Pareto Efficient Income Taxation with Stochastic Abilities[*]

Abstract

This paper studies Pareto efficient income taxation in an economy with infinitely-lived individuals whose income generating abilities evolve according to a two-state Markov process. The study yields two main results. First, when individuals are risk neutral, the fraction of individuals who face a positive marginal income tax rate is always positive but converges to zero. Moreover, the tax rate these individuals face also goes to zero. Second, Pareto efficient income tax systems can be time-consistent even when the degree of correlation in ability types is large.

Marco Battaglini
Department of Economics
Princeton University
Princeton NJ 08544
mbattagl@princeton.edu

Stephen Coate
Department of Economics
Yale University
New Haven CT 06511
stephen.coate@yale.edu

1 Introduction

A central problem in normative public economics is the design of income tax systems. The fundamental difficulty is that, while individuals' incomes may be observable, their abilities to earn income are unobservable (Mirrlees (1971)). Thus, if equity considerations demand that higher ability individuals should pay a larger share of government spending, those with higher incomes must pay more taxes. But this raises the possibility that higher ability individuals may avoid their obligations by reducing their earnings to masquerade as low ability individuals. To mitigate this possibility, income tax systems must optimally screen ability types which requires distorting the earnings of individuals downwards. These distortions imply a basic trade-off between equity and aggregate efficiency.

Much of the literature on the design of income taxation has taken a static perspective. While its lessons may apply in a dynamic context when individuals' income generating abilities are constant, the more relevant case is that in which abilities, while persistent, may vary over time. This case raises three new theoretical questions. When abilities have some persistence, the efficient screening of ability types may require that the tax system be non-stationary, making current taxes depend upon individuals' past earnings choices. The first question, therefore, is what is the pattern of distortions in individuals' earnings choices and how do these vary over time? Moreover, when abilities are variable, the tax system will impact the allocation of consumption across states and time. There will be a trade-off between the goals of smoothing consumption and providing incentives. The second question, therefore, is what are the pattern of distortions in the allocation of consumption across states and time? Finally, when abilities are variable, it is not clear if efficient tax systems are time consistent. In the constant ability case, optimal income tax systems are never time-consistent. Distortionary taxation is necessary to screen ability types, but after individuals have revealed their abilities, the government will find it optimal to eliminate such distortions, making the original tax system non credible (Roberts (1984)). However, when abilities are stochastic, residual uncertainty remains, because an individual may change type. Accordingly, the government must still screen types in the remaining periods. The third question, therefore, is under what circumstances are efficient tax systems time consistent?

This paper studies these questions in a dynamic version of the classic Mirrlees model. In any period, there are just two ability types - low and high - and individuals' abilities follow a

Markov process, meaning that an individual's current ability is a sufficient statistic for predicting his ability in the next period. Individuals' per period utility functions are additively separable in consumption and leisure. Generalizing the approach of Stiglitz (1982), we study the Pareto efficiency problem of maximizing the expected utility of those who are initially high ability subject to a given target utility for those who are initially low ability.[1] Following Diamond (1998), the bulk of our analysis focuses on the case in which individuals' per period utility functions are quasi-linear, implying that they are risk neutral. This makes the allocation of consumption across time and states payoff irrelevant and hence the second question moot.

Our analysis of the risk neutral case yields clean and striking results to the first and third questions. With regard to the dynamics of earnings distortions, when individuals' abilities are constant over time, Pareto efficient tax systems are stationary. Low ability individuals face a positive marginal tax rate and their earnings are distorted downwards. High ability individuals face a zero marginal tax rate and their earnings are undistorted. However, when ability types are not perfectly correlated over time, the results are dramatically different. Efficient tax systems are non-stationary and *the only individuals whose earnings are distorted are those who currently are and have always been low ability.* All other individuals face a zero marginal rate of taxation. Moreover, the degree to which these perpetual low types have their earnings distorted converges to zero, so that not only is the fraction of individuals who face a positive marginal tax rate converging to zero, but the tax rate these individuals are facing goes to zero. Thus, in a very strong sense, the distortions caused by efficient income tax schemes converge to zero over time.

With regard to time consistency, we establish a lower bound on the correlation in types such that below it the optimal tax system is time consistent. We also find that when the correlation of types is above this bound, it is governments with higher spending commitments and/or more ambitious redistributive objectives who find it harder to commit to implement efficient income tax systems. Accordingly, it is governments with more progressive agendas that will be forced to pursue their objectives with third best policies. Since these will lead to greater distortions and larger reductions in aggregate efficiency than second best policies, the result suggests that the equity-efficiency trade off will be steeper than suggested by static optimal tax theory.

We also study the case of risk averse individuals in a two period version of the model. We show

---

[1] This is distinct from the approach of Mirrlees (1971) who characterizes the problem of maximizing an additive social welfare function.

that individuals who are low ability in the second period face a positive marginal tax rate *even if they were previously high ability.* Thus, risk neutrality is a necessary condition for our result on earnings distortions described above. However, we also show that the distortion imposed on those who become low types in the second period is close to zero when individuals are only slightly risk averse and that the basic pattern of earnings under an efficient tax system is the same as in the risk neutral case. Thus, the broad features of efficient tax systems that we identify in our basic model are robust to introducing small amounts of risk aversion.

The paper contributes to a small but growing literature that approaches the problem of dynamic optimal taxation using the mechanism design approach of static optimal tax theory.[2] This literature includes papers by Berliant and Ledyard (2003), Brito et al (1991), Diamond and Mirrlees (1978), Golosov, Kocherlakota and Tsyvinsky (2003), Golosov and Tsyvinski (2003), Kocherlakota (2003), Roberts (1984), and Werning (2002). Our paper differs from these in both the model that we use and the focus of our analysis. First, we assume that individuals abilities evolve according to a Markov process, whereas the bulk of the literature assumes that either abilities are constant or that they are independent across periods.[3] Second, in all other respects, our model is much simpler than many of those in the literature. In particular, it has only two ability types, no capital, exogenous interest rates, and, for the most part, assumes that individuals are risk neutral. The benefit of these more restrictive assumptions is that they allow us to provide a full characterization of optimal allocations. Third, we characterize Pareto efficient tax systems rather than those that maximize a utilitarian social welfare function. Our risk neutrality assumption means that the utilitarian solution of our model simply involves lump sum taxation and is thus uninteresting. Fourth, our focus is primarily on exploring the dynamic pattern of distortions in earnings and time consistency issues (the first and third questions raised above) rather than the consumption smoothing question (the second question). Indeed, as far as we are aware, none of the papers in this new tax literature have results on the dynamic pattern of earnings distortions or on the time consistency of efficient tax systems.

In characterizing Pareto efficient allocations and studying their time consistency, our paper draws on the dynamic contracting literature. In particular, we follow the analytical approach

---

[2] The alternative approach to dynamic optimal taxation is to make the assumption that the government is constrained to use linear taxes (see Chari and Kehoe (1999) for a review).

[3] The exceptions are the papers by Golosov, Kocherlakota and Tsyvinsky (2003) and Kocherlakota (2003) that allow for rather general stochastic evolution of abilities.

employed by Battaglini (2003a) to study a monopoly pricing problem with long-lived consumers whose tastes evolve according to a Markov process. We show that his approach can be fruitfully applied to the problem of optimal income taxation. The taxation problem is somewhat more involved than the pricing problem, in part because it involves characterizing the entire Pareto frontier rather than simply finding the profit maximizing solution. Among other things, characterizing the entire frontier helps us understand the role of the government's initial spending commitments and redistributive objectives in determining the time consistency of efficient allocations. Our analysis also extends Battaglini's work by investigating the robustness of optimal policies to risk aversion.

The organization of the remainder of the paper is as follows. The next section presents the model. Section 3 explains how we characterize Pareto efficient allocations. Section 4 describes the properties of Pareto efficient allocations and draws out their implications for the efficient taxation of labor income. Section 5 studies a two period version of the model in which agents are risk averse and explains how risk aversion modifies the conclusions. Section 6 analyses the time consistency of Pareto efficient allocations and section 7 concludes.

2 The model

We study an economy populated by a continuum of infinitely-lived individuals. There are two goods - consumption and leisure. In each period t, individuals get utility from consumption x_t and work l_t according to the utility function $x_t - \varphi(l_t)$ where φ is increasing, strictly convex, and twice continuously differentiable. Individuals are endowed with $\bar{l}$ units of time in each period. To avoid having to worry about corner solutions, we assume that $\varphi'(0) = 0$ and that $\lim_{l \to \bar{l}} \varphi'(l) = \infty$. Individuals discount the future at rate $\delta < 1$ and can borrow or lend at the exogenously fixed interest rate $r = \frac{1}{\delta} - 1$.

Individuals differ in their income generating abilities. In period t, an individual with income generating ability θ_t earns income $y_t = \theta_t l_t$ if he works an amount l_t. In each period, there are just two ability levels, low and high, denoted by $\{\theta_L, \theta_H\}$ where $0 < \theta_L < \theta_H$. However, individuals' abilities may differ across periods. Specifically, each individual's ability follows a Markov process

4

with support $\{\theta_L, \theta_H\}$ and transition matrix:

$$\begin{bmatrix} \alpha_{LL} & \alpha_{LH} \\ \alpha_{HL} & \alpha_{HH} \end{bmatrix}.$$

We assume that types are correlated (i.e., that $\alpha_{HH} \geq \alpha_{LH}$), but we do not make assumptions on the degree of correlation. A fraction $\mu \in (0,1)$ of individuals are high types in period one.

The economy also has a government. In each period, this government spends an amount G. While this spending does not directly impact individuals' utilities, the government must raise the revenue necessary to finance it.

A *history* for an individual at time t consists of a list of his previous $t-1$ abilities; i.e., $h_t = \{\theta_1, ..., \theta_{t-1}\}$. Let $h_1 = \emptyset$ denote the history at time 1 and let H denote the set of all histories. Let the notation $h_{t+j} \succeq h_t$ mean that h_{t+j} follows h_t (i.e., its first $t-1$ components are equal to h_t).

An *allocation* in this economy is described by $(\mathbf{x}, \mathbf{y}) = \{(x_t(\theta_t; h_t), y_t(\theta_t; h_t))\}_{t=1}^{\infty}$. Here $(x_t(\theta_t; h_t), y_t(\theta_t; h_t))$ is the consumption-earnings bundle of those individuals who have ability θ_t in period t after history $h_t \in H$. To be *feasible*, an allocation must satisfy the aggregate resource constraint[4]

$$\sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t) + G] \leq \sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t)].$$

This says that the present value of consumption and government spending equals the present value of earnings. The expected lifetime utility of an individual with high ability in period one under such an allocation is

$$V(\theta_H \mid h_1) = \sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t) \mid \theta_1 = \theta_H],$$

while that of an individual with low ability in period one is

$$V(\theta_L \mid h_1) = \sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t) \mid \theta_1 = \theta_L].$$

An allocation is *efficient* if, for some $\underline{u}$, it solves the problem

$$\max_{(\mathbf{x}, \mathbf{y})} V(\theta_H \mid h_1)$$

$$s.t. \ V(\theta_L \mid h_1) \geq \underline{u} \quad (U_L)$$

---

[4] Throughout the paper, we will ignore non-negativity constraints on individuals' consumptions.

and

$$\sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t) + G] \leq \sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t)]. \quad (R)$$

Thus, it maximizes the expected utility of those who are high types in period one subject to giving a fixed level of utility to those who are low types and the aggregate resource constraint.

It is straightforward to show that an allocation $(\mathbf{x}, \mathbf{y})$ is efficient if and only if the earnings path $\mathbf{y}$ maximizes Marshallian aggregate surplus

$$\sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t)],$$

and the consumption path x is such that U_L and R hold with equality. The surplus maximizing earnings path has the property that individuals work up until the point at which their marginal disutility of work equals their marginal product. More precisely, in any period t after any history h_t, $y_t(\theta_H; h_t)$ must equal $y^*(\theta_H)$ and $y_t(\theta_L; h_t)$ must equal $y^*(\theta_L)$ where

$$y^*(\theta) = \arg\max\{y - \varphi(y/\theta) : y/\theta \in [0, \bar{l}]\} \qquad \theta \in \{\theta_L, \theta_H\}.$$

If the government can observe individuals' abilities, it can realize any utility allocation on the Pareto frontier through a very simple stationary tax system. All high ability individuals in any period are required to pay a lump sum tax T_H^* and all low ability individuals must pay a tax T_L^*. These taxes are chosen so that (i) the present value of tax receipts equals the present value of government expenditures and (ii) those who begin period one as low types have the required expected utility of $\underline{u}$. Given this system, high ability individuals will choose to earn $y^*(\theta_H)$ in any period and low ability individuals will choose to earn $y^*(\theta_L)$.

Efficient allocations will be difficult to realize when the government is unable to observe individuals' income generating abilities and hence unable to impose ability-specific lump sum taxes. Accounting for this unobservability, requires ensuring that it is always in individuals' interests to claim the bundles intended for them. This requires that the allocation satisfy the following set of incentive constraints. For all time periods t and histories h_t,

$$x_t(\theta_H; h_t) - \varphi(y_t(\theta_H; h_t)/\theta_H) + \delta E_{\theta_H} V(\theta_{t+1} | (h_t, \theta_H))$$

$$\geq x_t(\theta_L; h_t) - \varphi(y_t(\theta_L; h_t)/\theta_H) + \delta E_{\theta_H} V(\theta_{t+1} | (h_t, \theta_L)) \quad (IC_H(h_t))$$

and

$$x_t(\theta_L; h_t) - \varphi(y_t(\theta_L; h_t)/\theta_L) + \delta E_{\theta_L} V(\theta_{t+1} | (h_t, \theta_L))$$

$$\geq \quad x_t(\theta_H; h_t) - \varphi(y_t(\theta_H; h_t)/\theta_L) + \delta E_{\theta_L} V(\theta_{t+1} | (h_t, \theta_H)) \quad (IC_L(h_t)),$$

where $V(\theta_t | h_t)$ denotes the expected lifetime utility of an individual who has type θ_t in period t after history h_t. These constraints ensure that in any period t after any history h_t individuals are always better off with the bundle intended for them than the bundle intended for any other individual they could credibly claim to be.

We say that an allocation is *second best efficient* if, for some $\underline{u}$, it solves the problem

$$\max_{(\mathbf{x},\mathbf{y})} V(\theta_H | h_1) \qquad\qquad (\mathcal{P}^I)$$

$$s.t. \quad U_L, \ R, \ and \ IC_H(h_t) \ \& \ IC_L(h_t) \ for \ all \ t \ \& \ h_t$$

In the sequel, we refer to this as the *general problem*. Our interest lies in understanding what second best efficient allocations look like and how they may be decentralized via tax-transfer systems. To make the problem interesting, we restrict consideration to cases in which G and $\underline{u}$ are sufficiently high that any efficient allocation in which those who are low types in period one have expected utility $\underline{u}$ violates at least one of the high types' incentive constraints.

3   Solution procedure

To characterize second best efficient allocations, we study the following *relaxed problem*:

$$\max_{(\mathbf{x},\mathbf{y})} V(\theta_H | h_1) \qquad\qquad (\mathcal{P}^{II})$$

$$s.t. \ U_L, \ R, \ and \ IC_H(h_t) \ for \ all \ t \ \& \ h_t$$

The relaxed problem imposes the incentive constraints after any history only for those who are currently high types. We will first characterize the solution to the relaxed problem and then explain the relationship between the relaxed and general problems.

Our first observation about the relaxed problem is:

Lemma 1 *Let* $(\mathbf{x}, \mathbf{y})$ *solve the relaxed problem. Then both* U_L *and* $IC_H(h_1)$ *hold with equality.*

The reason why the period one incentive constraint is binding is that, if it were not, then by transferring resources forward in time as necessary, we could assure that none of the incentive

constraints were satisfied. But then the solution to the relaxed problem would be an efficient allocation and, by assumption, must violate at least one of the high type's incentive constraints.

Lemma 1 does not imply that all the incentive constraints are binding because the solution may involve giving those who are high types in the future sufficient consumption that they are strictly better off not masquerading as low types. It turns out, however, that this possibility can be ignored.

Lemma 2 *Let $(\mathbf{x}, \mathbf{y})$ be an allocation satisfying the constraints of the relaxed problem. Then there exists $\mathbf{x}'$ such that $(\mathbf{x}', \mathbf{y})$ satisfies all the constraints and yields the same value of the objective function as $(\mathbf{x}, \mathbf{y})$ but also satisfies $IC_H(h_t)$ with equality for all periods $t > 1$ and all histories h_t.*

To understand this result, suppose that under the allocation $(\mathbf{x}, \mathbf{y})$ an incentive constraint is not binding for individuals who are high types at some period $t > 1$ after some history $h_t = (h_{t-1}, \theta_{t-1})$. Then, we can make it bind by reducing the high types' consumption in that period and giving the expected present value to those with history h_{t-1} and ability θ_{t-1} in period $t - 1$. If $\theta_{t-1} = \theta_H$ then this has no implications for the incentive constraint of the high types in period $t - 1$ with history h_{t-1}. The gain in consumption in period $t - 1$ is exactly offset by the loss in expected consumption should they remain high types in period t. If $\theta_{t-1} = \theta_L$ then the transfer does have implications for the incentive constraint of the high types in period $t - 1$ with history h_{t-1}. On the one hand, masquerading as low types in period $t - 1$ now yields more consumption in period $t - 1$. On the other, it yields less consumption in period t if individuals remain high types. It turns out that because high types are more likely to remain high types than are low types to become high types, the cost of lower future consumption outweighs the benefit of higher current consumption so that the incentive constraint still holds. Indeed, the transfer leads the incentive constraint of the high type in period $t - 1$ with history h_{t-1} to be satisfied strictly. However, we can repeat the process by reducing the consumption of the high type in period $t - 1$ with history $h_{t-1} = (h_{t-2}, \theta_{t-2})$ and giving the expected present value to those with history h_{t-2} and ability θ_{t-2} in period $t - 2$. By repeating this process as many times as necessary, we find a consumption allocation $\mathbf{x}'$ that satisfies all the incentive constraints with equality except possibly the first period constraint.

