This revision October 2006

Pareto Efficient Income Taxation with Stochastic Abilities*

Abstract

This paper studies Pareto efficient income taxation in an economy with finitely-lived individuals whose income generating abilities evolve according to a two-state Markov process. The study yields three main results. First, when individuals are risk neutral, the fraction of individuals who face a positive marginal income tax rate is always positive but decreases over time, converging to zero if the time horizon is long enough. Moreover, the tax rate these individuals face also goes to zero. Second, the earnings distortions are continuous with respect to the degree of risk aversion at the risk neutral solution. Third, Pareto efficient income tax systems can be time-consistent even when the degree of correlation in ability types is large. The condition for time consistency suggests a novel theoretical reason why the classic equity-efficiency trade off may be steeper in a dynamic environment than previously thought.

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

Stephen Coate
Department of Economics
Cornell University
Ithaca, NY 14853
sc163@cornell.edu

*We are grateful to Narayana Kocherlakota for detailed comments and general enthusiasm for the topic. For helpful feedback, we also thank two anonymous referees, Christophe Chamley, V.V. Chari, Per Krusell, Emmanuel Saez, and seminar participants at Boston University, the Minneapolis Fed, NBER, Ohio State, Princeton, Rochester and Yale.

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 sheds light on the first and third of these questions. It analyzes a dynamic version of the classic Mirrlees model in which, in any period, there are two ability types - low and high - and individuals' abilities follow a Markov process. Following the approach of Stiglitz (1982, 1985a), the paper studies the Pareto efficiency problem of maximizing the expected utility of those who start out as high ability subject to a given target utility for those who are initially low ability.[1] It then studies whether these efficient allocations are time consistent in the sense that they cannot be Pareto dominated as information about individuals' ability types is revealed over time.

The paper begins by assuming that individuals' per period utility is a quasi-linear function of consumption and labor, implying that they are risk neutral.[2] This makes the consumption smoothing issue moot and permits a clean focus on the first and third questions. The analysis of this case yields clear and striking results. With regard to the dynamics of earnings distortions, 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 decreases over time, converging to zero if the time horizon is long enough. Thus, 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 vanish 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.

To assess the robustness of our results on earnings distortions, we also study the case of risk averse individuals. In a two period version of the model with risk aversion, we show 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

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

[2] The quasi-linear specification has also proved useful in the study of static optimal income taxation. See, for example, Diamond (1998), Besley and Coate (1995) and Salanie (2003).

distortions described above. However, we also provide a general continuity result (for any number of periods) showing that the distortion in the earnings of individuals who are either currently high types or who have been previously low types is small for small degrees of risk aversion. Thus, the basic insight that earnings distortions vanish over time is 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, the so-called *New Dynamic Public Finance*.[3] This literature was recently reviewed by Kocherlakota (2006).[4] Our paper differs from this recent literature in both focus and style. In terms of focus, the literature has been primarily concerned with the Utilitarian problem of maximizing aggregate expected utility rather than on characterizing Pareto efficient tax systems. Relatedly, it is the problem of consumption smoothing (the second question above) rather than the dynamics of earnings distortions or the problem of time inconsistency (the first and third questions) that has attracted the most attention. Most papers have assumed that ability types are serially uncorrelated which makes both the first and third questions less interesting, while those papers that have considered more general stochastic processes (Golosov, Kocherlakota and Tsyvinsky (2003), Kocherlakota (2005)), have focused their analysis on a study of the implications of the first order conditions for intertemporal consumption. In terms of style, our model is much simpler than those in the recent literature. While it does incorporate persistence in abilities, it has only two ability types, no capital, exogenous interest rates, and, for much of the analysis, risk neutral individuals. The advantage of these more restrictive assumptions is that they allow us both to provide a complete characterization of second best efficient allocations. Indeed, as far as we are aware, ours is the first paper to provide a full characterization of second best efficient allocations in a dynamic stochastic version of the Mirrlees model.

In characterizing second best efficient allocations and studying their time consistency, our paper draws on the dynamic contracting literature. In particular, we follow the analytical approach employed by Battaglini (2005a) to study a monopoly pricing problem with long-lived consumers

[3] This "new" approach is distinct from the "traditional" approach that makes the assumption that the government is constrained to use linear taxes (see Chari and Kehoe (1999) for a review).

[4] Earlier papers in this style include Brito et al (1991), Diamond and Mirrlees (1978), Ordover and Phelps (1979), Roberts (1984) and Stiglitz (1985b). See also Berliant and Ledyard (2003) who characterize time consistent taxation in a two period model with constant ability types.

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 explores the properties of second best efficient allocations under risk neutrality and draws out the implications for the efficient taxation of labor income. Section 4 studies how risk aversion modifies the conclusions. Section 5 analyses the time consistency of second best efficient allocations under risk neutrality and Section 6 concludes.

2 The model

We study an economy with a continuum of individuals that lasts for T periods. 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

$$\frac{x_t^{1-\sigma}}{1-\sigma} - \varphi(l_t),$$

where φ is increasing, strictly convex, and twice continuously differentiable. The parameter σ measures individuals' risk aversion. A special case of interest is when $\sigma = 0$ and individuals are risk neutral. 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 $\delta < 1$.

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. There are two ability levels, low and high, denoted by $\{\theta_L, \theta_H\}$ where $0 < \theta_L < \theta_H$. A fraction $\mu \in (0, 1)$ of individuals start out with high ability in period one. However, those who start out as high ability may become low ability and visa versa. Specifically, each individual's ability follows a Markov process with support

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

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

Ability types are correlated but not perfectly, so that $\alpha_{LL} > \alpha_{HL} > 0$ and $\alpha_{HH} > \alpha_{LH} > 0$.

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. The government can borrow or lend at the exogenously fixed interest rate which equals $\frac{1}{\delta} - 1$.

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_t denote the set of all histories at time t. 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 $(\mathbf{x}, \mathbf{y}) = \{(x_t(h_t, \theta_t), y_t(h_t, \theta_t))\}_{t=1}^T$. Here $(x_t(h_t, \theta_t), y_t(h_t, \theta_t))$ is the consumption-earnings bundle of those individuals who have ability θ_t in period t after history $h_t \in H_t$. To be *feasible*, an allocation must satisfy the aggregate resource constraint[5]

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

This says that the present value of consumption and government spending equals the present value of earnings. Under the allocation $(\mathbf{x}, \mathbf{y})$, the expected utility at time t of an individual with ability θ_t and history h_t is

$$V_t((\mathbf{x}, \mathbf{y}), h_t, \theta_t; \sigma) = \sum_{\tau=t}^{T} \delta^{\tau-t} E[\frac{x_\tau(h_\tau, \theta_\tau)^{1-\sigma}}{1-\sigma} - \varphi(\frac{y_\tau(h_\tau, \theta_\tau)}{\theta_\tau}) \mid \theta_t].$$

In addition to raising the revenue necessary to finance its spending, the government has the distributional objective of providing those citizens who start out with low ability a lifetime expected utility of at least $\underline{u}$. The government would like to achieve its distributional and revenue raising goals efficiently and hence would like to implement an allocation that solves the following

[5] Obviously, feasibility also demands that individuals' consumptions in each period be non-negative. However, we will ignore these constraints in what follows, effectively focusing on the properties of interior allocations.

problem

$$\max_{(\mathbf{x},\mathbf{y})} V_1((\mathbf{x},\mathbf{y}), h_1, \theta_H; \sigma)$$

$$s.t. \quad V_1((\mathbf{x},\mathbf{y}), h_1, \theta_L; \sigma) \geq \underline{u} \quad (U_L)$$

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

In the sequel we refer to this problem as the *Efficiency Problem* and to allocations that solve it as *efficient allocations.*

When individuals are risk averse ($\sigma > 0$), an allocation $(\mathbf{x},\mathbf{y})$ solves the Efficiency Problem if and only if three conditions are satisfied. First, individuals' consumption levels are constant across time and states. Second, individuals work up until the point at which their marginal disutility of work equals the marginal utility of the consumption that work produces. Third, the U_L and R constraints are satisfied with equality. The first condition requires that those who start out with high ability in period one have constant consumption x_H, while those who are low types get x_L. The second condition requires that those who are high types in period one earn an amount y_H^H in a period in which they have high ability and an amount y_H^L when they have low ability where $\theta_H x_H^{-\sigma} = \varphi'(y_H^H/\theta_H)$ and $\theta_L x_H^{-\sigma} = \varphi'(y_H^L/\theta_L)$. Similarly, those who are low types in period one earn an amount y_L^H in a period in which they have high ability and an amount y_L^L when they have low ability where $\theta_H x_L^{-\sigma} = \varphi'(y_L^H/\theta_H)$ and $\theta_L x_L^{-\sigma} = \varphi'(y_L^L/\theta_L)$.

In the case of risk neutrality, individuals are indifferent as to the allocation of consumption across time and states. Thus, for efficiency, all that is important is that individuals' work decisions are optimal. An allocation therefore solves the Efficiency Problem if and only if individuals work up until the point at which their marginal disutility of work equals their marginal product and the U_L and R constraints are satisfied with equality.

If the government can observe individuals' income generating abilities, it can implement an efficient allocation with a simple system of lump sum taxes.[6] However, we assume that the

[6] Let $y^*(T,\theta)$ denote the earnings level that would maximize the static utility of an individual with ability $\theta \in \{\theta_L, \theta_H\}$ if he had to pay a lump sum tax T; that is, $y^*(T,\theta)$ maximizes $\frac{(y-T)^{1-\sigma}}{1-\sigma} - \varphi(y/\theta)$ subject to the constraint that $y/\theta \in [0,\bar{l}]$. Then, any efficient allocation $(\mathbf{x},\mathbf{y})$ can be implemented as follows. Individuals who are start out with high ability in period one pay a lump sum tax T_H where $y^*(T_H,\theta_H) - T_H = x_H$. They also pay this tax in any future period in which they have high ability. In any period in which they have low ability, their tax burdens are reduced in such a way as to maintain their consumption at the same level. Thus, 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 start out with low ability in period one 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 any future period in which they are low types. In any period in which they experience high ability,

government is not able to observe individuals' income generating abilities. This unobservability constrains the allocations that the government might reasonably achieve. Specifically, allocations must now satisfy the following set of incentive constraints: for all time periods t and histories h_t,

$$V_t((\mathbf{x},\mathbf{y}),h_t,\theta_H;\sigma) \geq \frac{x_t(h_t,\theta_L)^{1-\sigma}}{1-\sigma} - \varphi(\frac{y_t(h_t,\theta_L)}{\theta_H}) + \delta E[V_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_{t+1};\sigma)\,|\theta_t = \theta_H\,] \quad (IC_H(h_t))$$

and

$$V_t((\mathbf{x},\mathbf{y}),h_t,\theta_L;\sigma) \geq \frac{x_t(h_t,\theta_H)^{1-\sigma}}{1-\sigma} - \varphi(\frac{y_t(h_t,\theta_H)}{\theta_L}) + \delta E[V_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_H),\theta_{t+1};\sigma)\,|\theta_t = \theta_L\,]. \quad (IC_L(h_t))$$

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.[7]

Given its informational constraints, the best the government can do is to achieve an allocation that solves the following incentive constrained problem

$$\max_{(\mathbf{x},\mathbf{y})} V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;\sigma)$$

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

We refer to this as the *Second Best Efficiency Problem* and to allocations that solve this as *second best efficient allocations.* Our interest lies in understanding what solutions to this problem look like and how the government may achieve them via tax-transfer systems.

It is important to be clear on the nature of the incentive problems created by the government's inability to observe its citizens' abilities. There are two distinct problems. The first is created by the government's desire to redistribute from those citizens who start out with high ability to those who are initially low ability. If the target level of utility for those who start out with low ability ($\underline{u}$) is sufficiently high, those who are initially high types will have an incentive to masquerade as low types. This is the incentive problem stressed in the literature on Pareto efficient taxation (see Stiglitz (1982), (1985a)).

The second incentive problem arises even if all individuals are ex ante identical and is created by the tension between the desire to provide insurance and the need to provide work incentives.

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.

[7] While these incentive constraints consider only one time deviations, the one-stage-deviation principle implies that they ensure that individuals cannot gain from more complex mis-reporting strategies.

When individuals are risk averse, efficiency requires that individuals have constant consumption across time and states. This means that they are fully insured from future ability shocks. But efficiency also requires that individuals who have higher productivity should provide more labor. These two goals are mutually inconsistent. This incentive problem is the major focus of the *New Dynamic Public Finance* literature.

Notice that the second incentive problem is not operative when individuals are risk neutral because then providing insurance is not necessary for efficiency. By contrast, the first incentive problem arises even when individuals are risk neutral, provided that the target level of utility for those who start out with low ability is high enough. To ensure that this target utility level is sufficiently high for the first incentive problem to arise, we make the following assumption. Consider the *Utilitarian Problem* of maximizing aggregate utility subject to the resource constraint and the incentive constraints; that is,

$$\max_{(\mathbf{x},\mathbf{y})} \mu V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;\sigma)+(1-\mu)V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;\sigma)$$
$$s.t. \quad R \text{ and } IC_H(h_t) \;\&\; IC_L(h_t) \text{ for all } t \;\&\; h_t.$$

Then we assume that any solution to this problem violates the utility maintenance constraint U_L; that is, if $(\mathbf{x},\mathbf{y})$ solves the Utilitarian Problem, then it must be the case that $V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;\sigma) < \underline{u}$.[8]

3 The case of risk neutrality

We begin our analysis of second best efficient allocations by studying the case of risk neutrality; that is, $\sigma = 0$. Note that, under this assumption, an allocation $(\mathbf{x},\mathbf{y})$ is efficient if and only if the earnings path $\mathbf{y}$ maximizes Marshallian aggregate surplus

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

and the consumption path $\mathbf{x}$ is such that U_L and R hold with equality. The surplus maximizing earnings path has the property that 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 $y^*(\theta)$ satisfies the first order condition $\theta=\varphi'(y/\theta)$.

[8] Thus, we are assuming that the government puts more weight on the utility of those who are initially low types than on those who are initially high types. This would emerge from any social welfare function that is a strictly concave function of citizen utilities.

