

**Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2006/26**



Incentive Problems with
Unidimensional Hidden
Characteristics: A Unified
Approach

Martin Hellwig



MAX PLANCK SOCIETY



Incentive Problems with Unidimensional Hidden Characteristics: A Unified Approach

Martin Hellwig

revised April 2010

Incentive Problems with Unidimensional Hidden Characteristics: A Unified Approach - Corrigendum

Martin F. Hellwig
Max Planck Institute for Research on Collective Goods
Kurt-Schumacher-Str. 10, D - 53113 Bonn, Germany
hellwig@coll.mpg.de

February 26, 2011

Sergei Vieira Silva, from the Instituto Nacional de Matematica Pura e Aplicada in Rio de Janeiro, has alerted me to an error in my article "Incentive Problems with Unidimensional Hidden Characteristics: A Unified Approach", *Econometrica* 78 (2010), 1201 – 1237. Fortunately, the error does not affect the validity of the analysis.

The analysis of the paper rests on replacing the notion of a type t in the usual sense by the notion of a pseudo-type x constructed so that the distribution G of pseudo-types has a density even though the distribution F of types does not. Absolute continuity of G is asserted in Lemma 3.1, p. 1215. For the given definition of G , however, this assertion is false; for it to be true, the definition must be modified.

The definition of G in the paper takes the map $t \rightarrow \xi(t) = t + F(t)$, from types to pseudo-types, and sets $G := F \circ \xi^{-1}$. With this definition, however, G is discontinuous whenever F is discontinuous. In the proof of Lemma 3.1, the assertion that equation (3.9) holds for all x is incorrect. I apologize for the error and thank Sergei Vieira for pointing it out.

To correct the error, replace the given definition of G by one that starts from equation (3.9), i.e., define G so that

$$G(x) = x - \tau(x)$$

for all x , where $\tau(x) := \sup_s \{s | \xi(s) \leq x\}$. With this definition, absolute continuity of G follows from the observation that τ is nondecreasing and Lipschitz with constant 1. Moreover, it is still true that $F = G \circ \tau^{-1}$, which provides the basis for the subsequent analysis.

Incentive Problems with Unidimensional Hidden Characteristics: A Unified Approach*

Martin F. Hellwig
Max Planck Institute for Research on Collective Goods
Kurt-Schumacher-Str. 10, D - 53113 Bonn, Germany
hellwig@coll.mpg.de

April 7, 2010

Abstract

The paper develops a technique for studying incentive problems with unidimensional hidden characteristics in a way that is independent of whether the type set is finite, the type distribution has a continuous density, or the type distribution has both mass points and an atomless part. By this technique, the proposition that optimal incentive schemes induce no distortion "at the top" and downward distortions "below the top" is extended to arbitrary type distributions. However, mass points in the interior of the type set require pooling with adjacent higher types and, unless there are other complications, a discontinuous jump in the transition from adjacent lower types.

Key Words: Incentive Problems, Principal-Agent Models, Hidden Characteristics, General Type Distributions,

JEL Classification: C61, D82, D86

1 Introduction

Incentive problems with unidimensional hidden characteristics have usually been analysed under the assumption that either the type set is finite or the type set is an interval and the type distribution has a continuous, strictly positive density. These assumptions permit the application of standard optimization techniques, calculus if the type set is finite, control theory if the type set is a continuum.

Both assumptions are special. In the space of distributions on the real line, distributions with finite supports and distributions with positive densities form a meager set. These distributions have the special property that the weights given to different types are *commensurate* in the sense that no one type is infinitely more important than any other type. This commensurateness property

*For helpful comments and suggestions, I thank Christoph Engel, Alia Gizatulina, Hendrik Hakenes, and Klaus Ritzberger, as well as a Co-Editor and three referees.

has played an important role in the analysis of such problems. It is therefore of interest to know to what extent the results and insights that have been obtained remain valid when commensurateness fails, in particular, when the type distribution has both mass points and a continuous part, so that some types (the mass points) are infinitely more important than others (the points at which the distribution has a positive density).

As an example, consider the theory of optimal utilitarian income taxation. In this theory, the positivity of the optimal marginal income tax for all but the very highest types is usually explained in terms of a local equity-efficiency tradeoff.¹ If, for some type, the marginal income tax were zero, then this type's labour-consumption pair would be efficient. At the margin, therefore, the efficiency loss induced by a small reduction of this type's labour and consumption would be negligible relative to the gains from the additional redistribution from higher types to lower types that is made possible by the induced slackening of incentive constraints for the higher types. This argument presumes that the different types are commensurate. If the type under consideration was a mass point of the type distribution, it would be "infinitely" more important than immediately adjacent higher types that are continuity points of the type distribution. Therefore, one could not presume that the efficiency loss induced by a small reduction of this type's labour and consumption is negligible relative to the gains from the additional redistribution.² Commensurateness of neighbouring types is also presumed in elasticities interpretations of optimal income tax formulae in models with a continuum of types.³

The analysis of incentive problems with arbitrary type distributions requires a new technique. This paper develops such a technique and uses it to provide a complete characterization of optimal contracts in a principal-agent problem with unidimensional hidden characteristics, with a single-crossing condition on preferences, when no restriction is imposed on the type distribution. The principal-agent problem is chosen because its simplicity facilitates the exposition. The new technique can, however, be applied to *any* incentive problem with unidimensional hidden characteristics and with a single-crossing condition on preferences. In particular, it can be applied to the optimal income tax problem.⁴

For distributions with mass points and a continuous part, two new properties of optimal incentive schemes are obtained, both of them illustrated in Figure

¹See Mirrlees (1971, 1976, 1986), Seade (1977, 1982), Hellwig (2007 a).