It follows from Lemma 1 and 2 that there is no loss of generality in assuming that in the solution to the relaxed problem $IC_H(h_t)$ holds with equality for all t and h_t. We can use this fact

to write the expected lifetime utility of an individual with high ability after history h_t as

$$V\left(\theta_H \mid h_t\right) = V\left(\theta_L \mid h_t\right) + \Phi\left(y_t(\theta_L; h_t)\right) + \Delta EV(\theta_{t+1} \mid (h_t, \theta_L)) \tag{1}$$

where $\Phi\left(y\right) = \varphi(y/\theta_L) - \varphi(y/\theta_H)$ and $\Delta EV(\theta_{t+1} \mid (h_t, \theta_L))$ is the difference in the continuation values for the two types $E_{\theta_H} V(\theta_{t+1} \mid (h_t, \theta_L)) - E_{\theta_L} V(\theta_{t+1} \mid (h_t, \theta_L))$. By successively using this equation, we can write the difference in the continuation values as solely a function of the earnings of an individual who is a low type in period t and remains one thereafter. Denote by $H^\circ\left(h_t\right)$ the set of histories following a history h_t in which in all the periods including and after t the individual has low ability. Let h_{t+j}° denote an element of $H^\circ\left(h_t\right)$. Then we can use (1) to show that:

Lemma 3 *Let* $(\mathbf{x}, \mathbf{y})$ *be an allocation satisfying* $IC_H(h_t)$ *with equality for all periods t and all histories h_t. Then, the utility of an individual with history h_t who is a high type in period t can be written as:*

$$V\left(\theta_H \mid h_t\right) = \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{t+j}(\theta_L; h_{t+j}^\circ)\right) + V\left(\theta_L \mid h_t\right). \tag{2}$$

This result can in turn be used to establish:

Lemma 4 *Let* $(\mathbf{x}, \mathbf{y})$ *solve the relaxed problem. Then the earnings levels* $\mathbf{y}$ *solve the problem:*

$$\max \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{1+j}(\theta_L; h_{1+j}^\circ)\right) + \underline{u}$$

$$s.t. \quad G \quad \leq \quad (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t)] \tag{3}$$

$$-(1-\delta)[\mu \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{1+j}(\theta_L; h_{1+j}^\circ)\right) + \underline{u}]$$

The problem described in Lemma 4 is straightforward to solve. Letting γ be the multiplier on the revenue constraint, the associated Lagrangian can be written as

$$\mathcal{L} \quad = \quad \sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t)] - G/(1-\delta) \tag{4}$$

$$-(\mu - \frac{1}{\gamma(1-\delta)})\{\sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{1+j}(\theta_L; h_{1+j}^\circ)\right) + \underline{u}\}$$

The first term is Marshallian aggregate surplus, while the second term represents the loss of surplus resulting from having to meet the incentive constraints. Letting $h_t^* = h_{1+(t-1)}^\circ$, the first order conditions are that for all t and $h_t \neq h_t^*$

$$\varphi'(y_t(\theta_t; h_t)/\theta_t) = \theta_t \tag{5}$$

and for all t and $h_t = h_t^*$

$$(1-\mu)[1 - \frac{\varphi'(y_t(\theta_L; h_t^*)/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{t-1} \Phi'(y_t(\theta_L; h_t^*))(\mu - \frac{1}{\gamma(1-\delta)}) \qquad (6)$$

As we show in the proof of Proposition 1, the value of the multiplier γ is such that $\mu > 1/\gamma(1-\delta)$, so that the right hand side of (6) is positive.

Before we study the implications of these conditions, we first clarify the relationship between the relaxed and general problems.

Lemma 5 *Let $(\mathbf{x}, \mathbf{y})$ be an allocation with the property that the earnings levels solve the problem described in Lemma 4 and the consumption levels are such as to make U_L and $IC_H(h_t)$ (for all t and h_t) hold with equality. Then, $(\mathbf{x}, \mathbf{y})$ is a second best efficient allocation. Conversely, if $(\mathbf{x}, \mathbf{y})$ is a second best efficient allocation, then the earnings levels must solve the problem described in Lemma 4.*

It follows from this result that if $(\mathbf{x}, \mathbf{y})$ is a second best efficient allocation then the earnings levels satisfy the first order conditions (5) and (6). In the next section, we use this to derive some results about the nature of second best efficient allocations. Before doing that, it is worth noting that the relationship between the relaxed and general problems is somewhat non-standard. In a standard problem, it is the case that any solution to the relaxed problem solves the general problem. In our problem, those solutions that do not satisfy the constraints with equality do not necessarily solve the general problem.

4   Second best efficient allocations

To set the stage for our first main result, we briefly describe the nature of second best efficient allocations in which ability types are perfectly correlated over time, so that $\alpha_{HH} = \alpha_{LL} = 1$. Under this assumption, the model is a replication of the Mirrlees model with two ability types. The following proposition will therefore come as no surprise.

Proposition 1 *Suppose that ability types are perfectly correlated over time. Then, in any second best efficient allocation, the earnings of individuals are constant over time. The earnings of those with high ability are undistorted (i.e., they earn $y^*(\theta_H)$ each period), while the earnings of those with low ability are distorted downward (i.e., they earn less than $y^*(\theta_L)$ each period).*

Proof: In any period t, high ability individuals have history $h_t = (\theta_H, ..., \theta_H)$. The first order condition (5) then implies that $y_t(\theta_H; h_t) = y^*(\theta_H)$. Low ability individuals, on the other hand, have history $h_t = (\theta_L, ..., \theta_L) = h_t^*$ so that (6) implies their earnings satisfy

$$(1-\mu)[1 - \frac{\varphi'(y_t(\theta_L; h_t^*)/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{t-1} \Phi'(y_t(\theta_L; h_t^*))(\mu - \frac{1}{\gamma(1-\delta)}).$$

Since $\alpha_{HL} = 0$, the latter implies that:

$$(1 - \mu)[1 - \frac{\varphi'(y_t(\theta_L; h_t^*)/\theta_L)}{\theta_L}] = \Phi'(y_t(\theta_L; h_t^*))(\mu - \frac{1}{\gamma(1 - \delta)}). \tag{7}$$

Assuming that $\mu > 1/\gamma(1 - \delta)$, (7) implies that low types earn a constant amount each period which is less than $y^*(\theta_L)$.

It only remains to prove that $\mu > 1/\gamma(1 - \delta)$. Assume, first that $\mu = 1/\gamma(1 - \delta)$. Then (7) implies that for all t, $y_t(\theta_L; h_t^*) = y^*(\theta_L)$. This means that the earnings levels that solve the problem described in Lemma 4, denoted $\widetilde{\mathbf{y}}$, are first best efficient. But we know by assumption that if $(\mathbf{x}, \mathbf{y})$ solves the relaxed problem then $\mathbf{y}$ is not first best efficient and hence cannot equal $\widetilde{\mathbf{y}}$. This contradicts Lemma 4.

Next suppose that $\mu < 1/\gamma(1 - \delta)$. Let $\widetilde{\mathbf{y}}$ denote the associated optimal earnings levels. Let $\widetilde{\mathbf{x}}$ be such as to make $IC_H(h_t)$ (for all t and h_t) and U_L hold with equality given $\widetilde{\mathbf{y}}$. Then, we will show that $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{y}})$ cannot solve the relaxed problem - a contradiction. To see this, consider a marginal reduction dy in $\widetilde{y}_1(\theta_L; h_1)$ and choose dx so as to keep the period one low type's utility constant; i.e., so that

$$\widetilde{x}_1(\theta_L; h_1) - dx - \varphi(\frac{\widetilde{y}_1(\theta_L; h_1) - dy}{\theta_L}) = \widetilde{x}_1(\theta_L; h_1) - \varphi(\frac{\widetilde{y}_1(\theta_L; h_1)}{\theta_L}).$$

Clearly,

$$dx = \frac{\varphi'(\widetilde{y}_1(\theta_L; h_1))}{\theta_L} dy.$$

Note that $IC_H(h_1)$ is now still satisfied, because the high type now finds the low type's bundle less attractive because it involves less earnings. However, the change in revenues is

$$dR = (1 - \mu)[dx - dy] = (1 - \mu)[\frac{\varphi'(\widetilde{y}_1(\theta_L; h_1))}{\theta_L} - 1]dy > 0.$$

This change is positive since $\mu < 1/\gamma(1 - \delta)$ implies that $1 - \frac{\varphi'(\widetilde{y}_1(\theta_L; h_1)/\theta_L)}{\theta_L} < 0$. Now take this revenue increase and divide it among those who are high types in period one; i.e., raise $\widetilde{x}_1(\theta_H; h_1)$ by dR/μ. Clearly, this change makes the high types strictly better off, which since it violates none of the constraints, means that $(\widetilde{\mathbf{x}}, \widetilde{\mathbf{y}})$ cannot solve the relaxed problem. $Q.E.D.$

Thus, the labor supply of high ability individuals is undistorted, while the labor supply of low ability individuals is distorted downwards. This is to counteract the incentives of high ability types to masquerade as low ability types. Any utility allocation on the (second best) Pareto frontier can be realized through a very simple *stationary* tax system. Let y_L^{**} denote the amount

earned by low ability individuals in the utility allocation in question. Let x_L^* be the constant level of consumption just sufficient to give low ability individuals utility level $\underline{u}$ given that they are working y_L^{**}/θ_L in each period; i.e.,

$$\frac{x_L^* - \varphi(y_L^{**}/\theta_L)}{1-\delta} = \underline{u}.$$

Let x_H^* be the constant level of consumption just sufficient to prevent high ability individuals from claiming to be low ability individuals given that they are working $y^*(\theta_H)/\theta_H$ in each period; i.e.,

$$\frac{x_H^* - \varphi(y^*(\theta_H)/\theta_H)}{1-\delta} = \frac{\varphi(y_L^{**}/\theta_L) - \varphi(y_L^{**}/\theta_H)}{1-\delta} + \underline{u}.$$

Then if the government requires that in each period individuals pay taxes according to the schedule $T(y)$ where $T(y_H^*) = y^*(\theta_H) - x_H^*$ and $T(y_{LL}^*) = y_{LL}^* - x_L^*$, it can achieve the utility allocation in question. This tax schedule has a zero marginal rate in the neighborhood of the income level $y^*(\theta_H)$ but a positive marginal rate in the neighborhood of y_{LL}^*.[5]

With this as background, we now present our first main result.

Proposition 2 *Suppose that ability types are not perfectly correlated over time. Then, in any second best efficient allocation, the earnings of individuals who are currently, or have at some point been, high types are undistorted (i.e., they earn $y^*(\theta_t)$ in period t when they have ability θ_t). The earnings of individuals who are currently and have always been low types are distorted downwards (i.e., they earn less than $y^*(\theta_L)$). However, the extent of this distortion decreases over time converging to 0 as $t \to \infty$.*

Proof: The first order conditions tell us that for all t and $h_t \neq h_t^*$

$$\varphi'(y_t(\theta_t; h_t)/\theta_t) = \theta_t \tag{8}$$

and for all t and $h_t = h_t^*$

$$(1-\mu)[1 - \frac{\varphi'(y_t(\theta_L; h_t^*)/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{t-1} \Phi'\left(y_t(\theta_L; h_t^*)\right)(\mu - \frac{1}{\gamma(1-\delta)}) \tag{9}$$

If an individual is currently or has at some point been a high type, then $h_t \neq h_t^*$ and, from (8), it can be seen that the first order conditions imply that they work up until the point at which their marginal disutility of work $\varphi'(y/\theta_t)$ equals their wage θ_t. If an individual is currently and has

always been a low type then $h_t = h_t^*$ and, from (9), it can be seen that the first order conditions imply that they work less than the amount at which their marginal disutility of work equals their wage. Since ability types are not perfectly correlated over time, $1 > \frac{\alpha_{HL}}{\alpha_{LL}}$ and $\left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{t-1}$ is decreasing in t and converges to zero as $t \to \infty$. The first order condition therefore implies that $y_t(\theta_L; h_t^*)/\theta_L$ is decreasing in t and converges to $y^*(\theta_L)/\theta_L$ as $t \to \infty$. $Q.E.D.$

The proposition implies that when ability types are not perfectly correlated over time, the fraction of individuals in any period whose labor supply is distorted in any second best efficient allocation is decreasing and converges to zero as $t \to \infty$. Moreover, the degree to which these individuals have their labor supply distorted also converges to zero. Thus, in a very strong sense, the distortion caused by imperfect observability of individuals' abilities goes to zero.

To understand the first part of the proposition, consider a group of individuals at some time t who share the same history h_t. Suppose that at some point in the past these individuals were high ability so that $h_t \neq h_t^*$. By Proposition 2, the earnings of these individuals are undistorted at time t. This is obviously optimal for those with high ability at time t, so consider those with low ability. Suppose, to the contrary, that the earnings of these individuals are distorted downwards. Then, if we were to increase their earnings slightly in period t we would make them better off. Of course, such a change would also necessitate an increase in the consumption of those who have high ability at time t to prevent them from masquerading as low types. This will reduce government revenues. However, this reduction in expected revenues can be financed by a concordant reduction in the consumption of these individuals in the period $\tau < t$ in which they were first high types. This reduction gives individuals with high ability in period τ and history $h_\tau \preceq h_t$ no incentive to masquerade as low types. The reason is that the reduction in current consumption is offset by the increase in expected future consumption at time t. This marginal change in the allocation would not cause any of the incentive constraints of low ability individuals to be violated since none of these are binding.

To understand the second part of the proposition, it is useful to contrast it with the result in Proposition 1. As noted, with constant types the earnings of low ability individuals are distorted downwards and the degree of distortion is constant over time. The size of the distortion is determined by a simple marginal cost - marginal benefit argument. A lower distortion increases the Marshallian surplus generated by an individual and therefore obviously increases welfare. How-

ever, it also increases the consumption that needs to be given to individuals with high ability. This reduces tax revenues for the government and increases the shadow cost of taxation γ. At the optimum, the marginal increase in surplus is exactly compensated by the marginal reduction in revenues. With constant abilities the marginal cost/benefit ratio is constant throughout periods. After any period t, the marginal benefit of a lower distortion is proportional to the fraction of low types (the constant $1 - \mu$), because types never change. Similarly, the marginal cost is constant: it is proportional to the fraction of high types whose consumption must be raised (the constant μ) and the shadow cost of taxation: $\mu - 1/\gamma(1 - \delta)$. When types change over time, the marginal cost/benefit ratio is not constant, because, depending on the realized history, there is a different composition of the population. The marginal benefit of a lower distortion in the earnings of those individuals who at time t are and have always been low types is proportional to the fraction of such individuals in the population: $(1 - \mu) \alpha_{LL}^{t-1}$. The marginal cost, evaluated at time 1, also depends on the time t of the change. At time t the consumption of high ability individuals who have previously been low types increases by, say, ΔR_t. At time $t - 1$ the expected utility of those who are and have always been low types increases because they can become high types at time t and benefit from the increase in consumption at that time. Part of this extra expected utility can be taxed away at $t - 1$, but not all since incentive compatibility must be satisfied at that time as well. At time $t - 1$ individuals who have high ability for the first time can not receive less than what they would receive if they chooses the option designed for those who remain low types. Even if we completely tax away the expected increase in consumption of those who, at time $t - 1$, are and have always been low types with a tax T_{t-1} such that $\alpha_{LH}\Delta R_t - T_{t-1} = 0$, those individuals who have high ability at time $t - 1$ after previously being low types must receive an increase in consumption equal to $\Delta R_{t-1} = (\alpha_{HH}\Delta R_t - T_{t-1}) - (\alpha_{LH}\Delta R_t - T_{t-1}) = (\alpha_{HH} - \alpha_{LH})\Delta R_t$. Repeating the same argument, if we try and tax away these gains at $t - 2$, we must provide an increase in consumption at time $t - 2$ for those individuals who have high ability for the first time equal to $\Delta R_{t-2} = (\alpha_{HH} - \alpha_{LH})\Delta R_{t-1} = (\alpha_{HH} - \alpha_{LH})^2 \Delta R_t$. Proceeding backward, we arrive to an increase in the consumption of those who are high ability at time 1 proportional to $(\alpha_{HH} - \alpha_{LH})^{t-1}$. Since these individuals make up a fraction μ of the population, the marginal cost of a lower distortion in the earnings of those individuals who at time t are and have always been low types is proportional to $\mu (\alpha_{HH} - \alpha_{LH})^{t-1}$. Accordingly, the marginal cost/benefit ratio at time t is now proportional to $\frac{\mu}{1-\mu} \left[\frac{\alpha_{HH} - \alpha_{LH}}{\alpha_{LL}} \right]^{t-1}$. As the cost/benefit ratio of a lower distortion

14

vanishes over time,[6] the distortion vanishes with it.

When ability types are not perfectly correlated, it is no longer the case that utility allocations on the (second best) Pareto frontier can be realized through a stationary tax system. Proposition 2 implies that the marginal tax rates individuals face should depend upon their entire history of earnings. What might such a non–stationary tax system look like? To provide a feel for this, we will describe a particular tax system that can realize utility allocations on the Pareto frontier. It should be stressed that this is not the only possibility. Given that individuals have constant marginal utility of consumption, the allocation of consumption across time or states is irrelevant for individuals' utility and this gives a great deal of flexibility in choosing consumption paths and hence tax systems.

Consider a particular utility allocation on the Pareto frontier and let $\mathbf{y}$ denote the associated earnings levels. These must solve the problem described in Lemma 4. Thus, $y_t(\theta_H; h_t) = y^*(\theta_H)$ for all t and h_t and $y_t(\theta_L; h_t) = y^*(\theta_L)$ for all t and all $h_t \neq h_t^*$. To simplify notation, let $y_{Lt}^* = y_t(\theta_L; h_t^*)$ for all t. Now, choose $\mathbf{x}$ as follows. First, let the consumption of high types in any period be constant, so that $x_t(\theta_H; h_t) = x_H^*$ for all t and h_t for some x_H^*. In addition, let the consumption of those who are currently low types but have previously been high types be constant, so that $x_t(\theta_L; h_t) = x_L^*$ for all t and $h_t \neq h_t^*$. Further, let this consumption be related to x_H^* in the following way:

$$x_L^* = x_H^* - (\varphi(y^*(\theta_H)/\theta_H) - \varphi(y^*(\theta_L)/\theta_H)).$$

Finally, for those who have always been low types let $x_t(\theta_L; h_t^*) = x_{Lt}^*$ where $(x_{Lt}^*)_{t=1}^\infty$ satisfy for all t

$$
\begin{aligned}
x_{Lt}^* &= x_L^* - (\varphi(y^*(\theta_L)/\theta_H) - \varphi(y_{Lt}^*/\theta_H)) + \delta\alpha_{HL}\{x_L^* - \varphi(y^*(\theta_L)/\theta_L) - (x_{Lt+1}^* - \varphi(y_{Lt+1}^*/\theta_L))\} \\
&\quad + \sum_{j=2}^\infty \delta^j \alpha_{LL}^{j-1}\{x_L^* - \varphi(y^*(\theta_L)/\theta_L) - (x_{Lt+j}^* - \varphi(y_{Lt+j}^*/\theta_L))\}.
\end{aligned}
$$

It may be verified that $\mathbf{x}$ so constructed is such as to make the incentive constraints $IC_H(h_t)$ (for all t and h_t) bind. To ensure U_L binds, let x_H^* be chosen so that when x_L^* and $(x_{Lt}^*)_{t=1}^\infty$ are defined by the above equations, then $V(\theta_L \mid h_1) = \underline{u}$.