3.1 Solution procedure

To characterize second best efficient allocations, we study the following *Relaxed Problem*:

$$\max_{(\mathbf{x},\mathbf{y})} V_1((\mathbf{x},\mathbf{y}), h_1, \theta_H; \sigma)$$

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

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 Second Best Problems.

Our first observation about the Relaxed Problem is:

Lemma 1 *Suppose that $\sigma = 0$ and 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 binding. But then the solution to the Relaxed Problem would involve the surplus maximizing earnings path and a consumption path that satisfied all the incentive constraints for the high type. From this allocation, by transferring consumption from the high type to the low type in each period and after every history as needed, we can construct an allocation that involves the surplus maximizing earnings levels and consumption levels such that all the incentive constraints of the high type hold with equality. This allocation can be shown to satisfy all the low types' incentive constraints and hence solves the Utilitarian Problem. But this is a contradiction since this allocation obviously satisfies the U_L constraint strictly.

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 *Suppose that $\sigma = 0$ and 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}, θ_{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 path $\mathbf{x}'$ 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

$$V_t\left((\mathbf{x},\mathbf{y}), h_t, \theta_H; \sigma\right) = V_t\left((\mathbf{x},\mathbf{y}), h_t, \theta_L; \sigma\right) + \Phi\left(y_t(h_t, \theta_L)\right) + \Delta EV_{t+1}((\mathbf{x},\mathbf{y}), (h_t, \theta_L), \theta_{t+1}; \sigma) \tag{1}$$

where $\Phi(y) = \varphi(y/\theta_L) - \varphi(y/\theta_H)$ and $\Delta EV_{t+1}((\mathbf{x},\mathbf{y}), (h_t, \theta_L), \theta_{t+1}; \sigma)$ is the difference in the continuation values for the two types.[9] 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 $H^{\circ}(h_t)$ 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 $H^{\circ}(h_t)$. Then we can use (1) to show:

[9] That is, the difference between $E[V_{t+1}((\mathbf{x},\mathbf{y}), (h_t, \theta_L), \theta_{t+1}; \sigma)\,|\theta_t = \theta_H\,]$ and $E[V_{t+1}((\mathbf{x},\mathbf{y}), (h_t, \theta_L), \theta_{t+1}; \sigma)\,|\theta_t = \theta_L\,]$.

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:*

$$V_t\left((\mathbf{x},\mathbf{y}),h_t,\theta_H;\sigma\right)=V_t\left((\mathbf{x},\mathbf{y}),h_t,\theta_L;\sigma\right)+\sum_{j=0}^{T-t}\delta^j\left[\alpha_{HH}-\alpha_{LH}\right]^j\Phi\left(y_{t+j}(h^\circ_{t+j},\theta_L)\right). \quad (2)$$

This result can in turn be used to establish:

Lemma 4 *Suppose that* $\sigma=0$ *and let* $(\mathbf{x},\mathbf{y})$ *solve the Relaxed Problem. Then the earnings path* $\mathbf{y}$ *solves the problem:*

$$\begin{aligned}&\max \textstyle\sum_{j=0}^{T-1}\delta^j\left[\alpha_{HH}-\alpha_{LH}\right]^j\Phi\left(y_{1+j}(h^\circ_{1+j},\theta_L)\right)+\underline{u}\\ &s.t.\quad G\le(1-\delta)\textstyle\sum_{t=1}^{T}\delta^{t-1}E[y_t(h_t,\theta_t)-\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})]\\ &-(1-\delta)[\mu\textstyle\sum_{j=0}^{T-1}\delta^j\left[\alpha_{HH}-\alpha_{LH}\right]^j\Phi\left(y_{1+j}(h^\circ_{1+j},\theta_L)\right)+\underline{u}].\end{aligned} \quad (3)$$

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

$$\begin{aligned}\mathcal{L}=&\textstyle\sum_{t=1}^{T}\delta^{t-1}E[y_t(h_t,\theta_t)-\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})]-G/(1-\delta)\\ &-(\mu-\tfrac{1}{\gamma(1-\delta)})\{\textstyle\sum_{j=0}^{T-1}\delta^j\left[\alpha_{HH}-\alpha_{LH}\right]^j\Phi\left(y_{1+j}(h^\circ_{1+j},\theta_L)\right)+\underline{u}\}.\end{aligned} \quad (4)$$

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 $h_t^*=h^\circ_{1+(t-1)}$, the first order conditions are that for all t and $h_t\neq h_t^*$

$$\varphi'(\frac{y_t(h_t,\theta_t)}{\theta_t})=\theta_t \quad (5)$$

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

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

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 Second Best Problems.

Lemma 5 *Suppose that* $\sigma=0$. *Let* $(\mathbf{x},\mathbf{y})$ *be an allocation with the property that the earnings path solves the problem described in Lemma 4 and the consumption path is 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 path 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 sub-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 Second Best Problems is somewhat non-standard. In a standard problem, it is the case that any solution to the Relaxed Problem solves the Second Best Problem. In our problem, those solutions that do not satisfy the constraints with equality do not necessarily solve the Second Best Problem.

3.2 Second best efficient allocations

We now present our first main result.

Proposition 1 *Suppose that $\sigma = 0$. 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 and converges to 0 if T is large enough.*

Proof: Let $(\mathbf{x}, \mathbf{y})$ be a second best efficient allocation. Then, by Lemma 5, the earnings path solves the problem described in Lemma 4. The first order conditions tell us that for all t and $h_t \neq h_t^*$

$$\varphi'(y_t(h_t, \theta_t)/\theta_t) = \theta_t \tag{7}$$

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

$$(1-\mu)[1 - \frac{\varphi'(y_t(h_t^*, \theta_L)/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{t-1} \Phi'\left(y_t(h_t^*, \theta_L)\right)(\mu - \frac{1}{\gamma(1-\delta)}) \tag{8}$$

If an individual is currently or has at some point been a high type, then $h_t \neq h_t^*$ and, from (7), 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 (8), 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 provided that $\mu > 1/\gamma(1-\delta)$. Since $\alpha_{LL} > \alpha_{HL}$, $\left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{t-1}$ is decreasing in t and converges to zero as $t \to \infty$. The first order condition therefore implies that $y_t(h_t^*, \theta_L)/\theta_L$ is decreasing in t and (since γ must be positive) converges to $y^*(\theta_L)/\theta_L$ as $t \to \infty$.

It only remains to prove that $\mu > 1/\gamma(1-\delta)$. Assume, first that $\mu = 1/\gamma(1-\delta)$. Then (8) implies that for all t, $y_t(h_t^*, \theta_L) = y^*(\theta_L)$. This means that the earnings levels that solve the problem described in Lemma 4 maximize Marshallian aggregate surplus. It follows from Lemma 5 that any second best efficient allocation must solve the Utilitarian Problem

$$\max_{(\mathbf{x},\mathbf{y})} \mu V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;0)+(1-\mu)V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;0)$$

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

But, by assumption, any solution to the Utilitarian Problem must violate the U_L constraint.

Next suppose that $\mu < 1/\gamma(1-\delta)$. Let $\mathbf{y}$ denote an earnings path that solves the problem in Lemma 4. 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 $\mathbf{y}$. Then, we will show that $(\widetilde{\mathbf{x}},\mathbf{y})$ cannot solve the Relaxed Problem - a contradiction. To see this, consider a marginal reduction dy in $y_1(h_1,\theta_L)$ and choose dx so as to keep the utility of those who are low ability in period one constant; i.e., so that

$$\widetilde{x}_1(h_1,\theta_L)-dx-\varphi(\frac{y_1(h_1,\theta_L)-dy}{\theta_L})=\widetilde{x}_1(h_1,\theta_L)-\varphi(\frac{y_1(h_1,\theta_L)}{\theta_L}).$$

Clearly,

$$dx=\frac{\varphi'(y_1(h_1,\theta_L)/\theta_L)}{\theta_L}dy.$$

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

$$dR=(1-\mu)[dx-dy]=(1-\mu)[\frac{\varphi'(y_1(h_1,\theta_L)/\theta_L)}{\theta_L}-1]dy>0.$$

This change is positive since $\mu < 1/\gamma(1-\delta)$ implies that $1-\frac{\varphi'(y_1(h_1,\theta_L)/\theta_L)}{\theta_L}<0$. Now take this revenue increase and divide it among those who are high types in period one; i.e., raise $\widetilde{x}_1(h_1,\theta_H)$ by dR/μ. Clearly, this change makes the high types strictly better off, which since it violates none of the constraints, means that $(\widetilde{\mathbf{x}},\mathbf{y})$ cannot solve the Relaxed Problem. *Q.E.D.*

The proposition implies that 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 $t \to T$ when T is large. Moreover, the degree to which these individuals have their labor supply distorted also converges to zero. Thus, in a very strong sense, the distortions caused by imperfect observability of individuals' abilities vanish over time.[10]

[10] The properties of second best efficient allocations described in Proposition 1 are similar to the properties of the monopolist's optimal contract in Battaglini's (2005a) pricing problem. Battaglini refers to the first as the "Generalized No-Distortion at the Top Principle" and the second as the "Vanishing Distortion at the Bottom Principle". Formally, Proposition 1 extends Battglini's result in two ways. First, due to more general functional forms, here we might have multiple solutions for a given target utility level for those who are initially low ability. Proposition 1 shows that the two properties are true for any solution. Second, and more importantly, Proposition 1 shows that the two properties hold for any distribution of utilities on the second best Pareto frontier.

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 1, 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 what would happen if ability types were constant. 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. However, 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 choose 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 vanishes over time,[11] the distortion vanishes with it.

Proposition 1 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 implement 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.

[11] 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.

Consider a particular utility allocation on the Pareto frontier and let $\mathbf{y}$ denote the associated earnings path. This must solve the problem described in Lemma 4. Thus, $y_t(h_t, \theta_H) = y^*(\theta_H)$ for all t and h_t and $y_t(h_t, \theta_L) = y^*(\theta_L)$ for all t and all $h_t \neq h_t^*$. To simplify notation, let $y_{Lt}^* = y_t(h_t^*, \theta_L)$ for all t. Now, choose $\mathbf{x}$ as follows. First, let the consumption of high types in any period be constant, so that $x_t(h_t, \theta_H) = x_H^*$ 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 $x_t(h_t, \theta_L) = x_L^*$ 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 $(x_{Lt}^*)_{t=1}^T$ satisfy for all t

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

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) hold with equality. To ensure U_L also holds with equality, let x_H^* be chosen so that when x_L^* and $(x_{Lt}^*)_{t=1}^T$ are defined by the above equations, then $V_1((\mathbf{x}, \mathbf{y}), h_1, \theta_L; 0) = \underline{u}$.

Now consider the features of a tax system that could implement 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 $y^*(\theta_H)$.[12]

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

[12] Suppose that the government is employing a smooth tax schedule $T_1(y)$ with the property that $T_1(y_H^*) = y^*(\theta_H) - x_H^*$ and $T_1(y_{L1}^*) = y_{L1}^* - x_{L1}^*$. Assuming that future taxes are locally invariant to individuals' first period incomes, the schedule must be such that y_{L1}^* is a local maximizer of $y - T_1(y) - \varphi(y/\theta_L)$ and y_H^* is a local maximizer of $y - T_1(y) - \varphi(y/\theta_H)$. Since $T_1(y)$ is smooth, this requires that $T_1'(y_{L1}^*)$ equal $1 - \varphi(y_{L1}^*/\theta_L)/\theta_L$ which is positive and that $T_1'(y_H^*)$ equal $1 - \varphi(y_H^*/\theta_H)/\theta_H$ which is zero. Of course, there is no reason that the government need use such a smooth schedule. It could, for example, set $T_1(y)$ equal to infinity for any y other than y_{L1}^* or y_H^*. In this case, the notion of a marginal rate of taxation is not well defined.

$T_2(y; y^*_{L1})$ 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}.[13] 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 $T_2(y; y^*(\theta_H))$.

4 Risk aversion

We now turn to the general case of risk aversion. The problem is significantly more complicated precisely because risk aversion introduces the incentive problem arising from the trade off between insurance and work incentives. These complications prevent us from presenting a full characterization of second best efficient allocations with risk aversion. Rather we restrict attention to the case in which risk aversion is small. We begin with a general continuity result that establishes that the distortions in earnings are small when the degree of risk aversion is small. We then provide a more detailed analysis of the two period case.

4.1 The general case

For all σ, let $\Psi(\sigma)$ be the set of solutions to the Second Best Problem corresponding to σ and let $V_1(\sigma)$ denote the value function for the problem; that is, $V_1(\sigma) = V_1((\mathbf{x}, \mathbf{y}), \theta_H, h_1; \sigma)$ for $(\mathbf{x}, \mathbf{y}) \in \Psi(\sigma)$. Our first result is:

Lemma 6 *The value function $V_1(\sigma)$ is continuous at $\sigma = 0$.*

[13] Again, this assumes that future taxes are locally invariant to second period earnings.

This result tells us that the problem is well-behaved as a function of the risk aversion parameter.

Next we show that as the degree of risk aversion gets smaller, the earnings levels converge to those that are optimal under risk neutrality.

Lemma 7 *For all $\varepsilon > 0$ there exists a $\sigma_\varepsilon > 0$ and an earnings path $\mathbf{y}^*$ which solves the problem described in Lemma 4 such that if $\sigma \in (0, \sigma_\varepsilon)$ and $(\mathbf{x}, \mathbf{y}) \in \Psi(\sigma)$ then:*

$$\frac{1}{T}[\sum_{t=1}^{T} \sup_{h_{t+1} \in H_{t+1}} |y_t(h_{t+1}) - y_t^*(h_{t+1})|] \leq \varepsilon.$$

Combining these results, we obtain our second main result:

Proposition 2 *For any $\varepsilon > 0$, there exists a $\sigma_\varepsilon > 0$ such that if $\sigma \in (0, \sigma_\varepsilon)$, in any second best efficient allocation, the distortion in the earnings of individuals who are either currently high types or who have previously been high types is less than ε. Furthermore, when T is sufficiently large, for any $\varepsilon > 0$ there exists a $\sigma_\varepsilon > 0$ and a $t_{\boldsymbol{\varepsilon}}$ such that if $\sigma \in (0, \sigma_\varepsilon)$, in any second best efficient allocation, the distortion in the earnings of individuals who have always been low types is less than ε in periods $t \in \{t_\varepsilon, ..., T\}$.*

Thus, the only individuals whose earnings are significantly distorted in any period t are those who were initally low types and have remained low types. This is a declining fraction of the population. Moreover, if T is large enough, then as $t \to T$ the distortion in the earnings of even these individuals vanishes. The bottom line then is that the basic lesson of the analysis of the risk neutral case - namely, that distortions vanish - is robust to the introduction of small amounts of risk aversion.