²On this point, a referee has commented: "you distort downwards to reduce the information rents of all types (not just the type that is immediately above)." While it is true that the weight of the mass point is commensurate with the weight of the set of all higher types, this observation is of little help when it comes to assessing whether one wants to use a distortion for type t or a distortion for type $t + \delta t$ in order to reduce the information rents of types above $t + \delta t$. Assessments of this sort underlie the Mirrlees formula for the optimal marginal income tax and the interpretations that this formula has been given.

³Roberts (2000), Saez (2001), Hellwig (2004/2008).

⁴Because the optimal income tax problem is encumbered by the difficulties involved in characterizing the welfare weights of different types (Hellwig 2007 a), it is not so well suited for purposes of illustration of the new technique for dealing with problems involving arbitrary type distributions.

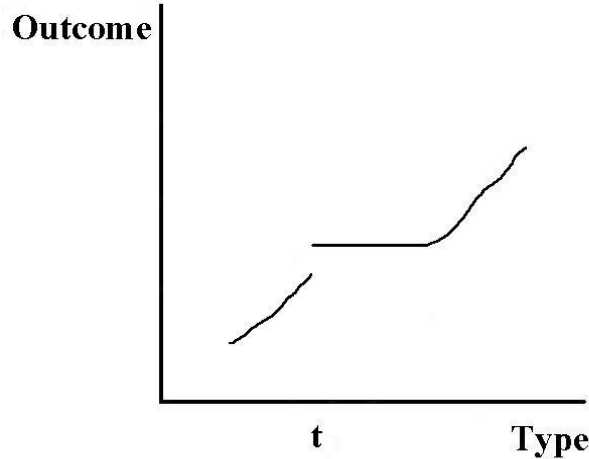


Figure 1: An Optimal Incentive Scheme When There is a Mass Point at t

1. First, any mass point below the top of the type set must be pooled with adjacent higher types. Second, a mass point between the top and the bottom of the type set is likely to give rise to a discontinuity in the mapping from types to outcomes.⁵ Both findings are due to the lack of commensurateness between the mass point and any adjacent higher or lower types that are continuity points of the type distribution. In terms of local tradeoffs between allocative and distributive concerns, e.g. the equity-efficiency tradeoff that is discussed in the literature on optimal income taxation, the difference in relative weights given to a mass point and to adjacent continuity points of the type distribution implies that, at the mass point, efficiency concerns are much more important than at adjacent continuity points of the type distribution. In an income tax model, this would suggest that the mass point should work significantly more and consume significantly more than the neighbouring types. With preferences satisfying a single-crossing condition, however, outcomes must be nondecreasing in types. Therefore, there cannot be a downward jump above the mass point. Instead, the monotonicity constraint is binding, and there is pooling of the mass point with adjacent higher types. By contrast, monotonicity does not preclude an upward jump as one moves from adjacent lower types to the mass point.⁶

⁵The effect can be neutralized by other, non-local considerations requiring the mass point to be pooled with lower as well as higher types. In the absence of such concerns, however, there *must* be an upward jump.

⁶However, there is no discontinuity in payoffs. Incentive compatibility precludes any discontinuity in the dependence of payoffs on types. At the discontinuity point, left-hand and right-hand limits of optimal outcomes lie on the same indifference curve for the critical type.

The paper also considers those properties of optimal incentive schemes that have been discussed in the literature. Regardless of the structure of the type distribution, optimal incentive schemes involve no distortion "at the top" and downward distortions "below the top" of the type distribution. In the more general setting of this paper, the latter result requires a new argument. Whereas the desirability of downward distortions "below the top" has traditionally been derived from tradeoffs imposed by the *first-order conditions* for incentive compatibility, at a mass point of the type distribution, the argument must rely on the *second-order conditions*. As explained above, the lack of commensurateness between a mass point and adjacent higher or lower continuity points of the type distribution implies that monotonicity conditions, i.e., second-order conditions for incentive compatibility are binding. The mass point must be pooled with adjacent higher types, and the analysis must show that, for all types in the pool, outcomes are distorted downwards from efficiency.

The paper builds on two technical innovations. First, for an arbitrary model with unidimensional hidden characteristics, a change of variables can be used to redefine the notion of "type" in such a way that the original incentive problem is transformed into a new one, where the distribution of the "redefined types" has a density. This density need not be continuous. However, from Clarke's (1976, 1983) version of the maximum principle under minimal hypotheses, we know that this is not a problem. The application of control theoretic methods does not require continuity of the Hamiltonian with respect to the exogenous parameter, i.e., the agent's type. It is important, however, to verify that the change of variables have no material effect on the solution to the incentive problem under consideration.

Second, for control problems with monotonicity constraints, a version of the maximum principle holds even if the map from types to outcomes is not continuous. According to this result, which is established in Hellwig (2008), one may think of the "slope" of the map from types to outcomes as a control variable even though this map may have a nontrivial singular component, and its "slope" may be unbounded. The maximum principle requires that, regardless of whether this "slope" is finite or infinite, it should not be possible to raise the value of the Hamiltonian by changing it. Thus, whenever the map from types to outcomes is strictly increasing, the associated costate variable must be zero.

Previous work on incentive problems with unidimensional hidden characteristics has assumed that the map from types to outcomes is piecewise continuously differentiable. This assumption facilitates the control theoretic treatment of monotonicity constraints. With piecewise continuous differentiability, the slope of the map from types to outcomes can be treated as a control variable; monotonicity of outcomes is equivalent to requiring this control variable to take nonnegative values.⁷ Because the map from types to outcomes is *endogenous*, however, the assumption of piecewise continuous differentiability is problematic. The technique developed in this paper provides a way of doing without it.