---

[6] This can be seen from the fact that the term in the square parenthesis is lower than one: indeed $\frac{\alpha_{HH}-\alpha_{LH}}{\alpha_{LL}} = 1 - \frac{\alpha_{HL}}{\alpha_{LL}} < 1$ because types are positively correlated.

Now consider the features of the tax system that would decentralize the allocation $(\mathbf{x}, \mathbf{y})$. In period 1, individuals would face a schedule $T_1(y)$ that requires them to pay a tax $T_1(y^*(\theta_H)) = y^*(\theta_H) - x_H^*$ if they earn $y^*(\theta_H)$ and a tax $T_1(y_{L1}^*) = y_{L1}^* - x_{L1}^*$ if they earn y_{L1}^*. This first period schedule has a positive marginal rate at income y_{L1}^* and a zero rate at income $y^*(\theta_H)$.

In the second period, the schedule individuals face depends upon their first period earnings. Those who earn $y^*(\theta_H)$ in the first period face a schedule $T_2(y; y^*(\theta_H))$ that requires them to pay a tax $T_2(y^*(\theta_H); y^*(\theta_H)) = y^*(\theta_H) - x_H^*$ if they earn $y^*(\theta_H)$ and a tax $T_2(y^*(\theta_L); y^*(\theta_H)) = y^*(\theta_L) - x_L^*$ if they earn $y^*(\theta_L)$. This tax schedule has zero marginal rates in the neighborhood of both the income levels $y^*(\theta_H)$ and $y^*(\theta_L)$. Those who earn y_{L1}^* in the first period, face a schedule $T_2(y; y_{L1}^*)$ that requires them to pay a tax $T_2(y^*(\theta_H); y_{L1}^*) = y^*(\theta_H) - x_H^*$ if they earn $y^*(\theta_H)$ and a tax $T_2(y_{L2}^*; y_{L1}^*) = y_{L2}^* - x_{L2}^*$ if they earn y_{L2}^*. This tax schedule has a zero marginal rate in the neighborhood of $y^*(\theta_H)$ but a positive marginal rate in the neighborhood of y_{L2}^*. Thus, the tax schedule $T_2(y; y_{L1}^*)$ has a *different marginal rate* in the neighborhood of $[y_{L2}^*, y^*(\theta_L)]$ than $T_2(y; y^*(\theta_H))$. Since $y_{L2}^* > y_{L1}^*$, the tax schedule $T_2(y; y_{L1}^*)$ has a lower marginal rate in the neighborhood of y_{L2}^* than the first period tax schedule.

In the third period, those who had earned $y^*(\theta_H)$ in the first period continue to face the schedule $T_2(y; y^*(\theta_H))$ as do those who earned $y^*(\theta_H)$ in the second period. Those who earned y_{L1}^* and y_{L2}^* in the first two periods, face the schedule $T_3(y; y_{L1}^*, y_{L2}^*)$ that requires them to pay a tax $T_3(y^*(\theta_H); y_{L1}^*, y_{L2}^*) = y^*(\theta_H) - x_H^*$ if they earn $y^*(\theta_H)$ and a tax $T_3(y_{L3}^*; y_{L1}^*, y_{L2}^*) = y_{L3}^* - x_{L3}^*$ if they earn y_{L3}^*. Since $y_{L3}^* > y_{L2}^*$, the tax schedule $T_3(y; y_{L1}^*, y_{L2}^*)$ involves a lower marginal rate in the neighborhood of y_{L3}^* than does the second period tax schedule. As time progresses, more and more individuals come under the tax schedule $T_2(y; y^*(\theta_H))$. Moreover, the schedule faced by those with an uninterrupted history of low earnings $T_t(y; y_{L1}^*, .., y_{Lt-1}^*)$ converges to the schedule $T_2(y; y^*(\theta_H))$.

5   Implications of risk aversion

The model studied so far has assumed that individuals are risk neutral. This implies that individuals are indifferent as to the allocation of consumption across time and states. This facilitates a clean focus on the problem of designing taxation to minimize distortions in labor supply. In this section, we extend the model to incorporate risk aversion and investigate how this impacts the results.

To incorporate risk aversion, we assume that:

$$u_t = (x_t)^\beta - \varphi(l_t) \tag{10}$$

where $\beta \in (0,1)$. As β approaches 1 individuals become risk neutral and we have the original model. For tractability, in this section we focus on a two-period version of the model.

In the two-period model, an allocation can be fully described by

$$(\mathbf{x}, \mathbf{y}) = \{(x_L, x_H, x_{LL}, x_{LH}, x_{HL}, x_{HH}); (y_L, y_H, y_{LL}, y_{LH}, y_{HL}, y_{HH})\}.$$

Thus, (x_L, y_L) is the consumption-earnings bundle intended for those individuals who have low ability in period one; (x_{LL}, y_{LL}) is the period two bundle intended for those who have low ability in both periods; and so on. As before, an allocation is *efficient* if, for some $\underline{u}$, it solves the problem

$$\max[(x_H)^\beta - \varphi(y_H/\theta_H)] + \delta[\alpha_{HH}((x_{HH})^\beta - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}((x_{HL})^\beta - \varphi(y_{HL}/\theta_L))].$$

s.t. $[(x_L)^\beta - \varphi(y_L/\theta_L)] + \delta[\alpha_{LL}((x_{LL})^\beta - \varphi(y_{LL}/\theta_L)) + \alpha_{LH}((x_{LH})^\beta - \varphi(y_{LH}/\theta_H))] \geq \underline{u}$ (U_L)

and

$$[\mu x_H + (1-\mu)x_L + G] + \delta[\mu(\alpha_{HH}x_{HH} + \alpha_{HL}x_{HL}) + (1-\mu)(\alpha_{LH}x_{LH} + \alpha_{LL}x_{LL}) + G]$$
$$\leq \quad [\mu y_H + (1-\mu)y_L] + \delta[\mu(\alpha_{HH}y_{HH} + \alpha_{HL}y_{HL}) + (1-\mu)(\alpha_{LH}y_{LH} + \alpha_{LL}y_{LL})]. \quad (R)$$

Efficient allocations now have two properties. First, individuals' consumption levels are constant across time and states. Thus, for the period one high types $x_H = x_{HL} = x_{HH}$ and for the low types $x_L = x_{LH} = x_{LL}$. Second, there is no distortion in individuals' labor supply decisions. This requires that each individual should work up until the point at which their marginal disutility of work equals the marginal utility of the consumption that work produces. Thus, $\theta_H \beta x_H^{\beta-1} = \varphi'(y_H/\theta_H)$, $\theta_L \beta x_{HL}^{\beta-1} = \varphi'(y_{HL}/\theta_L)$, and so on.

Again, if the government can observe individuals' abilities, it can realize efficient allocations with a simple tax system. Let $y^*(T, \theta)$ denote the earnings level that would maximize the static utility of an individual with ability θ if he had to pay a lump sum tax T; that is,

$$y^*(T, \theta) = \arg\max\{(y - T)^\beta - \varphi(y/\theta) : y/\theta \in [0, \bar{l}]\} \qquad \theta \in \{\theta_L, \theta_H\}.$$

Then, any efficient allocation $(\mathbf{x}, \mathbf{y})$ can be decentralized as follows. Individuals who are high types in the first period pay a lump sum tax T_H where $y^*(T_H, \theta_H) - T_H = x_H$. They also pay this

17

tax in the second period if they remain high types. If they become low types, their tax burdens are reduced in such a way as to maintain their consumption at the level it would be were they high types. More precisely, they pay a tax T_{HL} such that $y^*(T_{HL}, \theta_L) - T_{HL} = x_H$. In effect, the tax system completely insures them against any consumption loss resulting from a shock in their income generating ability. The story for those who are low types in the first period is similar. In the first period, they pay the lump sum tax T_L where $y^*(T_L, \theta_L) - T_L = x_L$. They also pay this tax in the second period if they remain low types. If they become high types, their tax burden is increased in such a way as to maintain their consumption at the level it would be were they low types. More precisely, they would pay a tax T_{LH} such that $y^*(T_{LH}, \theta_H) - T_{LH} = x_L$. Individuals have no incentive to save under this tax system, as it keeps their marginal utility of consumption constant across time and states.

An allocation is *second best efficient* if it solves the efficiency problem with the following additional set of incentive constraints:

$$(x_H)^\beta - \varphi(y_H/\theta_H) + \delta[\alpha_{HH}((x_{HH})^\beta - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}((x_{HL})^\beta - \varphi(y_{HL}/\theta_L))]$$
$$\geq \quad (x_L)^\beta - \varphi(y_L/\theta_H) + \delta[\alpha_{HH}((x_{LH})^\beta - \varphi(y_{LH}/\theta_H)) + \alpha_{HL}((x_{LL})^\beta - \varphi(y_{LL}/\theta_L))] \quad (IC(H))$$

$$(x_L)^\beta - \varphi(y_L/\theta_L) + \delta[\alpha_{LH}((x_{LH})^\beta - \varphi(y_{LH}/\theta_H)) + \alpha_{LL}((x_{LL})^\beta - \varphi(y_{LL}/\theta_L))]$$
$$\geq \quad (x_H)^\beta - \varphi(y_H/\theta_L) + \delta[\alpha_{LH}((x_{HH})^\beta - \varphi(y_{HH}/\theta_H)) + \alpha_{LL}((x_{HL})^\beta - \varphi(y_{HL}/\theta_L))] \quad (IC(L))$$

$$(x_{HH})^\beta - \varphi(y_{HH}/\theta_H) \geq (x_{HL})^\beta - \varphi(y_{HL}/\theta_H) \quad (IC(HH))$$

$$(x_{HL})^\beta - \varphi(y_{HL}/\theta_L) \geq (x_{HH})^\beta - \varphi(y_{HH}/\theta_L) \quad (IC(HL))$$

$$(x_{LH})^\beta - \varphi(y_{LH}/\theta_H) \geq (x_{LL})^\beta - \varphi(y_{LL}/\theta_H) \quad (IC(LH))$$

$$(x_{LL})^\beta - \varphi(y_{LL}/\theta_L) \geq (x_{LH})^\beta - \varphi(y_{LH}/\theta_L) \quad (IC(LL))$$

The first two constraints deal with the first period and the remainder the second period. We retain the assumption that when $\beta = 1$, G and $\underline{u}$ are sufficiently high that any efficient allocation in which those who are low types in period one have expected utility of at least $\underline{u}$ violates one or more of the high types' incentive constraints. Thus, the second best Pareto frontier lies strictly inside the first best frontier when $\beta = 1$.

To solve for second best efficient allocations we again consider the relaxed problem obtained by ignoring the incentive constraints for low types. Thus we study the problem:

$$\max_{(\mathbf{x},\mathbf{y})}[(x_H)^\beta - \varphi(y_H/\theta_H)] + \delta[\alpha_{HH}((x_{HH})^\beta - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}((x_{HL})^\beta - \varphi(y_{HL}/\theta_L))]$$

$$s.t. \ U_L, R, IC(H), IC(HH), IC(LH).$$

While we are no longer able to prove generally that second best efficient allocations must solve the relaxed problem, we can show that this is the case for β sufficiently large.

Lemma 6 *In the two period model with risk averse individuals, there exist a $\overline{\beta} < 1$ such that if $\beta \in (\overline{\beta}, 1)$, $(\mathbf{x},\mathbf{y})$ is a second best efficient allocation if and only if it solves the relaxed problem.*

By analyzing the first order conditions for the relaxed problem, we are able to establish:

Proposition 3 *In the two period model with risk averse individuals, there exist a $\overline{\beta} < 1$ such that if $\beta \in (\overline{\beta}, 1)$, in any second best efficient allocation, the earnings of individuals who are high types in either period are undistorted. The earnings of individuals who are low types in either period are distorted downwards. However, the degree of distortion in the earnings of those who become low types in the second period converges to 0 as $\beta \to 1$. Moreover, those who are low types in both periods earn more in the second period.*

This proposition shows that once we introduce risk aversion, the result that in any second best efficient allocation only those who remain low types have their labor supply decisions distorted no longer holds. Those who start out as high types and become low types in the second period, also work less than the efficient amount. However, as individuals become less and less risk averse the degree of this distortion converges to zero. Moreover, the basic pattern of earnings in any second best efficient allocation is the same as in the risk neutral case. In particular, the earnings of those who remain low types are increasing.

With risk aversion, the allocation of consumption across time and states is relevant for individuals' utility and this explains why the earnings of individuals with history HL are distorted downwards. Reducing the earnings level y_{HL} lessens the incentive of those with history HH to pretend they have history HL. In the risk neutral case, this problem could be handled by increasing the consumption of those with history HH and taking the expected discounted value from high ability individuals in the first period. But because individuals want to smooth their consumption this intertemporal reallocation is no longer without cost.

What can be established about the allocation of consumption across time and states in a second best efficient allocation? Our next proposition addresses this.

19

Proposition 4 *In the two period model with risk averse individuals, there exist a $\overline{\beta} < 1$ such that if $\beta \in (\overline{\beta}, 1)$, in any second best efficient allocation the consumption of individuals who are high types in the first period goes up if they remain high types in the second period and down if they become low types. Similarly, the consumption of individuals who are low types in the first period goes up if they are high types in the second period and down if they are low types. Moreover, for both low and high types, the marginal utility of consumption in the first period is less than the expected marginal utility of consumption in the second period.*

Thus, when compared with efficient allocations there are two distinct distortions in the allocation of consumption. First, the allocation of consumption across states is distorted in the sense that individuals are not fully insured. If they are low types in the second period, their consumption is lower than if they are high types. This is obviously a necessary condition for incentive compatibility. Second, the allocation of consumption across time is distorted in the sense that individuals consume more than is optimal in the first period. This is a particular application of the result first established by Rogerson (1985) and since generalized and applied to optimal taxation by Golosov, Kocherlakota, and Tsyvinsky (2003). The intuition is the following. Because of the incentive compatibility constraint, low types will supply less labor and enjoy lower consumption in each period. The marginal utility of consumption of low types, therefore, is higher than the marginal utility of high types in period two. Suppose, to the contrary, that the marginal utility of consumption in the first period were higher than the expected marginal utility in the second period for some type. If we reduce the second period consumption of high and low types by some amount Δ incentive compatibility is preserved, since the utility of low types is reduced by more than that of high types. This reduction frees Δ units of consumption that can be used to increase consumption in the first period. But then, since the marginal increase in utility at $t = 1$ is higher that the expected reduction at $t = 2$, the change creates a Pareto improvement: and we have a contradiction.

What would a tax system that could decentralize second best efficient allocations look like? Let $(\mathbf{x}, \mathbf{y})$ be the allocation we wish to decentralize. The first period tax schedule $T_1(y)$ is such that $T_1(y_H) = y_H - x_H$ and $T_1(y_L) = y_L - x_L$. The marginal rate of taxation is positive in the neighborhood of y_L and zero in the neighborhood of y_H. Those who earn y_H in the first period face a second period tax schedule $T_2(y; y_H)$ such that $T_2(y_{HH}; y_H) = y_{HH} - x_{HH}$ and $T_2(y_{HL}; y_H) = y_{HL} - x_{HL}$. The marginal rate of taxation is positive in the neighborhood of y_{HL} and zero in the neighborhood of y_{HH}. Those who earn y_L in the first period face a second period tax schedule $T_2(y; y_L)$ such that $T_2(y_{LH}; y_L) = y_{LH} - x_{LH}$ and $T_2(y_{LL}; y_L) = y_{LL} - x_{LL}$. The

20

marginal rate of taxation is positive in the neighborhood of y_{LL} and zero in the neighborhood of y_{LH}.

It is important to note that under this tax system, individuals have an incentive to save in the first period. As Proposition 4 demonstrates, $\mathbf{x}$ is such that for both types the marginal utility of consumption in the first period is less than the expected marginal utility of consumption in the second period. This provides individuals with an incentive to transfer consumption forward. To prevent this, it would be necessary to supplement the tax on earnings with a tax on capital income from savings. The tax rate on saving should be sufficient to deter individuals from saving given the consumptions $\mathbf{x}$. Obviously, there are many choices of capital tax rates that would achieve this goal.

6   Time consistency

Imagine that at the beginning of period one the government announces a tax/transfer system designed to implement a particular utility allocation on the second best Pareto frontier. Individuals' period one earnings choices would then reveal their period one types. If the government could use this information to design a new tax/transfer system that was better for all individuals and raised just as much revenue, one might imagine that it would be tempted to do so. In this case, the original tax/transfer system would not be *time consistent.*

Up to this point, we have ignored this time consistency problem, implicitly assuming that the government can credibly commit to the ex ante optimal tax/transfer system. The equilibrium characterized in Section 4, is therefore a *Ramsey equilibrium* (Ramsey (1927)): the government determines the optimal policy given individuals' reaction functions. However, it is well known that Ramsey optimal policies are often not time-consistent in the sense that even benevolent governments find it ex post optimal to depart from them.[7]   In a model like ours, distortionary taxation is necessary to extract individuals' private information; after individuals have revealed it, the government may find it optimal to eliminate such distortions: but then the original tax system would not be credible. This type of time-inconsistency in the context of taxation with asymmetric information was first observed by Roberts (1984).

In this section we show that when individuals' types can vary stochastically the time-inconsistency

---

[7] The classic reference is Kydland and Prescott (1977). See Chari, Kehoe and Prescott (1988) and Stokey (1989) for general discussion and surveys of the literature.

problem may not arise. To analyze the issue we return to the basic model with quasi-linear utilities and infinite periods.[8] We also impose the additional assumption that the disutility of labor function has a positive third derivative; i.e., $\varphi'''(l) \geq 0$.[9] This assumption guarantees that $\Phi'' > 0$ which in turn implies that the Lagrangian for the maximization problem in Lemma 4 (4) is strictly concave and that the efficient earnings levels are unique.