4.2 The two period case

To solve for second best efficient allocations we again consider the Relaxed Problem obtained by ignoring the incentive constraints for low types. This is tractable because there are only three incentive constraints. 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 σ sufficiently small.

Lemma 8 *Suppose that $T = 2$. Then, there exists a $\overline{\sigma} > 0$ such that if $\sigma \in (0, \overline{\sigma})$, $(\mathbf{x}, \mathbf{y})$ 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 *Suppose that $T = 2$. Then, there exists a $\overline{\sigma} > 0$ such that if $\sigma \in (0, \overline{\sigma})$, 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* $\sigma \to 0$*. 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, 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 these earnings level 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.

Proposition 4 *Suppose that* $T = 2$*. Then, there exists a* $\overline{\sigma} > 0$ *such that if* $\sigma \in (0, \overline{\sigma})$*, 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 strictly 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).[14] 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.

Constructing a tax system that can implement second best efficient allocations with risk aversion is a more complex problem because of the need to simultaneously provide correct earnings and savings incentives. Moreover, it is not the case that we can infer marginal rates of taxation from the distortions in earnings and consumption across periods. In particular, there is no general guarantee that, under a tax system in which marginal rates reflect the distortions associated with a particular second best efficient allocation, individuals will select the bundles intended for them. We refer the interested reader to Kocherlakota (2005) for a detailed analysis of this problem.

5 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, we will say that the original tax/transfer system is not *time consistent*.

This notion of time consistency is based on Pareto dominance. The underlying idea is that it

[14] It is worth noting that we establish that the inequality is strict, while Golosov, Kocherlakota and Tsyvinsky (2003) prove only a weak inequality. Following Rogerson (1985), their argument is based on Jensen's Inequality. However, to obtain a strict inequality this argument requires that consumption levels are state contingent. That this is indeed the case in every period of a general T period model is yet to be proven and is by no means an obvious result.

is likely to be politically difficult for a government to make policy changes that reduce the benefits previously promised to some group of society. This might be regarded as unfair even by those who are not affected. It would however be much harder to argue that a policy change that increases all citizens' utility should not be chosen. Accordingly, if a previously announced policy cannot be Pareto dominated, then there will exist political forces to help it survive. But in the case of Pareto dominance, a policy will be removed and hence would not be time consistent. While other notions of time consistency could doubtless be proposed, in our view, this is a natural way of modelling it.

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 even benevolent governments find it ex post optimal to depart from Ramsey optimal policies.[15] In a model like ours, distortionary taxation is necessary to extract individuals' private information but after individuals have revealed it, the government can improve their welfare by eliminating distortions (Roberts (1984)). This means that optimal tax systems in dynamic models with constant ability types can be ex post Pareto dominated and hence are not time consistent according to our definition.

In this section we show that when individuals' types can vary stochastically the time-inconsistency problem (as we have defined it) may not arise. To analyze the issue we return to the case of risk neutral individuals. We also impose the additional assumption that the marginal disutility of labor is convex; i.e., $\varphi'''(l) \geq 0$.[16] This assumption guarantees that $\Phi'' > 0$ which in turn implies that the Lagrangian for the maximization problem in Lemma 4 is strictly concave and that the efficient earnings levels are unique.

We begin by providing a formal 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

[15] 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.

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

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.,

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

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

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

Now consider the problem:

$$\begin{aligned} &\max_{(\mathbf{x}_{h_t}, \mathbf{y}_{h_t})} V_t((\mathbf{x}, \mathbf{y}), h_t, \theta_H; \sigma) \\ &s.t.\ V_t((\mathbf{x}, \mathbf{y}), h_t, \theta_L; \sigma) \geq V_t((\mathbf{x}^*, \mathbf{y}^*), h_t, \theta_L; \sigma) \\ &\textstyle\sum_{j=0}^{T-t} \delta^j E[y_{t+j}(h_{t+j}, \theta_{t+j}) - x_{t+j}(h_{t+j}, \theta_{t+j}) \,|h_t] \geq R^*_t(h_t) \\ &and\ IC_H(h_{t+j})\ \&\ IC_L(h_{t+j})\ for\ all\ h_{t+j} \succeq h_t\ and\ j = 0, 1, .. \end{aligned} \qquad (\mathcal{P}^I_{h_t})$$

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 $\mathcal{P}^I_{h_t}$.

We now have:

Lemma 9 *Suppose that $\sigma = 0$ and let $(\mathbf{x}^*, \mathbf{y}^*)$ be a second best efficient allocation. Then,* $(\mathbf{x}^*, \mathbf{y}^*)$

is time consistent if and only if

$$\frac{\alpha_{LH}}{\alpha_{HH}} \geq \frac{\mu - \frac{1}{\gamma(1-\delta)}}{1 - \frac{1}{\gamma(1-\delta)}} \tag{9}$$

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 (9) 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 (9) 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 9, we can derive our third main result:

Proposition 5 *Suppose that $\sigma = 0$ and let $(\mathbf{x}^*, \mathbf{y}^*)$ be a second best efficient allocation. Then (i) if $\alpha_{LH}/\alpha_{HH} \geq \mu$, $(\mathbf{x}^*, \mathbf{y}^*)$ is time consistent, and, (ii) 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, note that the right hand side of condition (9) is increasing in γ and converges to μ as γ approaches ∞. Accordingly, condition (9) is necessarily satisfied when α_{LH}/α_{HH} exceeds μ which implies part (i).[17] In the intermediate case, whether the condition is satisfied depends upon the precise value of the Lagrange multiplier γ associated

[17] It is interesting to note that when μ is equal to the fraction of high ability types in the stationary distribution of the Markov process describing the evolution of the individuals' income generating abilities, the condition in (i) is satisfied if $\alpha_{HH} \leq (1 + \alpha_{LH})/2$.

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 $\underline{u}$.[18]

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 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.[19] These will lead to greater distortions and larger reductions in the aggregate pie than suggested by the Mirrlees model.

[18] Proposition 5 is related to Battaglini's (2005a) result on the renegotiation proofness of the monopolist's optimal contract. Formally, it extends his result by showing which points on the second best frontier are time consistent. In particular, it shows that those involving a higher target utility for those who are initially low ability are less likely to be time consistent.

[19] 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 (2005b) 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.

6 Conclusion

The problem of optimal taxation in a world in which individuals' income generating abilities, while persistent, may vary over time raises three general theoretical questions. First, what is the pattern of distortions in individuals' earnings choices and how do these vary over time? Second, what are the pattern of distortions in the allocation of consumption across states and time? Third, under what circumstances are efficient tax systems time consistent? This paper has tried to shed light on these questions by investigated Pareto efficient income taxation in a simple dynamic economy with individuals whose income generating abilities evolve according to a two-state Markov process.

The bulk of the analysis has assumed that individuals are risk neutral which makes the problem of consumption smoothing moot and permits a clean focus on the first and third questions. With respect to earnings distortions, the paper shows that, in the risk neutral case, in any period the only individuals who face a positive marginal income tax rate are those who started with low ability and have always been low ability. This is a declining fraction of the population, converging to zero if the time horizon is long enough. In addition, the tax rate these individuals face decreases over time, also converging to zero if the time horizon is long enough. Thus, in an efficient income tax system, earnings distortions vanish over time.

The paper shows that with risk aversion, the result that any individual who is currently a high type or who has been previously a high type faces a zero marginal tax rate no longer holds. However, it also shows that the distortion in the earnings of these individuals is small for small degrees of risk aversion. Thus, the basic insight that earnings distortions vanish is robust to introducing small amounts of risk aversion. When risk aversion is high, the tendency toward efficiency that this paper identifies would still be a force in shaping optimal taxes but other considerations arise. Thus, our results should be considered a starting point for the understanding of the pattern of earnings distortions rather than a definitive account.

With respect to time consistency, the paper shows that, in the risk neutral case, 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.

References

Aliprantis, Charalambros and Kim Border, (1994), *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer-Verlag: Berlin.

Battaglini, Marco, (2005a), "Long Term Contracting with Markovian Consumers," *American Economic Review*, 95, 637-658.

Battaglini, Marco, (2005b), "Optimality and Renegotiation in Dynamic Contracting," forthcoming, *Games and Economic Behavior*.

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

Besley, Timothy and Stephen Coate, (1995), "The Design of Income Maintenance Programmes," *Review of Economic Studies*, 62, 187-221.

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

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

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 Tsyvinski, (2003), "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70, 569-588.

Kocherlakota, Narayana, (2006), "Advances in Dynamic Optimal Taxation" to appear in *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society.*

Kocherlakota, Narayana, (2005), "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73, 1587-1622.

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.

Ordover, Janusz and Edmund Phelps, (1979), "The Concept of Optimal Taxation in the Overlapping Generations Model of Capital and Wealth," *Journal of Public Economics*, 12, 1-26.

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.

Salanie, Bernard, (2003), *The Economics of Taxation*, MIT Press: Cambridge, MA.

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

Stiglitz, Joseph, (1985a), "The Theory of Pareto-Efficient and Optimal Redistributive Taxation," in *Handbook of Public Economics Volume II*, eds, Alan Auerbach and Martin Feldstein, Elsevier: Amsterdam.

Stiglitz, Joseph, (1985b), "Inequality and Capital Taxation," IMSSS Technical Report 457, Stanford University.

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

7 Appendix

Proof of Lemma 1: It is obvious that U_L holds with equality, so we will just show that $IC_H(h_1)$ is binding. Let $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ solve the relaxed problem and suppose that $IC_H(h_1)$ 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

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

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

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

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

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

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

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

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 $IC_H(h_1)$ 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 Efficiency Problem

$$\max_{(\mathbf{x},\mathbf{y})} V_1((\mathbf{x}, \mathbf{y}), h_1, \theta_H; 0)$$

$$s.t.\ \ U_L\ \ and\ R.$$

This means that the earnings path $\widehat{\mathbf{y}}$ maximizes Marshallian aggregate surplus

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

and the consumption path $\widehat{\mathbf{x}}$ is such that U_L and R hold with equality. Moreover, for all t and histories h_t the constraints $IC_H(h_t)$ are satisfied but slack. Now, by transferring consumption from the high type to the low type in each period and after every history, create an alternative consumption path $\widetilde{\mathbf{x}}$ that for all t and histories h_t makes the constraints $IC_H(h_t)$ bind when the earnings path is $\widehat{\mathbf{y}}$. The allocation $(\widetilde{\mathbf{x}}, \widehat{\mathbf{y}})$ so created can be shown to satisfy for all t and histories h_t the constraints $IC_L(h_t)$ (see the proof of Lemma 5 below). Moreover, it strictly satisfies the U_L constraint and (given constant marginal utility) solves the Utilitarian Problem

$$\max_{(\mathbf{x},\mathbf{y})} \mu V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;0) + (1-\mu)V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;0)$$
$$s.t. \quad R \; and \; IC_H(h_t) \; \& \; IC_L(h_t) \; for \; all \; t \; \& \; h_t.$$

But, by assumption, if $(\mathbf{x},\mathbf{y})$ solves the Utilitarian Problem, then it must be the case that $V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;0) < \underline{u}$ - a contradiction. *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 $t = 2,...,T$ 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:

$$V_2((\mathbf{x},\mathbf{y}),h_2,\theta_H;0) = x_2(h_2,\theta_L) - \varphi(\frac{y_2(h_2,\theta_L)}{\theta_H}) + \delta E[V_3((\mathbf{x},\mathbf{y}),h_2,\theta_L,\theta_3;0)\,|\theta_2 = \theta_H] + \varepsilon$$

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

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

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:

$$V_1((\mathbf{x}^2,\mathbf{y}),h_1,\theta_H;0) \geq x_1^2(h_1,\theta_L) - \varphi(\frac{y_1(h_1,\theta_L)}{\theta_H}) + \delta E[V_2((\mathbf{x}^2,\mathbf{y}),h_1,\theta_L,\theta_2;0)\,|\theta_1 = \theta_H].$$

We have that:

$$\begin{aligned}
V_1((\mathbf{x}^2,\mathbf{y}),h_1,\theta_H;0) &= x_1^2(h_1,\theta_H) - \varphi(\frac{y_1(h_1,\theta_H)}{\theta_H}) + \delta E[V_2((\mathbf{x}^2,\mathbf{y}),h_1,\theta_H,\theta_2;0)\,|\theta_1 = \theta_H]. \\
&= V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;0) \\
&\geq x_1(h_1,\theta_L) - \varphi(\frac{y_1(h_1,\theta_L)}{\theta_H}) + \delta E[V_2((\mathbf{x},\mathbf{y}),h_1,\theta_L,\theta_2;0)\,|\theta_1 = \theta_H]. \\
&= x_1^2(h_1,\theta_L) - \delta\varepsilon\alpha_{LH} - \varphi(\frac{y_1(h_1,\theta_L)}{\theta_H}) + \delta E[V_2((\mathbf{x}^2,\mathbf{y}),h_1,\theta_L,\theta_2;0)\,|\theta_1 = \theta_H] + \delta\varepsilon\alpha_{HH} \\
&\geq x_1^2(h_1,\theta_L) - \varphi(\frac{y_1(h_1,\theta_L)}{\theta_H}) + \delta E[V_2((\mathbf{x}^2,\mathbf{y}),h_1,\theta_L,\theta_2;0)\,|\theta_1 = \theta_H]
\end{aligned}$$

where the third equality follows from the fact that $(\mathbf{x},\mathbf{y})$ satisfies $IC_H(h_1)$ and the fifth 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:

$$V_2((\mathbf{x},\mathbf{y}),h_2,\theta_H;0) = x_2(h_2,\theta_L) - \varphi(\frac{y_2(h_2,\theta_L)}{\theta_H}) + \delta E[V_3((\mathbf{x},\mathbf{y}),h_2,\theta_L,\theta_3;0)\,|\theta_2 = \theta_H] + \varepsilon.$$