The generalization of the analysis to allow for type distributions with mass

⁷This approach was pioneered by Guesnerie and Laffont (1984).

points as well as a continuous part is not just a matter of mathematical generality. Such type distributions arise naturally in quasi-linear models in which the agent can get information about his type before he signs a contract. In such models, being uninformed is equivalent to having a type equal to the mean of the type distribution. Thus, if there is a positive probability that the agent does not learn his type at all, then, from the principal's perspective, the contracting problem can be treated as an incentive problem with hidden characteristics in which the type distribution has a mass point at its mean.⁸

In the following, Section 2 formulates the agency problem with hidden characteristics and states the main results. Section 3 uses a change of variables in order to make the problem amenable to control theoretic methods. Section 4 uses the maximum principle for control problems with monotonicity constraints to characterize the solutions to the agency problem and prove the main results.

2 A Principal-Agent Problem with Hidden Characteristics

2.1 Statement of the Problem

A principal wants an agent to produce some output $y \geq 0$ in return for a wage payment $w \geq 0$. The payoffs from the pair (w, y) are $y - w$ for the principal and $u(w, y, t)$ for the agent, where $t \in \mathbb{R}$ is a productivity parameter. The function u is assumed to be twice continuously differentiable, nondecreasing in w and t , nonincreasing in y , and strictly quasi-concave in w and y jointly. The agent's utility function also satisfies $u(0, 0, t) = 0$ for all t , $\lim_{y \rightarrow 0} u_y(w, y, t) = 0$ and $\lim_{y \rightarrow \infty} u_y(w, y, t) = -\infty$, uniformly in w , for all t , as well as

$$u_w(w, y, t) > 0, u_y(w, y, t) < 0, \quad (2.1)$$

and

$$\frac{\partial |u_y(w, y, t)|}{\partial t} \leq 0 \quad (2.2)$$

for all $w > 0, y > 0$, and t .⁹ The single-crossing condition (2.2) is imposed as a weak, rather than a strict inequality.

The principal is assumed to have all the bargaining power. If he offers the agent a contract (w, y) , the agent can only accept or reject this offer. The agent's payoff from rejecting the principal's offer is zero. Thus, under complete

⁸In the literature on information acquisition and incentive contracting, Szalay (2005) has the very technology considered in the text, but assumes that information is only acquired *after* the contract has been signed. Crémer et al. (1998 a, 1998 b) assume that information is acquired before the contract is offered, but they also have a further stage at which the agent learns and uses the information anyway; in their analysis therefore, being uninformed is *not* the same as having a type equal to the mean of the type distribution.

⁹This utility specification encompasses the commonly used quasi-linear specification $u = w - g(y)/t$; for the quasi-linear specification, the assumptions reduce to the conditions that $g(0) = g'(0) = 0, g'(y) > 0$ and $g''(y) > 0$ for $y > 0$, and $\lim_{y \rightarrow \infty} g'(y) = \infty$.

information, the principal would hire the agent at a wage that just compensates him for the disutility from working, without letting him share in the surplus from production.

However, there is incomplete information. Whereas the agent knows t , the principal thinks of t as the realization of a random variable \tilde{t} , to which he attributes a probability distribution F . The support T of the distribution F is assumed to be compact, with minimum t_0 and maximum t_1 .

Given his lack of information, the principal offers a menu of contracts and lets the agent choose a contract from the menu or reject the principal's offer altogether. A *contract menu* is a pair $(w(\cdot), y(\cdot))$ of integrable functions on T such that, for any $t \in T$, $(w(t), y(t))$ is the contract, i.e., the wage/output combination that is chosen by the agent when his productivity parameter is t . The *principal's problem* is to choose the contract menu $(w(\cdot), y(\cdot))$ so that his expected net payoff,

$$\int [y(t) - w(t)] dF(t), \quad (2.3)$$

is maximized subject to the incentive compatibility condition that

$$u(w(t), y(t), t) \geq u(w(t'), y(t'), t) \quad (\text{IC})$$

for all t and t' in T , and subject to the individual-rationality condition that

$$u(w(t), y(t), t) \geq 0 \quad (\text{IR})$$

for all t in T .¹⁰ A contract menu that satisfies the incentive compatibility and individual-rationality conditions is said to be *admissible*. A contract menu that maximizes the principal's expected net payoff (2.3) subject to the incentive compatibility and individual-rationality conditions is said to be *optimal*.

2.2 Distortions in Optimal Contracts

Conceptually, the principal's problem is a standard incentive problem with hidden characteristics. Textbook treatments are provided by Fudenberg and Tirole (1991), Mas-Colell, Whinston, Green (1995), or Laffont and Martimort (2001) under the assumptions that T is a finite set or that T is an interval and F has a density that is strictly positive and continuous on T . Here, I only assume that T is compact.

Let $(w(\cdot), y(\cdot))$ be an optimal contract menu and let

$$v(\cdot) := u(w(\cdot), y(\cdot), \cdot) \quad (2.4)$$

¹⁰Condition IR presumes that the principal does not prefer to make an offer which, for some t , the agent wants to reject. This presumption involves no loss of generality: Under the given assumptions, a contract menu $(w(\cdot), y(\cdot))$ with the property that, for some $t \in T$, the agent rejects the principal's offer is payoff-equivalent to the contract menu that is obtained if, for the rejecting types, the contract offers $(w(t), y(t))$ are replaced by $(0, 0)$; this latter contract menu satisfies the participation constraint IR for all t .

be the associated indirect utility function for the agent. To assess the efficiency properties of $(w(t), y(t))$, the literature compares $(w(t), y(t))$ to the pair $(w^*(t, v(t)), y^*(t, v(t)))$ that provides the person with productivity parameter t with the utility $v(t)$ at the lowest net resource cost; formally, for any t and v ,

$$(w^*(t, v), y^*(t, v)) = \arg \max_{y \geq 0, w \geq 0} \{y - w | u(w, y, n) \geq v\}. \quad (2.5)$$

The pair $(w^*(t, v), y^*(t, v))$ is fully characterized by the first-order condition

$$u_w(w^*(t, v), y^*(t, v), t) + u_y(w^*(t, v), y^*(t, v), t) = 0. \quad (2.6)$$

In those cases that have been treated in the literature, the optimal $(w(\cdot), y(\cdot))$ and $v(\cdot)$ have been shown to exhibit the following properties.