We begin by providing a more precise definition of time consistency. We will work directly with allocations rather than the tax-transfer systems that generate them. It is to be understood that a particular tax/transfer system is time consistent if and only if the allocation it generates is time consistent. Consider then a particular second best efficient allocation $(\mathbf{x}^*, \mathbf{y}^*)$ and imagine that we are at the beginning of some period $t \geq 2$. At that point, the government knows the histories of all the individuals in the economy but not their period t types. Consider a group of individuals with history h_t. We are interested in whether the government can change the future allocation intended for these individuals in such a way as to make them better off while still raising the same revenue from them.

Let $(\mathbf{x}_{h_t}, \mathbf{y}_{h_t})$ denote a future allocation for those individuals who at time t have history h_t; i.e.,

$$(\mathbf{x}_{h_t}, \mathbf{y}_{h_t}) = \{(x_{t+j}(\theta_{t+j}; h_{t+j}), y_{t+j}(\theta_{t+j}; h_{t+j})) \mid \forall\ h_{t+j} \succeq h_t\}_{j=0}^{\infty}.$$

The future allocation implied by the efficient allocation $(\mathbf{x}^*, \mathbf{y}^*)$ is denoted $(\mathbf{x}^*_{h_t}, \mathbf{y}^*_{h_t})$. Let $R^*(h_t)$ be the expected revenues raised from those individuals under $(\mathbf{x}^*_{h_t}, \mathbf{y}^*_{h_t})$; that is,

$$R^*(h_t) = \sum_{j=0}^{\infty} \delta^j E[y^*_{t+j}(\theta_{t+j}; h_{t+j}) - x^*_{t+j}(\theta_{t+j}; h_{t+j}) \mid h_t].$$

Finally, let $V^*(\theta \mid h_t)$ be the expected utility of a type θ after a history h_t under $(\mathbf{x}^*, \mathbf{y}^*)$.

---

[8] Our result also holds for a finite economy. This distinguishes it from Chari and Kehoe (1990) who show that in infinitely-lived economies Ramsey equilibria may be sustained by trigger strategies if the discount factor is high enough.

[9] This condition is satisfied by most common cost functions such as quadratic, logarithmic or exponential.

Now consider the problem:

$$\max_{(\mathbf{x}_{h_t},\mathbf{y}_{h_t})} V(\theta_H \mid h_t) \qquad\qquad (\mathcal{P}^I_{h_t})$$

$$s.t.\ V(\theta_L \mid h_t) \geq V^*(\theta_L \mid h_t)$$

$$\sum_{j=0}^{\infty} \delta^j E[y_{t+j}(\theta_{t+j}; h_{t+j}) - x_{t+j}(\theta_{t+j}; h_{t+j}) \mid h_t] \geq R^*(h_t)$$

$$and\ IC_H(h_{t+j})\ \&\ IC_L(h_{t+j})\ for\ all\ h_{t+j} \succeq h_t\ and\ j = 0,1,..$$

Thus, we seek to maximize the expected utility of those individuals with history h_t who are high types at time t, subject to the constraints that (i) those who are low types in period t with history h_t obtain at least as much utility as they obtain under $(\mathbf{x}^*_{h_t}, \mathbf{y}^*_{h_t})$; (ii) the same expected revenue is raised from these individuals as under $(\mathbf{x}^*_{h_t}, \mathbf{y}^*_{h_t})$; and (iii) the incentive compatibility constraints for these individuals in period t and beyond are satisfied. Clearly, $(\mathbf{x}^*_{h_t}, \mathbf{y}^*_{h_t})$ satisfies all the constraints of this problem. If $(\mathbf{x}^*_{h_t}, \mathbf{y}^*_{h_t})$ solves it, then the government cannot change the future allocation intended for individuals with history h_t in such a way as to make them better off while still raising the same revenue from them. Therefore, we say that $(\mathbf{x}^*, \mathbf{y}^*)$ is *time consistent* if for all periods $t \geq 2$ and all histories h_t, $(\mathbf{x}^*_{h_t}, \mathbf{y}^*_{h_t})$ is a solution to $\mathcal{P}^I_{h_t}$.[10]

We now have:

Lemma 7 *Let $(\mathbf{x}^*, \mathbf{y}^*)$ be a second best efficient allocation. Then, $(\mathbf{x}^*, \mathbf{y}^*)$ is time consistent if and only if*

$$\frac{\alpha_{LH}}{\alpha_{HH}} \geq \frac{\mu - \frac{1}{\gamma(1-\delta)}}{1 - \frac{1}{\gamma(1-\delta)}} \qquad\qquad (11)$$

where γ is the Lagrange multiplier associated with the maximization problem described in Lemma 4 that is solved by $\mathbf{y}^*$.

The intuition underlying this result is the following. When ability types are perfectly correlated, under the Ramsey tax system the government faces no residual uncertainty in period two and beyond. Because of this, it could impose lump sum taxes from that point on and eliminate all distortions in individuals' labor supply. Accordingly, Ramsey optimal taxes can never be time consistent. When types are stochastic, residual uncertainty remains because an individual may change type. Thus, the government must still screen types in the remaining periods. Condition

---

[10] This corresponds to the standard definition in the literature. See, for example, Kydland and Prescott (1977).

(11) guarantees that the ex post optimal distortion is the same as the ex ante optimal distortion. When this is the case, two competing forces offset each other. On the one hand, in order to create a Pareto improvement, the government must introduce a new tax system that involves less distortions than the original one. This necessitates increasing the earnings of those individuals whose earnings are distorted who, by Proposition 2, are those who currently are and always have been low types. On the other hand, increasing the earnings of these individuals requires compensating increases in consumption for those individuals with the same history who have become high types. When condition (11) is satisfied, these compensating increases in consumption are sufficient to offset the revenue gains from the higher earnings of those who are still low types and the net impact on revenue is negative.

From Lemma 7, we can derive our second main result:

Proposition 5 *Let* $(\mathbf{x}^*, \mathbf{y}^*)$ *be a second best efficient allocation. Then: (i) if* $\alpha_{LH} = 0$, $(\mathbf{x}^*, \mathbf{y}^*)$ *is not time consistent; (ii) if* $\alpha_{LH}/\alpha_{HH} \geq \mu$, $(\mathbf{x}^*, \mathbf{y}^*)$ *is time consistent; and (iii) if* $\alpha_{LH}/\alpha_{HH} \in (0, \mu)$ *there exists a threshold* Ω^* *such that* $(\mathbf{x}^*, \mathbf{y}^*)$ *is time consistent if and only if* $G + (1 - \delta)\underline{u} \leq \Omega^*$.

To understand how this result follows from the Lemma, recall that γ always exceeds $1/\mu(1-\delta)$ so that the right hand side of condition (11) is positive. This yields part (i). On the other hand, note that the right hand side of this inequality is increasing in γ and converges to μ as γ approaches ∞. Accordingly, condition (11) is necessarily satisfied when α_{LH}/α_{HH} exceeds μ which implies part (ii). In the intermediate case, whether the condition is satisfied depends upon the precise value of the Lagrange multiplier γ associated with $\mathbf{y}^*$. The smaller it is, the more likely is the condition to be satisfied. Since γ represents the marginal value of a unit relaxation in the government's revenue requirement, the degree to which it exceeds $1/\mu(1-\delta)$ will depend upon the tightness of the incentive constraints. This in turn will depend on the size of the revenue requirement G and on the amount of redistribution the government intends to do as measured by $\underline{u}$.

This proposition has two interesting implications. First, no matter how strong the correlation between types, if it is anything less than perfect, there are conditions under which the Ramsey optimal policy will be sustainable. This justifies our claim in the introduction that Pareto efficient income tax systems can be time-consistent even when the degree of correlation in ability types is large. Second, in the intermediate case in which $\alpha_{LH}/\alpha_{HH} \in (0, \mu)$, Pareto efficient tax systems will be time consistent only when the government's revenue requirement and its redistributive goals are "not too large". Thus, ceteris paribus, a government that starts with higher spending

commitments (for example, higher debt to repay) or more ambitious redistributive objectives will find it harder to implement second best optimal policies.

This second implication suggests a theoretical reason why the classic equity-efficiency trade off (see, for example, Okun (1975)) may be steeper than previously thought. A well-known lesson of public economics is that achieving stronger equity objectives requires more distortionary taxation which reduces the size of the aggregate pie. Indeed, the Mirrlees model is designed precisely to illustrate and quantify this trade off. Proposition 6 suggests that, in dynamic economies, stronger equity objectives might lead second best optimal policies to be time inconsistent. This will force governments to achieve their equity objectives with third best policies.[11]   These will lead to greater distortions and larger reductions in the aggregate pie than suggested by the Mirrlees model.

7   Conclusion

This paper has investigated Pareto efficient income taxation in an economy with infinitely-lived individuals whose income generating abilities evolve according to a two-state Markov process. The investigation has yielded two novel conclusions. First, when individuals are risk neutral, the fraction of individuals who face a positive marginal income tax rate converges to zero and, in addition, the tax rate these individuals face also goes to zero. Thus, in the long run, an efficient income tax system involves no distortions in labor supply and hence no excess burden. Second, Pareto efficient income tax systems can be time-consistent even when the degree of correlation in ability types is large. Moreover, time consistency is more likely when governments have less progressive policy agendas (i.e., lower spending and less redistribution). As we have argued, this provides a theoretical rationale for believing that the equity-efficiency trade off may be steeper than suggested by the static Mirrlees model.

It is well worth understanding how general these conclusions are. Our analysis of the two

---

[11] Understanding what these third best policies look like is a challenging problem because when the government cannot commit, the revelation principle does not hold. In a two period Principal-Agent model with variable types, Battaglini (2003b) fully characterizes the optimal renegotiation proof contract extending the revelation principle to this dynamic environment. He shows that when the second best optimal contract is not time consistent, the third best optimal contract involves the agent playing a mixed strategy. The optimal contract induces the high type agent to partially pool with the low type in the first period; and the degree of pooling monotonically increases with the level of types' persistence. Berliant and Ledyard (2003) also study third best policies in a two period optimal tax model with a continuum of constant ability types. They provide conditions under which the optimal tax scheme involves screening types with distortionary taxes in the first period and non-distortionary lump sum taxes in the second period. These second period lump sum taxes depend only upon first period earnings.

period model shows that risk neutrality is a necessary condition for our result about distortions. However, we also found that when individuals' risk aversion is small, the basic pattern of earnings under an efficient tax system is the same as in the risk neutral case. Thus, the broad features of efficient tax systems that we identify in our basic model are robust to introducing small amounts of risk aversion. There are many other directions that the model could usefully be extended. For example, it would be interesting to allow for more than two ability types. We leave this and other extensions for future work.

References

Battaglini, Marco, (2003a), "Long Term Contracting with Markovian Consumers," mimeo, Princeton University.

Battaglini, Marco, (2003b), "Optimality and Renegotiation in Dynamic Contracting," mimeo, Princeton University.

Berliant, Marcus and John Ledyard, (2003), "Optimal Dynamic Nonlinear Income Taxes with No Commitment," mimeo, California Institute of Technology.

Brito, Dagobert, Jonathan Hamilton, Steven Slutsky and Joseph Stiglitz, (1991), "Dynamic Optimal Taxation with Commitment," *Journal of Public Economics*, 44, 15-35.

Chari, V.V., Patrick Kehoe and Edward Prescott, (1989), "Time Consistency and Policy," in *Modern Business Cycle Theory*, eds. Robert Barro, Harvard University Press: Cambridge

Chari, V.V. and Patrick Kehoe, (1990), "Sustainable Plans," *Journal of Political Economy*, 98-4, 783-802.

Chari, V.V and Patrick Kehoe, (1999), "Optimal Fiscal and Monetary Policy," in *Handbook of Macroeconomics*, eds. John Taylor and Michael Woodford, Elsevier: New York.

Diamond, Peter, (1998), "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88, 83-95.

Diamond, Peter and James Mirrlees, (1978), "A Model of Social Insurance with Variable Retirement," *Journal of Public Economics*, 10, 295-336.

Golosov, Mikhail, Naryana Kocherlakota and Aleh Tsyvinsky, (2003), "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70, 569-588.

Golosov, Mikhail, and Aleh Tsyvinsky, (2003), "Designing Optimal Disability Insurance," mimeo, University of Minnesota.

Kocherlakota, Narayana, (2004), "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," mimeo, Federal Reserve Bank of Minneapolis.

Kydland, Finn and Edward Prescott, (1977), "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85-3, 473-491.

Mirrlees, James, (1971), "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175-208.

Okun, Arthur, (1975), *Equality and Efficiency: The Big Tradeoff*, The Brookings Institution: Washington DC.

Ramsey, Frank, (1927), "A Contribution to the Theory of Taxation," *Economic Journal*, 37, 47-61.

Roberts, Kevin, (1984), "The Theoretical Limits to Redistribution," *Review of Economics Studies*, LI, 177-195.

Rogerson, William, (1985), "Repeated Moral Hazard," *Econometrica*, 53, 69-76.

Stiglitz, Joseph, (1982), "Self-Selection and Pareto Efficient Taxation," *Journal of Public Economics*, 17, 213-240.

Stokey, Nancy, (1989), "Reputation and Time Consistency," *American Economic Review Papers and Proceedings*, 79-2, 134-139.

Werning, Ivan, (2002), *Optimal Dynamic Taxation and Social Insurance*, Doctoral Dissertation, University of Chicago.

8   Appendix

Proof of Lemma 1: It is obvious that U_L holds with equality, so we will just show that $IC_H\left(h_1\right)$ is binding. Let $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ solve the relaxed problem and suppose that $IC_H\left(h_1\right)$ is not binding. Without loss of generality, we can assume that for all $t > 1$ and histories h_t, the constraint $IC_H(h_t)$ is not binding. To see this, suppose that for some time period $\widehat{t}$ and some history $h_{\widehat{t}} = \{\theta_1, ..., \theta_{\widehat{t}-1}\}$, $IC_H(h_{\widehat{t}})$ were binding. Suppose first that $\theta_1 = \theta_H$. Then consider the allocation $(\mathbf{x}, \widehat{\mathbf{y}})$ in which

$$x_{\widehat{t}}(\theta_H; h_{\widehat{t}}) = \widehat{x}_{\widehat{t}}(\theta_H; h_{\widehat{t}}) + \varepsilon; \ x_{\widehat{t}}(\theta_L; h_{\widehat{t}}) = \widehat{x}_{\widehat{t}}(\theta_L; h_{\widehat{t}});$$

$$\text{and} \ \ x_1(\theta_H; h_1) = \widehat{x}_1(\theta_H; h_1) - \delta^{\widehat{t}-1} \varepsilon \Pr\left((h_{\widehat{t}}, \theta_H) \,|\theta_1 = \theta_H\right)$$

for $\varepsilon > 0$ and all the remaining consumptions are unchanged. Observe that the expected utility of a high type in period one under this allocation is exactly the same as under $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ because

$$E[\sum_{t=1}^{\infty} \delta^{t-1} x_t(\theta_t; h_t) \,|\theta_1 = \theta_H] = E[\sum_{t=1}^{\infty} \delta^{t-1} \widehat{x}_t(\theta_t; h_t) \,|\theta_1 = \theta_H].$$

It follows that $(\mathbf{x}, \widehat{\mathbf{y}})$ satisfies $IC_H(h_1)$ and yields the same value of the objective function as $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$. It is also satisfies R and U_L.

Next suppose that $\theta_1 = \theta_L$. Then consider the allocation $(\mathbf{x}, \widehat{\mathbf{y}})$ in which

$$x_{\widehat{t}}(\theta_H; h_{\widehat{t}}) = \widehat{x}_{\widehat{t}}(\theta_H; h_{\widehat{t}}) + \varepsilon; \ x_{\widehat{t}}(\theta_L; h_{\widehat{t}}) = \widehat{x}_{\widehat{t}}(\theta_L; h_{\widehat{t}});$$

$$\text{and} \ \ x_1(\theta_L; h_1) = \widehat{x}_1(\theta_L; h_1) - \delta^{\widehat{t}-1} \varepsilon \Pr\left((h_{\widehat{t}}, \theta_H) \,|\theta_1 = \theta_L\right)$$

for $\varepsilon > 0$ and all the remaining consumptions are unchanged. Observe that the expected utility of a low type in period one under this allocation is exactly the same as under $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ because

$$E[\sum_{t=1}^{\infty} \delta^{t-1} x_t(\theta_t; h_t) \,|\theta_1 = \theta_L] = E[\sum_{t=1}^{\infty} \delta^{t-1} \widehat{x}_t(\theta_t; h_t) \,|\theta_1 = \theta_L].$$

It follows that $(\mathbf{x}, \widehat{\mathbf{y}})$ satisfies U_L. It also yields the same value of the objective function as $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ and satisfies R. Moreover, since $IC_H\left(h_1\right)$ is not binding under $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ it will not be binding under $(\mathbf{x}, \widehat{\mathbf{y}})$ for ε sufficiently small.

It follows from this that $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ solves the problem

$$\max_{\{(x_t(\theta_t; h_t), y_t(\theta_t; h_t))\}_{t=1}^{\infty}} V(\theta_H \mid h_1)$$

$$s.t. \ U_L, \ R$$

This implies that $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ is simply an efficient allocation in which $V(\theta_L \mid h_1) = \underline{u}$. But, by assumption, $\underline{u}$ is sufficiently high that this fails to satisfy at least one of the high types' incentive constraints $IC_H(h_t)$. Q.E.D.

Proof of Lemma 2: Let $(\mathbf{x}, \mathbf{y})$ be an allocation satisfying the constraints of the relaxed problem. We will show that for all $t \geq 2$ we can find $\mathbf{x}^t$ such that the allocation $(\mathbf{x}^t, \mathbf{y})$ (i) satisfies all the constraints and yields the same value of the objective function as $(\mathbf{x}, \mathbf{y})$; (ii) satisfies $IC_H(h_\tau)$ with equality for all periods $\tau \in \{2, ..., t\}$ and all histories h_τ; and (iii) is identical to $(\mathbf{x}, \mathbf{y})$ for all periods $\tau > t$ and all histories h_τ. This implies the result.