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:

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

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: (i) 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 $(\widetilde{\mathbf{x}},\mathbf{y})$, (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 $(\widetilde{\mathbf{x}},\mathbf{y})$ for all periods $\tau > t-1$ and all histories h_τ. We can then let $\mathbf{x}^t = \widehat{\mathbf{x}}$. *Q.E.D.*

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

$$V_t\left((\mathbf{x},\mathbf{y}),h_t,\theta_H;\sigma\right) = V_t\left((\mathbf{x},\mathbf{y}),h_t,\theta_L;\sigma\right) + \Phi\left(y_t(h_t,\theta_L)\right) + \Delta EV_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_{t+1};\sigma)$$

In addition, we can write

$$\begin{aligned}\Delta EV_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_{t+1};\sigma) &= \delta\left(\alpha_{HH}-\alpha_{LH}\right)V_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_H;\sigma)\\ &\quad +\delta\left(\alpha_{HL}-\alpha_{LL}\right)V_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_L;\sigma).\end{aligned}$$

But from (1) we know that:

$$\begin{aligned}V_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_H;\sigma) &= V_{t+1}\left((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_L;\sigma\right) + \Phi\left(y_{t+1}((h_t,\theta_L),\theta_L)\right)\\ &\quad +\Delta EV_{t+2}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L),\theta_{t+2};\sigma)\\ &= V_{t+1}\left((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_L;\sigma\right) + \Phi\left(y_{t+1}((h_t,\theta_L),\theta_L)\right)\\ &\quad +\delta\left(\alpha_{HH}-\alpha_{LH}\right)V_{t+2}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L),\theta_H;\sigma)\\ &\quad +\delta\left(\alpha_{HL}-\alpha_{LL}\right)V_{t+2}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L),\theta_L;\sigma).\end{aligned}$$

Thus, it is the case that

$$\begin{aligned}\Delta EV_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_{t+1};\sigma) &= \delta\left(\alpha_{HH}-\alpha_{LH}\right)\Phi\left(y_{t+1}((h_t,\theta_L),\theta_L)\right)\\ &+\delta^2\left(\alpha_{HH}-\alpha_{LH}\right)^2 V_{t+2}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L),\theta_H;\sigma)\\ &+\delta^2\left(\alpha_{HH}-\alpha_{LH}\right)\left(\alpha_{HL}-\alpha_{LL}\right)V_{t+2}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L),\theta_L;\sigma).\end{aligned}$$

But again from (1) we have that

$$\begin{aligned}V_{t+2}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L),\theta_H;\sigma) &= V_{t+2}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L),\theta_L;\sigma)+\Phi\left(y_{t+2}((h_t,\theta_L,\theta_L),\theta_L)\right)\\ &+\delta\left(\alpha_{HH}-\alpha_{LH}\right)V_{t+3}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L,\theta_L),\theta_H;\sigma)\\ &+\delta\left(\alpha_{HL}-\alpha_{LL}\right)V_{t+3}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L,\theta_L),\theta_L;\sigma).\end{aligned}$$

So that

$$\begin{aligned}\Delta EV_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_{t+1};\sigma) &= \delta\left(\alpha_{HH}-\alpha_{LH}\right)\Phi\left(y_{t+1}((h_t,\theta_L),\theta_L)\right)\\ &+\delta^2\left(\alpha_{HH}-\alpha_{LH}\right)^2\Phi\left(y_{t+2}((h_t,\theta_L,\theta_L),\theta_L)\right)\\ &+\delta^3\left(\alpha_{HH}-\alpha_{LH}\right)^3 V_{t+3}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L,\theta_L),\theta_H;\sigma)\\ &+\delta^3\left(\alpha_{HH}-\alpha_{LH}\right)^2\left(\alpha_{HL}-\alpha_{LL}\right)V_{t+3}((\mathbf{x},\mathbf{y}),(h_t,\theta_L,\theta_L,\theta_L),\theta_L;\sigma).\end{aligned}$$

Repeated application of this argument yields

$$\Delta EV_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_{t+1};\sigma)=\sum_{j=1}^{T-t}\delta^j\left[\alpha_{HH}-\alpha_{LH}\right]^j\Phi\left(y_{t+j}(h^{\circ}_{t+j},\theta_L)\right),$$

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

$$V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;0)=\sum_{j=0}^{T-1}\delta^j\left[\alpha_{HH}-\alpha_{LH}\right]^j\Phi\left(y_{1+j}(h^{\circ}_{1+j},\theta_L)\right)+\underline{u} \tag{10}$$

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

$$V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;0)=\sum_{t=1}^{T}\delta^{t-1}E[x_t(h_t,\theta_t)-\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})\,|\theta_1=\theta_H],$$

and

$$V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;0)=\sum_{t=1}^{T}\delta^{t-1}E[x_t(h_t,\theta_t)-\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})\,|\theta_1=\theta_L\,].$$

Thus,

$$\begin{aligned}\sum_{t=1}^{T}\delta^{t-1}E[x_t(h_t,\theta_t)] &= \mu\sum_{t=1}^{T}\delta^{t-1}E[x_t(h_t,\theta_t)\,|\theta_1=\theta_H\,]+(1-\mu)\sum_{t=1}^{T}\delta^{t-1}E[x_t(h_t,\theta_t)\,|\theta_1=\theta_L\,]\\ &= \mu[V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;0)+\sum_{t=1}^{T}\delta^{t-1}E[\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})\,|\theta_1=\theta_H\,]\\ &\quad +(1-\mu)[V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;0)+\sum_{t=1}^{T}\delta^{t-1}E[\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})\,|\theta_1=\theta_L\,]\\ &= \mu V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;0)+(1-\mu)V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;0)+\sum_{t=1}^{T}\delta^{t-1}E[\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})].\end{aligned}$$

Substituting this into the resource constraint, yields

$$\begin{aligned}G &\leq (1-\delta)\sum_{t=1}^{T}\delta^{t-1}E[y_t(h_t,\theta_t)-\mu V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;0)\\ &\quad -(1-\mu)V_1((\mathbf{x},\mathbf{y}),h_1,\theta_L;0)-\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})].\end{aligned}$$

Using Lemma 3 and U_L we can write this as

$$\begin{aligned}G &\leq (1-\delta)\sum_{t=1}^{T}\delta^{t-1}E[y_t(h_t,\theta_t)-\varphi(\frac{y_t(h_t,\theta_t)}{\theta_t})] \\ &\quad -(1-\delta)[\mu\sum_{j=0}^{T}\delta^{j}\,[\alpha_{HH}-\alpha_{LH}]^{j}\,\Phi\left(y_{1+j}(h_{1+j}^{\circ},\theta_L)\right)+\underline{u}].\end{aligned} \tag{11}$$

Thus, it follows that the earnings path $\mathbf{y}$ must maximize the objective function (10) subject to the constraint (11). *Q.E.D.*

Proof of Lemma 5: Let $(\mathbf{x},\mathbf{y})$ be an allocation with the property that the earnings path solves the problem described in Lemma 4 and the consumption path is such as to make U_L and $IC_H(h_t)$ (for all t and h_t) hold with equality. We know that $(\mathbf{x},\mathbf{y})$ is a solution to the Relaxed Problem. To show that it solves the Second Best Problem, 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

$$\begin{aligned}& x_t(h_t,\theta_L)-\varphi(\frac{y_t(h_t,\theta_L)}{\theta_L})+\delta E[V_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_L),\theta_{t+1};0)\,|\theta_t=\theta_L\,]\\ \geq\;& x_t(h_t,\theta_H)-\varphi(\frac{y_t(h_t,\theta_H)}{\theta_L})+\delta E[V_{t+1}((\mathbf{x},\mathbf{y}),(h_t,\theta_H),\theta_{t+1};0)\,|\theta_t=\theta_L\,]\end{aligned}$$

or, equivalently, that

$$x_t(h_t, \theta_L) - x_t(h_t, \theta_H) \geq \varphi(\frac{y_t(h_t, \theta_L)}{\theta_L}) - \varphi(\frac{y_t(h_t, \theta_H)}{\theta_L}) \qquad (12)$$
$$+\delta E[V_{t+1}((\mathbf{x}, \mathbf{y}), (h_t, \theta_H), \theta_{t+1}; 0) - V_{t+1}((\mathbf{x}, \mathbf{y}), (h_t, \theta_L), \theta_{t+1}; 0) \, |\theta_t = \theta_L]$$

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

$$x_t(\theta_L, h_t) - x_t(\theta_H, h_t) = \varphi(\frac{y_t(\theta_L, h_t)}{\theta_H}) - \varphi(\frac{y_t(\theta_H, h_t)}{\theta_H}) \qquad (13)$$
$$\delta E[V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_{t+1}, (h_t, \theta_H), 0) - V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_{t+1}, (h_t, \theta_L), 0) \, |\theta_t = \theta_H].$$

We can use this to prove the desired inequality.

Note first that

$$\delta E[V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_{t+1}, (h_t, \theta_H), 0) - V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_{t+1}, (h_t, \theta_L), 0) \, |\theta_t = \theta_H]$$

is at least as big as

$$\delta E[V_{t+1}((\mathbf{x}, \mathbf{y}), (h_t, \theta_H), \theta_{t+1}; 0) - V_{t+1}((\mathbf{x}, \mathbf{y}), (h_t, \theta_L), \theta_{t+1}; 0) \, |\theta_t = \theta_L].$$

To see this, denote the difference between the former and the latter by Δ. Computing this difference yields

$$\Delta = (\alpha_{HH} - \alpha_{LH})[V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_H, (h_t, \theta_H); 0) - V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_H, (h_t, \theta_L); 0)]$$
$$+(\alpha_{HL} - \alpha_{LL})[V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_L, (h_t, \theta_H); 0) - V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_L, (h_t, \theta_L); 0)].$$

By assumption under the allocation $(\mathbf{x}, \mathbf{y})$ all the incentive constraints for the high type are binding. Thus, by the same argument used to establish Lemma 3, we can write

$$V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_H, (h_t, \theta_H); 0) - V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_H, (h_t, \theta_L); 0)$$
$$= V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_L, (h_t, \theta_H); 0) + \sum_{j=0}^{T-(t+1)} \delta^j [\alpha_{HH} - \alpha_{LH}]^j \Phi(y_{t+1+j}(\theta_L, (h_t, \theta_H)^o_{t+1+j}))$$
$$-V_{t+1}((\mathbf{x}, \mathbf{y}), \theta_L, (h_t, \theta_L); 0) - \sum_{j=0}^{T-(t+1)} \delta^j [\alpha_{HH} - \alpha_{LH}]^j \Phi(y_{t+1+j}((\theta_L, (h_t, \theta_L)^o_{t+1+j})).$$

Substituting this expression into the expression for Δ yields

$$\Delta = (\alpha_{HH} - \alpha_{LH}) \sum_{j=0}^{T-(t+1)} \delta^j [\alpha_{HH} - \alpha_{LH}]^j (\Phi(y_{t+1+j}((h_t, \theta_H)^o_{t+1+j}, \theta_L)) - \Phi(y_{t+1+j}((h_t, \theta_L)^o_{t+1+j}, \theta_L))$$

Observe now from the first order conditions that for all j

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

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

$$\varphi'(\frac{y_{t+1+j}((h_t,\theta_L)^{\circ}_{t+1+j},\theta_L)}{\theta_L}) \leq \theta_L.$$

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

It is also the case that

$$\varphi(\frac{y_t(h_t,\theta_L)}{\theta_H}) - \varphi(\frac{y_t(h_t,\theta_H)}{\theta_H}) \geq \varphi(\frac{y_t(h_t,\theta_L)}{\theta_L}) - \varphi(\frac{y_t(h_t,\theta_H)}{\theta_L})$$

To see this, note first that for all t and histories h_t, $y_t(h_t,\theta_L)$ is less than $y_t(h_t,\theta_H)$. Second, note that when $y_t(h_t,\theta_L)$ is less than $y_t(h_t,\theta_H)$ the function

$$f(\theta) = \varphi(\frac{y_t(h_t,\theta_L)}{\theta}) - \varphi(\frac{y_t(h_t,\theta_H)}{\theta})$$

is increasing in θ. It follows from these two claims and from (12) and (13) that $IC_L(h_t)$ is satisfied.

Conversely, let $(\mathbf{x},\mathbf{y})$ be a solution to the Second Best Problem. We need to show that the earnings path $\mathbf{y}$ solves the problem described in Lemma 4. 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 $\mathbf{y}' \neq \mathbf{y}$. Then by Lemma 4, we know that $\mathbf{y}'$ solves the problem described in Lemma 4. 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: Recall that an *allocation* in this economy is described by $(\mathbf{x},\mathbf{y}) = \{(x_t(h_t,\theta_t), y_t(h_t,\theta_t))\}_{t=1}^{T}$ or, equivalently, $(\mathbf{x},\mathbf{y}) = \{(x_t(h_{t+1}), y_t(h_{t+1}))\}_{t=1}^{T}$. Let $\digamma$ denote the set of all allocations. We can define a metric on $\digamma$ in the following way. For each t, let

$$d_t((x_t,y_t),(x'_t,y'_t)) = \sup_{h_{t+1}\in H_{t+1}} \|(x_t(h_{t+1}),y_t(h_{t+1})),(x'_t(h_{t+1}),y'_t(h_{t+1}))\|$$

where $\|\cdot\|$ is the standard Euclidean norm on $\Re^2$. Then, define

$$d((\mathbf{x},\mathbf{y}),(\mathbf{x}',\mathbf{y}')) = \frac{1}{T}[\sum_{t=1}^{T} d_t((x_t,y_t),(x'_t,y'_t))].$$

The set of allocations $\mathcal{F}$ together with the metric $d(\cdot,\cdot)$ is a metric space.