Property A *There is no distortion at the top:*

If $F(\{t_1\}) > 0$, then

$$(w(t_1), y(t_1)) = (w^*(t_1, v(t_1)), y^*(t_1, v(t_1))); \quad (2.7)$$

if $F(\{t_1\}) = 0$, then

$$\lim_{k \rightarrow \infty} (w(t^k), y(t^k)) = (w^*(t_1, v(t_1)), y^*(t_1, v(t_1))) \quad (2.8)$$

for any sequence $\{t^k\}$ in T that converges to t_1 from below.

Property B *There are downward distortions below the top:*

For any $t \in T \cap [t_0, t_1)$,

$$(w(t), y(t)) \ll (w^*(t, v(t)), y^*(t, v(t))). \quad (2.9)$$

The following theorem extends the results in the literature to the more general setting considered here. Some care must be taken with the formulation because the choice of an optimal contract menu involves a certain arbitrariness, due to the fact that the principal's objective function (2.3) is unchanged if the contract menu $(w(\cdot), y(\cdot))$ is modified on a set of probability zero. The arbitrariness is inessential, however, because the joint distribution of wage/output combinations and types is unaffected by such a modification. I will say that two contract menus $(w(\cdot), y(\cdot))$ and $(w'(\cdot), y'(\cdot))$ are *equivalent* if $(w(t), y(t)) = (w'(t), y'(t))$ for F -almost all t . Two contract menus $(w(\cdot), y(\cdot))$ and $(w'(\cdot), y'(\cdot))$ are said to be *strongly equivalent* if they are equivalent and, in addition, $u(w(t), y(t), t) = u(w'(t), y'(t), t)$ for all t , i.e., they yield the same payoff to every type of the agent.

Theorem 2.1 *For any optimal contract menu $(w(\cdot), y(\cdot))$, with associated indirect utility function $v(\cdot)$ for the agent, there exists a strongly equivalent contract menu $(\bar{w}(\cdot), \bar{y}(\cdot))$ (which is also optimal) and there exists $\hat{t} \in [t_0, t_1]$ such that $(\bar{w}(\cdot), \bar{y}(\cdot))$ is nondecreasing and, moreover, the following hold:*

(a.1) If $F([\hat{t}, t_1]) > 0$, then, for all $t \in [\hat{t}, t_1]$,

$$(\bar{w}(t), \bar{y}(t)) = (w^*(t, v(t)), y^*(t, v(t))) = (\bar{w}(t_1), \bar{y}(t_1)) \quad (2.10)$$

(a.2) If $F([\hat{t}, t_1]) = 0$, i.e., if $\hat{t} = t_1$ and $F(\{t_1\}) = 0$, then

$$\lim_{k \rightarrow \infty} (\bar{w}(t^k), \bar{y}(t^k)) = (w^*(t_1, v(t_1)), y^*(t_1, v(t_1))) \quad (2.11)$$

for any sequence $\{t^k\}$ in T that converges to t_1 from below.

(b) For any $t \in [t_0, \hat{t}]$,

$$(\bar{w}(t), \bar{y}(t)) \ll (w^*(t, v(t)), y^*(t, v(t))). \quad (2.12)$$

Corollary 2.2 *If, at the point $(w, y, t) = (w^*(t_1, v(t_1)), y^*(t_1, v(t_1)), t_1)$, the single-crossing condition (2.2) holds with a strict inequality, then the critical \hat{t} in Theorem 2.1 coincides with t_1 , and the contract menus $(w(\cdot), y(\cdot))$ and $(\bar{w}(\cdot), \bar{y}(\cdot))$ exhibit Properties A and B.*

For utility functions satisfying the single-crossing condition (2.2) with a strict inequality, Corollary 2.2 shows that, regardless of the form of the type distribution, optimal contract menus must have Properties A and B, no distortion at the top and downward distortions below the top of the type distribution. If the single-crossing condition (2.2) holds only as a weak inequality, optimal contract menus need not literally have Properties A and B. In this case, however, one still gets a decomposition of the type set into an upper part, $T \cap [\hat{t}, t_1]$, where the optimal contract is efficient, and a lower part, $T \cap [t_0, \hat{t}]$, where the optimal contract is distorted downwards from efficiency. We still have no distortion at the top, but "the top" now can be an entire interval $[\hat{t}, t_1]$. If this is the case, then, by statement (a.1) of the theorem, the contract $(\bar{w}(t), \bar{y}(t))$ is the same for all t in the interval $[\hat{t}, t_1]$. Because this contract is efficient for all these types, the single-crossing condition (2.2) must locally hold as an equation, i.e., one must have

$$\frac{\partial}{\partial t} \frac{|u_y(w, y, t)|}{u_w(w, y, t)}(w^*(t_1, v(t_1)), y^*(t_1, v(t_1)), t) = 0 \quad (2.13)$$

for all $t \in [\hat{t}, t_1]$. Conversely, if, locally, at $(w^*(t_1, v(t_1)), y^*(t_1, v(t_1)), t)$, the single-crossing condition (2.2) is strict, one must have $\hat{t} = t_1$, so that the optimal contract menu satisfies Properties A and B as specified. The corollary makes this converse explicit.