We prove our claim by induction. Consider the claim for $t = 2$. Suppose that $IC_H(h_2)$ is not binding after some history h_2. Suppose first that $h_2 = \{\theta_L\}$, so that the high type was a low type in period 1. Since $IC_H(h_2)$ is not binding, there must exist some $\varepsilon > 0$ such that:

$$V(\theta_H \mid h_2) = x_2(\theta_L; h_2) - \varphi(y_2(\theta_L; h_2)/\theta_H) + \delta E_{\theta_H} V(\theta_3 \mid \{h_2, \theta_L\}) + \varepsilon$$

Now let $\mathbf{x}^2$ satisfy

$$
\begin{aligned}
x_2^2(\theta_H; h_2) &= x_2(\theta_H; h_2) - \varepsilon; \ x_2^2(\theta_L; h_2) = x_2(\theta_L; h_2) \\
x_1^2(\theta_L; h_1) &= x_1(\theta_L; h_1) + \delta\alpha_{LH}\varepsilon; \ x_1^2(\theta_H; h_1) = x_1(\theta_H; h_1)
\end{aligned}
$$

and otherwise equals $\mathbf{x}$. Thus, all we have done is to take consumption away from the high type after history h_2 and give the expected discounted value to the low type in period one. Clearly, this does not effect the value of the objective function. Nor does it effect the (R) or (U_L) constraints. It satisfies the $IC_H(h_2)$ constraint with equality by construction. We need to check that the $IC_H(h_1)$ constraint is satisfied; i.e., that:

$$V^2(\theta_H \mid h_1) \geq x_1^2(\theta_L; h_1) - \varphi(y_1(\theta_L; h_1)/\theta_H) + \delta E_{\theta_H} V^2(\theta_2 \mid \{\theta_L\})$$

where $V^2(\theta_t \mid h_t)$ is the value function corresponding to the allocation $(\mathbf{x}^2, \mathbf{y})$. We have that:

$$
\begin{aligned}
V^2(\theta_H \mid h_1) &= x_1^2(\theta_H; h_1) - \varphi(y_1(\theta_H; h_1)/\theta_H) + \delta E_{\theta_H} V^2(\theta_2 \mid \{\theta_H\}). \\
&= V(\theta_H \mid h_1) \\
&\geq x_1(\theta_L; h_1) - \varphi(y_1(\theta_L; h_1)/\theta_H) + \delta E_{\theta_H} V(\theta_2 \mid \{\theta_L\}). \\
&= x_1^2(\theta_L; h_1) - \delta\varepsilon\alpha_{LH} - \varphi(y_1(\theta_L; h_1)/\theta_H) + \delta E_{\theta_H} V^2(\theta_2 \mid \{\theta_L\}) + \delta\varepsilon\alpha_{HH} \\
&\geq x_1^2(\theta_L; h_1) - \varphi(y_1(\theta_L; h_1)/\theta_H) + \delta E_{\theta_H} V^2(\theta_2 \mid \{\theta_L\})
\end{aligned}
$$

30

where the third equality follows from the fact that $(\mathbf{x}, \mathbf{y})$ satisfies $IC_H(h_1)$ and the fourth follows from the fact that $\alpha_{HH} \geq \alpha_{LH}$.

Next suppose that $h_2 = \{\theta_H\}$ so that the high type was also a high type in period 1. Again, there must exist some $\varepsilon > 0$ such that:

$$V(\theta_H \,|h_2) = x_2(\theta_L; h_2) - \varphi(y_2(\theta_L; h_2)/\theta_H) + \delta E_{\theta_H} V(\theta_3 \,|\{h_2, \theta_L\}) + \varepsilon$$

Again, we will show that we can find an alternative allocation that yields at least the same value of the objective function, satisfies all the constraints of the relaxed problem and has the property that $IC_H(h_2)$ is binding. Now let $\mathbf{x}^2$ be defined by:

$$
\begin{aligned}
x_2^2(\theta_H; h_2) &= x_2(\theta_H; h_2) - \varepsilon; \, x_2^2(\theta_L; h_2) = x_2(\theta_L; h_2) \\
x_1^2(\theta_H; h_1) &= x_1(\theta_H; h_1) + \delta\alpha_{LH}\varepsilon; \, x_1^2(\theta_L; h_1) = x_1(\theta_L; h_1)
\end{aligned}
$$

and equals $\mathbf{x}$ otherwise. Thus, all we have done is to take consumption away from the high type after history h_2 and give the expected discounted value to the high type in period 1. Clearly, this does not effect the value of the objective function. Nor does it effect the R or U_L constraints. It satisfies the $IC_H(h_2)$ constraint with equality by construction and has no effect on the $IC_H(h_1)$ constraint.

Now suppose that the claim is true for $\tau = 2, ..., t - 1$ and consider the claim for t. Since the claim is true for $t - 1$, we can find $\mathbf{x}^{t-1}$ such that the allocation $(\mathbf{x}^{t-1}, \mathbf{y})$ satisfies all the constraints and yields the same value of the objective function as $(\mathbf{x}, \mathbf{y})$; (ii) satisfies $IC_H(h_\tau)$ with equality for all periods $\tau \in \{2, ..., t - 1\}$ and all histories h_τ; and (iii) is identical to $(\mathbf{x}, \mathbf{y})$ for all periods $\tau > t - 1$ and all histories h_τ. If $(\mathbf{x}^{t-1}, \mathbf{y})$ is such that $IC_H(h_t)$ is binding for all histories h_t then we can simply let $\mathbf{x}^t = \mathbf{x}^{t-1}$.

If this is not the case, there must exist some history h_t such that $IC_H(h_t)$ is not binding. Again, there are two possibilities, $h_t = \{h_{t-1}, \theta_L\}$ and $h_t = \{h_{t-1}, \theta_H\}$. In either case, in the same manner as above, we can find $\widetilde{\mathbf{x}}$ such that the allocation $(\widetilde{\mathbf{x}}, \mathbf{y})$ (i) yields the same value of the objective function as $(\mathbf{x}^{t-1}, \mathbf{y})$ (and hence $(\mathbf{x}, \mathbf{y})$); (ii) satisfies $IC_H(h_t)$ with equality; and (iii) equals $(\mathbf{x}^{t-1}, \mathbf{y})$ (and hence $(\mathbf{x}, \mathbf{y})$) for all periods $\tau > t$ and all histories h_τ. If $h_t = \{h_{t-1}, \theta_H\}$ then $\widetilde{\mathbf{x}}$ will also satisfy $IC_H(h_\tau)$ with equality for all periods $\tau \in \{2, ..., t-1\}$ so we can let $\mathbf{x}^t = \widetilde{\mathbf{x}}$.

If $h_t = \{h_{t-1}, \theta_L\}$ and $\alpha_{HH} > \alpha_{LH}$, then $IC_H(h_{t-1})$ will hold strictly. However, in this case, since the claim is true for $\tau = t - 1$ we can find $\widehat{\mathbf{x}}$ such that the allocation $(\widehat{\mathbf{x}}, \mathbf{y})$ (i) satisfies all

the constraints and yields the same value of the objective function as $(\tilde{\mathbf{x}}, \mathbf{y})$; (ii) satisfies $IC_H(h_\tau)$ with equality for all periods $\tau \in \{2, ..., t-1\}$ and all histories h_τ; and (iii) is identical to $(\tilde{\mathbf{x}}, \mathbf{y})$ for all periods $\tau > t-1$ and all histories h_τ. We can then let $\mathbf{x}^t = \hat{\mathbf{x}}$. *Q.E.D.*

Proof of Lemma 3: From (1) we have that:

$$V\left(\theta_H \left| h_t\right.\right) = V\left(\theta_L \left| h_t\right.\right) + \Phi\left(y_t(\theta_L; h_t)\right) + \Delta EV(\theta_{t+1} | \{h_t, \theta_L\}).$$

We have that

$$\Delta EV(\theta_{t+1} | \{h_t, \theta_L\}) = \delta\left(\alpha_{HH} - \alpha_{LH}\right) V(\theta_H | \{h_t, \theta_L\}) + \delta\left(\alpha_{HL} - \alpha_{LL}\right) V(\theta_L | \{h_t, \theta_L\})$$

where $V(\theta | \{h_t, \theta_L\})$ is the expected lifetime utility of an individual who has type θ in period $t+1$ after history $\{h_t, \theta_L\}$. But by (1) we have that:

$$
\begin{aligned}
V(\theta_H | \{h_t, \theta_L\}) &= V(\theta_L | \{h_t, \theta_L\}) + \Phi\left(y_{t+1}(\theta_L; \{h_t, \theta_L\})\right) + \Delta EV(\theta_{t+2} | \{h_t, \theta_L, \theta_L\}) \\
&= V(\theta_L | \{h_t, \theta_L\}) + \Phi\left(y_{t+1}(\theta_L; \{h_t, \theta_L\})\right) + \delta\left(\alpha_{HH} - \alpha_{LH}\right) V(\theta_H | \{h_t, \theta_L, \theta_L\}) \\
&\quad + \delta\left(\alpha_{HL} - \alpha_{LL}\right) V(\theta_L | \{h_t, \theta_L, \theta_L\}).
\end{aligned}
$$

Thus, we have that

$$
\begin{aligned}
\Delta EV(\theta_{t+1} | \{h_t, \theta_L\}) &= \delta\left(\alpha_{HH} - \alpha_{LH}\right) \Phi\left(y_{t+1}(\theta_L; \{h_t, \theta_L\})\right) \\
&\quad + \delta^2\left(\alpha_{HH} - \alpha_{LH}\right)^2 V(\theta_H | \{h_t, \theta_L, \theta_L\}) \\
&\quad + \delta^2\left(\alpha_{HH} - \alpha_{LH}\right)\left(\alpha_{HL} - \alpha_{LL}\right) V(\theta_L | \{h_t, \theta_L, \theta_L\})
\end{aligned}
$$

But again from (1) we have that

$$
\begin{aligned}
V(\theta_H | \{h_t, \theta_L, \theta_L\}) &= V(\theta_L | \{h_t, \theta_L, \theta_L\}) + \Phi\left(y_{t+2}(\theta_L; \{h_t, \theta_L, \theta_L\})\right) \\
&\quad + \delta\left(\alpha_{HH} - \alpha_{LH}\right) V(\theta_H | \{h_t, \theta_L, \theta_L, \theta_L\}) + \delta\left(\alpha_{HL} - \alpha_{LL}\right) V(\theta_L | \{h_t, \theta_L, \theta_L, \theta_L\}).
\end{aligned}
$$

So that

$$
\begin{aligned}
\Delta EV(\theta_{t+1} | \{h_t, \theta_L\}) &= \delta\left(\alpha_{HH} - \alpha_{LH}\right) \Phi\left(y_{t+1}(\theta_L; \{h_t, \theta_L\})\right) \\
&\quad + \delta^2\left(\alpha_{HH} - \alpha_{LH}\right)^2 \Phi\left(y_{t+2}(\theta_L; \{h_t, \theta_L, \theta_L\})\right) \\
&\quad + \delta^3\left(\alpha_{HH} - \alpha_{LH}\right)^3 V(\theta_H | \{h_t, \theta_L, \theta_L, \theta_L\}) \\
&\quad + \delta^3\left(\alpha_{HH} - \alpha_{LH}\right)^2\left(\alpha_{HL} - \alpha_{LL}\right) V(\theta_L | \{h_t, \theta_L, \theta_L, \theta_L\}).
\end{aligned}
$$

Repeated application of this argument yields

$$\Delta EV(\theta_{t+1}\,|\{h_t,\theta_L\}) = \sum_{j=1}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{t+j}(\theta_L; h_{t+j}^\circ)\right).$$

and we have the claimed expression. *Q.E.D.*

Proof of Lemma 4: Let $(\mathbf{x}, \mathbf{y})$ solve the relaxed problem. Then by Lemmas 1 and 2 we may assume with no loss of generality that $\mathbf{x}$ is such that $(\mathbf{x}, \mathbf{y})$ satisfies U_L and all the incentive constraints with equality. Thus, by Lemma 3 we can write the value of the objective function as

$$V(\theta_H\,|h_1) = \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{1+j}(\theta_L; h_{1+j}^\circ)\right) + \underline{u} \tag{12}$$

The resource constraint can be written as $G \leq (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - x_t(\theta_t; h_t)]$. By definition, we know that

$$V(\theta_H\,|h_1) = \sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t)\,|\theta_1 = \theta_H],$$

and

$$V(\theta_L\,|h_1) = \sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t)\,|\theta_1 = \theta_L].$$

Thus,

$$
\begin{aligned}
\sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t)] &= \mu \sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t)\,|\theta_1 = \theta_H] + (1-\mu)\sum_{t=1}^{\infty} \delta^{t-1} E[x_t(\theta_t; h_t)\,|\theta_1 = \theta_L] \\
&= \mu[V(\theta_H\,|h_1) + \sum_{t=1}^{\infty} \delta^{t-1} E[\varphi(y_t(\theta_t; h_t)/\theta_t)\,|\theta_1 = \theta_H] \\
&\quad + (1-\mu)[V(\theta_L\,|h_1) + \sum_{t=1}^{\infty} \delta^{t-1} E[\varphi(y_t(\theta_t; h_t)/\theta_t)\,|\theta_1 = \theta_L] \\
&= \mu V(\theta_H\,|h_1) + (1-\mu)V(\theta_L\,|h_1) + \sum_{t=1}^{\infty} \delta^{t-1} E[\varphi(y_t(\theta_t; h_t)/\theta_t)].
\end{aligned}
$$

Substituting this into the resource constraint, yields,

$$G \leq (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - \mu V(\theta_H\,|\,h_1) - (1-\mu)V(\theta_L\,|\,h_1) - \varphi(y_t(\theta_t; h_t)/\theta_t)].$$

Using Lemma 3 and U_L we can write this as

$$
\begin{aligned}
G \;\leq\; & (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t)] \tag{13} \\
& -(1-\delta)[\mu \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{1+j}(\theta_L; h_{1+j}^\circ)\right) + \underline{u}]
\end{aligned}
$$

Thus, it follows that the earnings levels $\mathbf{y}$ must maximize the objective function (12) subject to the constraint (13). $Q.E.D.$

Proof of Lemma 5: We know that $(\mathbf{x}, \mathbf{y})$ is a solution to Problem $\mathcal{P}^{II}$. To show that it solves Problem $\mathcal{P}^{I}$, all we need show is that the low type's incentive constraint $IC_L(h_t)$ is satisfied for all t and h_t. For a given period t and history h_t, this requires showing that

$$x_t(\theta_L; h_t) - \varphi(y_t(\theta_L; h_t)/\theta_L) + \delta E_{\theta_L} V(\theta_{t+1} | \{h_t, \theta_L\})$$
$$\geq \quad x_t(\theta_H; h_t) - \varphi(y_t(\theta_H; h_t)/\theta_L) + \delta E_{\theta_L} V(\theta_{t+1} | \{h_t, \theta_H\})$$

or, equivalently, that

$$x_t(\theta_L; h_t) - x_t(\theta_H; h_t) \quad \geq \quad \varphi(y_t(\theta_L; h_t)/\theta_L) - \varphi(y_t(\theta_H; h_t)/\theta_L) \tag{14}$$
$$+\delta[E_{\theta_L} V(\theta_{t+1} | \{h_t, \theta_H\}) - \delta E_{\theta_L} V(\theta_{t+1} | \{h_t, \theta_L\})]$$

From the fact that $IC_H(h_t)$ holds with equality, we have that

$$x_t(\theta_L; h_t) - x_t(\theta_H; h_t) \quad = \quad \varphi(y_t(\theta_L; h_t)/\theta_H) - \varphi(y_t(\theta_H; h_t)/\theta_H) \tag{15}$$
$$+\delta E_{\theta_H} [V(\theta_{t+1} | \{h_t, \theta_H\}) - E_{\theta_H} V(\theta_{t+1} | \{h_t, \theta_L\})])$$

We can use this to prove the desired inequality. Note first that

$$[E_{\theta_H} V(\theta_{t+1} | \{h_t, \theta_H\}) - E_{\theta_H} V(\theta_{t+1} | \{h_t, \theta_L\})] \geq [E_{\theta_L} V(\theta_{t+1} | \{h_t, \theta_H\}) - E_{\theta_L} V(\theta_{t+1} | \{h_t, \theta_L\})].$$

To see this, note that

$$[E_{\theta_H} V(\theta_{t+1} | \{h_t, \theta_H\}) - E_{\theta_H} V(\theta_{t+1} | \{h_t, \theta_L\})] - [E_{\theta_L} V(\theta_{t+1} | \{h_t, \theta_H\}) - E_{\theta_L} V(\theta_{t+1} | \{h_t, \theta_L\})]$$
$$= \quad (\alpha_{HH} - \alpha_{LH}) [V(\theta_H | \{h_t, \theta_H\}) - V(\theta_H | \{h_t, \theta_L\})]$$
$$+(\alpha_{HL} - \alpha_{LL})[V(\theta_L | \{h_t, \theta_H\})) - V(\theta_L | \{h_t, \theta_L\})]$$

Using Lemma 3, this difference equals

$$(\alpha_{HH} - \alpha_{LH}) \sum_{j=0}^{\infty} \delta^j [\alpha_{HH} - \alpha_{LH}]^j (\Phi (y_{t+1+j}(\theta_L; \{h_t, \theta_H\}_{+j}^{\circ})) - \Phi (y_{t+1+j}(\theta_L; \{h_t, \theta_L\}_{+j}^{\circ}))).$$

Observe now that from the first order conditions that for all j

$$\varphi'(y_{t+1+j}(\theta_L; \{h_t, \theta_H\}_{+j}^{\circ})/\theta_L) = \theta_L,$$

while (since $\gamma\mu\,(1-\delta) > 1$)

$$\varphi'(y_{t+1+j}(\theta_L; \{h_t, \theta_L\}^\circ_{+j})/\theta_L) \le \theta_L.$$

Thus, for all j, $y_{t+1+j}(\theta_L; \{h_t, \theta_H\}^\circ_{+j}) \ge y_{t+1+j}(\theta_L; \{h_t, \theta_L\}^\circ_{+j})$. Since $\Phi' \ge 0$, it follows that the above difference is non-negative.

It is also the case that

$$\varphi(y_t(\theta_L; h_t)/\theta_H) - \varphi(y_t(\theta_H; h_t)/\theta_H) \ge \varphi(y_t(\theta_L; h_t)/\theta_L) - \varphi(y_t(\theta_H; h_t)/\theta_L).$$

To see this, note that the first order conditions imply that $y_t(\theta_L; h_t) < y_t(\theta_H; h_t)$. In addition, it is the case that

$$\frac{\varphi'(y/\theta_H)}{\theta_H} - \frac{\varphi'(y/\theta_L)}{\theta_L} < 0.$$

It follows from these two claims and from (14) and (15) that $IC_L(h_t)$ is satisfied.