Let $\Omega(\sigma)$ denote, for each $\sigma \in [0,1)$, the subset of allocations satisfying the constraints of the Second Best Problem and let $\Psi(\sigma)$ be the set of solutions corresponding to σ. Let $V_1(\sigma)$ denote the value function for the problem. We first note the following useful fact.

Fact A.1: *The constraint set correspondence $\Omega(\sigma)$ of the Second Best Problem is upper hemi continuous.*

Proof: Since all the constraints of the problem are expressed as weak inequalities and involve continuous functions, each of them defines a compact valued, upper hemi continuous correspondence. By Theorem 14.24 in Aliprantis and Border [1994] the intersection of these correspondences is also upper hemi continuous. ■

We can now prove the Lemma. Let $\varepsilon > 0$ be given. Then we must show that there exists $\sigma_\varepsilon > 0$ such that for any $\sigma \in (0, \sigma_\varepsilon)$ we have that

$$|V_1(\sigma) - V_1(0)| < \varepsilon.$$

We begin by demonstrating the existence of $\sigma_\varepsilon > 0$ such that for any $\sigma \in (0, \sigma_\varepsilon)$ we have that $V_1(\sigma) \geq V_1(0) - \varepsilon$. This is accomplished by showing the existence of $\overline{\sigma} > 0$ and an allocation $(\mathbf{x}(\sigma), \mathbf{y}(\sigma))$ with the property that for any $\sigma \in [0, \overline{\sigma})$: (i) $V_1((\mathbf{x}(\sigma), \mathbf{y}(\sigma)), \theta_H, h_1; \sigma)$ is continuous in σ, (ii) $(\mathbf{x}(\sigma), \mathbf{y}(\sigma)) \in \Omega(\sigma)$, and (iii) $V_1((\mathbf{x}(0), \mathbf{y}(0)), \theta_H, h_1; 0) = V_1(0)$.

We begin by constructing the allocation $(\mathbf{x}(\sigma), \mathbf{y}(\sigma))$. As a building block, we take an allocation $(\mathbf{x}^*, \mathbf{y}^*)$ which is optimal with risk neutrality (i.e., $(\mathbf{x}^*, \mathbf{y}^*) \in \Psi(0)$) and is such that the consumption levels make U_L and $IC_H(h_t)$ (for all t and h_t) hold with equality. Working backwards, we start the construction with period T. For all σ and all h_T choose $(x_T(h_{T+1};\sigma), y_T(h_{T+1};\sigma))$ such that: (i) $(x_T(h_T,\theta_L;\sigma), y_T(h_T,\theta_L;\sigma)) = (x_T^*(h_T,\theta_L), y_T^*(h_T,\theta_L))$, (ii) $y_T(h_T,\theta_H;\sigma) = y_T^*(h_T,\theta_H)$, and, (iii)

$$\frac{x_T(h_T,\theta_H;\sigma)^{1-\sigma}}{1-\sigma} = \frac{x_T(h_T,\theta_L;\sigma)^{1-\sigma}}{1-\sigma} + \varphi(\frac{y_T(h_T,\theta_H;\sigma)}{\theta_H}) - \varphi(\frac{y_T(h_T,\theta_L;\sigma)}{\theta_H}).$$

Thus, $(x_T(h_{T+1};\sigma), y_T(h_{T+1};\sigma))$ is equal to $(x_T^*(h_{T+1}), y_T^*(h_{T+1}))$ except for the high type's consumption which is designed to maintain the incentive constraint $IC_H(h_T)$ with equality. The function $(x_T(h_{T+1};\sigma), y_T(h_{T+1};\sigma))$ is continuous in σ and $(x_T(h_{T+1};0), y_T(h_{T+1};0)) = (x_T^*(h_{T+1}), y_T^*(h_{T+1}))$.

Now go to period $T-1$. For all σ and all h_{T-1} choose $(x_{T-1}(h_T;\sigma), y_{T-1}(h_T;\sigma))$ such that: (i) $(x_{T-1}(h_{T-1},\theta_L;\sigma), y_{T-1}(h_{T-1},\theta_L;\sigma)) = (x^*_{T-1}(h_{T-1},\theta_L), y^*_{T-1}(h_{T-1},\theta_L))$, (ii) $y_{T-1}(h_{T-1},\theta_H;\sigma) = y^*_{T-1}(h_{T-1},\theta_H)$, and, (iii)

$$\begin{aligned}
\frac{x_{T-1}(h_{T-1},\theta_H;\sigma)^{1-\sigma}}{1-\sigma} &= \frac{x_{T-1}(h_{T-1},\theta_L;\sigma)^{1-\sigma}}{1-\sigma} + \varphi(\frac{y_{T-1}(h_{T-1},\theta_H;\sigma)}{\theta_H}) - \varphi(\frac{y_{T-1}(h_{T-1},\theta_L;\sigma)}{\theta_H}) \\
&+\delta[\alpha_{HH}(\frac{x_T(h_{T-1},\theta_H,\theta_H;\sigma)^{1-\sigma}}{1-\sigma}) - \varphi(\frac{y_T(h_{T-1},\theta_H,\theta_H;\sigma)}{\theta_H})) \\
&+\alpha_{HL}(\frac{x_T(h_{T-1},\theta_H,\theta_L;\sigma)^{1-\sigma}}{1-\sigma}) - \varphi(\frac{y_T(h_{T-1},\theta_H,\theta_L;\sigma)}{\theta_L})) \\
&-\alpha_{HH}(\frac{x_T(h_{T-1},\theta_L,\theta_H;\sigma)^{1-\sigma}}{1-\sigma}) - \varphi(\frac{y_T(h_{T-1},\theta_L,\theta_H;\sigma)}{\theta_H})) \\
&-\alpha_{HL}(\frac{x_T(h_{T-1},\theta_L,\theta_L;\sigma)^{1-\sigma}}{1-\sigma}) - \varphi(\frac{y_T(h_{T-1},\theta_L,\theta_L;\sigma)}{\theta_L}))]
\end{aligned}$$

That is, $x_{T-1}(h_{T-1},\theta_H;\sigma)$ is chosen to make the high type's incentive constraint bind given the other period $T-1$ choices and what is going to happen in period T.

Keep going this way through period 2. Let $R(\sigma)$ denote the expected present value of revenues at the beginning of period 2 under the allocation so constructed; that is,

$$R(\sigma) = \sum_{t=2}^{T} \delta^{t-2} E[y_t(h_t,\theta_t;\sigma)] - \sum_{t=2}^{T} \delta^{t-2} E[x_t(h_t,\theta_t;\sigma) + G].$$

Similarly, let $V(\sigma)$ denote the expected utility at the beginning of period 2 of an individual whose first period ability was θ_L under this allocation; that is,

$$V(\sigma) = \sum_{t=2}^{T} \delta^{t-2} E[\frac{x_\tau(h_\tau,\theta_\tau;\sigma)^{1-\sigma}}{1-\sigma} - \varphi(\frac{y_\tau(h_\tau,\theta_\tau;\sigma)}{\theta_\tau}) \mid \theta_1 = \theta_L].$$

Finally, let $S(\sigma)$ denote the expected gain in utility for an individual who was high ability in period 1 from truthfully reporting as opposed to masquerading as a low type; that is,

$$\begin{aligned}
S(\sigma) = \textstyle\sum_{t=2}^{T} \delta^{t-2} E[\frac{x_t(\theta_H,\theta_2,\ldots,\theta_t;\sigma)^{1-\sigma}}{1-\sigma} - \varphi(\frac{y_t(\theta_H,\theta_2,\ldots,\theta_t;\sigma)}{\theta_t}) \mid \theta_1 = \theta_H] \\
- \textstyle\sum_{t=2}^{T} \delta^{t-2} E[\frac{x_t(\theta_L,\theta_2,\ldots,\theta_t;\sigma)^{1-\sigma}}{1-\sigma} - \varphi(\frac{y_\tau(\theta_L,\theta_2,\ldots,\theta_t;\sigma)}{\theta_t}) \mid \theta_1 = \theta_H].
\end{aligned}$$

For all σ choose $(x_1(h_1,\theta;\sigma), y_1(h_1,\theta;\sigma))$ such that: (i) $y_1(h_1,\theta_L;\sigma) = y^*_1(h_1,\theta_L)$, and, (ii) the triple $y_1(h_1,\theta_H;\sigma)$, $x_1(h_1,\theta_H;\sigma)$, and $x_1(h_1,\theta_L;\sigma)$ satisfy the following three equalities:

$$\mu(y_1(h_1,\theta_H;\sigma) - x_1(h_1,\theta_H;\sigma)) + (1-\mu)(y^*_1(h_1,\theta_L) - x_1(h_1,\theta_L;\sigma)) - G + \delta R(\sigma) = 0,$$

$$\frac{x_1(h_1,\theta_L;\sigma)^{1-\sigma}}{1-\sigma} - \varphi(\frac{y^*_1(h_1,\theta_L)}{\theta_L}) + \delta V(\sigma) = \underline{u},$$

and,

$$\frac{x_1(h_1,\theta_H;\sigma)^{1-\sigma}}{1-\sigma}-\varphi(\frac{y_1(h_1,\theta_H;\sigma)}{\theta_H})+\delta S(\sigma)-\frac{x_1(h_1,\theta_L;\sigma)^{1-\sigma}}{1-\sigma}+\varphi(\frac{y_1^*(h_1,\theta_L)}{\theta_L})=0.$$

These equalities represent a system of three equations in the three unknowns $y_1(h_1,\theta_H;\sigma)$, $x_1(h_1,\theta_H;\sigma)$, and $x_1(h_1,\theta_L;\sigma)$. We know that at $\sigma=0$ the triple $y_1^*(h_1,\theta_H)$, $x_1^*(h_1,\theta_H)$, and $x_1^*(h_1,\theta_L)$ satisfies the above three equalities. Moreover, the Jacobian matrix associated with this system at $\sigma=0$ is:

$$J=\begin{bmatrix}\mu & -\mu & -(1-\mu)\\ 1 & 0 & 0\\ 1 & -\varphi'(\frac{y_1(h_1,\theta_H;\sigma)}{\theta_H})/\theta_H & 1\end{bmatrix}$$

The determinant of the Jacobian is

$$\det J=\{\mu+(1-\mu)\varphi'(\frac{y_1^*(h_1,\theta_H)}{\theta_H})/\theta_H\}=1$$

Thus, by the *Implicit Function Theorem* there exists some $\overline{\sigma}$ such that for all $\sigma\in[0,\overline{\sigma}]$ there exists a solution $(y_1(h_1,\theta_H;\sigma),x_1(h_1,\theta_H;\sigma),x_1(h_1,\theta_L;\sigma))$ which is continuous in σ.

Thus, we have constructed for all $\sigma\in[0,\overline{\sigma}]$ an allocation $(\mathbf{x}(\sigma),\mathbf{y}(\sigma))$ that satisfies U_L, R, and $IC_H(h_t)$ (for all t and h_t) with equality and that is continuous in σ. In addition, using an argument similar to that presented in Lemma 5, we can show that it satisfies $IC_L(h_t)$ (for all t and h_t).

The next step is to demonstrate the existence of $\sigma_\varepsilon>0$ such that for any $\sigma\in(0,\sigma_\varepsilon)$ we have that

$$V_1(\sigma)\leq V_1(0)+\varepsilon.$$

Suppose to the contrary that there did not exist such a σ_ε. Then for all n there would exist $\sigma_n\in(0,1/n)$ such that $V_1(\sigma_n)>V_1(0)+\varepsilon$. Let $(\mathbf{x}(\sigma_n),\mathbf{y}(\sigma_n))$ denote the optimal allocation associated with σ_n and consider the sequence $\langle(\mathbf{x}(\sigma_n),\mathbf{y}(\sigma_n))\rangle$. Since the constraint set correspondence $\Omega(\sigma)$ is upper hemi continuous by Fact A.1 and $\lim_{n\to\infty}\sigma_n=0$, there exists a convergent subsequence $\langle(\mathbf{x}(\sigma_k),\mathbf{y}(\sigma_k))\rangle$ whose limit point $(\widehat{\mathbf{x}},\widehat{\mathbf{y}})$ is in $\Omega(0)$. This implies that

$$V_1((\widehat{\mathbf{x}},\widehat{\mathbf{y}}),h_1,\theta_H;0)\leq V_1(0).$$

But on the other hand the function $V_1((\mathbf{x},\mathbf{y}),h_1,\theta_H;\sigma)$ is continuous in $(\mathbf{x},\mathbf{y})$ and σ and hence

$$V_1((\widehat{\mathbf{x}},\widehat{\mathbf{y}}),h_1,\theta_H;0)=\lim_{k\to\infty}V_1((\mathbf{x}(\sigma_k),\mathbf{y}(\sigma_k)),h_1,\theta_H;\sigma_k)\geq V_1(0)+\varepsilon.$$

■

Proof of Lemma 7: Let $\varepsilon > 0$ be given and suppose that the claim does not hold. Then for all n there would exist $\sigma_n \in (0, 1/n)$ and $(\mathbf{x}(\boldsymbol{\sigma}_n), \mathbf{y}(\sigma_n)) \in \Psi(\sigma_n)$ such that

$$\frac{1}{T}[\sum_{t=1}^{T} \sup_{h_{t+1} \in H_{t+1}} |y_t(h_{t+1}, \sigma_n) - y_t^*(h_{t+1})|] > \varepsilon.$$

for any earnings path $\mathbf{y}^*$ that solves the problem described in Lemma 4. Consider the sequence $\langle(\mathbf{x}(\sigma_n), \mathbf{y}(\sigma_n))\rangle$. Since the constraint set correspondence $\Omega(\sigma)$ is upper hemi continuous and $lim_{n\to\infty}\sigma_n = 0$ there exists a convergent subsequence $\langle(\mathbf{x}(\sigma_k), \mathbf{y}(\sigma_k))\rangle$ whose limit point $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ is in $\Omega(0)$. Moreover, since $V_1(\sigma)$ is continuous at $\sigma = 0$ by Lemma 6, we know that

$$\begin{aligned} \lim_{k\to\infty} V_1((\mathbf{x}(\sigma_k), \mathbf{y}(\sigma_k)), h_1, \theta_H; \sigma_k) &= V_1((\widehat{\mathbf{x}}, \widehat{\mathbf{y}}), h_1, \theta_H; 0) \\ &= V_1(0) \end{aligned}$$

