher amounts of taxes can only be collected if the identity of the marginal type is decreased to $t = i - 2$, and so on. As long as $T < T^{\max}$, the identity of the marginal type is a weakly decreasing function of T, and h_t is a saw-tooth shaped function of T.

The allocation is second-best optimal for $P_{i-1}(i) < T \leq T^{\max}$. First, collecting more and more taxes pushes more and more types out of the market, that is $t < i$. Second, for a fixed marginal type, the time the marginal type spends in the market is generically sub-optimally low so as to allow the government to collect more taxes from the inframarginal types, so

14

$h_t < 1$. Finally, a simple implication of Proposition 1 is that it is never optimal to subsidize (impose a negative tax on) a subset of individuals among the considered group and tax the others.

3.2 The Problem of Optimally Subsidizing

The incentive problems for a group that receives subsidies are diametrically opposed to the incentive problems of those who are taxed. Heuristically, the government must remove the incentive to claim that one's productivity in the informal sector is lower than it in fact is. The government wishes to help those with a lower productivity in the informal sector but it does not want to give its money to those who fare well on their own. In what follows we use the term "exclusion" as a synonym for not paying subsidies to an agent, and "inclusion" in the opposite sense.

The government's problem can be understood as a combined problem of exclusion and redistribution. The optimum is characterized by the government wishing to subsidize all agents whose productivity in the informal sector is less than or equal to s, where s is a variable of its choice. Equivalently, the government wishes to exclude from subsidization all agents whose non-market productivity is greater than or equal to $s+1$. For any given s, the government's problem is to distribute the available income to the agents that are included in the redistribution program. Henceforth, we call type (i,s) the marginal type. We begin again with the statement of an equivalent, but more tractable form of the government's problem.

Lemma 2 *The government's problem of distributing a subsidy S to individuals with observable market productivity i is equivalent to the following program:*

$$\tilde{W}_i(S) \equiv \max_{\{h_{ij}\}_{j=1,\dots,m}, v_{i1}, s} \sum_{j=1}^{s} p_j(i) \cdot U\left(j + v_{i1} - \sum_{k=1}^{j-1} h_{ik}\right) + \sum_{j=s+1}^{m} p_j(i) \cdot U(j) \ s.t.$$

$$\sum_{j=1}^{s} p_j(i) \cdot \left(h_{ij} \cdot (i-j) - \left(v_{i1} - \sum_{k=1}^{j-1} h_{ik}\right)\right) + S = 0,$$

$$v_{i1} - \sum_{k=1}^{s-1} h_{ik} \geq 0 \quad and \quad v_{i1} - \sum_{k=1}^{s} h_{ik} \leq 0,$$

$$h_{ij} \geq h_{ij+1} \ for \ all \ j < m-1 \quad and \quad 0 \leq h_{ij} \leq 1.$$

At the first-best allocation, only incentive constraints to mimic types with a lower productivity in the informal sector are binding. Hence, these constraints must be binding in the solution to the government's problem as well. Individuals are divided into two groups, one group of individuals with $j \leq s$ receiving subsidies, the other with $j > s$ not. Again a simple optimization argument shows that for any type (i,j) with $j \leq s$ the constraint ensuring that there is no incentive to mimic his left-wards adjacent type $(i, j-1)$ must be binding.

Solving recursively for the after-tax excess incomes, we find that

$$v_{ij} = v_{i1} - \sum_{k=1}^{j-1} h_{ik} \text{ for all } j \leq s. \tag{12}$$

Only monotonic allocations of working time are incentive compatible, and condition (12) in conjunction with the monotonicity constraint and $v_{i1} - \sum_{k=1}^{s-1} h_{ik} \geq 0$ are jointly sufficient for the incentive compatibility and participation constraints of all types with $j \leq s$. Since the types with value of non-market time of $s + 1$ and higher are excluded, they get zero excess incomes. To make sure that all these types are indeed excluded, the allocation must be such that type $(i, s + 1)$ would obtain a non-positive excess income if he mimicked the marginal type. The allocation must satisfy the exclusion constraint $v_{i1} - \sum_{k=1}^{s} h_{ik} \leq 0.$[22] Since types with value of non-market time higher than $s + 1$ find it even more costly to work in the formal sector than type $(i, s + 1)$ does, the exclusion constraint excludes all types (i, j) with $j \geq s + 1$.

Assumption 1 implies that the trade-off between an individual's contribution to the resources available for redistribution and the rents left to inframarginal types changes monotonically as we increase nonmarket productivity. Therefore, there is a single marginal type, who divides his fellow types into two groups. Those who have a lower value of non-market time work full time and those with higher opportunity costs of time do not work at all. All types who are subsidized receive the same amount of total income, so their marginal utilities are equalized, which allows us to characterize the optimal allocation by the first-order condition

$$U'(s + h_s) = \lambda_i \cdot \left(1 - \frac{p_s(i)}{P_s(i)}(i - s)\right) \tag{13}$$

where λ_i is the multiplier on the group's resource constraint. To determine s and h_s, we recover the subsidies paid to inframarginal and marginal types, respectively, and substitute these expressions into the government's budget constraint

$$P_{s-1}(i) \cdot [s - i + h_s] + p_s(i) \cdot h_s \cdot [s + 1 - i] = S. \tag{14}$$

s is then the highest $j \geq i$ such that equation (14) has a solution for some $h_s \in (0, 1]$. We summarize this discussion:

Proposition 2 *If the government wants to distribute a total subsidy S to a group of individuals with market productivity i, then the optimal allocation takes the following form. Individuals are partitioned into three groups. There is a marginal type with nonmarket productivity $j = s$, who works part time, $h_{is}^* = h_s$, receives a subsidy equal to $-T_{is} = h_s \cdot (s + 1 - i)$, and*

---

[22]The exclusion and inclusion constraints make this model different from standard problems and from the previous taxation problem. For a general analysis of participation constraints in adverse selection models, see Jullien (2000).

has an after-tax income of $s+h_s$. Types with nonmarket productivity smaller than s work full time, $h_{ij}^* = 1$ for $j < s$, receive subsidies equal to $-T_{ij} = s - i + h_s$ and have total incomes equal to $s + h_s$. Types with nonmarket productivities larger than s do not work, $h_{ij}^* = 0$ for $j > s$, do not receive subsidies, and have incomes equal to j.

The intuition for this result is similar to the one for the taxation case, but there are nonetheless important differences. First, as we have explained above, the binding incentive constraints are left-ward looking; individuals want to claim that their productivity in the informal market is less than it actually is. Second, there is an additional exclusion constraint and this must be binding. If this constraint were slack, then there would be no reason to put the marginal type to work in the formal sector. But $S > 0$ implies that the marginal type's productivity in the formal sector must be lower than his productivity in the informal sector, so absent any binding constraint this type would spend no time in the formal sector. But then the allocation cannot be incentive compatible, because all types with higher productivity in the informal sector than the marginal type, and who should be excluded from receiving subsidies, can claim the subsidy targeted at the marginal type without cost. Put differently, the allocation forces some individuals to spend (part of) their time inefficiently in the market. This working requirement serves as a screening device that discourages individuals with a high value of nonmarket time from claiming the subsidies targeted at those with lower values of nonmarket time.

It is again instructive to investigate how the solution depends on the level of subsidies distributed, S. Suppose first that S is rather small, in particular smaller than $P_i(i)$. Then, the marginal type is $s = i$. Since the marginal type is indifferent as to where to work, we can allocate him to the formal sector. h_s is increasing in S for S smaller than $P_i(i)$. For $S = P_i(i)$, the marginal type works full time in the formal sector. If we want to distribute a higher amount of subsidies, we have to adjust the identity of the marginal type to $s = i + 1$. For $P_i(i) < S \le 2P_{i+1}(i)$, the hours spent by the marginal type in the formal sector are again increasing in S. So, the identity of the marginal type is a weakly increasing function of S and the hours spent by the marginal type in the formal sector is increasing in S. By the binding exclusion constraint the variable h_s serves to screen out types that receive a higher income by not working than the subsidized individuals who work receive. h_s is strictly positive unless S is so large that there is no type left to screen out. This happens when $S \ge P_{m-1}(m - i)$, so that all types have an income that is at least equal to m. In that case there is no need to screen anybody out, and the marginal type need not work in the formal sector. However, this case can never arise at an overall optimum, when we endogeneize the amount of redistribution, so we can safely neglect this case.

Note that the allocation is such that employment decisions are weakly upward distorted. Since $s \ge i$, with a strict inequality when $S > P_i(i)$, some types are allocated to market activities although they would be more productive when allocated to non-market activities. Moreover, the marginal type spends a positive amount of time in the market although, for $s > i$, he would be more productive at informal activities. Finally, we observe again that it is never optimal to collect taxes from some individuals with market productivity i if some other individuals with the same market productivity receive subsidies.

17

3.3 Optimal Redistribution with 2 Observable Market Types

We now consider the case where there are two market types $i \in \{i', i''\}$ and many unobservable non-market types. Let $p_{i'}$ denote the probability that $i = i'$ and let $p_{i''}$ denote the probability that $i = i''$.[23] The government chooses T, the amount of taxes collected from one group, and S the amount to give to the other group, in order to maximize the expected utility of the entire population. In this case, it is obvious that it is optimal to levy a tax on the group with the higher market productivity, and subsidize the group with $i = i'$. The budget constraint links T and S through the condition $p_{i''}T = p_{i'}S$. We can write the government's problem as

$$\max_{T,S} \left\{ p_{i''} W_{i''}(T) + p_{i'} \tilde{W}_{i'}(S) \right\} \quad s.t. \quad p_{i''}T = p_{i'}S \quad \text{and} \quad T \leq T^{\max}.$$

Whenever an optimum exists, then the optimal levels of T and S, denoted T^* and S^*, satisfy either

$$-\frac{\partial W_{i''}(T^*)}{\partial T} = \frac{\partial \tilde{W}_{i'}(S^*)}{\partial S} \tag{15}$$

or $T^* = T^{\max}$, $S^* = \frac{p_{i''}}{p_{i'}} T^{max}$ and $-\frac{\partial W_{i''}(T^{\max})}{\partial T} < \frac{\partial \tilde{W}_{i'}(\frac{p_{i''}}{p_{i'}} T^{max})}{\partial S}$.