Theorem 2.1 and Corollary 2.2 provide a positive answer to the question, which has been raised in the literature on optimal taxation,¹¹ whether Property A holds if the type distribution has a density and the value of the density "at the top" is equal to zero. There is no need to assume that the density is strictly positive at t_1 . To understand the underlying logic, consider the case where F has a continuous density f and u takes the form

$$u(w, y, t) = w - \gamma\left(\frac{y}{t}\right), \quad (2.14)$$

¹¹Brett and Weymark (2003).

which was used by Mirrlees (1971). If a first-order approach to incentive compatibility is valid, an optimal contract menu must satisfy the optimality condition

$$f(t)(1 + u_y(\bar{w}(t), \bar{y}(t), t)) = (1 - F(t))u_{yt}(\bar{w}(t), \bar{y}(t), t), \quad (2.15)$$

i.e.,

$$f(t)\left(1 - \frac{1}{t}\gamma'\left(\frac{y}{t}\right)\right) = (1 - F(t))\frac{1}{t^2}\left(\gamma''\left(\frac{y}{t}\right)\frac{y}{t} + \gamma'\left(\frac{y}{t}\right)\right). \quad (2.16)$$

For any sequence $\{t^k\}$ that converges to t_1 from below, the ratio $\frac{1-F(t^k)}{f(t^k)}$ converges to zero.¹² Along such a sequence, therefore, $\frac{1}{t^k}\gamma'\left(\frac{y}{t^k}\right)$ must converge to one, i.e., the wedge distorting the output level of type t^k vanishes. If one cannot just rely on a first-order approach, the argument is more complicated, but the economic logic is the same: Near t_1 , distortions are kept small because the ratio $\frac{1-F(t)}{f(t)}$ of the weight of the set of types above t from whom the principal can extract more rents as a result of a distortion at t and the density of the type t that is affected by the distortion is close to zero. As t converges to t_1 , therefore, the tradeoff between the efficiency loss and the rent extraction gain from the distortion at t becomes degenerate.

A referee has commented that Property A depends on the assumption that t_1 is known and finite. If $t_1 = \infty$ and $F(\{t_1\}) = 0$, the efficient pair in (2.8) is not well defined, but, in the formulation of Property A, (2.8) can be replaced by the requirement that

$$\lim_{t \uparrow t_1} \frac{|u_y(\bar{w}(t), \bar{y}(t), t)|}{u_w(\bar{w}(t), \bar{y}(t), t)} = 1. \quad (2.17)$$

Condition (2.17) is equivalent to (2.8) if $t_1 < \infty$ and can also be applied when $t_1 = \infty$.

However, with $t_1 = \infty$, the ratio $\frac{1-F(t)}{f(t)}$ need not go to zero as t becomes large. Adapting an example of Diamond (1998), suppose that, above some threshold, F is a Pareto distribution, i.e., $1 - F(t) = t^{-\alpha}$ for some $\alpha > 0$, so that the ratio $\frac{1-F(t)}{f(t)} = \frac{t}{\alpha}$ goes out of bounds with t . If u takes the form (2.14) with $\gamma\left(\frac{y}{t}\right) = \left(\frac{y}{t}\right)^q$ for some $q > 1$, (2.16) can be shown to imply

$$\frac{1}{t}\gamma'\left(\frac{y}{t}\right) = \frac{\alpha}{q + \alpha}. \quad (2.18)$$

For the given utility specification, the marginal rate of substitution $\frac{|u_y(\bar{w}(t), \bar{y}(t), t)|}{u_w(\bar{w}(t), \bar{y}(t), t)}$ is just equal to the term $\frac{1}{t}\gamma'\left(\frac{y}{t}\right)$ on the left-hand side. Condition (2.18) requires this term to be constant and less than one, which is incompatible with (2.17). In this case, Property A does not hold.

¹²Because $1 - F(t_1) = 0$, this claim is trivial if $f(t_1) > 0$. More generally, $1 - F(t_1) = 0$ implies $\lim_{t \uparrow t_1} \ln(1 - F(t)) = -\infty$. Therefore, $\lim_{t \uparrow t_1} \frac{d}{dt} \ln(1 - F(t)) = -\infty$ and, hence, $\lim_{t \uparrow t_1} \frac{1-F(t)}{f(t)} = 0$,

The crucial difference between this example and Theorem 2.1 is in the behaviour of the ratio $\frac{1-F(t)}{f(t)}$ when t converges to t_1 . In the example, this ratio goes out of bounds, in the setting of Theorem 2.1, with a compact type set, it necessarily goes to zero as t converges to t_1 . For the utility specification (2.14), the arguments of Werning (2007) imply that, even with $t_1 = \infty$, Property A, with (2.17) replacing (2.8), is obtained if $\frac{1-F(t)}{tf(t)}$ goes to zero as $t \rightarrow \infty$ and if the relative curvature $\frac{\gamma'' \frac{y}{t}}{\gamma'}$ of the effort cost function is bounded. I conjecture that this is generally true whenever $\frac{u_{yt}}{u_y}$ is uniformly bounded and $\frac{1-F(t)}{f(t)}$ goes to zero as t becomes large. Property A is properly understood as the statement that, near the top of the type distribution, the tradeoff between distributive and allocative concerns is degenerate if the ratio $\frac{1-F(t)}{f(t)}$ is close to zero. Relative to the density of the type t that is affected by a distortion at t , the weight of the set of types above t from whom the principal can extract more rents as a result of the distortion at t is close to zero. Therefore, it is undesirable to have any significant distortion.