This implies that $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}) \in \Psi(0)$ and by Lemma 5 we know that this implies that $\widehat{\mathbf{y}} = \mathbf{y}^*$ for some $\mathbf{y}^*$ that solves the problem described in Lemma 4. But since $\widehat{\mathbf{y}} = \lim_{k\to\infty} \mathbf{y}(\sigma_k)$ we know that

$$\frac{1}{T}[\sum_{t=1}^{T} \sup_{h_{t+1} \in H_{t+1}} |\widehat{y}_t(h_{t+1}) - y_t^*(h_{t+1})|] = \lim_{k\to\infty} \frac{1}{T}[\sum_{t=1}^{T} \sup_{h_{t+1} \in H_{t+1}} |y_t(h_{t+1}, \sigma_k) - y_t^*(h_{t+1})|] \geq \varepsilon.$$

This is a contradiction. ■

Proof of Proposition 2: Beginning with the first part of the Proposition, we first demonstrate that for any $\varepsilon > 0$ there exists $\sigma_\varepsilon > 0$ such that if $\sigma < \sigma_\varepsilon$ and $(\mathbf{x}, \mathbf{y}) \in \Psi(\sigma)$, then, for any time period $t \geq 2$ and history $h_t \neq (\theta_L, ..., \theta_L)$

$$\left| \theta \left[x_t \left(h_t, \theta \right) \right]^{-\sigma} - \varphi' \left(\frac{y_t \left(h_t, \theta \right)}{\theta} \right) \right| < \varepsilon \quad \text{for } \theta \in \{\theta_L, \theta_H\}.$$

Let $\varepsilon > 0$ and suppose that the result does not hold. Then, for all n there exists some $\sigma_n < 1/n$, an allocation $(\mathbf{x}(\sigma_n), \mathbf{y}(\sigma_n)) \in \Psi(\sigma_n)$, a time period $t^n \geq 2$ and a history $h_{t^n} \neq (\theta_L, ..., \theta_L)$ such that

$$\left| \theta x_t(h_{t^n}, \theta; \sigma_n)^{-\sigma_n} - \varphi'(\frac{y_t(h_{t^n}, \theta; \sigma_n)}{\theta}) \right| \geq \varepsilon,$$

for some $\theta \in \{\theta_L, \theta_H\}$. We know from the fact that $\Omega(\sigma)$ is upper hemi continuous that there exists a convergent sub-sequence of the sequence of allocations $\langle(\mathbf{x}(\sigma_n), \mathbf{y}(\sigma_n))\rangle$ whose limit point belongs to $\Omega(0)$. Denote this convergent sub-sequence $\langle(\mathbf{x}(\sigma_k), \mathbf{y}(\sigma_k))\rangle$ and let its limit point be

$(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$. Since $V_1(\sigma)$ is continuous at $\sigma = 0$, we know that $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}) \in \Psi(0)$. It follows from Lemma 5 that for all k

$$\theta = \varphi'(\frac{\widehat{y}_{t^k}(h_{t^k}, \theta)}{\theta}).$$

But we have that for all k

$$\left| \frac{\varphi'(\frac{y_{t^k}(h_{t^k}, \theta, \sigma_k)}{\theta})}{x_{t^k}(h_{t^k}, \theta, \sigma_k)^{\sigma_k}} - \varphi'(\frac{y_{t^k}(h_{t^k}, \theta, \sigma_k)}{\theta}) \right| \geq \varepsilon.$$

Since $\lim_{k \to \infty} x_{t^k}(h_{t^k}, \theta, \sigma_k)^{\sigma_k} = 1$, this implies that there must exist $\varsigma > 0$ such that for sufficiently large k

$$|y_{t^k}(h_{t^k}, \theta, \sigma_k) - \widehat{y}_{t^k}(h_{t^k}, \theta))| \geq \varsigma.$$

The fact that $\langle (\mathbf{x}(\sigma_k), \mathbf{y}(\sigma_k)) \rangle$ converges to $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}})$ implies that

$$\frac{1}{T}[\sum_{t=1}^{T} \sup_{h_{t+1} \in H_{t+1}} |y_t(h_{t+1}, \sigma_k) - \widehat{y}_t(h_{t+1})|] \to 0.$$

Thus, for sufficiently large k,

$$\frac{1}{T}[\sum_{t=1}^{T} \sup_{h_{t+1} \in H_{t+1}} |y_t(h_{t+1}, \sigma_k) - \widehat{y}_t(h_{t+1})|] < \frac{\varsigma}{T}.$$

By definition

$$\frac{1}{T}[\sum_{t=1}^{T} \sup_{h_{t+1} \in H_{t+1}} |y_t(h_{t+1}, \sigma_k) - \widehat{y}_t(h_{t+1})|] \geq \frac{1}{T}[|y_{t^k}(h_{t^k}, \theta, \sigma_k) - \widehat{y}_{t^k}(h_{t^k}, \theta)|].$$

But, for sufficiently large k

$$\frac{1}{T}[\sum_{t=1}^{T} \sup_{h_{t+1} \in H_{t+1}} |y_t(h_{t+1}, \sigma_k) - \widehat{y}_t(h_{t+1})|] \geq \frac{1}{T} |y_{t^k}(h_{t^k}, \theta, \sigma_k) - \widehat{y}_{t^k}(h_{t^k}, \theta))| \geq \frac{\varsigma}{T}.$$

This is a contradiction.

To complete the proof of the first part of the Proposition, it only remains to show that for any $\varepsilon > 0$ there exists $\sigma_\varepsilon > 0$ such that if $\sigma < \sigma_\varepsilon$ and $(\mathbf{x}, \mathbf{y}) \in \Psi(\sigma)$, then

$$\left| \theta_H \left[x_1 \left(h_1, \theta_H\right)\right]^{-\sigma} - \varphi' \left(\frac{y_1 \left(h_1, \theta_H\right)}{\theta_H} \right) \right| < \varepsilon.$$

This can be done by following the exact same steps.

For the second part of the Proposition, we need to show that, when T is sufficiently large, for any $\varepsilon > 0$ there exists a $\sigma_\varepsilon > 0$ and a $t_{\boldsymbol{\varepsilon}}$ such that if $\sigma \in (0, \sigma_\varepsilon)$, then for any $t \in \{t_\varepsilon, ..., T\}$ and history $h_t^* = (\theta_L, ..., \theta_L)$:

$$\left| \theta_L \left[x_t \left(h_t^*, \theta_L; \sigma\right)\right]^{-\sigma} - \varphi' \left(\frac{y_t \left(h_t^*, \theta_L; \sigma\right)}{\theta_L} \right) \right| < \varepsilon.$$

Let $\varepsilon > 0$. From Proposition 1, we know that for T sufficiently large, for any $\varepsilon_0 > 0$ there exists a t_{ε_0} such that if $(\mathbf{x}, \mathbf{y}) \in \Psi(0)$, then, for time periods $t \in \{t_{\varepsilon_o}, ..., T\}$ and history $h_t^* = (\theta_L, ..., \theta_L)$

$$|y^*(\theta_L) - y_t\left(h_t^*, \theta_L\right)| < \varepsilon_0.$$

By Lemma 7, for any $\varepsilon_1 > 0$ and any T there exists a $\sigma_{\varepsilon_1} > 0$ and earnings path $\mathbf{y}^*$ which solves the problem described in Lemma 4 such that if $\sigma \in (0, \sigma_{\varepsilon_1})$ and $(\mathbf{x}, \mathbf{y}) \in \Psi(\sigma)$ then for all $t \in \{1, ..., T\}$

$$|y_t(h_t^*, \theta_L; \sigma) - y_t^*(h_t^*, \theta_L)| < \varepsilon_1.$$

Combining these implies that for T sufficiently large, when $\sigma \in (0, \sigma_{\varepsilon_1})$

$$|y^*(\theta_L) - y_t(h_t^*, \theta_L; \sigma)| < \varepsilon_0 + \varepsilon_1$$

for $t \in \{t_{\varepsilon_0}, ..., T\}$. Noting that $\theta_L = \varphi'\left(y^*(\theta_L)\right)$ and choosing ε_0 and ε_1 appropriately yields the result. ■

Proof of Lemma 8: 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. The Second Best Problem can be written as:

$$\max[\frac{(x_H)^{1-\sigma}}{1-\sigma} - \varphi(y_H/\theta_H)] + \delta[\alpha_{HH}(\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}(\frac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_L))].$$

$$s.t. \quad [\frac{(x_L)^{1-\sigma}}{1-\sigma} - \varphi(y_L/\theta_L)] + \delta[\alpha_{LL}(\frac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L)) + \alpha_{LH}(\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H))] \geq \underline{u} \quad (U_L)$$

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

$$\begin{aligned} & \frac{(x_H)^{1-\sigma}}{1-\sigma} - \varphi(y_H/\theta_H) + \delta[\alpha_{HH}(\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}(\frac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_L))] \\ \geq \quad & \frac{(x_L)^{1-\sigma}}{1-\sigma} - \varphi(y_L/\theta_H) + \delta[\alpha_{HH}(\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H)) + \alpha_{HL}(\frac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L))] \quad (IC(H)) \end{aligned}$$

$$\begin{aligned} & \frac{(x_L)^{1-\sigma}}{1-\sigma} - \varphi(y_L/\theta_L) + \delta[\alpha_{LH}(\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H)) + \alpha_{LL}(\frac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L))] \\ \geq \quad & \frac{(x_H)^{1-\sigma}}{1-\sigma} - \varphi(y_H/\theta_L) + \delta[\alpha_{LH}(\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H)) + \alpha_{LL}(\frac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_L))] \quad (IC(L)) \end{aligned}$$

$$\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H) \geq \frac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_H) \quad (IC(HH))$$

$$\frac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_L) \geq \frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_L) \quad (IC(HL))$$

$$\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H) \geq \frac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_H) \quad (IC(LH))$$

$$\frac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L) \geq \frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_L) \quad (IC(LL))$$

The Relaxed Problem eliminates the incentive constraints $IC(L)$, $IC(HL)$ and $IC(LL)$. Let $(\mathbf{x}, \mathbf{y})$ solve the Relaxed Problem. To prove the Lemma it suffices to show that there exists a $\overline{\sigma} > 0$ such that if $\sigma \in (0, \overline{\sigma})$ then the eliminated constraints are satisfied.

It is straightforward to show using similar arguments to those used in the proof of Lemma 5 that if an allocation $(\mathbf{x}, \mathbf{y})$ satisfies the constraints $IC(H)$, $IC(HH)$ and $IC(LH)$ with equality and if the earnings levels are such that $y_H \geq y_L$, $y_{HH} \geq y_{HL}$ and $y_{LH} \geq y_{LL}$, then the allocation satisfies the constraints $IC(L)$, $IC(HL)$ and $IC(LL)$. By a similar argument used in the proof of Lemmata 6 and 7, we know that in the solution to the Relaxed Problem the earnings levels converge to those that solve the problem in Lemma 4 and hence the earnings monotonicity conditions will be satisfied for σ sufficiently small. Thus, to prove the Lemma it suffices to show that if $(\mathbf{x}, \mathbf{y})$ solves the Relaxed Problem, then for sufficiently small σ, the constraints $IC(H)$, $IC(HL)$ and $IC(LL)$ bind.

Fact A.2: *Let $(\mathbf{x}, \mathbf{y})$ solve the Relaxed Problem. Then, the constraints $IC(HL)$ and $IC(LL)$ bind.*

Proof: The Lagrangian for the Relaxed Problem is

$$\begin{aligned}
\mathcal{L} \;=\; & \frac{(x_H)^{1-\sigma}}{1-\sigma} - \varphi(y_H/\theta_H) + \delta[\alpha_{HH}(\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}(\frac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_L))] \quad (14) \\
& +\lambda_U\{\frac{(x_L)^{1-\sigma}}{1-\sigma} - \varphi(y_L/\theta_L) + \delta[\alpha_{LH}(\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H)) + \alpha_{LL}(\frac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L))]\} \\
& +\lambda_H\{\frac{(x_H)^{1-\sigma}}{1-\sigma} - \varphi(y_H/\theta_H) + \delta[\alpha_{HH}(\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}(\frac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_L))] \\
& -\frac{(x_L)^{1-\sigma}}{1-\sigma} + \varphi(y_L/\theta_H) - \delta[\alpha_{HH}(\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H)) + \alpha_{HL}(\frac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L))]\} \\
& +\lambda_{HH}\{\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H) - \frac{(x_{HL})^{1-\sigma}}{1-\sigma} + \varphi(y_{HL}/\theta_H)\} \\
& +\lambda_{LH}\{\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H) - \frac{(x_{LL})^{1-\sigma}}{1-\sigma} + \varphi(y_{LL}/\theta_H)\} \\
& +\lambda_R\{\mu[(y_H - x_H) + \delta(\alpha_{HH}(y_{HH} - x_{HH}) + \alpha_{HL}(y_{HL} - x_{HL}))] \\
& +(1-\mu)[(y_L - x_L) + \delta(\alpha_{LH}(y_{LH} - x_{LH}) + \alpha_{LL}(y_{LL} - x_{LL}))]\}.
\end{aligned}$$

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

$$(x_H)^{-\sigma} = \frac{\lambda_R \mu}{1+\lambda_H}, \tag{15}$$

$$(x_{HH})^{-\sigma} = \frac{\lambda_R \mu}{1+\lambda_H+\lambda_{HH}/\delta\alpha_{HH}}, \tag{16}$$

and

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

Those for the low type's consumptions imply that:

$$(x_L)^{-\sigma} = \frac{\lambda_R(1-\mu)}{\lambda_U-\lambda_H}, \tag{17}$$

$$(x_{LL})^{-\sigma} = \frac{\lambda_R(1-\mu)}{\lambda_U-\frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H-\lambda_{LH}/\delta\alpha_{LL}}, \tag{18}$$

and

$$(x_{LH})^{-\sigma} = \frac{\lambda_R(1-\mu)}{\lambda_U-\frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H+\lambda_{LH}/\delta\alpha_{LH}}. \tag{19}$$