Invoking the envelope theorem, we observe that condition (15) requires the equality of the shadow costs of taxation and subsidization. Recalling our previous propositions, we can express the condition in terms of the variables t, h_t, s, and h_s. The result is formally stated as

Proposition 3 *The solution of the government's redistribution problem when there are two observable market productivities must satisfy either*

$$\frac{U'(t-1+h_t)}{U'(s+h_s)} = \frac{1 - \frac{p_t(i'')}{P_{t-1}(i'')} \cdot (i''-t)}{1 - \frac{p_s(i')}{P_s(i')} \cdot (i'-s)}$$

where t, s, h_t and h_s satisfy the conditions of Propositions 1 and 2 or $T = T^{\max}$ and

$$\frac{U'(t-1+h_t)}{U'(s+h_s)} < \frac{1 - \frac{p_t(i'')}{P_{t-1}(i'')} \cdot (i''-t)}{1 - \frac{p_s(i')}{P_s(i')} \cdot (i'-s)}.$$

We omit a formal proof of this result, because it directly follows from the discussion preceding the Proposition. The four objects t, s, h_t and h_s are all functions of the total tax to be levied on the high market productivity group. Hence, this proposition implicitly defines the optimal level T to levy on these types, and the government's budget constraint indicates how much to subsidize in total the low market productivity group. Then, given the

---

[23]Alternatively, in terms of our general notation, this corresponds to the case where $i \in \{1, 2, ..., n\}$ and the probability distribution puts zero probability on all $i \notin \{i', i''\}$.

optimal levels of total taxes and total subsidies, Propositions 1 and 2 indicate the associated individual levels of taxes and the individual levels of subsidies that support the optimal allocation. Hence, Proposition 3, in conjunction with Propositions 1 and 2, offers a complete characterization of the optimal redistribution problem with two observable market types.[24] We illustrate the properties of the optimal allocation in Figure 2.

Figure 2: The second-best allocation for individuals with market productivity i' (left panel) and i'' (right panel) are depicted below. The dots are pre-tax incomes $(ih + (1 - h)j)$ and the circles are total income $(ih + (1 - h)j - T)$. The triangles are hours worked in market activities.
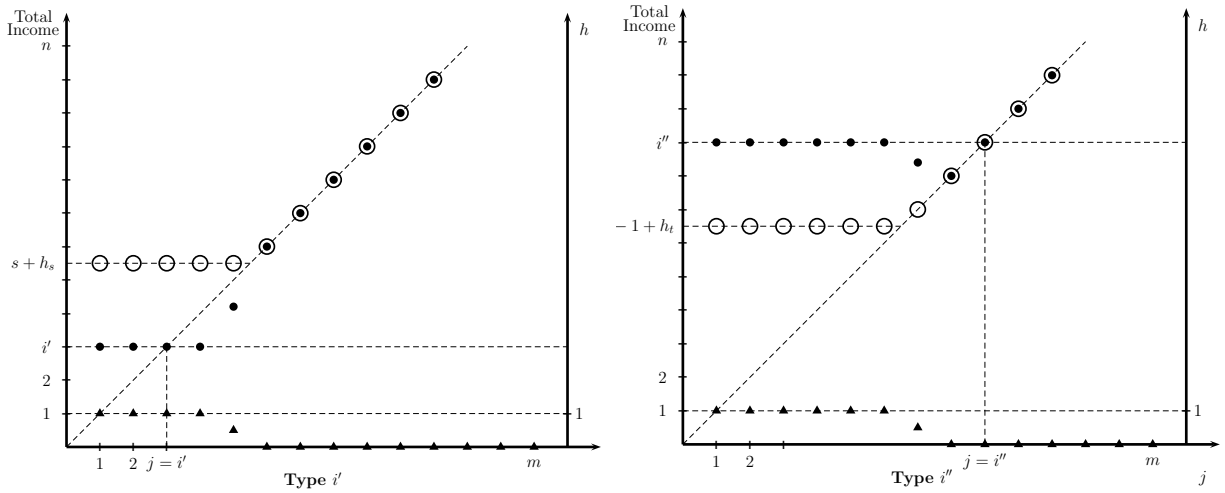
The second best optimal amount of redistribution is clearly smaller than the first-best amount. To see this, recall that the first-best allocation equalizes the incomes of all individuals who work in the formal sector. This is too costly with asymmetric information. The cost of taxation in the high productivity group is that more and more individuals are driven out of the market the higher is T. The cost of subsidization in the low productivity group is that more and more people who would be relatively more productive in the informal sector have to work in the formal sector the higher is T (and thus the higher is S). As a result, the incomes of the working individuals with market productivity i'' are strictly higher than the incomes of the individuals with productivity i'.

3.4 Many Observable Market Characteristics

We now generalize our findings to the case of many observed market characteristics and many unobservable non-market productivities. This problem can be stated as finding a sequence of

---

[24]In general our solution is second-best optimal, but it is easy to verify that in the special case where $i'' = i' + 1$, the first best allocation is actually implemented by this solution.

total taxes, T_i for $i = \{1, \ldots, n\}$, where an element T_i represents the total tax levied on the group of individuals with market productivity i, a negative value of T_i represents a subsidy; T_i equals zero when $p_i = 0$. For any two groups for which the maximal tax capacity is not attained, it must be that the marginal cost of taxation is equalized. In particular, if groups i'' and i' are subsidized (negative value of T_i), it must be the case that

$$\frac{U'\left(s\left(i'',T_{i''}\right) + h_s\left(i'',T_{i''}\right)\right)}{\left(1 - \frac{p_{s\left(i'',T_{i''}\right)}^{(i'')}}{P_{s\left(i'',T_{i''}\right)}^{(i'')}} \left(i'' - s\left(i'',T_{i''}\right)\right)\right)} = \frac{U'\left(s\left(i',T_{i'}\right) + h_s\left(i',T_{i'}\right)\right)}{\left(1 - \frac{p_{s\left(i',T_{i'}\right)}^{(i')}}{P_{s\left(i',T_{i'}\right)}^{(i')}} \left(i' - s\left(i',T_{i'}\right)\right)\right)}. \tag{16}$$

In (16), we make explicit the dependence of s and h_s on i and the total tax paid by a group. Similarly, if group i'' has a positive value of total taxes, $T_i > 0$, and group i' is subsidized, it must be the case that either

$$\frac{U'\left(t\left(i'',T_{i''}\right) - 1 + h_t\left(i'',T_{i''}\right)\right)}{\left(1 - \frac{p_{t\left(i'',T_{i''}\right)}^{(i'')}}{P_{t\left(i'',T_{i''}\right)-1}^{(i'')}} \left(i'' - t\left(i'',T_{i''}\right)\right)\right)} = \frac{U'\left(s\left(i',T_{i'}\right) + h_s\left(i',T_{i'}\right)\right)}{\left(1 - \frac{p_{s\left(i',T_{i'}\right)}^{(i')}}{P_{s\left(i',T_{i'}\right)}^{(i')}} \left(i' - s\left(i',T_{i'}\right)\right)\right)} \tag{17}$$

or $T_i = T^{\max}(i'') \equiv \max_{t \leq i''} (i'' - (t-1)) P_{t-1}(i'')$ and condition (17) holds as an inequality. A similar condition applies if two groups are taxed. These conditions, in addition to the government's budget constraint, $\sum_{i=1}^{n} p_i \cdot T_i = 0$, determine the optimal level of total taxes and subsidies for each group. To get further insights into the structure of the solution, we assume from now on that the optimum is described by an interior solution. One interesting question is who gets taxed and who gets subsidized. This is addressed in Proposition 4.

Proposition 4 *There is a critical market productivity $\underline{i}$ such that individuals are taxed only if their market productivity satisfies $i \geq \underline{i}$. Individuals are subsidized only if their market productivity satisfies $i < \underline{i}$.*

This proposition indicates that an optimal redistribution plan has the property that individuals are taxed or subsidized depending on whether their market productivity falls short of or exceeds a critical value. In other words, Proposition 4 indicates that the determinant of whether individuals should be taxed or subsidized is not their market income but instead it is their market wage rate. The intuition for the proof is as follows. Suppose the result were not true, and there were two groups $i'' > i'$ where group i'' is subsidized and group i' is taxed. We know from Propositions 1 and 2 that the employment decisions are distorted upwards in the subsidized group, so $s \geq i''$, and that the employment decisions are distorted downwards in the group that is taxed, so $i' \geq t$. The after-tax incomes of the individuals in group i'' are equal to $\max\{s + h_s, j\}$ and the after-tax incomes of group i' are equal to $\max\{t - 1 + h_t, j\}$. So, the effect of such a policy is to increase the dispersion in incomes, which is obviously suboptimal. We show in the appendix that there is also an incentive

20

compatible and budget neutral way to improve such an allocation, so it can never be part of an optimum.

More generally, one can ask how the after-tax incomes of individuals change with their market productivity. Said differently, how many individuals are pushed out of market activities due to taxation and are allocated to market activities due to subsidization? To perform these comparative statics exercises, we impose another regularity condition on the distribution of types. In particular, we assume that $p_i > 0$ for all i, and moreover we impose:

Assumption 2: For any j, $\frac{P_j(i)}{p_j(i)} \geq \frac{1}{2}\frac{P_j(i-1)}{p_j(i-1)} + \frac{1}{2}\frac{P_j(i+1)}{p_j(i+1)}$ for all i such that $1 < i < n$.

Similar to Assumption 1, we require that the inverse hazard rate has non-increasing increments, but here we require the property to hold for the conditioning variable i. A simple and natural case that satisfies Assumption 2 is when market and non-market productivities are independent of each other. However, Assumption 2 holds more generally when—heuristically—market and non-market productivities are more strongly correlated the higher the level of market productivity. More precisely, we show in the appendix that the inverse hazard rate has non-increasing increments in i if the strength of affiliation between market and non-market productivity is non-decreasing in the level of market productivity. We have the following result:

Proposition 5 *The after-tax incomes of individuals who work in the formal sector, both within the groups that are taxed and within groups that are subsidized, are non-decreasing in i. Moreover, the identity of the marginal types is non-decreasing in i.*

To understand the role of Assumption 2 for the result, recall the first-order conditions (10) and (13), that determine the after-tax incomes of the inframarginal types in taxed and in subsidized groups, respectively. If the benefit-cost ratios on the right-hand sides of these expressions change monotonically in i then the marginal utility expressions on the left-hand sides will adjust the same way at the optimum. Assumption 2 implies precisely this.

To understand this in more detail, consider a special case when i and j are independent. In this case, for any two individuals with non-market productivity j and market productivity i and $i+1$, respectively, the relative frequency of types with nonmarket productivity j relative to types with smaller nonmarket productivities is the same. However, for the individual in group i the productivity gap between the two activities is $i - j$, while the productivity gap in group $i + 1$ is $i + 1 - j$. Hence, increasing market productivity implies that the benefit of allocating the individual to market activities increases while the cost of doing so - in terms of foregone rents that have to be given away to types with smaller nonmarket productivities - is held constant. Therefore, at the optimum all individuals with market productivity $i + 1$ will work at least as much in the formal sector as those do with market productivity i. By incentive compatibility, through conditions (7) or (12), respectively, depending on whether the groups are taxed or subsidized, the allocation of working times determines the net excess incomes of these individuals. Hence, if individuals with higher market productivity spend weakly more time in the formal sector, the resulting total incomes and the identity of the

marginal individual must be weakly higher in the group with the higher market productivity. Distributions that satisfy Assumption 2 are sufficiently regular that any changes in the inverse hazard rate due to a change in market productivity do not outweigh the positive effect on the productivity gap.