2.3 Mass Points, Pooling and Discontinuities in Optimal Contract Menus

The following results establish some additional properties of optimal contract menus. These properties arise *only* when the type distribution has both mass points and a continuous part.

Theorem 2.3 *Let $(w(\cdot), y(\cdot))$ be any optimal contract menu, and let $(\bar{w}(\cdot), \bar{y}(\cdot))$ and $\hat{t} \in [t_0, t_1]$ be the associated strongly equivalent contract menu and critical type as given by Theorem 2.1. If $F(\{t\}) > 0$ for some $t \in [t_0, \hat{t}]$, there exists $\bar{t} \in (t, \hat{t})$ such that, on the interval $[t, \bar{t}]$, the functions $\bar{w}(\cdot), \bar{y}(\cdot)$, and $\bar{y}(\cdot) - \bar{w}(\cdot)$ are constant; in particular, if $F((t, \bar{t})) > 0$, the menus $(w(\cdot), y(\cdot))$ and $(\bar{w}(\cdot), \bar{y}(\cdot))$ both provide for pooling of type t with adjacent higher types.*

The rationale for this result has been sketched in the introduction: If there was no pooling with higher types, the usual tradeoff between the distributive effects and the allocative effects of a downward distortion in the contract $(\bar{w}(t), \bar{y}(t))$ for a type t that has positive mass would be degenerate. Type t would be deemed to have so much weight that a distortion in $(\bar{w}(t), \bar{y}(t))$, away from efficiency, would seem undesirable. However, for a type t' just above t that does not have positive mass, standard arguments imply that the contract $(\bar{w}(t'), \bar{y}(t'))$ is distorted downward from efficiency. The resulting contract menu though, with an efficient outcome for t and a downward distortion for $t' > t$ would be decreasing and would violate incentive compatibility. The assumption that type t is not pooled with higher types thus leads to a contradiction.

By contrast, the monotonicity requirement does not preclude *upward* jumps in the optimal contract menu. The observation that a mass point is incomparably more important than any continuity point of the type distribution

implies that, as one moves from immediately adjacent *lower* types to the mass point, the weights given to losses from distortions of efficiency and to gains from alleviating incentive constraints change discontinuously. As illustrated in Figure 1 in the introduction, this induces an upward jump in the contract menu.

Proposition 2.4 *Assume that, for any w and y , the function $t \rightarrow \ln \frac{|u_y(w,y,t)|}{u_w(w,y,t)}$ is convex.¹³ Assume also that the density f_a of the absolutely continuous component F_a of the type distribution F is strictly positive and nondecreasing on $[t_0, t_1]$.*

Let $(w(\cdot), y(\cdot))$ be any optimal contract menu, and let $(\bar{w}(\cdot), \bar{y}(\cdot))$ and $\hat{t} \in [t_0, t_1]$ be the associated strongly equivalent contract menu and critical type as given by Theorem 2.1. If $F(\{t\}) > 0$ for some $t \in [t_0, \hat{t})$, consider the lowest type that gets the same outcome as t , i.e.,

$$\underline{t} := \inf\{t' | (\bar{w}(t'), \bar{y}(t')) = (\bar{w}(t), \bar{y}(t))\}.$$

If $\underline{t} > 0$, the contract menu $\bar{w}(\cdot)$ and $\bar{y}(\cdot)$ is discontinuous at \underline{t} . If $\underline{t} > 0$ and $F((\underline{t}-\Delta, \underline{t}-\Delta)) > 0$ for all $\Delta > 0$, the contract menu $(w(\cdot), y(\cdot))$ is also discontinuous at \underline{t} .

In Proposition 2.4, the additional assumptions, log-convexity of $\frac{|u_y(w,y,t)|}{u_w(w,y,t)}$ in t and monotonicity of the density f_a , are introduced to eliminate the possibility that the mass point might belong to the interior of an ironing interval à la Guesnerie-Laffont (1984). Such an interval would provide for a pooling of types even when the type distribution has a continuous density. The additional assumptions imply that the contracts offered to higher types necessarily involve strictly greater outcomes than the contracts offered to lower *unless* the higher types are pooled with mass points. Without mass points, these additional assumptions imply that there is no pooling. This is formally stated as:

Proposition 2.5 *Assume that, for any w and y , the function $t \rightarrow \ln \frac{|u_y(w,y,t)|}{u_w(w,y,t)}$ is convex. Assume also that the type distribution has a density, and that this density is strictly positive and nondecreasing on $[t_0, t_1]$.*

If $(w(\cdot), y(\cdot))$ is any optimal contract menu, and if $(\bar{w}(\cdot), \bar{y}(\cdot))$ is the associated strongly equivalent contract menu and critical type as given by Theorem 2.1, then both menus $(w(\cdot), y(\cdot))$ and $(\bar{w}(\cdot), \bar{y}(\cdot))$ are strictly increasing on $[t_0, \hat{t}]$ where \hat{t} is the critical type above which outcomes involve no distortion away from efficiency.

Whereas log-convexity of $\frac{|u_y(w,y,t)|}{u_w(w,y,t)}$ in t and monotonicity of the density provide for *monotonicity* of optimal contract menus, *continuity* requires a different set of assumptions. Adapting an argument of Mirrlees (1986, pp. 1231 f.), one obtains

¹³This assumption is satisfied, for example, by the quasi-linear specification $u(w, y, t) = w - \frac{1}{t}g(y)$. For this specification, $\ln \frac{|u_y(w,y,t)|}{u_w(w,y,t)} = \ln g'(y) - \ln t$ is strictly convex in t .