Conversely, let $(\mathbf{x}, \mathbf{y})$ be a solution to Problem P^I. We need to show that the earnings $\mathbf{y}$ solve problem (3). Suppose not. Then $(\mathbf{x}, \mathbf{y})$ cannot solve the relaxed problem. Let $(\mathbf{x}', \mathbf{y}')$ be a solution to the relaxed problem with $\mathbf{y}' \ne \mathbf{y}$. Then by Lemma 4, we know that $\mathbf{y}'$ solves problem (3). Moreover, we can assume by Lemmas 1 and 2 without loss of generality that $\mathbf{x}'$ is such that $(\mathbf{x}', \mathbf{y}')$ satisfies $IC_H(h_t)$ with equality for all h_t and that U_L binds. But then it follows by the above argument that $(\mathbf{x}', \mathbf{y}')$ satisfies $IC_L(h_t)$ for all t and h_t. This is a contradiction. Q.E.D.

Proof of Lemma 6: Let $(\mathbf{x}, \mathbf{y})$ solve the relaxed problem. It suffices to show that there exists a $\overline{\beta} < 1$ such that if $\beta \in (\overline{\beta}, 1)$ then the eliminated constraints IC_L, $IC_L(H)$ and $IC_L(L)$ are satisfied. To do this, we must first establish some properties of the solution to the relaxed problem.

The Lagrangian for the relaxed problem is

$$
\begin{aligned}
\mathcal{L} \;=\; & (x_H)^\beta - \varphi(y_H/\theta_H) + \delta[\alpha_{HH}((x_{HH})^\beta - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}((x_{HL})^\beta - \varphi(y_{HL}/\theta_L))] \quad (16) \\
& + \lambda_U\{(x_L)^\beta - \varphi(y_L/\theta_L) + \delta[\alpha_{LH}((x_{LH})^\beta - \varphi(y_{LH}/\theta_H)) + \alpha_{LL}((x_{LL})^\beta - \varphi(y_{LL}/\theta_L))]\} \\
& + \lambda_H\{(x_H)^\beta - \varphi(y_H/\theta_H) + \delta[\alpha_{HH}((x_{HH})^\beta - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}((x_{HL})^\beta - \varphi(y_{HL}/\theta_L))] \\
& - (x_L)^\beta + \varphi(y_L/\theta_H) - \delta[\alpha_{HH}((x_{LH})^\beta - \varphi(y_{LH}/\theta_H)) + \alpha_{HL}((x_{LL})^\beta - \varphi(y_{LL}/\theta_L))]\} \\
& + \lambda_{HH}\{(x_{HH})^\beta - \varphi(y_{HH}/\theta_H) - (x_{HL})^\beta + \varphi(y_{HL}/\theta_H)\} \\
& + \lambda_{LH}\{(x_{LH})^\beta - \varphi(y_{LH}/\theta_H) - (x_{LL})^\beta + \varphi(y_{LL}/\theta_H)\} \\
& + \lambda_R\{\mu[(y_H - x_H) + \delta(\alpha_{HH}(y_{HH} - x_{HH}) + \alpha_{HL}(y_{HL} - x_{HL}))] \\
& + (1 - \mu)[(y_L - x_L) + \delta(\alpha_{LH}(y_{LH} - x_{LH}) + \alpha_{LL}(y_{LL} - x_{LL}))]\}.
\end{aligned}
$$

The first order conditions for the high type's consumptions imply that

$$\beta(x_H)^{\beta-1} = \frac{\lambda_R \mu}{1 + \lambda_H}, \tag{17}$$

$$\beta(x_{HH})^{\beta-1} = \frac{\lambda_R \mu}{1 + \lambda_H + \lambda_{HH}/\delta\alpha_{HH}}, \tag{18}$$

and

$$\beta(x_{HL})^{\beta-1} = \frac{\lambda_R \mu}{1 + \lambda_H - \lambda_{HH}/\delta\alpha_{HL}}.$$

Those for the low type's consumptions imply that:

$$\beta(x_L)^{\beta-1} = \frac{\lambda_R(1-\mu)}{\lambda_U - \lambda_H}, \tag{19}$$

$$\beta(x_{LL})^{\beta-1} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL}}, \tag{20}$$

and

$$\beta(x_{LH})^{\beta-1} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H + \lambda_{LH}/\delta\alpha_{LH}}. \tag{21}$$

With respect to earnings levels, the first order conditions for the high type's earnings imply:

$$\frac{\varphi'(y_H/\theta_H)}{\theta_H} = \frac{\lambda_R \mu}{1 + \lambda_H},$$

$$\frac{\varphi'(y_{HH}/\theta_H)}{\theta_H} = \frac{\lambda_R \mu}{1 + \lambda_H + \lambda_{HH}/\delta\alpha_{HH}},$$

and

$$\frac{\varphi'(y_{HL}/\theta_L)}{\theta_L} = \frac{\lambda_R \mu - \frac{\varphi'(y_{HL}/\theta_H)}{\theta_H}\frac{\lambda_{HH}}{\delta\alpha_{HL}}}{(1 + \lambda_H)}. \tag{22}$$

Those for the low type imply that:

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} = \frac{\frac{\varphi'(y_L/\theta_H)}{\theta_H}\lambda_H + \lambda_R(1-\mu)}{\lambda_U}, \tag{23}$$

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} = \frac{\frac{\lambda_{LH}}{\delta\alpha_{LL}}\frac{\varphi'(y_{LL}/\theta_H)}{\theta_H} + \lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H}, \tag{24}$$

36

and

$$\frac{\varphi'(y_{LH}/\theta_H)}{\theta_H} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H + \lambda_{LH}/\delta\alpha_{LH}}. \tag{25}$$

We can use these conditions to establish some facts about the solution to the relaxed problem.

Claim 1: $\lambda_H > 0$ implies $\lambda_{HH} > 0$ and $\lambda_{LH} > 0$.

We first show that $\lambda_{HH} > 0$. Suppose, to the contrary, that $\lambda_{HH} = 0$. Then it follows from the first order conditions for the high type's consumptions that $x_{HL} = x_H = x_{HH}$. But from the conditions describing the high type's earnings levels if $\lambda_{HH} = 0$, then, since $\lambda_H > 0$, we have that:

$$\frac{\varphi'(y_{HL}/\theta_L)}{\theta_L} = \frac{\varphi'(y_{HH}/\theta_H)}{\theta_H}$$

which implies that $y_{HL}/\theta_L < y_{HH}/\theta_H$. But then, since $x_{HH} = x_{HL}$, it is clear that $IC(HH)$ would be violated.

We now show that $\lambda_{LH} > 0$. Again, suppose to the contrary, that $\lambda_{LH} = 0$. Then it follows from the first order conditions for the low type's consumptions and the fact that $\frac{\alpha_{HH}}{\alpha_{LH}} \geq \frac{\alpha_{HL}}{\alpha_{LL}}$ that

$$\alpha(x_{LL})^{\alpha-1} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H} \leq \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H} = \alpha(x_{LH})^{\alpha-1}.$$

It follows that $x_{LH} \leq x_{LL}$. But from our analysis of earnings levels, if $\lambda_{LH} = 0$ then

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H} \leq \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H} = \frac{\varphi'(y_{LH}/\theta_H)}{\theta_H}$$

which implies that $y_{LL}/\theta_L < y_{LH}/\theta_H$. But then it is clear that $IC(LH)$ would be violated. ∎

Claim 2: $\lambda_H > 0$ implies that $y_{HH} \geq y_{HL}$ and $y_{LH} \geq y_{LL}$

We first show that $y_{HH} \geq y_{HL}$. From the previous claim, we know that $\lambda_{HH} > 0$ and hence that

$$(x_{HH})^{\beta} - \varphi(y_{HH}/\theta_H) = (x_{HL})^{\beta} - \varphi(y_{HL}/\theta_H).$$

But the first order conditions for the high type's consumptions together with the fact that $\lambda_{HH} > 0$ imply that $x_{HL} < x_{HH}$. Thus, y_{HH} must exceed y_{HL}.

We next show that $y_{LH} \geq y_{LL}$. From the previous claim, we know that $\lambda_{LH} > 0$ and hence that

$$(x_{LH})^{\beta} - \varphi(y_{LH}/\theta_H) = (x_{LL})^{\beta} - \varphi(y_{LL}/\theta_H).$$

It follows that if $y_{LH} < y_{LL}$ we must have that $x_{LL} > x_{LH}$. From the first order conditions for y_{LH} and x_{LH} we know that

$$\beta(x_{LH})^{\beta-1} = \frac{\varphi'(y_{LH}/\theta_H)}{\theta_H}.$$

From the first order conditions for y_{LL}, we know that

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L}(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \frac{\lambda_{LH}}{\delta\alpha_{LL}}) + \frac{\lambda_{LH}}{\delta\alpha_{LL}}\{\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} - \frac{\varphi'(y_{LL}/\theta_H)}{\theta_H}\} = \lambda_R(1-\mu).$$

Thus,

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} + \frac{\lambda_{LH}/\delta\alpha_{LL}}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL}}\{\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} - \frac{\varphi'(y_{LL}/\theta_H)}{\theta_H}\} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL}}.$$

Since

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} > \frac{\varphi'(y_{LL}/\theta_H)}{\theta_H},$$

it follows from the first order condition for x_{LL} that

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} < \beta(x_{LL})^{\beta-1}.$$

Thus, if $x_{LL} > x_{LH}$ we have that

$$\frac{\varphi'(y_{LH}/\theta_H)}{\theta_H} = \beta(x_{LH})^{\beta-1} > \beta(x_{LL})^{\beta-1} > \frac{\varphi'(y_{LL}/\theta_L)}{\theta_L}.$$

But this inequality implies that $y_{LH} > y_{LL}$ - a contradiction. ∎

Claim 3: There exists a $\varepsilon > 0$ and a $\beta_1 < 1$ such that for all $\beta \in (\beta_1, 1)$, $\lambda_H \geq \varepsilon$.

Observe first that there must be a $\kappa > 0$ and a $\beta_0 < 1$ such that for any $\beta \in (\beta_0, 1)$, $max\{\lambda_H, \lambda_{HH}, \lambda_{LH}\} > \kappa$. If this were not to be true, then the solution of the problem would be arbitrarily near to the efficient solution for some β close to 1. But, by assumption, the second best Pareto frontier is strictly below the first best frontier at $\beta = 1$ and, by the Theorem of the Maximum, the value function of the maximization of a continuous function with continuous constraint correspondence (as (16)) is continuous: therefore we would have a contradiction.

Now suppose, contrary to the claim, that for any $\varepsilon > 0$ and any $\beta_1 < 1$ there is a $\beta \in [\beta_1, 1)$ such that $\lambda_H < \varepsilon$. Then, since $\frac{\lambda_R \mu}{1 - \lambda_{HH}/\delta\alpha_{HL}} \to 1$ as $\beta \to 1$, there must be a sequence β^n converging to one, such that the multiplier λ_{HH} corresponding to β^n is less than κ. We can therefore assume that the multiplier λ_{LH} corresponding to β^n exceeds κ. But then we have:

$$\begin{aligned}
\beta^n(x_{LL})^{\beta^n-1} &= \frac{\lambda_R(1-\mu)}{\lambda_U - \lambda_{LH}/\delta\alpha_{LL}} > \frac{\lambda_R(1-\mu)}{\lambda_U - \kappa/\delta\alpha_{LL}} \\
&= \left(\frac{\lambda_U}{\lambda_U - \kappa/\delta\alpha_{LL}}\right)\frac{\lambda_R(1-\mu)}{\lambda_U}.
\end{aligned}$$

We know that, as $\beta^n \to 1$, $\frac{\lambda_R(1-\mu)}{\lambda_U}$ converges to one and λ_U converges to $\frac{1-\mu}{\mu}$. Thus, we can find an $\overline{n}$ and some $\epsilon > 0$, such that for $n > \overline{n}$,

$$\beta^n (x_{LL})^{\beta^n - 1} > 1 + \epsilon.$$

But this contradicts the fact that $\beta^n \to 1$. ∎

Claim 4: There exists a $\beta_2 \in (\beta_1, 1)$ such that for all $\beta \in (\beta_2, 1)$, $y_H \geq y_L$.

From (23) we can write:

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} \left(1 - \frac{\lambda_H}{\lambda_U}\right) + \frac{\lambda_H}{\lambda_U} \left[\frac{\varphi'(y_L/\theta_L)}{\theta_L} - \frac{\varphi'(y_L/\theta_H)}{\theta_H}\right] = \frac{\lambda_R (1 - \mu)}{\lambda_U}$$

It can be verified from (23) that $y_L > 0$ for any β. By Claim 3 it follows that there exists some $\varepsilon > 0$ such that

$$\lambda_H \left[\frac{\varphi'(y_L/\theta_L)}{\theta_L} - \frac{\varphi'(y_L/\theta_H)}{\theta_H}\right] > \varepsilon$$

for any $\beta \in (\beta_1, 1)$. Thus, for any $\beta \in (\beta_1, 1)$

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} \left(1 - \frac{\lambda_H}{\lambda_U}\right) + \frac{\varepsilon}{\lambda_U} < \frac{\lambda_R (1 - \mu)}{\lambda_U}$$

which implies that

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} < \frac{\lambda_R (1 - \mu)}{\lambda_U - \lambda_H} - \varepsilon$$

But we know that

$$\lim_{\beta \to 1} \frac{\lambda_R (1 - \mu)}{\lambda_U - \lambda_H} = \lim_{\beta \to 1} \beta x_L^{\beta - 1} \simeq 1 = \frac{\varphi'(y_H/\theta_H)}{\theta_H}.$$

Therefore, it must be the case that for sufficiently large β

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} < \frac{\varphi'(y_H/\theta_H)}{\theta_H},$$

which implies that $y_H \geq y_L$. This implies the result. ∎

Claim 5: There exists some $\beta_3 < 1$ such that if $\beta \in (\beta_3, 1)$, then $y_{HL} \geq y_{LL}$.

The first order conditions for the high type's consumptions imply that $\lim_{\beta \to 1} \lambda_{HH} = 0$. From the first order conditions for x_{HL} and y_{HL} we have that

$$\lim_{\beta \to 1} \frac{\varphi'(y_{HL}/\theta_L)}{\theta_L} = \lim_{\beta \to 1} \frac{\lambda_R \mu - \frac{\varphi'(y_{HL}/\theta_H)}{\theta_H} \frac{\lambda_{HH}}{\delta \alpha_{HL}}}{(1 + \lambda_H)} = 1. \tag{26}$$

We now prove that $\lim_{\beta \to 1} \frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} < \lim_{\beta \to 1} \beta(x_{LL})^{\beta-1} = 1$. From the first order condition (24) we have:

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} = \frac{\lambda_{LH}}{\delta \alpha_{LL} \left(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}} \lambda_H \right)} \frac{\varphi'(y_{LL}/\theta_H)}{\theta_H} + \frac{\lambda_R (1 - \mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}} \lambda_H}$$

But from the first order condition (20) we also know that:

$$\frac{\lambda_R (1 - \mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}} \lambda_H} = \beta (x_{LL})^{\beta-1} \left[1 - \frac{\lambda_{LH}}{\delta \alpha_{LL} \left(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}} \lambda_H \right)} \right]$$

So:

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} - 1 = \beta (x_{LL})^{\beta-1} - 1 + \frac{\lambda_{LH}}{\delta \alpha_{LL} \left(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}} \lambda_H \right)} \left[\frac{\varphi'(y_{LL}/\theta_H)}{\theta_H} - \beta (x_{LL})^{\beta-1} \right]$$

$$(27)$$

It is easy to verify that λ_{LH} must be strictly positive for any β in $[0,1]$ (otherwise (20) and (21) can not be satisfied simultaneously); and, as shown above, λ_U and λ_H are also strictly positive. If $\lim_{\beta \to 1} \left(\frac{\varphi'(y_{LL}/\theta_H)}{\theta_H} - 1 \right) < 0$, therefore, (27) implies that $\lim_{\beta \to 1} \frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} - 1 < 0$. If $\lim_{\beta \to 1} \left(\frac{\varphi'(y_{LL}/\theta_H)}{\theta_H} - 1 \right) \geq 0$, then we have:

$$\begin{aligned}
\frac{\varphi'(\lim_{\beta \to 1} y_{LL}/\theta_L)}{\theta_L} - 1 &= \left(\lim_{\beta \to 1} \frac{\lambda_{LH}}{\delta \alpha_{LL} \left(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}} \lambda_H \right)} \right) \left[\frac{\varphi'(\lim_{\beta \to 1} y_{LL}/\theta_H)}{\theta_H} - 1 \right] \\
&< \frac{\varphi'(\lim_{\beta \to 1} y_{LL}/\theta_L)}{\theta_L} - 1
\end{aligned}$$

since, from (20) and (21), $\lim_{\beta \to 1} \frac{\lambda_{LH}}{\delta \alpha_{LL} \left(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}} \lambda_H \right)} < 1$. But this is obviously impossible. We conclude that

$$\lim_{\beta \to 1} \frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} < 1. \tag{28}$$

We now show that there exists a $\beta_3 < 1$ such that for any $\beta \in (\beta_3, 1)$ any solution of the program implies:

$$\frac{\varphi'(y_{HL}/\theta_L)}{\theta_L} > \frac{\varphi'(y_{LL}/\theta_L)}{\theta_L}$$

Assume not. Then we can construct a sequence $\langle \beta_n \rangle_{n=1}^{\infty}$ converging to one such that for any element of the sequence we have

$$\frac{\varphi'(y_{HL}^n/\theta_L)}{\theta_L} \leq \frac{\varphi'(y_{LL}^n/\theta_L)}{\theta_L}$$

where y_{HL}^n and y_{HL}^n are solutions corresponding to β_n. Let $y'_{HL} = \lim_{\beta_n \to 1} y_{HL}^n$ and $y'_{LL} = \lim_{\beta_n \to 1} y_{LL}^n$ be limits of the solution along this sequence (it is immediate to verify that the limits exist without loss of generality). Then, by the continuity of φ', it must be that $\frac{\varphi'(y'_{HL}/\theta_L)}{\theta_L} \leq \frac{\varphi'(y'_{LL}/\theta_L)}{\theta_L}$. This implies that y'_{HL} and y'_{LL} are not solutions of the program, otherwise we would violate (28) and (26). But it is impossible that the limit of solutions is not a solution since, by the Theorem of the Maximum, the solution of the program is an upper-hemicontinuous correspondence. ∎

Now let $\overline{\beta} = \max\{\beta_2, \beta_3\} < 1$. Then we claim that if $\beta \in (\overline{\beta}, 1)$ then $IC(L)$, $IC(HL)$ and $IC(LL)$ are satisfied. The last two follow from the standard static argument given that (by Claims 1 and 3) $IC(HH)$ and $IC(LH)$ hold with equality and (by Claims 2 and 3) $y_{HH} \geq y_{HL}$ and $y_{LH} \geq y_{LL}$. For $IC(L)$ to be satisfied we need that

$$(x_L)^\beta - \varphi(y_L/\theta_L) + \delta[\alpha_{LH}((x_{LH})^\beta - \varphi(y_{LH}/\theta_H)) + \alpha_{LL}((x_{LL})^\beta - \varphi(y_{LL}/\theta_L))]$$
$$\geq (x_H)^\beta - \varphi(y_H/\theta_L) + \delta[\alpha_{LH}((x_{HH})^\beta - \varphi(y_{HH}/\theta_H)) + \alpha_{LL}((x_{HL})^\beta - \varphi(y_{HL}/\theta_L))]$$

or equivalently that

$$(x_L)^\beta - (x_H)^\beta \geq \varphi(y_L/\theta_L) - \varphi(y_H/\theta_L) + \delta[\alpha_{LH}((x_{HH})^\beta - \varphi(y_{HH}/\theta_H)) + \alpha_{LL}((x_{HL})^\beta - \varphi(y_{HL}/\theta_L))]$$
$$- \delta[\alpha_{LH}((x_{LH})^\beta - \varphi(y_{LH}/\theta_H)) + \alpha_{LL}((x_{LL})^\beta - \varphi(y_{LL}/\theta_L))]$$
$$= \varphi(y_L/\theta_L) - \varphi(y_H/\theta_L) + \delta[\alpha_{LH}(U_{HH} - U_{LH}) + \alpha_{LL}(U_{HL} - U_{LL})]$$

where $U_{HH} = (x_{HH})^\beta - \varphi(y_{HH}/\theta_H)$, etc.