4   Unobservable Market and Non-Market Productivities

Throughout the previous section we assumed that market productivities were observable by the government. In this section, we relax this assumption and consider the main case of interest where individuals can claim to have market productivities lower than that given by their innate ability as well as being able to lie about their informal market productivities.

First, we prove formally that it is indeed optimal to allocate each individual working in the formal sector to his most productive market task.

Proposition 6 *For any incentive compatible, individually rational, and budget balanced allocation in which $w_{ij} < i$ for some (i, j), there exists an incentive compatible, individually rational, and budget balanced allocation that satisfies $w_{ij} = i$ for all (i, j) that Pareto dominates the former allocation.*

The idea is simply to change the taxes without changing the allocation of working times and without violating any of the constraints.

Adding a second dimension of asymmetric information usually increases the number of incentive constraints dramatically, because individuals can mimic others who differ from them both in the value of market and non-market time. However, our problem differs from the usual problem of multidimensional screening. If we had a problem of the usual sort, an individual's excess income would depend on the vector $(i, \hat{i}, j, \hat{j})$, which contains both preference parameters and messages in both dimensions. Here, the agent's after-tax excess incomes $h_{i\hat{j}} \cdot (\hat{i} - j) - T_{i\hat{j}}$ depend only on the vector $(\hat{i}, j, \hat{j})$ but only implicitly on the true ability parameter i through the fact that an individual can mimic only those individuals who are assigned to market tasks that require at most productiviy i. The fact that the after-tax excess income is independent of i implies that when the individual mimics a less qualified person with the same productivity in the informal sector, he obtains exactly the same excess income as that person obtains. This insight allows us to prove the following result:

Proposition 7 *The two-dimensional incentive constraint* (3) *is satisfied if and only if this pair of one-dimensional constraints is satisfied for all (i, j)*

$$h_{ij} \cdot (i - j) - T_{ij} \geq h_{\hat{i}j} \cdot (\hat{i} - j) - T_{\hat{i}j} \ \forall \hat{i} \leq i \tag{18}$$

and

$$h_{ij} \cdot (i - j) - T_{ij} \geq h_{i\hat{j}} \cdot (i - j) - T_{i\hat{j}} \ \forall \hat{j}. \tag{19}$$

Proof. The only if part is trivial. So, consider the sufficiency part. Suppose (18) and (19) are satisfied and consider type (i, j) which mimics type $(\hat{i}, \hat{j})$ for $\hat{i} \leq i$. The excess income he obtains this way is exactly the excess income that type $(\hat{i}, j)$ obtains from mimicking type $(\hat{i}, \hat{j})$. But by (19) applied to type $(\hat{i}, j)$, it would be better for type (i, j) to mimic type $(\hat{i}, j)$; but then, by (18) it would even be better to state the true type (i, j), which implies that incentive compatibility is satisfied for an arbitrary $\hat{j}$ and $\hat{i} \leq i$. ∎

Proposition 7 is an important simplification, since it reduces the number of relevant incentive constraints dramatically.[25] Moreover, it allows us to show that all of our results found for the case of observable market types carry over to the current case.

Proposition 8 *For distributions that satisfy Assumptions 1 and 2, e.g., when nonmarket productivities are independent and uniformly distributed, the optimal allocation with observable market characteristics remains incentive compatible when market characteristics are not observable.*

Proof. The proof of Proposition 8 consists of pulling together the implications of all the preceding propositions. Our solution procedure for the case of observable market productivites ensured that constraint (19) is taken care of. Given Proposition 7, it suffices to check whether our solution satisfies also condition (18). From Propositions 1 through 3 we know that with two market productivities, incomes in the group that is taxed are weakly higher than incomes in the group that is subsidized, because redistribution is suboptimally low at the optimum. Proposition 4 states that with many market productivities, there is a group with the smallest market productivity $i = \underline{i}$ among all the groups that are taxed. Since $t(\underline{i}, T_{\underline{i}}) \leq \underline{i}$ for $T_{\underline{i}} \geq 0$ and $s(\underline{i} - 1, T_{\underline{i}-1}) \geq \underline{i}$ for $T_{\underline{i}-1} \leq 0$, the incomes in group $\underline{i}$ are weakly higher than the incomes in group $\underline{i} - 1$. Proposition 5 showed that the incomes of those who work some time in the formal sector are non-decreasing in i for $i \geq \underline{i}$ and for $i < \underline{i}$. Taken together, these results imply that the incomes of those who work are monotonic in i for all i. For individuals who do not work at the optimum, they receive their outside option, j, regardless of whether they are honest or whether they mimic an individual with a lower market productivity. This follows from the fact that the identity of the marginal type is non-decreasing in i. Finally, observe that the government can never do better when it cannot observe i than when it can observe i. Hence, when it is feasible to implement the same solution as with observed market productivities, then it is optimal to do so. Hence the solution to the family of screening problems conditional on i that we have found characterized in Section 3 represents the solution to the overall screening problem. ∎ [26]

---

[25]Notice that Propositions 6 and 7 do not depend on either Assumption 1 or 2.

[26]As an example that violates Assumption 2 and where it is not possible to solve the entire problem by solving a family of screening problems conditional on i, suppose there are two groups i'' and $i'' + 1$ that are taxed and let almost all individuals in group $i'' + 1$ have $j = 1$ and almost all individuals in group i'' have $j = i''$. Clearly, individuals in group i'' drop out of formal market activities quickly when taxed, while individuals in group $i'' + 1$ have less attractive outside options *within* their group. But if the government wishes to levy a large $T_{i''+1}$, then the outside options *across* groups become relevant in this example. Hence, the allocation for group i'' places a constraint on the feasible allocations for group $i'' + 1$.

5   The Structure of Income Taxes and Subsidies

Our analysis of the informationally constrained redistribution problem has allowed us to derive properties of an optimal tax system in the form of a direct revelation mechanism. However, in practice, tax systems do not take this form. Instead, tax systems are more akin to indirect revelation mechanism. For example, in the introduction, we discussed a simplified tax system that depended on one's income and one's wage rate. We can call such a system a wage contingent income tax system, and denote such as system by the function $\hat{T}(w, y)$ where $y = w \cdot h$ is income. In this section, we describe the properties of the wage contingent income tax system that implements the solution to our optimal redistribution problem.[27] In this case, we are assuming the government can observe both workers' incomes, and their wage rate. Obviously, this is equivalent to assuming the government can observe income and hours worked.

The first notable property, which follows directly from Proposition 4, is that there exists a critical wage $\underline{w}$, such that $\hat{T}(w, y) \geq 0$ for all y if $w > \underline{w}$, and $\hat{T}(w, y) \leq 0$ for all y if $w \leq \underline{w}$. This observation emphasizes that being taxed versus subsidized depends first and foremost on one's wage, not on income. The second striking property is that $\hat{T}(w, y) = 0$ if $y = 0$, that is, individuals that choose not to work do not get any subsidies. This result, which is implied by the nature of the direct tax functions derived in Propositions 1 and 2, implies the absence of welfare payments for employable individuals. This is in stark contrast to the traditional optimal tax literature which generally prescribes positive welfare payments to individuals who do not choose to work. In our setup, it is always better to use wage contingent employment subsidies to redistribute income since this allows the government to target workers with poor options both within and outside the market.[28]

The third property relates to the nature of marginal tax rates and marginal subsidies. In particular, one can show that a wage contingent income tax system implements the optimal allocation only if it is convex in income over all levels of income that are achieved by some type in equilibrium. This indicates that an optimal wage-contingent income tax system has the property that as an individual increases his income (by increasing his hours worked), he faces either weakly increasing marginal tax rates if his wage rate is high, or alternatively faces weakly decreasing marginal income subsidies if his wage rate is low. In other words, negative marginal tax rates are weakly increasing as an individual increases his income. As an example, the following piece-wise linear tax schedule could be used to implement the optimal allocation.

---

[27]Since the function h_{ij} which prescribes the optimal allocations is monotonic in j (for a given i), it is easy to verify that a wage-contingent income tax schedule can be used to implement the optimal redistribution problem.

[28]The result that the optimal program prescribes an absence of transfers to individuals that choose not to work may appear extreme, especially since it runs counter to common practice in most countries. Some qualifications apply to this result and should be kept in mind. For instance, the analysis is meant to apply to employable individuals with no observable characteristics which would favor an allocation of time toward non-market activities. Hence the analysis is not structured to handle individuals that have dependent children, or individuals with handicaps which make employment difficult.

For an individual being paid a wage above the critical level $w > \underline{w}$, and earning income y, then taxes are given by

$$\hat{T}(w, y) = \begin{cases} \left(1 - \frac{t}{w}\right) \cdot y & \text{for } y \leq \overline{y} \equiv h_t \cdot w \\ \left(1 - \frac{t}{w}\right) \cdot \overline{y} + (y - \overline{y}) \cdot \left(1 - \frac{t-1}{w}\right) & \text{for } y > \overline{y}. \end{cases}$$

For an individual begin paid a wage below the critical level $w \leq \underline{w}$, then subsidies are given by

$$-\hat{T}(w, y) = \begin{cases} \left(\frac{s+1}{w} - 1\right) \cdot y & \text{for } y \leq \underline{y} \equiv h_s \cdot w \\ \left(\frac{s+1}{w} - 1\right) \cdot \underline{y} + \left(y - \underline{y}\right) \cdot \left(\frac{s}{w} - 1\right) & \text{for } y > \underline{y} \end{cases}$$

In the above, the indices t, s and the work hours h_t, h_s are a function of the market productivity i and are determined as in Propositions 1 and 2, in conjunction with the conditions presented in Section 3.4. As can be seen, this tax schedule has the property that marginal taxes are weakly increasing for individuals with $w > \overline{w}$ since $\left(1 - \frac{t}{w}\right) < \left(1 - \frac{t-1}{w}\right)$. Similarly, marginal subsidies are decreasing for low wage individuals since $\left(\frac{s+1}{w} - 1\right) > \left(\frac{s}{w} - 1\right)$.

In summary, our analysis implies that a wage-contingent tax system has the following four properties: (1) the existence of a cutoff wage, where individuals with wages above the cutoff are taxed and individuals with wages below the cutoff are subsidized, (2) individuals below the cutoff wage face wage-contingent marginal income subsidies that decrease as income increases, (3) individuals above the cutoff wage face positive and increasing marginal tax rates as income increases, and (4) individuals who choose not to work receive no income transfer.