Proposition 2.6 *Assume that, for any v and t , the function $y \rightarrow u_t(c(v, y, t), y, t)$ is convex, where $c(v, y, t)$ is defined so that $u(c(v, y, t), y, t) = v$ for any v, y, t . If the type distribution has a continuous density, any optimal contract menu is continuous.*

Convexity of the function $y \rightarrow u_t(c(v, y, t), y, t)$ is implied by convexity of the function $(v, y) \rightarrow u_t(c(v, y, t), y, t)$. As discussed in Hellwig (2004, 2007 b), this latter condition is equivalent to the assumption that consumption-specific risk aversion be weakly decreasing in t , i.e., that for any random pair (\tilde{w}, \tilde{y}) the amount of consumption that the agent is willing to sacrifice in order to eliminate the uncertainty in (\tilde{w}, \tilde{y}) be nonincreasing in t . This assumption also implies that, up to equivalence, the optimal contract menu is unique. Moreover, the optimal deterministic contract menu remains optimal if randomization is allowed; randomization is undesirable.

2.4 Three Preliminary Lemmas

The remainder of the paper provides formal proofs of the theorems and propositions. I begin by stating three lemmas that allow me to replace incentive compatibility and individual-rationality constraints by analytically tractable conditions on the indirect utility function $v(\cdot)$ and by a monotonicity condition on outcomes, along the lines of Mirrlees (1976). These lemmas are proved in the online Appendix.

The first result shows that, even if the support of the type distribution is a strict subset of the interval $[t_0, t_1]$, one can always formulate the principal's problem in terms of contract menus that are defined on the interval $[t_0, t_1]$. Refer to a contract menu with domain X as *incentive-compatible on X* if condition IC holds for all t and t' in X ; refer to it as *individually rational on X* if condition IR holds for all t in X . Then one obtains:

Lemma 2.7 *A contract menu $(w(\cdot), y(\cdot))$ that is defined on $T \subset [t_0, t_1]$ is incentive-compatible and individually rational on T if and only if there exists an extension of $(w(\cdot), y(\cdot))$ to the interval $[t_0, t_1]$ that is incentive-compatible and individually rational on $[t_0, t_1]$.*

This lemma implies that there is no loss of generality in assuming that all contract menus are defined on the interval $[t_0, t_1]$ and have to satisfy incentive compatibility and individual-rationality conditions on this interval. For $t \in [t_0, t_1] \setminus T$, contracts $(w(t), y(t))$ can always be chosen so that conditions IC and IR hold. Incentive-compatibility and individual-rationality requirements for such types do not add materially to the principal's constraints. These types' contracts are of course irrelevant for the principal's payoff expectations.

For contract menus that are defined on $[t_0, t_1]$, the arguments of Mirrlees (1976) are easily adapted to yield the following characterization of incentive compatibility.

Lemma 2.8 *A nondecreasing contract menu $(w(\cdot), y(\cdot))$ is incentive-compatible and individually rational on $[t_0, t_1]$ if and only if the induced indirect utility function $v(\cdot)$ satisfies the integral equation*

$$v(t) = v(t_0) + \int_{t_0}^t u_t(w(\tau), y(\tau), \tau) d\tau \quad (2.19)$$

for $t \in [t_0, t_1]$ and, moreover,

$$v(t_0) \geq 0. \quad (2.20)$$

In Lemma 2.8, weak monotonicity of the contract menu is *assumed*. Under a *strict* single-crossing condition, weak monotonicity is in fact known to be necessary for incentive compatibility. Here, with only a *weak* single-crossing condition, this is not the case. However, with strictly convex indifference curves of the agent, the principal does not want to implement a non-monotonic contract menu.¹⁴ This is the point of the following lemma.

Lemma 2.9 *For any incentive-compatible contract menu $(w(\cdot), y(\cdot))$, there exists a nondecreasing incentive-compatible contract menu $(\bar{w}(\cdot), \bar{y}(\cdot))$ that provides the agent with the same payoff $v(t) = u(w(t), y(t), t)$ for all t and that satisfies*

$$\int [\bar{y}(t) - \bar{w}(t)] dF(t) \geq \int [y(t) - w(t)] dF(t); \quad (2.21)$$

moreover, the inequality in (2.21) is strict unless the contract menus $(w(\cdot), y(\cdot))$ and $(\bar{w}(\cdot), \bar{y}(\cdot))$ are equivalent.

3 A Reformulation of the Principal's Problem

Lemmas 2.7 - 2.8 imply that the principal's problem is equivalent to the problem of choosing $w(\cdot), y(\cdot)$, and $v(\cdot)$ so as to maximize (2.3) under the constraints that $y(\cdot)$ be nondecreasing and that $v(\cdot) = u(w(\cdot), y(\cdot), \cdot)$ satisfy the integral equation (2.19) and the boundary condition (2.20). With a slight abuse of language, I will refer to this problem also as *the principal's problem*.

The integral equation (2.19) is equivalent to the requirement that $v(\cdot)$ be absolutely continuous, with Radon-Nikodym derivative

$$v'(t) = u_t(w(t), y(t), t) \quad (3.1)$$

for almost all t . One is therefore tempted to treat the principal's problem as a problem of optimal control with state variable v and control variables w and y . This would be the natural way to proceed if $y(\cdot)$ wasn't required to be nondecreasing and if F had a density. Here, however, a direct application of control theoretic methods is precluded by the monotonicity requirement on $y(\cdot)$ and the lack of any structure on F . I will therefore reformulate the principal's problem so as to circumvent these difficulties.

¹⁴The logic is the same as in Figure 2 below.

For this purpose, I change the variable of integration in (2.3), using a new variable x , rather than t , as the argument of the functions that are to be chosen. In a sense, this amounts to a redefinition of the notion of "type". The new "pseudotype" is constructed so that its distribution has a density. This density, however, need not be continuous.