By Claim 1, we know that $IC(H)$ is satisfied with equality in the solution to the relaxed problem, so that

$$(x_H)^\beta - \varphi(y_H/\theta_H) + \delta[\alpha_{HH}(U_{HH}) + \alpha_{HL}(U_{HL})] = (x_L)^\beta - \varphi(y_L/\theta_H) + \delta[\alpha_{HH}(U_{LH}) + \alpha_{HL}(U_{LL})]$$

implying that

$$(x_L)^\beta - (x_H)^\beta = \varphi(y_L/\theta_H) - \varphi(y_H/\theta_H) + \delta[\alpha_{HH}(U_{HH} - U_{LH}) + \alpha_{HL}(U_{HL} - U_{LL})].$$

Thus, all we need to show is that

$$\Phi(y_H) + \delta[\alpha_{HH}(U_{HH} - U_{LH}) + \alpha_{HL}(U_{HL} - U_{LL})]$$
$$\geq \Phi(y_L) + \delta[\alpha_{LH}(U_{HH} - U_{LH}) + \alpha_{LL}(U_{HL} - U_{LL})]$$

By Claim 4 we know that $y_L \leq y_H$ which, since $\Phi'(y) > 0$, implies that $\Phi(y_H) \geq \Phi(y_L)$. Thus it suffices to show that

$$(\alpha_{HH} - \alpha_{LH})(U_{HH} - U_{LH}) + (\alpha_{HL} - \alpha_{LL})(U_{HL} - U_{LL}) \geq 0.$$

But since

$$
\begin{aligned}
U_{HH} - U_{LH} &= (x_{HH})^\beta - \varphi(y_{HH}/\theta_H) - ((x_{LH})^\beta - \varphi(y_{LH}/\theta_H)) \\
&= (x_{HL})^\beta - \varphi(y_{HL}/\theta_H) - ((x_{LL})^\beta - \varphi(y_{LL}/\theta_H)) \\
&= U_{HL} - U_{LL} + \Phi(y_{HL}) - \Phi(y_{LL})
\end{aligned}
$$

this requires that

$$(\alpha_{HH} - \alpha_{LH})(U_{HL} + \Phi(y_{HL}) - U_{LL} - \Phi(y_{LL})) + (\alpha_{HL} - \alpha_{LL})(U_{HL} - U_{LL}) \geq 0$$

or, equivalently that $(\alpha_{HH} - \alpha_{LH})(\Phi(y_{HL}) - \Phi(y_{LL})) \geq 0$. This follows from the fact that, by Claim 5, $y_{HL} \geq y_{LL}$.

Now suppose that $(\mathbf{x}, \mathbf{y})$ is a second best efficient allocation. Suppose that it did not solve the relaxed problem. Then there would exist another allocation $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ that yields a higher value of the objective function and satisfies $U_L, R, IC(H), IC(LH)$ and $IC(HH)$ all with equality. It must be the case therefore that it does not satisfy one or more of the eliminated constraints $IC(L)$, $IC(HL)$, and $IC(LL)$. But for $\beta \in (\overline{\beta}, 1)$ that cannot be true as we have just shown. *Q.E.D.*

Proof of Proposition 3: It follows from the first order conditions for the high types consumptions and earnings derived in the proof of the previous Lemma and the fact that $\lambda_{HH} > 0$ that y_H and y_{HH} are set efficiently, while y_{HL} is distorted downwards. It is also clear from the first order conditions that y_{LH} is set efficiently. To prove that y_L is distorted downwards, we need to show that

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} = \frac{\frac{\varphi'(y_L/\theta_H)}{\theta_H}\lambda_H + \lambda_R(1-\mu)}{\lambda_U} < \beta(x_L)^{\beta-1} = \frac{\lambda_R(1-\mu)}{\lambda_U - \lambda_H}$$

From the condition that y_L satisfies, we know that

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L}(\lambda_U - \lambda_H) + \lambda_H\{\frac{\varphi'(y_L/\theta_L)}{\theta_L} - \frac{\varphi'(y_L/\theta_H)}{\theta_H}\} = \lambda_R(1-\mu)$$

Thus,

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} + \frac{\lambda_H}{(\lambda_U - \lambda_H)}\{\frac{\varphi'(y_L/\theta_L)}{\theta_L} - \frac{\varphi'(y_L/\theta_H)}{\theta_H}\} = \frac{\lambda_R(1-\mu)}{(\lambda_U - \lambda_H)}$$

and we have that
$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} > \frac{\varphi'(y_L/\theta_H)}{\theta_H},$$
which yields the result since $\lambda_U - \lambda_H > 0$. To prove that y_{LL} is distorted downwards, we need to show that
$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} = \frac{\frac{\lambda_{LH}}{\delta\alpha_{LL}}\frac{\varphi'(y_{LL}/\theta_H)}{\theta_H} + \lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H} < \beta(x_{LL})^{\beta-1} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL}}$$

From the condition that y_{LL} satisfies, we know that
$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L}(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \frac{\lambda_{LH}}{\delta\alpha_{LL}}) + \frac{\lambda_{LH}}{\delta\alpha_{LL}}\{\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} - \frac{\varphi'(y_{LL}/\theta_H)}{\theta_H}\} = \lambda_R(1-\mu)$$

Thus,
$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} + \frac{\lambda_{LH}/\delta\alpha_{LL}}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL}}\{\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} - \frac{\varphi'(y_{LL}/\theta_H)}{\theta_H}\} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL}}$$

and we have that
$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} > \frac{\varphi'(y_{LL}/\theta_H)}{\theta_H},$$
which yields the result since $\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL} > 0$.

We next show that the degree of distortion in the earnings of those who becomes low types in the second period converges to 0 as $\beta \to 1$. As noted in the proof of Claim 5 of Lemma 6, the first order conditions for the high type's consumptions imply that $\lim_{\beta\to1}\lambda_{HH} = 0$. From the first order conditions for x_{HL} and y_{HL} we have that
$$\lim_{\beta\to1}\frac{\varphi'(y_{HL}/\theta_L)}{\theta_L} = \lim_{\beta\to1}\frac{\lambda_R\mu - \frac{\varphi'(y_{HL}/\theta_H)}{\theta_H}\frac{\lambda_{HH}}{\delta\alpha_{HL}}}{(1+\lambda_H)} = 1 = \lim_{\beta\to1}\beta(x_{HL})^{\beta-1}.$$

Finally, we show that $y_{LL} > y_L$. From the first order conditions for the low type's earnings, we know that
$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L}(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H) - \frac{\lambda_{LH}}{\delta\alpha_{LL}}\frac{\varphi'(y_{LL}/\theta_H)}{\theta_H} = \lambda_R(1-\mu)$$

and that
$$\frac{\varphi'(y_L/\theta_L)}{\theta_L}\lambda_U - \frac{\varphi'(y_L/\theta_H)}{\theta_H}\lambda_H = \lambda_R(1-\mu).$$

It will be shown in the next proposition that $x_L > x_{LL}$. This implies from the first order conditions for x_L and x_{LL} that $[\alpha_{LL} - \alpha_{HL}]\lambda_H\delta < \lambda_{LH}$. Thus,
$$\begin{aligned}\lambda_R(1-\mu) &= \frac{\varphi'(y_{LL}/\theta_L)}{\theta_L}(\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H) - \frac{\lambda_{LH}}{\delta\alpha_{LL}}\frac{\varphi'(y_{LL}/\theta_H)}{\theta_H} \\ &< \frac{\varphi'(y_{LL}/\theta_L)}{\theta_L}(\lambda_U - \lambda_H) + \left(1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right)\lambda_H\Phi'(y_{LL}/\theta_H)\end{aligned}$$

43

Since, by (19), as $\beta \to 1$ we have that $(\lambda_U - \lambda_H) \to \lambda_R(1-\mu)$, we can write:

$$1 - \frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} < \left(1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right) \frac{\lambda_H}{\lambda_R(1-\mu)} \Phi'(y_{LL}/\theta_H)$$

Consider now (23), again (19) implies

$$
\begin{aligned}
1 - \frac{\varphi'(y_L/\theta_L)}{\theta_L} &\simeq \frac{\lambda_H}{\lambda_R(1-\mu)} \Phi'(y_{LL}/\theta_H) \\
&> \left(1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right) \frac{\lambda_H}{\lambda_R(1-\mu)} \Phi'(y_L/\theta_H)
\end{aligned}
$$

Therefore:

$$\frac{\varphi'(y_{LL}/\theta_L)}{\theta_L} > \frac{\varphi'(y_L/\theta_L)}{\theta_L}$$

and the result follows by the convexity of φ. Q.E.D.

Proof of Proposition 4: For the first statement we need to show that $x_H \in (x_{HL}, x_{HH})$ and $x_L \in (x_{LL}, x_{LH})$. The first claim follows immediately from the first order conditions for the high types consumption and the fact that (as shown in the proof of Lemma 6) λ_{HH} is positive. For the second claim, note first that since (as shown in the proof of Lemma 6) $y_{LL} < y_{LH}$ the incentive constraint $IC(LH)$ implies that $x_{LL} < x_{LH}$. Thus, if $x_L \notin (x_{LL}, x_{LH})$, then either it is the case that $x_L \leq x_{LL} < x_{LH}$ or it is the case that $x_{LL} < x_{LH} \leq x_L$.

Suppose the former. Then, from the first order conditions for x_L and x_{LL},

$$\frac{\lambda_R(1-\mu)}{\lambda_U - \lambda_H} \geq \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL}}.$$

This implies that $[\alpha_{LL} - \alpha_{HL}]\lambda_H\delta \geq \lambda_{LH}$. But this means that

$$\frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H - \lambda_{LH}/\delta\alpha_{LH} \geq \frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H - \lambda_H[\alpha_{LL} - \alpha_{HL}]/\alpha_{LH} = \lambda_H$$

and hence that

$$\frac{\lambda_R(1-\mu)}{\lambda_U - (\frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H - \lambda_{LH}/\delta\alpha_{LH})} \geq \frac{\lambda_R(1-\mu)}{\lambda_U - \lambda_H}.$$

From the first order conditions for x_L and x_{LH} this implies that $\beta(x_{LH})^{\beta-1} \geq \beta(x_L)^{\beta-1}$ which means that $x_{LH} \leq x_L$ - a contradiction.

Suppose then that $x_{LL} < x_{LH} \leq x_L$. From the first order conditions for x_L and x_{LH},

$$\frac{\lambda_R(1-\mu)}{\lambda_U - \lambda_H} \leq \frac{\lambda_R(1-\mu)}{\lambda_U - (\frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H - \lambda_{LH}/\delta\alpha_{LH})}.$$

This implies that $\lambda_{LH} \leq \delta(\alpha_{HH} - \alpha_{LH})\lambda_H$. But this means that

$$\frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H + \lambda_{LH}/\delta\alpha_{LL} \leq \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H + (\alpha_{HH} - \alpha_{LH})\lambda_H/\alpha_{LL} = \lambda_H$$

and hence that

$$\frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL}} \leq \frac{\lambda_R(1-\mu)}{\lambda_U - \lambda_H}.$$

From the first order conditions for x_L and x_{LL} this implies that $\beta(x_{LL})^{\beta-1} \leq \beta(x_L)^{\beta-1}$ which means that $x_{LL} \geq x_L$ - a contradiction.

For the second statement, we need to show that for $K \in \{L, H\}$

$$\beta(x_K)^{\beta-1} < \alpha_{KH}\beta(x_{KH})^{\beta-1} + \alpha_{KL}\beta(x_{KL})^{\beta-1}.$$

Define $\upsilon(x) = (x)^\beta$ and $\upsilon(x_i) = v_i$ for $i = K, KL, KH$. Consider a decrease in v_K by ϕ (which can be positive or negative) and a contextual increase of v_{KL} and v_{KH} by $\frac{\phi}{\delta}$. After this change the utility maintenance constraint and the incentive compatibility constraints at $t = 1$ and 2 are obviously satisfied, since utilities at $t = 2$ change by the same amounts and the net present value of the expected utility of reporting K at $t = 1$ is unchanged. It must be that this change does not relax the resources constraint, therefore:

$$\frac{\partial}{\partial\phi}\left[\upsilon^{-1}(v_K - \phi) + \delta\left(\alpha_{KH}\upsilon^{-1}\left(v_{KL} + \frac{\phi}{\delta}\right) + \alpha_{KH}\upsilon^{-1}\left(v_{KH} + \frac{\phi}{\delta}\right)\right)\right] = 0 \qquad (29)$$

where $\upsilon^{-1}(\cdot)$ is the inverse of υ. By Jensen's inequality, we have:

$$0 = \left(\frac{\alpha_{KH}}{\upsilon'(v_{KH})} + \frac{\alpha_{KH}}{\upsilon'(v_{KL})}\right) - \frac{1}{\upsilon'(v_K)} > \left(\frac{1}{\alpha_{KH}\upsilon'(v_{KH}) + \alpha_{KH}\upsilon'(v_{KL})}\right) - \frac{1}{\upsilon'(v_K)}$$

which implies $\beta(x_K)^{\beta-1} < \alpha_{KH}\beta(x_{KH})^{\beta-1} + \alpha_{LL}\beta(x_{KL})^{\beta-1}$. Q.E.D.

Proof of Lemma 7: Consider a particular period $t \geq 2$ and some history h_t. We are interested in knowing when $(\mathbf{x}_{h_t}^*, \mathbf{y}_{h_t}^*)$ will be a solution to $\mathcal{P}_{h_t}^I$. This is clearly the case if $h_t \neq h_t^* = \{\theta_L, ...\theta_L\}$ since Proposition 2 tells us that $(\mathbf{x}_{h_t}^*, \mathbf{y}_{h_t}^*)$ is first best efficient. Therefore we focus attention on the history h_t^*.

Observe that the program $\mathcal{P}_{h_t^*}^I$ is identical to $\mathcal{P}^I$, but for two exceptions. On the one hand, the reservation value of those with history h_t^* who are low types at time t is their expected continuation value $V^*(\theta_L \mid h_t^*)$ under $(\mathbf{x}^*, \mathbf{y}^*)$ instead of $\underline{u}$. On the other hand, the revenue requirement is not $G/(1-\delta)$, but the expected revenue generated from individuals with history h_t^* by $(\mathbf{x}^*, \mathbf{y}^*)$. We

45

will exploit this similarity to solve the program $\mathcal{P}^I_{h^*_t}$ in the same way as we did the program $\mathcal{P}^I$. However, a certain amount of work is necessary to show that the equivalent of Lemma 1 holds for the relaxed program corresponding to $\mathcal{P}^I_{h^*_t}$.