Throughout this analysis, we have been assuming that the government can observe both a worker's income and his wage rate (or hours worked). For many individuals this appears to be a reasonable assumption since many social programs in industrialized countries are based on such information and these programs appear to function properly. However, for some individuals, especially many high market productivity individuals, this assumption is unlikely to hold in practice. It is therefore relevant to ask how our results would need to be modified if governments could not observe hours worked for individuals paid at high wage rates. Without providing a full analysis here, such a modification would not change the flavor of our main results if the unobservability of hours or wages arose (mainly) for individuals with market productivity above the critical level associated with subsidization. In this case, the government could run a standard income tax system (based only on income) plus a separate earned-income subsidy system where individuals would need to have verifiable income and hours (or wages) statements to be eligible for a subsidy. While the tax system would be less efficient in the absence of information on hours worked, the subsidy system could still avoid transferring income to individuals with high value of time outside the market by requiring them to prove that they are working at low paying jobs. What is crucial for most of our results is the observability of hours (or wages) for potentially subsidized jobs; the observability of hours for high paying jobs is less critical.

6 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 programs that target transfers more effectively than simple negative income tax schemes.[29] 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 no transfers to non-working individuals.

How should these results be interpreted? In our view, these results are not a call for redesigning income tax systems to include a dependence on wages. 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.[30] For example, these results provide potential support for programs, such as the UK Working Tax Credit and Canadian Self-Sufficiency Project, which supplement the income of low wage earners who choose to work, making transfers contingent on both income and time worked.[31] More generally, we view our results as suggesting the use of work time information to implement phased-out wage subsidies as a means of redistributing income to low wage earners, that is, wage subsidies that decrease in intensity 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.

7 Appendix

Proof of Lemma 1. We begin showing that the monotonicity condition, $h_{ij} \geq h_{i,j+1}$ for $j \leq m - 1$, is necessary for incentive compatibility. Consider type (i, j) and a deviation to $(i, \hat{j})$:

$$h_{ij} \cdot (i - j) - T_{ij} \geq h_{i\hat{j}} \cdot (i - j) - T_{i\hat{j}}.$$

---

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

[30] 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.

[31] The result that governments may want to include work requirements in the design of transfer programs is derived here under the assumption of a utilitarian/welfarist government. Moffitt (2006) derives a similar result under the assumption that the government has preferences over the work allocation of transfer recipients.

Now, consider type $(i, \hat{\jmath})$ (interchanging type and message) and a deviation to (i, j) :

$$h_{i\hat{\jmath}} \cdot (i - \hat{\jmath}) - T_{i\hat{\jmath}} \geq h_{ij} \cdot (i - \hat{\jmath}) - T_{ij}.$$

Rearranging, we have

$$(h_{ij} - h_{i\hat{\jmath}}) (\hat{\jmath} - j) \geq 0$$

which proves the claim.

Next we argue that $v_{ij} = v_{im} + \sum_{k=j+1}^{m} h_{ik}$ and h_{ij} non-increasing in j are sufficient for incentive compatibility. To ease notation, define

$$V(i, \hat{\jmath}, j) \equiv h_{i\hat{\jmath}} \cdot (i - j) - T_{i\hat{\jmath}}$$

The condition $v_{ij} = v_{im} + \sum_{k=j+1}^{m} h_{ik}$ results from imposing the right-wards adjacent incentive constraints with equality and solving recursively. To see this, suppose the right-ward adjacent constraint holds with equality. Then,

$$\begin{aligned} h_{ij} \cdot (i - j) - T_{ij} &= h_{i,j+1} \cdot (i - j) - T_{i,j+1} \\ &= h_{i,j+1} \cdot (i - (j+1)) - T_{i,j+1} + h_{i,j+1}. \end{aligned}$$

So, $V(i, j, j) = V(i, j+1, j+1) + h_{i,j+1}$. Applying this logic repeatedly and solving recursively, gives expression (7). We wish to show that (7) and monotonicity of h_{ij} jointly imply that any deviation from truth-telling is suboptimal. Notice that the excess income that type (i, j) obtains from mimicking type (i, l) is given by

$$\begin{aligned} V(i, l, j) &= h_{il} \cdot (i - j) - T_{il} \\ &= h_{il} \cdot (i - l) - T_{il} - (j - l) \cdot h_{il} \\ &= V(i, l, l) - (j - l) \cdot h_{il}. \end{aligned}$$

Thus, $V(i, j, j) \geq V(i, l, j)$ for any l and j if

$$V(i, m, m) + \sum_{k=j+1}^{m} h_{ik} \geq V(i, m, m) + \sum_{k=l+1}^{m} h_{ik} - (j - l) \cdot h_{il}.$$

Consider first any $l > j$. We can write the comparison as

$$\begin{aligned} & V(i, m, m) + \sum_{k=l+1}^{m} h_{ik} + h_{i,j+1} + ... + h_{il} \\ \geq \quad & V(i, m, m) + \sum_{k=l+1}^{m} h_{ik} - (j - l) \cdot h_{il} \end{aligned}$$

Cancelling equal terms on both sides we can simplify the condition to

$$h_{i,j+1} + ... + h_{il} \geq (l - j) \cdot h_{il}.$$

Since the number of terms on each side is the same, and h_j is non-increasing in j, the inequality is satisfied. The proof for the case where $l < j$ is similar and therefore omitted.

Consider now the participation constraints. From the right-wards adjacent incentive constraints, $V(i,j,j) = V(i,j+1,j) \geq V(i,j+1,j+1)$, and from the participation constraint of type (i,m), $V(i,m,m) \geq 0$, all the participation constraints are satisfied.

Next, we show that all the incentive constraints must hold with equality. To see this, suppose there is a type (i,j) such that

$$V(i,j,j) > V(i,j+1,j+1) + h_{i,j+1}.$$

Then we can change the incentive system as follows. We can find $\varepsilon_1, \varepsilon_2 > 0$ to change the taxes to

$$\tilde{T}_{ij} = T_{ij} + \varepsilon_1 \quad \text{and} \quad \tilde{T}_{i,j+1} = T_{i,j+1} - \varepsilon_2.$$

The effect is to reduce type $(i,j)'s$ after-tax excess income and to increase type $(i,j+1)'s$ after-tax excess income. Let $\tilde{V}(i,j,j)$ and $\tilde{V}(i,j+1,j+1)$, respectively, denote the resulting after-tax excess incomes. Recall that $p_j(i)$ denotes the conditional probability that the non-market productivity takes value j conditional on i. Since we do not change the allocation of types' (i,j) and $(i,j+1)$ working time, we have to respect the condition $p_j(i) \cdot \varepsilon_1 = p_{j+1}(i) \cdot \varepsilon_2$. By construction, $\begin{pmatrix} V(i,j,j) \\ V(i,j+1,j+1) \end{pmatrix}$ can be viewed as generated from $\begin{pmatrix} \tilde{V}(i,j,j) \\ \tilde{V}(i,j+1,j+1) \end{pmatrix}$ by a mean-preserving spread. Since $U(\cdot)$ is concave, the latter gives the objective function a higher value.

Finally, we can recover the taxes collected from the allocation of work time and the optimal after-tax excess incomes using the relation $v_{ij} = h_{ij} \cdot (i - j) - T_{ij}$ and (7). We obtain

$$T_{ij} = h_{ij} \cdot (i - j) - v_{im} - \sum_{k=j+1}^{m} h_{ik} \tag{20}$$

Substituting condition (20) into the objective function and the resource constraint gives the representation of the problem in the lemma. $\blacksquare$

Proof of Proposition 1. The proof is given in two parts. In the first part, we characterize the optimal allocation. In the second part, we use the structure of the optimal allocation to derive the budget constraint.

Part i: the structure of the allocation

The Lagrangian for our problem takes the form

$$
\begin{aligned}
L_i &= \sum_{j=1}^{m} p_j(i) \cdot U\left(j + v_{im} + \sum_{k=j+1}^{m} h_{ik}\right) \\
&\quad + \lambda_i \cdot \left(\sum_{j=1}^{m} p_j(i) \cdot \left(h_{ij} \cdot (i - j) - v_{im} - \sum_{k=j+1}^{m} h_{ik}\right) - T\right).
\end{aligned}
$$

28

For notational ease in this proof, let the marginal utility of type (i, j) be

$$u_{ij} \equiv U' \left(j + v_{im} + \sum_{k=j+1}^{m} h_{ik} \right).$$

The derivative of L_i with respect to h_{i1} is equal to

$$\frac{\partial L_i}{\partial h_{i1}} = \lambda_i \cdot (i - 1) \cdot p_1(i)$$

which implies directly that $h_{i1}^* = 1$ since $i - 1 \geq 0$.

The derivative of L_i with respect to h_{iz} is equal to

$$\frac{\partial L_i}{\partial h_{iz}} = \sum_{j=1}^{z-1} p_j(i) \cdot u_{ij} + \lambda_i \cdot (p_z(i) \cdot (i - z) - P_{z-1}(i)).$$

In what follows, we will make repeated use of a convenient transformation. Define

$$E\left[u_{ij} \mid j \leq z - 1\right] \equiv \sum_{j=1}^{z-1} \frac{p_j(i)}{P_{z-1}(i)} \cdot u_{ij}.$$

We prove that our problem admits an interior solution for at most one h_{iz}. The derivative of L_i with respect to h_{iz} for $z > 1$ is proportional to

$$\frac{\frac{\partial L_i}{\partial h_{iz}}}{P_{z-1}(i)} = E\left[u_{ij} \mid j \leq z - 1\right] + \lambda_i \cdot \left(\frac{p_z(i)}{P_{z-1}(i)} \cdot (i - z) - 1 \right). \tag{21}$$

Suppose (21) admits an interior solution for $z = t$, so the first-order condition holds:

$$E\left[u_{ij} \mid j \leq t - 1\right] = \lambda_i \cdot \left(1 - \frac{p_t(i)}{P_{t-1}(i)} \cdot (i - t) \right).$$

$E\left[u_{ij} \mid j \leq z - 1\right]$ is non-increasing in z. To see this, note that incomes are non-decreasing in opportunity costs, since by incentive compatibility

$$j + 1 + \sum_{k=j+2}^{m} h_{ik} - \left(j + \sum_{k=j+1}^{m} h_{ik} \right) = 1 - h_{i,j+1} \geq 0. \tag{22}$$

Hence, to prove our claim, it suffices to show that the expression $\frac{p_z(i)}{P_{z-1}(i)} \cdot (i - z)$ is decreasing in z, because that implies that the first-order condition cannot hold for $z > t$. So we want to show that

$$\frac{p_z(i)}{P_{z-1}(i)} \cdot (i - z) > \frac{p_{z+1}(i)}{P_z(i)} \cdot (i - (z + 1)).$$

29

Let $a \equiv i - z$. With these definitions, the condition is equivalent to

$$a \frac{p_{i-a}(i)}{P_{i-a-1}(i)} > (a-1) \frac{p_{i-a+1}(i)}{P_{i-a}(i)}.$$

The condition is trivially satisfied for $a = 1$; so assume that $a > 1$. Multiplying both sides by $\frac{P_{i-a}(i)}{p_{i-a+1}(i)} \cdot \frac{P_{i-a-1}(i)}{p_{i-a}(i)}$ and rearranging we have the equivalent condition

$$\frac{P_{i-a-1}(i)}{p_{i-a}(i)} > a \cdot \left(\frac{P_{i-a-1}(i)}{p_{i-a}(i)} - \frac{P_{i-a}(i)}{p_{i-a+1}(i)} \right). \tag{23}$$

Notice that the condition is trivially satisfied if the inverse hazard rate is non-decreasing in j, because that makes the expression on the right-hand side become negative, while the expression on the left-hand side is strictly positive. However, suppose the inverse hazard rate is decreasing so that the right-hand side is strictly positive. In that case, (23) is still satisfied, provided that

$$\frac{P_{i-a-1}(i)}{p_{i-a}(i)} \geq \frac{P_{i-1}(i)}{p_i(i)} + a \cdot \left(\frac{P_{i-a-1}(i)}{p_{i-a}(i)} - \frac{P_{i-a}(i)}{p_{i-a+1}(i)} \right).$$

Rearranging, we have

$$\frac{P_{i-a-1}(i)}{p_{i-a}(i)} + a \left(\frac{P_{i-a}(i)}{p_{i-a+1}(i)} - \frac{P_{i-a-1}(i)}{p_{i-a}(i)} \right) \geq \frac{P_{i-1}(i)}{p_i(i)}. \tag{24}$$

Assumption 1 implies that $\frac{P_j(i)}{p_{j-1}(i)} - \frac{P_{j-1}(i)}{p_j(i)}$ is non-increasing, which in turn implies that (24) is satisfied. Hence, the solution for $z > t$ is $h_{iz}^* = 0$.