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

$$\frac{\varphi'(y_{HL}/\theta_L)}{\theta_L} = \frac{\lambda_R\mu-\frac{\varphi'(y_{HL}/\theta_H)}{\theta_H}\frac{\lambda_{HH}}{\delta\alpha_{HL}}}{(1+\lambda_H)}. \tag{20}$$

Those for the low type imply that:

$$\frac{\varphi'(y_L/\theta_L)}{\theta_L} = \frac{\frac{\varphi'(y_L/\theta_H)}{\theta_H}\lambda_H+\lambda_R(1-\mu)}{\lambda_U}, \tag{21}$$

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

and

$$\frac{\varphi'(y_{LH}/\theta_H)}{\theta_H} = \frac{\lambda_R(1-\mu)}{\lambda_U - \frac{\alpha_{HH}}{\alpha_{LH}}\lambda_H + \lambda_{LH}/\delta\alpha_{LH}}. \tag{23}$$

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 $x_{HL} = x_{HH}$. But from the conditions describing the high type's earnings levels if $\lambda_{HH} = 0$, 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

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

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. ■

Now consider the *Relaxed Utilitarian Problem* given by

$$\max \mu[\tfrac{(x_H)^{1-\sigma}}{1-\sigma} - \varphi(y_H/\theta_H)] + \delta[\alpha_{HH}(\tfrac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}(\tfrac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_L))]$$
$$+(1-\mu)[\tfrac{(x_L)^{1-\sigma}}{1-\sigma} - \varphi(y_L/\theta_L)] + \delta[\alpha_{LL}(\tfrac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L)) + \alpha_{LH}(\tfrac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H))]$$
$$s.t. \quad R,\ IC(H),\ IC(HH),\ \&\ IC(LH).$$

Following the logic of Lemma 7, it is straightforward to show that as σ converges to 0, the earnings levels that solve this problem converge to those that solve the Utilitarian Problem when $\sigma = 0$; namely, the surplus maximizing levels. In addition, we have that:

Fact A.3: *Let* $(\mathbf{x}, \mathbf{y})$ *solve the Relaxed Utilitarian Problem. Then, for sufficiently small* σ, *the constraints* $IC(L)$, $IC(HL)$ *and* $IC(LL)$ *bind.*

Proof: Showing that $IC(HL)$ and $IC(LL)$ bind follows the proof of Fact A.2. Suppose then that $IC(L)$ does not bind. Then, we first note that it must be that for any $\varepsilon > 0$ there is a σ_ε such that if $\sigma < \sigma_\varepsilon$ then $x_{HK} - x_{LK} < \varepsilon$ for $K \in \{H, L\}$. To see this assume by contradiction that for any $\sigma > 0$ either $x_{HH} - x_{LH} \geq \varepsilon$, or $x_{HL} - x_{LL} \geq \varepsilon$. Then, for sufficiently small σ it must be the case that *both* $x_{HH} - x_{LH} \geq \varepsilon$ and $x_{HL} - x_{LL} \geq \varepsilon$. This follows from the fact that the incentive compatibility constraints are binding in period 2 and hence

$$\begin{aligned}\tfrac{(x_{HH})^{1-\sigma}}{1-\sigma} - \tfrac{(x_{HL})^{1-\sigma}}{1-\sigma} &= \varphi(y_{HH}/\theta_H) - \varphi(y_{HL}/\theta_H) \\ \tfrac{(x_{LH})^{1-\sigma}}{1-\sigma} - \tfrac{(x_{LL})^{1-\sigma}}{1-\sigma} &= \varphi(y_{LH}/\theta_H) - \varphi(y_{LL}/\theta_H)\end{aligned}$$

But as σ becomes small $\varphi(y_{HH}/\theta_H)$ converges to $\varphi(y_{LH}/\theta_H)$ and $\varphi(y_{HL}/\theta_H)$ converges to $\varphi(y_{LL}/\theta_H)$. Thus, $\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \frac{(x_{HL})^{1-\sigma}}{1-\sigma}$ converges to $\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \frac{(x_{LL})^{1-\sigma}}{1-\sigma}$ which implies that $x_{HH} - x_{HL}$ converges to $x_{LH} - x_{LL}$.

Given this, consider a marginal decrease in x_{HH} by $\frac{\Delta}{\mu(x_{HH})^{-\sigma}}$, a decrease in x_{HL} by $\frac{\Delta}{\mu(x_{HL})^{-\sigma}}$, and a marginal increase in x_{LH} by $\frac{\Delta}{(1-\mu)(x_{LH})^{-\sigma}}$, and in x_{LL} by $\frac{\Delta}{(1-\mu)(x_{LL})^{-\sigma}}$. This change maintains the incentive constraints at time 2. For example, the incentive constraint after history (θ_H, θ_H) is given by

$$\frac{(x_{HH} - \frac{\Delta}{\mu(x_{HH})^{-\sigma}})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H) = \frac{(x_{HL} - \frac{\Delta}{\mu(x_{HL})^{-\sigma}})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_H)$$

which is maintained for small Δ. In addition, this change keeps expected utility at time 1 constant. Expected utility at time 1 as a function of Δ is

$$\begin{aligned}W(\Delta) \;\; = \;\; & \mu\{\frac{(x_H)^{1-\sigma}}{1-\sigma} - \varphi(y_H/\theta_H) + \alpha_{HH}(\frac{(x_{HH} - \frac{\Delta}{\mu(x_{HH})^{-\sigma}})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H)) \\ & + \alpha_{HL}(\frac{(x_{HL} - \frac{\Delta}{\mu(x_{HL})^{-\sigma}})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_H))\} + (1-\mu)\{\frac{(x_L)^{1-\sigma}}{1-\sigma} - \varphi(y_L/\theta_L) \\ & + \alpha_{LH}(\frac{(x_{LH} + \frac{\Delta}{(1-\mu)(x_{LH})^{-\sigma}})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H)) + \alpha_{LL}(\frac{(x_{LL} + \frac{\Delta}{(1-\mu)(x_{LL})^{-\sigma}})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L))\}.\end{aligned}$$

Differentiating, we obtain

$$\begin{aligned}W'(0) &= -\mu\{\alpha_{HH}(\frac{1}{\mu}) + \alpha_{HL}(\frac{1}{\mu})\} + (1-\mu)\{\alpha_{LH}(\frac{1}{1-\mu}) + \alpha_{LL}(\frac{1}{1-\mu})\} \\ &= 0.\end{aligned}$$

However, expected consumption at time 1 decreases. Expected consumption at time 1 as a function of Δ is

$$\begin{aligned}C(\Delta) \;\; = \;\; & \mu\{x_H + \alpha_{HH}(x_{HH} - \frac{\Delta}{\mu\,(x_{HH})^{-\sigma}}) + \alpha_{HL}(x_{HL} - \frac{\Delta}{\mu\,(x_{HL})^{-\sigma}})\} \\ & + (1-\mu)\{x_L + \alpha_{LH}(x_{LH} + \frac{\Delta}{(1-\mu)\,(x_{LH})^{-\sigma}}) + \alpha_{LL}(x_{LL} + \frac{\Delta}{(1-\mu)\,(x_{LL})^{-\sigma}})\}.\end{aligned}$$

Differentiating, we obtain

$$C'(0) = -\{\alpha_{HH}(\frac{1}{(x_{HH})^{-\sigma}}) + \alpha_{HL}(\frac{1}{(x_{HL})^{-\sigma}})\} + \{\alpha_{LH}(\frac{1}{(x_{LH})^{-\sigma}}) + \alpha_{LL}(\frac{1}{(x_{LL})^{-\sigma}})\}.$$

We know that $x_{HH} > x_{LH}$ and $x_{HL} > x_{LL}$ and hence

$$C'(0) < -\{(\alpha_{HH} - \alpha_{LH})(\frac{1}{(x_{LH})^{-\sigma}}) + (\alpha_{HL} - \alpha_{LL})(\frac{1}{(x_{LL})^{-\sigma}})\}.$$

Moreover, since $y_{LH} > y_{LL}$ for sufficiently small σ, we have that $x_{LH} > x_{LL}$ and hence

$$C'(0) < -\{\frac{(\alpha_{HH} - \alpha_{LH}) + (\alpha_{HL} - \alpha_{LL})}{(x_{LL})^{-\sigma}}\} = 0.$$

This implies that the resources constraint can be relaxed without violating the other constraints: a contradiction.

Now consider the incentive constraint $IC(L)$. This is given by

$$\begin{aligned} & \frac{(x_H)^{1-\sigma}}{1-\sigma} - \varphi(y_H/\theta_H) + \alpha_{HH}(\frac{(x_{HH})^{1-\sigma}}{1-\sigma} - \varphi(y_{HH}/\theta_H)) + \alpha_{HL}(\frac{(x_{HL})^{1-\sigma}}{1-\sigma} - \varphi(y_{HL}/\theta_H)) \\ \geq \; & \frac{(x_L)^{1-\sigma}}{1-\sigma} - \varphi(y_L/\theta_H) + \alpha_{LH}(\frac{(x_{LH})^{1-\sigma}}{1-\sigma} - \varphi(y_{LH}/\theta_H)) + \alpha_{LL}(\frac{(x_{LL})^{1-\sigma}}{1-\sigma} - \varphi(y_{LL}/\theta_L))). \end{aligned}$$

If the incentive constraint $IC(L)$ is not binding then we know that $x_H = x_L$ and $y_H > y_L$. Moreover, as σ converges to 0 we know that y_{HH} converges to y_{LH} and y_{HL} converges to y_{LL}. In addition, as we have argued, x_{HH} converges to x_{LH} and x_{HL} converges to x_{LL}. It follows that the incentive constraint $IC(L)$ must be violated - a contradiction. ■

As noted above, if an allocation $(\mathbf{x}, \mathbf{y})$ satisfies the constraints $IC(H)$, $IC(HH)$ and $IC(LH)$ with equality and the earnings levels are monotonic, then it satisfies the constraints $IC(L)$, $IC(HL)$ and $IC(LL)$. We can therefore use Fact A.3 to deduce that there exists a $\overline{\sigma} > 0$ such that if $\sigma \in (0, \overline{\sigma})$, $(\mathbf{x}, \mathbf{y})$ solves the Utilitarian Problem if and only if it solves the Relaxed Utilitarian Problem.

Fact A.4: *Let $(\mathbf{x}, \mathbf{y})$ solve the Relaxed Problem. Then, for sufficiently small σ, the constraint $IC(L)$ binds.*

Proof: Suppose that for sufficiently small σ the constraint $IC(L)$ in the Relaxed Problem does not bind. Let $(\mathbf{x}^*, \mathbf{y}^*)$ solve the Relaxed Problem. Then we know from the utility maintenance constraint that

$$V_1((\mathbf{x}^*, \mathbf{y}^*), h_1, \theta_L; \sigma) \geq \overline{u}$$

In addition, since the incentive constraint is not binding, we have that

$$V_1((\mathbf{x}^*,\mathbf{y}^*),h_1,\theta_H;\sigma) \geq \frac{(x_L^*)^{1-\sigma}}{1-\sigma} - \varphi(y_L^*/\theta_H) + \delta E[V_2((\mathbf{x}^*,\mathbf{y}^*),(h_1,\theta_L),\theta_2;\sigma)\,|\theta_1 = \theta_H].$$

Now let $(\mathbf{x}^o,\mathbf{y}^o)$ solve the Relaxed Utilitarian Problem. Then we know *by assumption* that $V_1((\mathbf{x}^o,\mathbf{y}^o),h_1,\theta_L;\sigma) < \overline{u}$. In addition, since all the incentive constraints are binding, we have by Lemma 3

$$V_1((\mathbf{x}^o,\mathbf{y}^o),h_1,\theta_H;\sigma) = V_1((\mathbf{x}^o,\mathbf{y}^o),h_1,\theta_L;\sigma) + \Phi\left(y_L^o\right) + \delta(\alpha_{HH}-\alpha_{LH})\Phi\left(y_{LL}^o\right).$$

But, on the other hand, we know that since the second period incentive constraints are binding in the Relaxed Problem and the first period constraint is not binding, then we have:

$$\begin{aligned} V_1((\mathbf{x}^*,\mathbf{y}^*),h_1,\theta_H;\sigma) &\geq V_1((\mathbf{x}^*,\mathbf{y}^*),h_1,\theta_L;\sigma) + \Phi\left(y_L^*\right) + \delta(\alpha_{HH}-\alpha_{LH})\Phi\left(y_{LL}^*\right) \\ &\geq \overline{u} + \Phi\left(y_L^*\right) + \delta(\alpha_{HH}-\alpha_{LH})\Phi\left(y_{LL}^*\right). \end{aligned}$$

But given that $IC(L)$ is not binding the earnings levels converge to those that solve the Relaxed Utilitarian Problem; namely, the surplus maximizing levels. This implies that

$$V_1((\mathbf{x}^*,\mathbf{y}^*),h_1,\theta_H;\sigma) > V_1((\mathbf{x}^o,\mathbf{y}^o),h_1,\theta_H;\sigma).$$

Thus,

$$\begin{aligned} &\mu V_1((\mathbf{x}^*,\mathbf{y}^*),h_1,\theta_H;\sigma) + (1-\mu)V_1((\mathbf{x}^*,\mathbf{y}^*),h_1,\theta_L;\sigma) \\ > \;&\mu V_1((\mathbf{x}^o,\mathbf{y}^o),h_1,\theta_H;\sigma) + (1-\mu)V_1((\mathbf{x}^o,\mathbf{y}^o),h_1,\theta_L;\sigma), \end{aligned}$$

which contradicts the fact that $(\mathbf{x}^o,\mathbf{y}^o)$ solves the Relaxed Utilitarian Problem. ■

The result now follows from Facts A.2 and A.4. *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

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

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

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

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 $\lambda_U - \frac{\alpha_{HL}}{\alpha_{LL}}\lambda_H - \lambda_{LH}/\delta\alpha_{LL} > 0$.