To this end, consider first the following revenue maximization problem:

$$\max_{(\mathbf{x}_{h^*_t}, \mathbf{y}_{h^*_t})} \sum_{j=0}^{\infty} \delta^j E[y_{t+j}(\theta_{t+j}; h_{t+j}) - x_{t+j}(\theta_{t+j}; h_{t+j}) \,|h^*_t]$$

$$s.t. \; V(\theta_L \mid h^*_t) \geq V^*(\theta_L \mid h^*_t)$$

$$and \; IC_H(h_{t+j}) \; \& \; IC_L(h_{t+j}) \; \forall \; h_{t+j} \succeq h^*_t \; \; \forall \; j = 0, 1, ..$$

Thus, we maximize the expected present value of revenues that can be extracted from individuals with history h^*_t at time t subject to the constraint that those with low ability at time t have at least as much utility as under $(\mathbf{x}^*_{h^*_t}, \mathbf{y}^*_{h^*_t})$ and the incentive constraints. Let $(\mathbf{x}^R_{h^*_t}, \mathbf{y}^R_{h^*_t})$ denote the solution to the revenue maximizing problem. We can immediately apply Lemmata 1-5 to this problem and conclude that the earnings path $\mathbf{y}^R_{h^*_t}$ solves the problem

$$\max_{\mathbf{y}_{h^*_t}} (1 - \delta) \sum_{j=0}^{\infty} \delta^j E[y_{t+j}(\theta_{t+j}; h_{t+j}) - \varphi(y_{t+j}(\theta_{t+j}; h_{t+j})/\theta_{t+j}) \,|h^*_t]$$

$$-(1 - \delta)[\Pr(\theta_t = \theta_H \,|h^*_t) \sum_{j=0}^{\infty} \delta^j \,[\alpha_{HH} - \alpha_{LH}]^j \, \Phi \left(y_{t+j}(\theta_L; h^*_{t+j})\right) + V^*(\theta_L \mid h^*_t)]$$

We can now establish:

Lemma 8 *If (11) holds, then for any h^*_{t+j}, $y^*_{t+j}(\theta_L; h^*_{t+j}) > y^R_{t+j}(\theta_L; h^*_{t+j})$.*

Proof: Since $\Pr(\theta_t = \theta_H \,|h^*_t) = \alpha_{LH}$, for any history h^*_{t+j}, $y^R_{t+j}(\theta_L; h^*_{t+j})$ satisfies the first order condition

$$\frac{\alpha_{LL}}{\alpha_{LH}}[1 - \frac{\varphi'(y^R_{t+j}(\theta_L; h^*_{t+j})/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^j \Phi' \left(y^R_{t+j}(\theta_L; h^*_{t+j})\right). \tag{30}$$

Under our assumption that $\varphi''' \geq 0$ the revenues are a strictly concave function of each y_{t+j}, implying that revenues are decreasing in $y_{t+j}(\theta_L; h^*_{t+j})$ on the interval $[y^R_{t+j}(\theta_L; h^*_{t+j}), \infty)$. From (6) we have that $y^*_{t+j}\left(\theta_L; h^*_{t+j}\right)$ solves:

$$\left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t} \frac{\gamma \, (1 - \delta) \, (1 - \mu)}{\gamma \mu (1 - \delta) - 1}[1 - \frac{\varphi'(y^*_{t+j}(\theta_L; h^*_{t+j})/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^j \Phi' \left(y^*_{t+j}(\theta_L; h^*_{t+j})\right).$$

$$\tag{31}$$

Using (11):

$$\left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t} \frac{\gamma\,(1-\delta)\,(1-\mu)}{\gamma\mu(1-\delta)-1}$$

$$\geq \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{-1} \frac{\gamma\,(1-\delta)\,(1-\mu)}{\gamma\mu(1-\delta)-1} > \frac{\alpha_{LL}}{\alpha_{LH}}$$

So the right hand side of (31) is larger than the right hand side of (30), and concavity of the revenue function implies that $y^*_{t+j}(\theta_L; h^*_{t+j}) > y^R_{t+j}(\theta_L; h^*_{t+j})$.   Q.E.D.

We can now show that the equivalent of Lemma 1 holds for the relaxed program corresponding to $\mathcal{P}^I_{h^*_t}$.

Lemma 9 *Let* $(\mathbf{x}_{h^*_t}, \mathbf{y}_{h^*_t})$ *solve the relaxed problem corresponding to* $\mathcal{P}^I_{h^*_t}$ *in which the incentive compatibility constraints for the low types are ignored. Then* $IC_H(h^*_t)$ *holds with equality.*

Proof: Assume, by contradiction, that $IC_H(h^*_t)$ is not binding. Following the same argument as in Lemma 1, it follows that $(\mathbf{x}_{h^*_t}, \mathbf{y}_{h^*_t})$ must be efficient starting from h^*_t. Therefore, using Lemma 7 and Proposition 2

$$y_{t+j}(\theta_L; h^*_{t+j}) > y^*_{t+j}(\theta_L; h^*_{t+j}) > y^R_{t+j}(\theta_L; h^*_{t+j})$$

for any $j \geq 0$, while for all histories $h_{t+j} \neq h^*_{t+j}$

$$y_{t+j}(\theta_L; h^*_{t+j}) = y^*_{t+j}(\theta_L; h^*_{t+j}) = y^R_{t+j}(\theta_L; h^*_{t+j}).$$

Since revenues are strictly decreasing on the interval $[y^R_{t+j}(\theta_L; h^*_{t+j}), \infty)$, it follows that the tax revenues generated by $(\mathbf{x}_{h^*_t}, \mathbf{y}_{h^*_t})$ must be strictly lower than the revenues generated by the ex ante optimal solution $(\mathbf{x}^*_{h_t}, \mathbf{y}^*_{h_t})$ starting from h^*_t: but this is a contradiction because then the revenues constraint would be violated. *Q.E.D.*

Given this last lemma, we can apply Lemmata 2-5 and conclude that $(\mathbf{x}^*_{h^*_t}, \mathbf{y}^*_{h^*_t})$ will be a solution to $\mathcal{P}^I_{h^*_t}$ if and only if $\mathbf{y}^*_{h^*_t}$ is a solution to the problem

$$\max_{\mathbf{y}_{h_t}} \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{t+j}(\theta_L; h^*_{t+j})\right) + V^*(\theta_L \mid h^*_t) \qquad (\mathcal{P}^S_{h^*_t})$$

$$s.t. \quad R^*(h^*_t)(1-\delta) \leq (1-\delta)\sum_{j=0}^{\infty} \delta^j E[y_{t+j}(\theta_{t+j}; h_{t+j}) - \varphi(y_{t+j}(\theta_{t+j}; h_{t+j})/\theta_{t+j}) \mid h^*_t]$$

$$(32)$$

$$-(1-\delta)[\Pr(\theta_t = \theta_H \mid h^*_t)\sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{t+j}(\theta_L; h^*_{t+j})\right) + V^*(\theta_L \mid h^*_t)]$$

47

Accordingly, to prove the result we need to show that $\mathbf{y}_{h_t^*}^*$ is a solution to the problem if and only if (11) holds.

Since $\Pr(\theta_t = \theta_H \,|h_t^*) = \alpha_{LH}$, the Lagrangian of $\mathcal{P}_{h_t^*}^S$ is:

$$
\mathcal{L}_{h_t^*} \;=\; \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{t+j}(\theta_L; h_{t+j}^*)\right) \tag{33}
$$

$$
+ \frac{\left[\gamma_S \left(1 - \delta\right)\right]}{\left(1 - \gamma_S \alpha_{LH}(1 - \delta)\right)} \sum_{j=0}^{\infty} \delta^j E[y_{t+j}(\theta_{t+j}; h_{t+j}) - \varphi(y_{t+j}(\theta_{t+j}; h_{t+j})/\theta_{t+j}) \,|h_t^*]
$$

where we have divided through by $1 - \gamma_S \alpha_{LH}(1 - \delta)$ and omitted the constants $V^*(\theta_L \mid h_t^*)$ and $R^*(h_t^*)$. We denote the Lagrange multiplier γ_S to distinguish it from the analogous multiplier γ for the program solved by $\mathbf{y}^*$. Let $\mathbf{y}_{h_t^*}^S$ denote the solution to this program. Under our assumption that $\varphi''' \geq 0$ the Lagrangian is a strictly concave function of each y_{t+j} and the solution is unique.

We can now prove the Proposition:

Sufficient condition: If (11) is satisfied, then $\mathbf{y}_{h_t^*}^*$ is a solution to problem $\mathcal{P}_{h_t^*}^S$. We proceed in three steps.

Step 1. We first show that

$$
\frac{\gamma_S \left(1 - \delta\right) \alpha_{LL}}{\gamma_S \alpha_{LH}(1 - \delta) - 1} \leq \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t} \frac{\gamma \left(1 - \delta\right)\left(1 - \mu\right)}{\gamma \mu (1 - \delta) - 1} \tag{34}
$$

Assume, by contradiction, that this inequality is not true. The optimal solution $y_{t+j}^S \left(\theta_L; h_{t+j}^*\right)$ satisfies the first order condition:

$$
\frac{\gamma_S \left(1 - \delta\right) \alpha_{LL}}{\gamma_S \alpha_{LH}(1 - \delta) - 1}[1 - \frac{\varphi'(y_{t+j}^S(\theta_L; h_{t+j}^*)/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^j \Phi'\left(y_{t+j}^S(\theta_L; h_{t+j}^*)\right) \tag{35}
$$

if $h_{t+j} \neq h_{t+j}^*$; and it would be fully efficient otherwise. If (34) is not true, then after any history h_{t+j}^*, concavity of the Lagrangian implies that the solution $y_{t+j}^S(\theta_L; h_{t+j}^*)$ of (35) is strictly larger than the ex ante optimal solution $y_{t+j}^*(\theta_L; h_{t+j}^*)$. Moreover, from Lemma 7 we know that that the ex ante optimal solution $y_{t+j}^*(\theta_L; h_{t+j}^*)$ is larger than the revenue maximizing solution $y_{t+j}^R(\theta_L; h_{t+j}^*)$. Accordingly,

$$
y_{t+j}^S(\theta_L; h_{t+j}^*) > y_{t+j}^*(\theta_L; h_{t+j}^*) > y_{t+j}^R(\theta_L; h_{t+j}^*)
$$

Since, as proven in Lemma 7, tax revenues are strictly decreasing on the interval $[y_{t+j}^R(\theta_L; h_{t+j}^*), \infty)$, this would imply that the revenues corresponding to the earnings path $\mathbf{y}_{h_t^*}^S$ are strictly lower than

under the ex ante optimal solution $\mathbf{y}^*_{h^*_t}$: it follows that the revenue constraint (32) is not satisfied - a contradiction.

Step 2. Next we show that:

$$\frac{\gamma_S(1-\delta)\alpha_{LL}}{\gamma_S\alpha_{LH}(1-\delta)-1} \geq \left[1-\frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t}\frac{\gamma(1-\delta)(1-\mu)}{\gamma\mu(1-\delta)-1} \tag{36}$$

Assume, by contraction, that (36) is not true. We know by the analogue of Lemma 4 that if $\mathbf{y}^S_{h^*_t}$ solves problem $\mathcal{P}^S_{h^*_t}$ and the consumption levels $\mathbf{x}^S_{h^*_t}$ are such as to make $IC_H(h_{t+j})$ for all j and $h_{t+j} \succeq h^*_t$ and the low type's utility constraint hold with equality given $\mathbf{y}^S_{h^*_t}$, then $(\mathbf{x}^S_{h^*_t}, \mathbf{y}^S_{h^*_t})$ must solve problem $\mathcal{P}^I_{h^*_t}$. But if (36) is not true, then after any history h^*_{t+j} the solution $y^S_{t+j}(\theta_L; h^*_{t+j})$ of (35) is smaller and hence more distorted than the ex ante optimal solution $y^*_{t+j}(\theta_L; h^*_{t+j})$. Since the solution on any other history would be efficient both under $\mathbf{y}^*_{h^*_t}$ and $\mathbf{y}^S_{h^*_t}$, we would have that aggregate surplus is lower under $(\mathbf{x}^S_{h^*_t}, \mathbf{y}^S_{h^*_t})$ than under $(\mathbf{x}^*_{h^*_t}, \mathbf{y}^*_{h^*_t})$. But this is a contradiction since in this case it is impossible that all the constraints of program $\mathcal{P}^I_{h^*_t}$ are satisfied and its value is strictly larger than $V^*(\theta_H \mid h^*_t)$.

Step 3. From Steps 1 and 2 it follows that

$$\frac{\gamma_S(1-\delta)\alpha_{LL}}{\gamma_S\alpha_{LH}(1-\delta)-1} = \left[1-\frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t}\frac{\gamma(1-\delta)(1-\mu)}{\gamma\mu(1-\delta)-1}.$$

From (35) and (31) it follows that $\mathbf{y}^*_{h^*_t}$ solves $\mathcal{P}^S_{h^*_t}$ as claimed.

Necessary condition: If (11) is not satisfied, then $\mathbf{y}^*_{h^*_t}$ is not a solution to problem $\mathcal{P}^S_{h^*_t}$.

If (11) does not hold, then (31) and (30) imply that after any history h^*_{t+j}, $y^*_{t+j}(\theta_L; h^*_{t+j})$ is smaller than the revenue maximizing solution $y^R_{t+j}(\theta_L; h^*_{t+j})$. Let $\mathbf{y}_{h^*_t}$ be an earnings path such that for any history h^*_{t+j}, $y_{t+j}(\theta_L; h^*_{t+j}) \in (y^*_{t+j}(\theta_L; h^*_{t+j}), \min\{y^R_{t+j}(\theta_L; h^*_{t+j}), y^*(\theta_L)\})$ but otherwise equals $\mathbf{y}^*_{h^*_t}$. Then this earnings path raises strictly more revenue and yields a strictly higher level of the objective function than does $\mathbf{y}^*_{h^*_t}$. Accordingly, $\mathbf{y}^*_{h^*_t}$ is not a solution to problem $\mathcal{P}^S_{h^*_t}$. *Q.E.D.*

Proof of Proposition 6: Let $\Omega = G + (1-\delta)\underline{u}$. Then, if $(\mathbf{x}^*, \mathbf{y}^*)$ is a second best efficient

allocation, Lemma 5 tells us that $\mathbf{y}^* = \mathbf{y}^*(\Omega)$ where $\mathbf{y}^*(\Omega)$ solves the problem

$$\max \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi \left(y_{1+j}(\theta_L; h_{1+j}^\circ)\right) + \underline{u}$$

$$s.t. \quad \Omega \ \leq \ (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t)]$$

$$-(1-\delta)[\mu \sum_{j=0}^{\infty} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi \left(y_{1+j}(\theta_L; h_{1+j}^\circ)\right)].$$

Let $\gamma(\Omega)$ denote the associated Lagrange multiplier. Condition (11) of Proposition 5 implies that, when $\frac{\alpha_{LH}}{\alpha_{HH}} \in (0, \mu)$, there is a threshold $\overline{\gamma} > \frac{1}{\mu(1-\delta)}$ such that $(\mathbf{x}^*, \mathbf{y}^*)$ is time consistent if and only if $\gamma(\Omega) \leq \overline{\gamma}$. Let $\overline{\Omega}$ be the maximum value of Ω such that the above problem has a solution. Let $\underline{\Omega}$ be the largest value of Ω such that there exists an efficient allocation in which those who are low types in period one have expected utility $\underline{u}$ and none of the high types' incentive constraints are violated. To prove the Proposition, we will demonstrate that as Ω increases from $\underline{\Omega}$ to $\overline{\Omega}$, $\gamma(\Omega)$ increases from $\frac{1}{\mu(1-\delta)}$ to ∞.

Define the functions

$$\Lambda_1 (y) = \frac{\Phi'(y)}{1 - \frac{\varphi'(y/\theta_L)}{\theta_L}}$$

which is increasing in y; and

$$\Lambda_2 (t, \gamma) = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t} \frac{(1-\mu)}{1 - \frac{1}{\gamma(1-\delta)}}$$

which is decreasing in γ. We know from our characterization of $\mathbf{y}^*(\Omega)$ that, for any t:

$$\Lambda_1 (y_t^*(\theta_L; h_t^*; \Omega)) = \Lambda_2 (t, \gamma(\Omega)).$$

For all other histories, $y_t^*(\theta_L; h_t; \Omega)$ is efficient.

Now consider the pure revenue maximization problem:

$$\max_{(\mathbf{x}, \mathbf{y})} \sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - x_t(\theta_t; h_t)]$$

$$s.t. \ V(\theta_L \mid h_1) \geq \underline{u}$$

$$\text{and } IC_H(h_t) \ \& \ IC_L(h_t) \quad \text{for all } t \ \& \ h_t.$$

Let $(\mathbf{x}^R, \mathbf{y}^R)$ denote the solution to this problem. As in the proof of the previous proposition, we can immediately apply Lemmata 1-5 to this problem and conclude that the revenue maximizing

earnings path earnings path $\mathbf{y}^R$ solves the problem:

$$\max_{\mathbf{y}} (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} E[y_t(\theta_t; h_t) - \varphi(y_t(\theta_t; h_t)/\theta_t)]$$

$$-(1-\delta)[\mu \sum_{j=0}^{\infty} \delta^j [\alpha_{HH} - \alpha_{LH}]^j \Phi(y_{1+j}(\theta_L; h^\circ_{1+j})) + \underline{u}].$$

It is easy to see that:

$$\Lambda_1\left(y_t^R(\theta_L; h_t^*)\right) = \lim_{\lambda \to \infty} \Lambda_2(t, \lambda) = 1 - \mu,$$

while for all other histories $y_t^R(\theta_L; h_t)$ is efficient. Note also that $\mathbf{y}^R$ is completely independent of Ω.

Now consider two values of Ω, $\widetilde{\Omega}, \Omega' \in (\underline{\Omega}, \overline{\Omega})$ such that $\widetilde{\Omega} > \Omega'$. We claim that $\gamma(\widetilde{\Omega}) > \gamma(\Omega')$. Suppose, to the contrary, that $\gamma(\widetilde{\Omega}) \le \gamma(\Omega')$. Then, we have that along the history h_t^* for any time period t,

$$
\begin{aligned}
\left\| y_t^R(\theta_L; h_t^*) - y_t^*(\theta_L; h_t^*; \widetilde{\Omega}) \right\| &= \left\| \Lambda_1^{-1}\Lambda_2\left(t, \gamma(\widetilde{\Omega})\right) - \Lambda_1^{-1}(1-\mu) \right\| \\
&\ge \left\| \Lambda_1^{-1}\Lambda_2\left(t, \gamma(\Omega')\right) - \Lambda_1^{-1}(1-\mu) \right\| \\
&= \left\| y_t^R(\theta_L; h_t^*) - y_t^*(\theta_L; h_t^*; \Omega') \right\|
\end{aligned}
$$

so that the difference between the revenue maximizing income level $y_t^R(\theta_L; h_t^*)$ and the constrained efficient income $y_t^*(\theta_L; h_t^*; \widetilde{\Omega})$ is not smaller than the difference between $y_t^R(\theta_L; h_t^*)$ and $y_t^*(\theta_L; h_t^*; \Omega')$. Since these differences must have the same sign and since revenues are concave in y_t, it follows that $\mathbf{y}^*(\widetilde{\Omega})$ cannot generate more revenues than $\mathbf{y}^*(\Omega')$. This is a contradiction since $\widetilde{\Omega} > \Omega'$.

To see that as $\Omega \to \underline{\Omega}$, $\gamma(\Omega) \to \frac{1}{\mu(1-\delta)}$ note that as $\Omega \to \underline{\Omega}$ the incentive compatibility constraints for the high types become non-binding, so taxation becomes efficient. This implies that $\gamma(\Omega) \to \frac{1}{\mu(1-\delta)}$. Similarly as Ω converges to $\overline{\Omega}$, $\gamma(\Omega)$ must converge to infinity, otherwise some resources would be left to the high type and tax revenues would not be maximized. $Q.E.D.$