Part ii: derivation of the resource constraint

Using the particular allocation, the tax paid by the marginal type satisfies $h_t \cdot (i - t) = T_{it}$ because this type's participation constraint is binding. The excess income of type $(i, t - 1)$ satisfies $V(i, t-1, t-1) = V(i, t, t-1) = h_t$, so his total income is equal to $t - 1 + h_t$. The taxes he pays satisfy the relation $i - (t - 1) - T_{i,t-1} = h_t$, so

$$T_{i,t-1} = i - ((t-1) + h_t).$$

Since the marginal utilities of all inframarginal types are the same, all their incomes are the same, so the taxes paid by all inframarginal types are the same. Summing the taxes together we obtain the expression in the text:

$$T = P_{t-1}(i) \cdot (i - ((t-1) + h_t)) + p_t(i) \cdot h_t \cdot (i - t).$$

Let $T^{\max} \equiv \max_{t \leq i} (i - (t-1)) P_{t-1}(i)$ and $\tau \equiv \arg\max_{t \leq i} (i - (t-1)) P_{t-1}(i)$. Given Assumption 1, τ exists and is unique. We can compute t and h_t as follows, generalizing the procedure described in the text. For $x \in \mathbb{N}_+ \cup 0$ and T satisfying

$$x P_{i-a} < T \leq (x+1) P_{i-(x+1)}$$

we set

$$t = i - x.$$

This solution is feasible if and only if $t \geq \tau$. The time spent by the marginal type in the formal sector satisfies

$$h_t = \frac{T - (x+1) P_{i-(x+1)}(i)}{x p_{i-x}(i) - P_{i-(x+1)}(i)}.$$

∎

Proof of Lemma 2. In contrast to Lemma 1, the left-wards adjacent constraints must bind whenever the left-wards neighbor is working in the formal sector. Imposing these constraints and solving recursively, we find that

$$v_{ij} = v_{i1} - \sum_{k=1}^{j-1} h_{ik} \qquad (25)$$

for any type who is supposed to be included in the redistribution program.

It can be shown that the left-ward adjacent incentive constraints plus monotonicity imply that there is no profitable deviation from truth-telling. Since this is standard, it is omitted. Second, following the same proof as in Lemma 1, one can show that the adjacent constraints must be tight for all types that work in the formal sector. To avoid repetition, this step is omitted as well.

If the government wishes to include type (i, s), then

$$v_{is} = v_{i1} - \sum_{k=1}^{s-1} h_{ik} \geq 0.$$

The participation constraint of type (i, s) implies that all types (i, j) for $j < s$ also want to participate. On the other hand, the exclusion constraint for type $(i, s+1)$,

$$v_{is+1} = v_{i1} - \sum_{k=1}^{s} h_{ik} \leq 0$$

implies that all types (i, j) for $j > s+1$ are also excluded. In particular, if type $(i, s+2)$ mimics the marginal type (i, s) he obtains a net excess income of

$$V(i, s, s+2) = V(i, 1, 1) - \sum_{k=1}^{s} h_{ik} - h_{is} = -h_{is} < 0$$

An analogous argument can be given for any type (i, j) for $j > s+1$.

Finally, using

$$v_{ij} = \begin{cases} v_{i1} - \sum\limits_{k=1}^{j-1} h_{ik} & \text{for } j \leq s \\ 0 & \text{otherwise,} \end{cases}$$

one can recover the subsidies paid to types (i, j) for $j \leq s$. Substituting the resulting expressions into the objective and the resource constraint gives the representation of the problem in the lemma.

∎

Proof of Proposition 2. The proof is given in two parts. In part i, we derive the structure

31

of the allocation. In part ii) we use this structure to derive the representation of the resource constraint.

Part i: structure of the allocation

The Lagrangian function for our problem takes the form

$$
\begin{aligned}
L_i\left(s\right) \;=\;& \sum_{j=1}^{s} p_j\left(i\right) \cdot U\left(j + v_{i1} - \sum_{k=1}^{j-1} h_{ik}\right) + \sum_{j=s+1}^{m} p_j\left(i\right) \cdot U\left(j\right) \\
&+\lambda_i\left(s\right)\cdot\left(\sum_{j=1}^{s} p_j\left(i\right)\cdot\left(h_{ij}\cdot\left(i-j\right) - \left(v_{i1} - \sum_{k=1}^{j-1} h_{ik}\right)\right) + S\right) \\
&+\alpha\left(v_{i1} - \sum_{k=1}^{s-1} h_{ik}\right) - \beta\left(v_{i1} - \sum_{k=1}^{s} h_{ik}\right).
\end{aligned}
$$

To ease notation in this proof, we define the marginal utility of type (i,j) as

$$
u_{ij} \equiv U'\left(j + v_{i1} - \sum_{k=1}^{j-1} h_{ik}\right).
$$

We begin by stating the derivatives of the objective function with respect to the relevant choice variables. The derivative with respect to h_{iz} for $z < s$ is equal to

$$
\frac{\partial L_i\left(s\right)}{\partial h_{iz}} = -\sum_{j=z+1}^{s} p_j\left(i\right)\cdot u_{ij} + \lambda_i\left(s\right)\cdot\left(p_z\left(i\right)\cdot\left(i-z\right) + \sum_{j=z+1}^{s} p_j\left(i\right)\right) - \alpha + \beta. \tag{26}
$$

The derivative with respect to h_{is} is equal to

$$
\frac{\partial L_i\left(s\right)}{\partial h_{is}} = \lambda_i\left(s\right)\cdot p_s\left(i\right)\cdot\left(i-s\right) + \beta. \tag{27}
$$

The derivative with respect to v_{i1} is equal to

$$
\frac{\partial L_1\left(s\right)}{\partial v_{i1}} = \sum_{j=1}^{s} p_j\left(i\right)\cdot u_{ij} - \lambda_i\left(s\right)\cdot\sum_{j=1}^{s} p_j\left(i\right) + \alpha - \beta. \tag{28}
$$

We analyze the case where $S < P_{m-1}(m-i)$, because -as we show in sections 3.3 an 3.4 - the case $S > P_{m-1}(m-i)$ cannot be part of an overall optimum. At the optimum, for $S > 0$ we must have $i - s < 0$. This implies by (27) that $\beta > 0$. To see this, suppose that $\beta = 0$. Then, by (27), we would have $h_{is}^* = 0$. But then, type $(i, s+1)$ is not excluded, so the allocation is not incentive compatible. Next, notice that $\alpha = 0$ at the optimum. If both β and α were strictly positive, then - since both constraints must hold with equality - we would have again that $h_{is}^* = 0$, which means that effectively type $(i, s-1)$ is the marginal type. Finally, at any optimum both h_{is} and v_{i1} must be at stationary points, so $\frac{\partial L_i(s)}{\partial h_{is}} = 0$ and $\frac{\partial L_1(s)}{\partial v_{i1}} = 0$. From (28) we have the first-order condition

for v_{i1}

$$\sum_{j=1}^{s} p_j(i) \cdot u_{ij} - \lambda_i(s) \cdot \sum_{j=1}^{s} p_j(i) = \beta. \tag{29}$$

Substituting (29) into the condition (27), we obtain

$$\frac{\partial L_i(s)}{\partial h_{is}} = \sum_{j=1}^{s} p_j(i) \cdot u_{ij} + \lambda_i(s) \cdot \left(p_s(i) \cdot (i-s) - \sum_{j=1}^{s} p_j(i) \right). \tag{30}$$

Substituting (29) into (26), we obtain

$$\begin{aligned}
\frac{\partial L_i(s)}{\partial h_{iz}} &= -\sum_{j=z+1}^{s} p_j(i) \cdot u_{ij} + \lambda_i(s) \cdot \left(p_z(i) \cdot (i-z) + \sum_{j=z+1}^{s} p_j(i) \right) \\
&\quad + \sum_{j=1}^{s} p_j(i) \cdot u_{ij} - \lambda_i(s) \cdot \sum_{j=1}^{s} p_j(i) \\
&= \sum_{j=1}^{z} p_j(i) \cdot u_{ij} + \lambda_i(s) \cdot \left(p_z(i) \cdot (i-z) - \sum_{j=1}^{z} p_j(i) \right). \tag{31}
\end{aligned}$$

Dividing by $P_z(i)$, we can write both (30) and (31) as

$$\frac{\frac{\partial L_i(s)}{\partial h_{iz}}}{P_z(1)} = \sum_{j=1}^{z} \frac{p_j(i)}{P_z(i)} \cdot u_{ij} + \lambda_i(s) \cdot \left(\frac{p_z(i)}{P_z(i)} \cdot (i-z) - 1 \right) \tag{32}$$

for $z \le s$. From our derivation above, the right-hand side of (32) is equal to zero at $z = s$, so

$$\sum_{j=1}^{s} \frac{p_j(i)}{P_s(i)} \cdot u_{ij} + \lambda_i(s) \cdot \left(\frac{p_s(i)}{P_s(i)} \cdot (i-s) - 1 \right) = 0. \tag{33}$$

To prove our proposition, it suffices to show that (33), in conjunction with Assumption 1 implies that

$$\sum_{j=1}^{z} \frac{p_j(i)}{P_z(i)} \cdot u_{ij} + \lambda_i(s) \cdot \left(\frac{p_z(i)}{P_z(i)} \cdot (i-z) - 1 \right) > 0$$

for all $z < s$. Letting $E[u_{ij}|j \le z] \equiv \sum_{j=1}^{z} \frac{p_j(i)}{P_z(i)} \cdot u_{ij}$ this inequality can be written as

$$E[u_{ij}|j \le z] > \lambda_i(s) \cdot \left(1 - \frac{p_z(i)}{P_z(i)}(i-z) \right)$$

for all $z < s$. We note that type (i, z) receives a weakly higher total income than type $(i, z-1)$,

since

$$z + v_{i1} - \sum_{k=1}^{z-1} h_{ik} - \left(z - 1 + v_{i1} - \sum_{k=1}^{z-2} h_{ik} \right) = 1 - h_{iz-1} \geq 0.$$

Therefore, $E\left[u_{ij}\mid j \leq z\right]$ is non-increasing in z. Hence, $E\left[u_{ij}\mid j \leq z\right] \geq E\left[u_{ij}\mid j \leq s\right]$ for all $z < s$. To complete the argument, it suffices to show that $\left(1 - \frac{p_z(i)}{P_z(i)} \cdot (i - z) \right) < \left(1 - \frac{p_s(i)}{P_s(i)} \cdot (i - s) \right)$ for all $z < s$. This is equivalent to

$$\frac{p_z(i)}{P_z(i)} \cdot (z - i) < \frac{p_s(i)}{P_s(i)} \cdot (s - i). \tag{34}$$

It is easy to see that this condition is verified for all z such that $z \leq i$. We now prove that, under Assumption 1, the condition is also verified for any z such that $i < z < s$.