That the degree of distortion in the earnings of those who becomes low types in the second period converges to 0 as $\sigma \to 0$ follows from Proposition 2. Thus, it only remains to 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,

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

Since, by (17), as $\sigma \to 0$ 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 (21), again (17) 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 (see the proof of Fact A.2) and the fact that (as shown in the proof of Fact A.2) λ_{HH} is positive. For the second claim, note first that since (as shown in the proof of Proposition 3) $y_{LL} \leq y_{LH}$ 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 $(x_{LH})^{-\sigma} \geq (x_L)^{-\sigma}$ 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 $(x_{LL})^{-\sigma} \leq (x_L)^{-\sigma}$ which means that $x_{LL} \geq x_L$ - a contradiction.

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

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

Define $\upsilon(x) = (x)^{1-\sigma}/(1-\sigma)$ and $\upsilon(x_i) = \upsilon_i$ for $i = K, KL, KH$. Consider a decrease in υ_K by ϕ (which can be positive or negative) and a contextual increase of υ_{KL} and υ_{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:

$$\frac{\partial}{\partial\phi}\left[\upsilon^{-1}(\upsilon_K - \phi) + \delta\left(\alpha_{KL}\upsilon^{-1}\left(\upsilon_{KL} + \frac{\phi}{\delta}\right) + \alpha_{KH}\upsilon^{-1}\left(\upsilon_{KH} + \frac{\phi}{\delta}\right)\right)\right] = 0 \qquad (24)$$

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

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

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

Proof of Lemma 9: 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 Problem $\mathcal{P}^I_{h_t}$. 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 $\mathcal{P}^I_{h^*_t}$ is identical to the Second Best Problem, 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 $V_t((\mathbf{x}^*, \mathbf{y}^*), h^*_t, \theta_L; 0)$ 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 will exploit this similarity to solve the program $\mathcal{P}^I_{h_t^*}$ in the same way as we did the Second Best Problem. However, a certain amount of work is necessary to show that the equivalent of Lemma 1 holds for the Relaxed Problem corresponding to $\mathcal{P}^I_{h_t^*}$.

To this end, consider first the following *Revenue Maximization Problem*:

$$\max_{(\mathbf{x}_{h_t^*}, \mathbf{y}_{h_t^*})} \sum_{j=0}^{T-t} \delta^j E[y_{t+j}(h_{t+j}, \theta_{t+j}) - x_{t+j}(h_{t+j}, \theta_{t+j}) \, |h_t^*]$$
$$s.t. \; V_t \left((\mathbf{x}_{h_t^*}, \mathbf{y}_{h_t^*}), h_t^*, \theta_L; 0\right) \geq V_t \left((\mathbf{x}^*, \mathbf{y}^*), h_t^*, \theta_L; 0\right)$$
$$and \; IC_H(h_{t+j}) \; \& \; IC_L(h_{t+j}) \; \forall \; h_{t+j} \succeq h_t^* \;\; \forall \; j = 0, 1, ..$$

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

$$\max_{\mathbf{y}_{h_t^*}} (1-\delta) \sum_{j=0}^{T-t} \delta^j E[y_{t+j}(h_{t+j}, \theta_{t+j}) - \varphi(y_{t+j}(h_{t+j}, \theta_{t+j})/\theta_{t+j}) \, |h_t^*]$$
$$-(1-\delta)[\Pr(\theta_t \;\; = \;\; \theta_H \, |h_t^*) \sum_{j=0}^{T-t} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{t+j}(h^*_{t+j}, \theta_L)\right) + V_t \left((\mathbf{x}^*, \mathbf{y}^*), h_t^*, \theta_L; 0\right)]$$

We can now establish:

Fact A.5: *If (9) holds, then for any h^*_{t+j}, $y^*_{t+j}(h^*_{t+j}, \theta_L) > y^R_{t+j}(h^*_{t+j}, \theta_L)$.*

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

$$\frac{\alpha_{LL}}{\alpha_{LH}}[1 - \frac{\varphi'(y^R_{t+j}(h^*_{t+j}, \theta_L)/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^j \Phi'\left(y^R_{t+j}(h^*_{t+j}, \theta_L)\right). \qquad (25)$$

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 $y_{t+j}(h^*_{t+j}, \theta_L)$ on the interval $[y^R_{t+j}(h^*_{t+j}, \theta_L), \infty)$. From (6) we have that $y^*_{t+j}\left(h^*_{t+j}, \theta_L\right)$ solves:

$$\left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t} \frac{\gamma\,(1-\delta)\,(1-\mu)}{\gamma\mu(1-\delta)-1}[1 - \frac{\varphi'(y^*_{t+j}(h^*_{t+j}, \theta_L)/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^j \Phi'\left(y^*_{t+j}(h^*_{t+j}, \theta_L)\right). \qquad (26)$$

Using (9):

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

So the right hand side of (26) is larger than the right hand side of (25), and concavity of the revenue function implies that $y^{*}_{t+j}(h^{*}_{t+j},\theta_L)>y^{R}_{t+j}(h^{*}_{t+j},\theta_L)$. ■

We can now show that the equivalent of Lemma 1 holds for the Relaxed Problem corresponding to $\mathcal{P}^{I}_{h^*_t}$.

Fact A.6: *Let* $(x_{h^*_t},y_{h^*_t})$ *solve the Relaxed Problem corresponding to* $P^{I}_{h^*_t}$ *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

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

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

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

Since revenues are strictly decreasing on the interval $[y^{R}_{t+j}(h^{*}_{t+j},\theta_L),\infty)$, 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. ■

Given Fact A.6, 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

$$\max_{\mathbf{y}_{h_t}}\sum_{j=0}^{T-t}\delta^j\left[\alpha_{HH}-\alpha_{LH}\right]^j\Phi\left(y_{t+j}(h^{*}_{t+j},\theta_L)\right)+V_t\left((\mathbf{x}^*,\mathbf{y}^*),h^*_t,\theta_L;0\right)$$

$$\begin{aligned}s.t.\quad & R^*\left(h^*_t\right)(1-\delta)\leq(1-\delta)\sum_{j=0}^{T-t}\delta^j E[y_{t+j}(h_{t+j},\theta_{t+j})-\varphi(y_{t+j}(h_{t+j},\theta_{t+j})/\theta_{t+j})\,|h^*_t\,]\\ & -(1-\delta)\left[\begin{array}{c}\Pr(\theta_t=\theta_H\,|h^*_t\,)\sum_{j=0}^{T-t}\delta^j\left[\alpha_{HH}-\alpha_{LH}\right]^j\Phi\left(y_{t+j}(h^{*}_{t+j},\theta_L)\right)\\ +V_t\left((\mathbf{x}^*,\mathbf{y}^*),h^*_t,\theta_L;0\right)\end{array}\right]\end{aligned}\qquad(\mathcal{P}^{S}_{h^*_t})$$

Accordingly, to prove the result we need to show that $\mathbf{y}^*_{h_t^*}$ is a solution to the problem if and only if (9) holds.

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

$$\begin{aligned}\mathcal{L}_{h_t^*} \quad &= \sum_{j=0}^{T-t} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{t+j}(h^*_{t+j}, \theta_L)\right) \\ &+ \frac{[\gamma_S (1-\delta)]}{(1-\gamma_S \alpha_{LH}(1-\delta))} \sum_{j=0}^{T-t} \delta^j E[y_{t+j}(h_{t+j}, \theta_{t+j}) - \varphi(y_{t+j}(h_{t+j}, \theta_{t+j})/\theta_{t+j}) \,|h_t^*]\end{aligned}$$

where we have divided through by $1-\gamma_S\alpha_{LH}(1-\delta)$ and omitted the constants $V_t\left((\mathbf{x}^*, \mathbf{y}^*), h_t^*, \theta_L; 0\right)$ and $R^*\left(h_t^*\right)$. We denote the Lagrange multiplier γ_S to distinguish it from the analogous multiplier γ for the program solved by $\mathbf{y}^*$. Let $\mathbf{y}^S_{h_t^*}$ 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 Lemma:

Sufficient condition: If (9) is satisfied, then $\mathbf{y}^*_{h_t^*}$ is a solution to problem $\mathcal{P}^S_{h_t^*}$. We proceed in three steps.

Step 1. We first show that

$$\frac{\gamma_S (1-\delta)\,\alpha_{LL}}{\gamma_S \alpha_{LH}(1-\delta) - 1} \leq \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t} \frac{\gamma (1-\delta)(1-\mu)}{\gamma\mu(1-\delta) - 1} \tag{27}$$

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

$$\frac{\gamma_S (1-\delta)\,\alpha_{LL}}{\gamma_S \alpha_{LH}(1-\delta) - 1}[1 - \frac{\varphi'(y^S_{t+j}(h^*_{t+j}, \theta_L)/\theta_L)}{\theta_L}] = \left[1 - \frac{\alpha_{HL}}{\alpha_{LL}}\right]^j \Phi'\left(y^S_{t+j}(h^*_{t+j}, \theta_L)\right) \tag{28}$$

if $h_{t+j} \neq h^*_{t+j}$; and it would be fully efficient otherwise. If (27) is not true, then after any history h^*_{t+j}, concavity of the Lagrangian implies that the solution $y^S_{t+j}(h^*_{t+j}, \theta_L)$ of (28) is strictly larger than the ex ante optimal solution $y^*_{t+j}(h^*_{t+j}, \theta_L)$. Moreover, from Fact A.5 we know that that the ex ante optimal solution $y^*_{t+j}(h^*_{t+j}, \theta_L)$ is larger than the revenue maximizing solution $y^R_{t+j}(h^*_{t+j}, \theta_L)$. Accordingly,

$$y^S_{t+j}(h^*_{t+j}, \theta_L) > y^*_{t+j}(h^*_{t+j}, \theta_L) > y^R_{t+j}(h^*_{t+j}, \theta_L).$$

Since tax revenues are strictly decreasing on the interval $[y^R_{t+j}(h^*_{t+j}, \theta_L), \infty)$, this would imply that the revenues corresponding to the earnings path $\mathbf{y}^S_{h_t^*}$ are strictly lower than under the ex ante optimal solution $\mathbf{y}^*_{h_t^*}$: it follows that the revenue constraint $(\mathcal{P}^S_{h_t^*})$ is not satisfied - a contradiction.

Step 2. Next we show that:

$$\frac{\gamma_S\,(1-\delta)\,\alpha_{LL}}{\gamma_S\alpha_{LH}(1-\delta)-1} \geq \left[1-\frac{\alpha_{HL}}{\alpha_{LL}}\right]^{1-t} \frac{\gamma\,(1-\delta)\,(1-\mu)}{\gamma\mu(1-\delta)-1} \tag{29}$$

Assume, by contraction, that (29) 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 (29) is not true, then after any history h^*_{t+j} the solution $y^S_{t+j}(h^*_{t+j}, \theta_L)$ of (28) is smaller and hence more distorted than the ex ante optimal solution $y^*_{t+j}(h^*_{t+j}, \theta_L)$. 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 $V_t\left((\mathbf{x}^*, \mathbf{y}^*), h^*_t, \theta_H; 0\right)$.

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 (28) and (26) it follows that $\mathbf{y}^*_{h^*_t}$ solves $\mathcal{P}^S_{h^*_t}$ as claimed.

Necessary condition: If (9) is not satisfied, then $\mathbf{y}^*_{h^*_t}$ is not a solution to problem $\mathcal{P}^S_{h^*_t}$.

If (9) does not hold, then (26) and (25) imply that after any history h^*_{t+j}, $y^*_{t+j}(h^*_{t+j}, \theta_L)$ is smaller than the revenue maximizing solution $y^R_{t+j}(h^*_{t+j}, \theta_L)$. Let $\mathbf{y}_{h^*_t}$ be an earnings path such that for any history h^*_{t+j}, $y_{t+j}(h^*_{t+j}, \theta_L) \in (y^*_{t+j}(h^*_{t+j}, \theta_L), \min\{y^R_{t+j}(h^*_{t+j}, \theta_L), y^*(\theta_L)\})$ 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

$$\begin{aligned} &\max \sum_{j=0}^{T-1} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{1+j}(h^\circ_{1+j}, \theta_L)\right) + \underline{u} \\ s.t. \quad \Omega \quad \leq \quad &(1-\delta) \sum_{t=1}^{T} \delta^{t-1} E[y_t(h_t, \theta_t) - \varphi(y_t(h_t, \theta_t)/\theta_t)] \\ &-(1-\delta)[\mu \sum_{j=0}^{T-1} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{1+j}(h^\circ_{1+j}, \theta_L)\right)]. \end{aligned}$$

Let $\gamma(\Omega)$ denote the associated Lagrange multiplier. Condition (9) of Lemma 9 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:

$$\max_{(\mathbf{x},\mathbf{y})} \sum_{t=1}^{T} \delta^{t-1} E[y_t(h_t, \theta_t) - x_t(h_t, \theta_t)]$$
$$s.t.\ V_1((\mathbf{x}, \mathbf{y}), h_1, \theta_L; 0) \geq \underline{u}$$
$$and\ IC_H(h_t)\ \&\ IC_L(h_t) \quad for\ all\ t\ \&\ h_t.$$

Let $(\mathbf{x}^R, \mathbf{y}^R)$ denote the solution to this problem. As in the proof of Lemma 9, 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:

$$\max_{\mathbf{y}} (1-\delta) \sum_{t=1}^{T} \delta^{t-1} E[y_t(h_t, \theta_t) - \varphi(y_t(h_t, \theta_t)/\theta_t)]$$
$$-(1-\delta)[\mu \sum_{j=0}^{T-1} \delta^j \left[\alpha_{HH} - \alpha_{LH}\right]^j \Phi\left(y_{1+j}(h_{1+j}^{\circ}, \theta_L)\right) + \underline{u}].$$

It is easy to see that:

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

while for all other histories $y_t^R(h_t, \theta_L)$ 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 $\gamma(\widetilde{\Omega}) \leq \gamma(\Omega')$. Then, we have that along the history h_t^* for any time period t,

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

so that the difference between the revenue maximizing income level $y_t^R(h_t^*, \theta_L)$ and the constrained efficient income $y_t^*(h_t^*, \theta_L; \widetilde{\Omega})$ is not smaller than the difference between $y_t^R(h_t^*, \theta_L)$ and $y_t^*(h_t^*, \theta_L; \Omega')$. 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.*