In particular, we show that Assumption 1 implies that for all $z \leq s$

$$\frac{p_{z-1}(i)}{P_{z-1}(i)} \cdot (z - 1 - i) < \frac{p_z(i)}{P_z(i)} \cdot (z - i)$$

which in turn implies (34). To see this, it proves convenient to normalize this monotonicity condition around i. Let $a \equiv z - i$. With that definition, the condition is equivalent to

$$(a - 1) \cdot \frac{p_{i+a-1}(i)}{P_{i+a-1}(i)} < a \cdot \frac{p_{i+a}(i)}{P_{i+a}(i)}.$$

This condition is trivially satisfied for $a = 1$. So consider the case where $a > 1$. Manipulating this condition the same way as we did in the case of taxation, we have the equivalent condition that

$$\frac{P_{i+a}(i)}{p_{i+a}(i)} > a \cdot \left(\frac{P_{i+a}(i)}{p_{i+a}(i)} - \frac{P_{i+a-1}(i)}{p_{i+a-1}(i)} \right). \tag{35}$$

Finally, notice that Assumption 1 implies condition (35). To see this, observe simply that Assumption 1 implies that

$$\frac{P_{z+a}(i)}{p_{z+a}(i)} \geq \frac{P_z(i)}{p_z(i)} + a \cdot \left(\frac{P_{z+a}(i)}{p_{z+a}(i)} - \frac{P_{z+a-1}(i)}{p_{z+a-1}(i)} \right) \tag{36}$$

for any z and any $a \geq 0$. Since $\frac{P_z(i)}{p_z(i)} > 0$, (36) implies (35).

Part ii: Derivation of the Resource Constraint

With a binding exclusion constraint we have $v_{i1} - \sum\limits_{k=1}^{s} h_{ik} = 0$. Therefore, the excess income of all types who receive subsidies are given by

$$v_{ij} = v_{i1} - \sum_{k=1}^{j-1} h_{ik} = \sum_{k=1}^{s} h_{ik} - \sum_{k=1}^{j-1} h_{ik} = \sum_{k=j}^{s} h_{ik}.$$

We can calculate the individual subsidies, $S_{ij} = -T_{ij}$, using the relation

$$v_{ij} = h_{ij} \cdot (i - j) + S_{ij}$$

34

Hence,

$$S_{ij} = \sum_{k=j}^{s} h_{ik} - h_{ij} \cdot (i - j)$$

Using the structure of the allocation, we get

$$S_{ij} = \begin{cases} s - i + h_s & for \ j < s \\ h_s (s + 1 - i) & for \ j = s \\ 0 & for \ j > s \end{cases}$$

Summing these individual subsidies up we obtain

$$P_{s-1} (i) \cdot (s - i + h_s) + p_s (i) \cdot h_s \cdot (s + 1 - i) = S. \tag{37}$$

The marginal type is chosen optimally if s is as large as possible. If s can still be increased, this means that we can find a Pareto improvement as follows. By raising s, fewer types are excluded. All types that are included receive the same level of income. Hence, by raising s we raise all the incomes of all types that are included. The incomes of those who are and remain excluded are unchanged.

Generalizing the procedure described in the text, one can check that for $x \in \mathbb{N}_+ \cup 0$ and for $S \in \left(x P_{i+(x-1)} (i), (x + 1) P_{i+x} (i) \right]$ the marginal type is

$$s = i + x$$

and h_s is determined by the condition

$$h_s = \frac{S - x P_{i+x-1} (i)}{x p_{i+x} (i) + P_{i+x} (i)}.$$

∎

Proof of Proposition 4. Suppose the contrapositive were true and there were a productivity group i'' that is subsidized and a productivity group i' that is taxed, where $i'' > i'$. Based on this assumption, we will construct a budget balanced, incentive compatible redistribution scheme between these two groups. It follows that the initial allocation was not optimal.

The idea of the redistribution scheme is as follows. Given $i' \geq 1$, there is in each productivity group a set of individuals with low opportunity costs of time who will work full time at the optimal allocation. In groups that are taxed, the right-wards incentive constraints are tight. It follows that the marginal type, who works part time, has a strict preference for his own allocation relative to mimicking his left-ward neighbor who works full time. To ease notation in this proof, we shall write $t (i')$ for $t (i', T_{i'})$ and $s (i'')$ for $s (i'', T_{i''})$.

To see these arguments formally, recall that the optimal allocation satisfies

$$V \left(i', t (i') - 1, t (i') - 1 \right) = V \left(i', t (i'), t (i') - 1 \right) = V \left(i', t (i'), t (i') \right) + h_{t(i')}.$$

35

Hence, we can write

$$V\left(i',t\left(i'\right),t\left(i'\right)\right) = V\left(i',t\left(i'\right)-1,t\left(i'\right)-1\right) - h_{t(i')}.$$

If the marginal type mimics his left-wards neighbor, he obtains excess income

$$V\left(i',t\left(i'\right)-1,t\left(i'\right)\right) = V\left(i',t\left(i'\right)-1,t\left(i'\right)-1\right) - 1.$$

But then it follows that

$$\begin{aligned} V\left(i',t\left(i'\right)-1,t\left(i'\right)\right) &= V\left(i',t\left(i'\right)-1,t\left(i'\right)-1\right) - 1 \\ &< V\left(i',t\left(i'\right)-1,t\left(i'\right)-1\right) - h_{t(i')} = V\left(i',t\left(i'\right),t\left(i'\right)\right). \end{aligned}$$

Hence, we can decrease the taxes paid by all individuals who work full time by an identical amount, say $\varepsilon_{i'}$, without violating incentive compatibility.

In the group that is subsidized, we can decrease the subsidies paid to all individuals who work full time by an amount $\varepsilon_{i''}$ without affecting incentive compatibility and the exclusion constraint. To see this, recall that we have imposed the left-wards constraint for the marginal type so that

$$V\left(i'',s\left(i''\right),s\left(i''\right)\right) = V\left(i'',s\left(i''\right)-1,s\left(i''\right)\right) = V\left(i'',s\left(i''\right)-1,s\left(i''\right)-1\right) - 1$$

Hence,

$$V\left(i'',s\left(i''\right)-1,s\left(i''\right)-1\right) = V\left(i'',s\left(i''\right),s\left(i''\right)\right) + 1.$$

If type $\left(i'',s\left(i''\right)-1\right)$ mimics his right-wards neighbor, then he would obtain an excess income of

$$V\left(i'',s\left(i''\right),s\left(i''\right)-1\right) = V\left(i'',s\left(i''\right),s\left(i''\right)\right) + h_{s(i'')}.$$

Hence, it follows that

$$V\left(i'',s\left(i''\right)-1,s\left(i''\right)-1\right) = V\left(i'',s\left(i''\right),s\left(i''\right)\right)+1 > V\left(i'',s\left(i''\right),s\left(i''\right)\right)+h_{s(i'')} = V\left(i'',s\left(i''\right),s\left(i''\right)-1\right).$$

Choose $\varepsilon_{i''}$ and $\varepsilon_{i'}$ such that

$$p_{i''} \cdot P_{s(i'')-1}\left(i''\right) \cdot \varepsilon_{i''} + p_{i'} \cdot P_{t(i')-1}\left(i'\right) \cdot \varepsilon_{i'} = 0.$$

By construction, the new allocation and the initial allocation generate the same expected level of income for all groups together. However, the distributions differ by a mean preserving spread. Hence, the new allocation is preferred. ■

The following definition and Lemma are used in the proof of Proposition 5.

Definition *The parameters i and j are affiliated if for any $j > j'$ and any integer $b > 0$*

$$\frac{p_{ij'}}{p_{ij}} - \frac{p_{i+b,j'}}{p_{i+b,j}} \geq 0. \tag{38}$$

36

The parameters $-i$ and j are affiliated if for any $j > j'$ and any integer $b > 0$

$$\frac{p_{ij'}}{p_{ij}} - \frac{p_{i+b,j'}}{p_{i+b,j}} \leq 0. \tag{39}$$

We say that the degree of affiliation is non-decreasing in i if for any $j > j'$ and any integer $b > 0$

$$\frac{p_{ij'}}{p_{ij}} - \frac{p_{i+b,j'}}{p_{i+b,j}} \geq \frac{p_{i'j'}}{p_{i'j}} - \frac{p_{i'+b,j'}}{p_{i'+b,j}} \text{ for any } i > i'. \tag{40}$$

Lemma 3 *If the degree of affiliation is non-decreasing in i then for any j $\frac{P_j(i)}{p_j(i)} \geq \frac{1}{2}\frac{P_j(i-1)}{p_j(i-1)} + \frac{1}{2}\frac{P_j(i+1)}{p_j(i+1)}$ for all $1 < i < n$.*

Proof of Lemma 3. With $b = 1$ and $i' = i - 1$ we have from (40)

$$\frac{p_{ij'}}{p_{ij}} - \frac{p_{i+1,j'}}{p_{i+1,j}} \geq \frac{p_{i-1,j'}}{p_{i-1,j}} - \frac{p_{i,j'}}{p_{i,j}}.$$

Multiplying the first ratio by $\frac{p_i}{p_i}$, the second by $\frac{p_{i+1}}{p_{i+1}}$, and so on, we can write

$$\frac{p_i}{p_i} \cdot \frac{p_{ij'}}{p_{ij}} - \frac{p_{i+1}}{p_{i+1}} \cdot \frac{p_{i+1,j'}}{p_{i+1,j}} \geq \frac{p_{i-1}}{p_{i-1}} \cdot \frac{p_{i-1,j'}}{p_{i-1,j}} - \frac{p_i}{p_i} \cdot \frac{p_{ij'}}{p_{ij}}.$$

Substituting for $\frac{p_{ij'}}{p_i} = p_{j'}(i)$, and for analogous terms in the remaining ratios, we have

$$\frac{p_{j'}(i)}{p_j(i)} - \frac{p_{j'}(i+1)}{p_j(i+1)} \geq \frac{p_{j'}(i-1)}{p_j(i-1)} - \frac{p_{j'}(i)}{p_j(i)}.$$

Summing for $j' = 1, ..., j - 1$, we can write

$$\sum_{j'=1}^{j-1} \frac{p_{j'}(i)}{p_j(i)} - \sum_{j'=1}^{j-1} \frac{p_{j'}(i+1)}{p_j(i+1)} \geq \sum_{j'=1}^{j-1} \frac{p_{j'}(i-1)}{p_j(i-1)} - \sum_{j'=1}^{j-1} \frac{p_{j'}(i)}{p_j(i)}.$$

Performing this summation, we have

$$\frac{P_{j-1}(i)}{p_j(i)} - \frac{P_{j-1}(i+1)}{p_j(i+1)} \geq \frac{P_{j-1}(i-1)}{p_j(i-1)} - \frac{P_{j-1}(i)}{p_j(i)}.$$

Adding $\frac{p_j(i)}{p_j(i)} - \frac{p_j(i+1)}{p_j(i+1)} = 0$ on the left-hand side and $\frac{p_j(i-1)}{p_j(i-1)} - \frac{p_j(i)}{p_j(i)} = 0$ on the right-hand side, we obtain

$$\frac{P_j(i)}{p_j(i)} - \frac{P_j(i+1)}{p_j(i+1)} \geq \frac{P_j(i-1)}{p_j(i-1)} - \frac{P_j(i)}{p_j(i)}.$$

Rearranging this condition, we have

$$\frac{P_j(i)}{p_j(i)} \geq \frac{1}{2}\frac{P_j(i-1)}{p_j(i-1)} + \frac{1}{2}\frac{P_j(i+1)}{p_j(i+1)}.$$

37

■

Proof of Proposition 5. The proof proceeds by direct comparison of the marginal utilities expressions in three cases that cover all the possibilities. First, we compare the marginal utilities across groups that are taxed, second across the group with the lowest market productivity that is still taxed and the group with the highest market productivity that already receives a subsidy, and third across groups that are subsidized. In each case we conclude that the marginal utilities are monotonic in i, and that the identity of the marginal type is monotonic in i.

Case 1: two groups that are taxed. Consider two groups with market productivity i and $i-1$, where $i > \underline{i}$. Suppose the equation

$$U'(z - 1 + h) = \lambda \cdot \left(1 - \frac{p_z(i-1)}{P_{z-1}(i-1)} \cdot (i - 1 - z)\right)$$

has a solution at $z = t(i-1)$ for some $h = h_{t(i-1)} \in [0,1]$, so

$$U'\left(t(i-1) - 1 + h_{t(i-1)}\right) = \lambda \cdot \left(1 - \frac{p_{t(i-1)}(i-1)}{P_{t(i-1)-1}(i-1)} \cdot (i - 1 - t(i-1))\right).$$

Consider now the derivative of the payoff function with respect to $h_{it(i-1)}$. Fix the allocation for the group with market productivity i at the optimal allocation for group $i-1$, that is at $h_{ij} = 1$ for $j < t(i-1)$ and $h_{it(i-1)} = h_{t(i-1)}$, and consider the derivative of the payoff with respect to $h_{it(i-1)}$, evaluated at this allocation. At this allocation, all individuals with opportunity costs less than or equal to $t(i-1)$ receive the same total income; so the derivative of the payoff function is equal to

$$U'\left(t(i-1) - 1 + h_{t(i-1)}\right) - \lambda \cdot \left(1 - \frac{p_{t(i-1)}(i)}{P_{t(i-1)-1}(i)} \cdot \left(i - h_{t(i-1)}\right)\right).$$

This expression is strictly positive if and only if

$$\lambda \cdot \left(1 - \frac{p_{t(i-1)}(i)}{P_{t(i-1)-1}(i)} \cdot (i - t(i-1))\right) < \lambda \cdot \left(1 - \frac{p_{t(i-1)}(i-1)}{P_{t(i-1)-1}(i-1)} \cdot (i - 1 - t(i-1))\right).$$

In turn, this condition is equivalent to

$$\frac{p_{t(i-1)}(i)}{P_{t(i-1)-1}(i)} \cdot (i - t(i-1)) > \frac{p_{t(i-1)}(i-1)}{P_{t(i-1)-1}(i-1)} \cdot (i - 1 - t(i-1)).$$

Rearranging this expression we have

$$\frac{P_{t(i-1)-1}(i-1)}{p_{t(i-1)}(i-1)} \cdot (i - t(i-1)) > \frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} \cdot (i - 1 - t(i-1)).$$

We can reformulate this expression to

$$\frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} > \left(\frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} - \frac{P_{t(i-1)-1}(i-1)}{p_{t(i-1)}(i-1)} \right) \cdot (i - t(i-1)). \tag{41}$$

Since $i - t(i-1) > i - 1 - t(i-1) \geq 0$ (41) is always satisfied if $\frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} - \frac{P_{t(i-1)-1}(i-1)}{p_{t(i-1)}(i-1)} \leq$ 0. So, suppose that $\frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} - \frac{P_{t(i-1)-1}(i-1)}{p_{t(i-1)}(i-1)} > 0$. Note that Assumption 2 implies then that $\frac{P_{t(i-1)-1}(i')}{p_{t(i-1)}(i')} - \frac{P_{t(i-1)-1}(i'-1)}{p_{t(i-1)}(i'-1)} > 0$ for all $i' < i$. Moreover, we can bound the expression on the right-hand side, noting that

$$i - t(i-1) \leq i - 1.$$

So, it is true that

$$\frac{P_{t(i-1)-1}(1)}{p_{t(i-1)}(1)} + \left(\frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} - \frac{P_{t(i-1)-1}(i-1)}{p_{t(i-1)}(i-1)} \right) \cdot (i-1) > \left(\frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} - \frac{P_{t(i-1)-1}(i-1)}{p_{t(i-1)}(i-1)} \right) \cdot (i - t(i-1)).$$

Thus, a sufficient condition for (41) is

$$\frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} \geq \frac{P_{t(i-1)-1}(1)}{p_{t(i-1)}(1)} + \left(\frac{P_{t(i-1)-1}(i)}{p_{t(i-1)}(i)} - \frac{P_{t(i-1)-1}(i-1)}{p_{t(i-1)}(i-1)} \right) \cdot (i-1),$$

which is implied by Assumption 2.

So, we have shown that

$$U'\left(t(i-1) - 1 + h_{t(i-1)} \right) - \lambda \cdot \left(1 - \frac{p_{t(i-1)}(i)}{P_{t(i-1)-1}(i)} \cdot (i - t(i-1)) \right)$$

$$> \quad U'\left(t(i-1) - 1 + h_{t(i-1)} \right) - \lambda \cdot \left(1 - \frac{p_{t(i-1)}(i-1)}{P_{t(i-1)-1}(i-1)} \cdot (i - 1 - t(i-1)) \right) = 0.$$

It follows that

$$t(i) + h_{t(i)} > t(i-1) + h_{t(i-1)}. \tag{42}$$

Therefore, the after-tax incomes of the inframarginal types in group i are strictly higher at the optimum than the incomes of the inframarginal types in group $i-1$. (42) implies directly that

$$t(i) \geq t(i-1). \tag{43}$$

To see this, suppose that we had $t(i) < t(i-1)$, contrary to what we just claimed. Substracting $t(i)$ on both sides of inequality (42), we have

$$h_{t(i)} > t(i-1) - t(i) + h_{t(i-1)}.$$

But if $t(i) < t(i-1)$ then $t(i-1) - t(i) \geq 1$, so $t(i-1) - t(i) + h_{t(i-1)} \geq 1 + h_{t(i-1)}$. However, that can only hold if

$$h_{t(i)} > 1 + h_{t(i-1)}.$$

However, this contradicts the fact that both $h_{t(i)}$ and $h_{t(i-1)}$ belong to the unit interval. Hence, we

39

have shown that $t\left(i\right) \geq t\left(i-1\right).$

Case 2: $i = \underline{i}$, two groups where one with market productivity i is taxed and the other with market productivity $i - 1$ is subsidized.

For the group that is taxed, the marginal utility of the marginal type satisfies

$$U'\left(t\left(i\right) - 1 + h_{t(i)}\right) = \lambda \cdot \left(1 - \frac{p_{t(i)}\left(i\right)}{P_{t(i)-1}\left(i\right)} \cdot \left(i - t\left(i\right)\right)\right).$$

For the group that is subsidized, the marginal utility of the marginal type satisfies

$$U'\left(s\left(i-1\right) + h_{s(i-1)}\right) = \lambda \cdot \left(1 - \frac{p_{s(i-1)}\left(i-1\right)}{P_{s(i-1)}\left(i-1\right)} \cdot \left(i - 1 - s\left(i-1\right)\right)\right).$$

These solutions satisfy $U'\left(t\left(i\right) - 1 + h_{t(i)}\right) \leq U'\left(s\left(i-1\right) + h_{s(i-1)}\right)$ if and only if

$$\lambda \cdot \left(1 - \frac{p_{t(i)}\left(i\right)}{P_{t(i)-1}\left(i\right)} \cdot \left(i - t\left(i\right)\right)\right) \leq \lambda \cdot \left(1 - \frac{p_{s(i-1)}\left(i-1\right)}{P_{s(i-1)}\left(i-1\right)} \cdot \left(i - 1 - s\left(i-1\right)\right)\right).$$

It is easy to see that this condition must always hold, since we have $i - t\left(i\right) \geq 0$ and $i - 1 - s\left(i-1\right) \leq 0$. The former property is necessary since group i is taxed; the latter property is optimal since incomes of individuals in group $i - 1$ are raised. It follows that the after-tax incomes of the inframarginal types in group i are strictly higher than the after-tax incomes of all types in group $i - 1$ who receive any subsidies. Formally, we have

$$t\left(i\right) - 1 + h_{t(i)} \geq s\left(i-1\right) + h_{s(i-1)}. \tag{44}$$

This implies further that

$$t\left(i\right) - 1 \geq s\left(i-1\right). \tag{45}$$

To see this, suppose again that the contrary were true, so that $t\left(i\right) - 1 < s\left(i-1\right).$ This, however, would imply that

$$h_{t(i)} \geq s\left(i-1\right) - \left(t\left(i\right) - 1\right) + h_{s(i-1)} \geq 1 + h_{s(i-1)},$$

which cannot be the case, because both $h_{t(i)}$ and $h_{s(i-1)}$ belong to the unit interval and $h_{s(i-1)} > 0$.

Case 3: $i < \underline{i}$, two groups with market productivity i and $i - 1$ that are both subsidized.

Subsidized individuals in the lower productivity group receive incomes equal to $s\left(i-1\right) + h_{s(i-1)}$; subsidized individuals in the higher productivity group receive incomes equal to $s\left(i\right) + h_{s(i)}$. Suppose the equation

$$U'\left(z + h\right) = \lambda \cdot \left(1 - \frac{p_z\left(i-1\right)}{P_z\left(i-1\right)} \cdot \left(i - 1 - z\right)\right)$$

has a solution for $z = s\left(i-1\right)$ and some $h = h_{s(i-1)} \in \left(0, 1\right]$, so

$$U'\left(s\left(i-1\right) + h_{s(i-1)}\right) = \lambda \cdot \left(1 - \frac{p_{s(i-1)}\left(i-1\right)}{P_{s(i-1)}\left(i-1\right)} \cdot \left(i - 1 - s\left(i-1\right)\right)\right).$$

As for case 1) we evaluate the derivative of the payoff function for the group with market produc-

tivity i at the optimal allocation for group $i - 1$, to show that

$$U'\left(s\left(i-1\right)+h_{s(i-1)}\right) - \lambda \cdot \left(1 - \frac{p_{s(i-1)}\left(i\right)}{P_{s(i-1)}\left(i\right)} \cdot \left(i - s\left(i-1\right)\right)\right) > 0.$$

This statement is true if and only if

$$\left(1 - \frac{p_{s(i-1)}\left(i\right)}{P_{s(i-1)}\left(i\right)} \cdot \left(i - s\left(i-1\right)\right)\right) < \left(1 - \frac{p_{s(i-1)}\left(i-1\right)}{P_{s(i-1)}\left(i-1\right)} \cdot \left(i - 1 - s\left(i-1\right)\right)\right)$$

or, equivalently, if and only if

$$\left(\frac{P_{s(i-1)}\left(i-1\right)}{p_{s(i-1)}\left(i-1\right)} - \frac{P_{s(i-1)}\left(i\right)}{p_{s(i-1)}\left(i\right)}\right) \cdot \left(s\left(i-1\right) - i\right) < \frac{P_{s(i-1)}\left(i\right)}{p_{s(i-1)}\left(i\right)}. \tag{46}$$

Note that $s\left(i-1\right) \geq i - 1$. In fact, it is easy to show that (46) is always satisfied if $s\left(i-1\right) \leq i$, so consider the case where $s\left(i-1\right) > i$. Then, (46) is always satisfied if $\frac{P_{s(i-1)}(i-1)}{p_{s(i-1)}(i-1)} - \frac{P_{s(i-1)}(i)}{p_{s(i-1)}(i)} \leq 0$, so suppose that $\frac{P_{s(i-1)}(i-1)}{p_{s(i-1)}(i-1)} - \frac{P_{s(i-1)}(i)}{p_{s(i-1)}(i)} > 0$. In this case, Assumption 2 implies that $\frac{P_{s(i-1)}(i'-1)}{p_{s(i-1)}(i'-1)} - \frac{P_{s(i-1)}(i')}{p_{s(i-1)}(i')} > 0$ for any $i' > i$. Moreover, we can derive an upper bound for the expression on the left-hand side of (46); noting that $s\left(i-1\right) \leq m$, we have

$$s\left(i-1\right) - i \leq m - i,$$

so we can write

$$\left(\frac{P_{s(i-1)}\left(i-1\right)}{p_{s(i-1)}\left(i-1\right)} - \frac{P_{s(i-1)}\left(i\right)}{p_{s(i-1)}\left(i\right)}\right) \cdot \left(s\left(i-1\right) - i\right) \leq \left(\frac{P_{s(i-1)}\left(i-1\right)}{p_{s(i-1)}\left(i-1\right)} - \frac{P_{s(i-1)}\left(i\right)}{p_{s(i-1)}\left(i\right)}\right) \cdot \left(m - i\right) + \frac{P_{s(i-1)}\left(m\right)}{p_{s(i-1)}\left(m\right)}$$

Thus, our condition is satisfied if we have

$$\frac{P_{s(i-1)}\left(m\right)}{p_{s(i-1)}\left(m\right)} + \left(\frac{P_{s(i-1)}\left(i-1\right)}{p_{s(i-1)}\left(i-1\right)} - \frac{P_{s(i-1)}\left(i\right)}{p_{s(i-1)}\left(i\right)}\right) \cdot \left(m - i\right) < \frac{P_{s(i-1)}\left(i\right)}{p_{s(i-1)}\left(i\right)}.$$

Rearranging appropriately, we have

$$\frac{P_{s(i-1)}\left(i\right)}{p_{s(i-1)}\left(i\right)} + \left(\frac{P_{s(i-1)}\left(i\right)}{p_{s(i-1)}\left(i\right)} - \frac{P_{s(i-1)}\left(i-1\right)}{p_{s(i-1)}\left(i-1\right)}\right) \cdot \left(m - i\right) > \frac{P_{s(i-1)}\left(m\right)}{p_{s(i-1)}\left(m\right)}, \tag{47}$$

which is implied by Assumption 2. Hence we have shown that

$$s\left(i\right) + h_{s(i)} > s\left(i-1\right) + h_{s(i-1)}.$$

Moreover, by the now familiar argument, this implies also that

$$s\left(i\right) \geq s\left(i-1\right).$$

∎

41

Proof of Proposition 6. The incentive constraint for an arbitrary allocation rule w_{ij} is that for all (i,j)

$$h_{ij} \cdot (w_{ij} - j) - T_{ij} \geq h_{\hat{i}\hat{j}} \cdot (w_{\hat{i}\hat{j}} - j) - T_{\hat{i}\hat{j}} \quad \forall (\hat{i},\hat{j}) \ s.t. w_{\hat{i}\hat{j}} \leq i. \tag{48}$$

Adjust the allocation for all (i,j) as follows: change w_{ij} to $\tilde{w}_{ij} = i$, leave the allocation of working times unchanged, $h_{ij} = \tilde{h}_{ij}$, and change the taxes from T_{ij} to $\tilde{T}_{ij} = T_{ij} + h_{ij} \cdot (i - w_{ij})$. Substituting the adjusted rules, $\left\{ h_{ij}, \tilde{w}_{ij}, \tilde{T}_{ij} \right\}$ for all i, j into (48), we find that

$$h_{ij} \cdot (i - j) - \tilde{T}_{ij} \geq h_{\hat{i}\hat{j}} \cdot (\hat{i} - j) - \tilde{T}_{\hat{i}\hat{j}} \quad \forall (\hat{i},\hat{j}) \ s.t. w_{\hat{i}\hat{j}} \leq i. \tag{49}$$

By construction, the left-hand side of (48) and (49) are equal for all (i,j) and the right-hand sides of (48) and (49) are equal for all $\hat{i}, \hat{j}$. The incentive constraint under the new allocation takes the form

$$h_{ij} \cdot (i - j) - \tilde{T}_{ij} \geq h_{\hat{i}\hat{j}} \cdot (\hat{i} - j) - \tilde{T}_{\hat{i}\hat{j}} \quad \forall (\hat{i},\hat{j}) \ s.t. \hat{i} \leq i, \tag{50}$$

because $\tilde{w}_{ij} = i$. (50) is implied by (49). The reason is that feasibility of the initial allocation requires that the rule w_{ij} satisfies $w_{ij} \leq i$ for all i, j. Hence, the set $\{\hat{i}, \hat{j} : \hat{i} \leq i\}$ is a subset of the set $\{\hat{i}, \hat{j} : w_{\hat{i}\hat{j}} \leq i\}$. So, the new rule is incentive compatible. Finally consider budget balancedness and individual rationality. Since the initial allocation was budget balanced, the new allocation runs a surplus, which can be distributed in lumpsum fashion to all individuals. This redistribution does not affect any participation constraints. ∎

REFERENCES

Armstrong, M., Multiproduct Nonlinear Pricing, *Econometrica* **64**, 1996, 51-76.

Balestrino, A., A. Cigno and A. Petini, 2003, Doing Wonders with an Egg: Optimal Re-Distribution When Households Differ in Market and Non-Market Abilities, *Journal of Public Economic Theory*, **5**, 479-498.

Beaudry, P. and C. Blackorby, Taxes and Employment Subsidies in Optimal Redistribution Programs, The Warwick Economics Research Paper Series, 2004, 766.

Besley, T. and S. Coate, 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.

Boone, J. and A. L. Bovenberg, 2004, The optimal Taxation of Unskilled Labor with Job Search and Social Assistance, *Journal of Public Economics* **88**, 2227-2258.

Brett, C., 1998, Who Should Be on Workfare? The Use of Work Requirement as Part of an Optimal Tax Mix, *Oxford Economic Papers* **50**, 607-622.

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

Choné, P. and G. Laroque, 2005, Optimal Incentives for Labor Force Participation, *Journal of Public Economics* **89**, 395-425.

Choné, P. and G. Laroque, 2006, Should low skilled work be subsidized?, INSEE-CREST.

Cuff, K., 2000, Optimality of Workfare with Heterogeneous Preferences *Canadian Journal of Economics* **33**, 149-174.

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

Diamond, P., 1980, Income Taxation with Fixed Hours of Work, *Journal of Public Economics* **13**, 101-110.

Diamond, P., 1997, Optimal Income Taxation: an Example with a U-Shaped Pattern of Optimal Marginal Tax Rates, *American Economic Review* **88**, 1998, 83-95.

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

Grogger, J., and L. Karoly, Welfare Reform: Effects of a Decade of Change, *Harvard University Press, Cambridge USA*.

Jullien, B., 2000, Participation Constraints in Adverse Selection Models, *Journal of Economic Theory* **93**, 1-47.

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.

Laroque, G., 2005, Income Maintenance and labor force participation, *Econometrica* **73**, 341-376.

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.

Maskin and Riley, 1985, "Input versus Output Incentive Schemes", *Journal of Public Economics*, 28, 1-23.

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

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

Moffitt, R., 2004, Welfare Programs and Labor Supply, in Handbook of Public Economics vol 4, *Elsevier, North Holland*

Moffitt, R., 2006, Welfare Work Requirements with Paternalistic Government Preferences, *The Economic Journal*, **116**, F441-F458.

Rochet, J.-C. and Choné, P., 1998, Ironing, Sweeping, and Multidimensional Screening, *Econometrica*, **66**, 783-826.

Saez, E., 2002, Optimal Income Transfer Programmes: Intensive versus Extensive Labor Supply Responses, *Quarterly Journal of Economics*, **117**, 1039-73.