For any $t \in [t_0, t_1]$, set

$$\xi(t) := t + F(t). \quad (3.2)$$

The function ξ is strictly increasing and has an inverse $\tau = \xi^{-1}$. The inverse is defined on the range of ξ , a subset of the interval $[x_0, x_1] := [t_0, t_1 + 1]$. Using the fact that $F(\cdot)$ and therefore also $\xi(\cdot)$ is right-continuous, as well as increasing, one can extend the inverse τ to the entire interval $[x_0, x_1]$ by setting

$$\tau(x) = \sup_s \{s | \xi(s) \leq x\}. \quad (3.3)$$

If one sets $\tilde{x} = \xi(\tilde{t})$, one has $\tilde{t} = \tau(\tilde{x})$. The distribution of \tilde{x} is $G := F \circ \xi^{-1}$, and the distribution F of \tilde{t} satisfies $F = G \circ \tau^{-1} = G \circ \xi$. By the change-of-variable formula, it follows that, for any function h on $[t_0, t_1]$,

$$\int_{t_0}^{t_1} h(t) dF(t) = \int_{x_0}^{x_1} h(\tau(x)) dG(x). \quad (3.4)$$

The following lemma shows that G has a density so that (3.4) can actually be written in the form

$$\int_{t_0}^{t_1} h(t) dF(t) = \int_{x_0}^{x_1} h(\tau(x)) g(x) dx. \quad (3.5)$$

Lemma 3.1 *The function $\tau(\cdot)$ that is defined by (3.2) and (3.3) is absolutely continuous. Its derivative $\tau'(\cdot)$ satisfies*

$$\tau'(x) = \frac{1}{1 + f(\tau(x))} \quad (3.6)$$

if, at $t = \tau(x)$, the derivative $F'(t) = f(t)$ is well defined, and

$$\tau'(x) = 0 \quad (3.7)$$

otherwise. The distribution function $G = F \circ \xi^{-1}$ is also absolutely continuous. Its density $g(x) = G'(x)$ satisfies

$$g(x) = 1 - \tau'(x) \quad (3.8)$$

for all x .

Proof. From (3.2), (3.3), and the definition of G , one has

$$x = \tau(x) + F(\tau(x)) = \tau(x) + G(x) \quad (3.9)$$

for all $x \in [x_0, x_1]$. Since $\tau(\cdot)$ and $G(\cdot)$ are both nondecreasing, it follows that both are Lipschitz continuous, hence absolutely continuous. Moreover, their slopes must add to one.

For any x , $\tau(x + \Delta) > \tau(x - \Delta)$ for all $\Delta > 0$ implies

$$1 = \frac{\tau(x + \Delta) - \tau(x - \Delta)}{2\Delta} + \frac{F(\tau(x + \Delta)) - F(\tau(x - \Delta))}{2\Delta},$$

hence

$$\lim_{\Delta \rightarrow 0} \frac{\tau(x + \Delta) - \tau(x - \Delta)}{2\Delta} = \frac{1}{1 + \lim_{\Delta \rightarrow 0} \frac{F(\tau(x + \Delta)) - F(\tau(x - \Delta))}{\tau(x + \Delta) - \tau(x - \Delta)}},$$

which yields (3.6) if

$$\lim_{\Delta \rightarrow 0} \frac{F(\tau(x + \Delta)) - F(\tau(x - \Delta))}{\tau(x + \Delta) - \tau(x - \Delta)} = f(\tau(x))$$

is well defined and (3.7) if

$$\lim_{\Delta \rightarrow 0} \frac{F(\tau(x + \Delta)) - F(\tau(x - \Delta))}{\tau(x + \Delta) - \tau(x - \Delta)} = \infty.$$

Trivially, (3.7) holds also if $\tau(x + \Delta) = \tau(x - \Delta)$ for some $\Delta > 0$. For this case, (3.3) indicates that ξ and F are discontinuous at $t = \tau(x + \Delta) = \tau(x - \Delta)$. The derivative $F'(t)$ is then not well defined at t . ■

For any contract menu $(w(\cdot), y(\cdot))$, (3.3) and (3.5) imply that the principal's payoff (2.3) can be rewritten as

$$\int_{x_0}^{x_1} [y(\tau(x)) - w(\tau(x))]g(x)dx. \quad (3.10)$$

Moreover, if \hat{u} is defined so that

$$\hat{u}(w, y, x) := u(w, y, \tau(x)) \quad (3.11)$$

for all w, y, x , (2.4), (2.19), and (2.20) are equivalent to the conditions

$$v(\tau(x)) = \hat{u}(w(\tau(x)), y(\tau(x)), x), \quad (3.12)$$

$$v'(\tau(x))\tau'(x) = \hat{u}_x(w(\tau(x)), y(\tau(x)), x), \quad (3.13)$$

and

$$v(\tau(t_0)) \geq 0. \quad (3.14)$$

Now (3.10) - (3.14) can be rewritten as

$$\int_{x_0}^{x_1} [\hat{y}(x) - \hat{w}(x)]g(x)dx, \quad (3.15)$$

$$\hat{v}(x) = \hat{u}(\hat{w}(x), \hat{y}(x), x), \quad (3.16)$$

$$\hat{v}'(x) = \hat{u}_x(\hat{w}(x), \hat{y}(x), x), \quad (3.17)$$

and

$$\hat{v}(t_0) \geq 0, \quad (3.18)$$

where

$$\hat{w} := w \circ \tau, \hat{y} := y \circ \tau, \text{ and } \hat{v} := v \circ \tau. \quad (3.19)$$

Consider the problem of maximizing (3.15) under the constraints that $\hat{y}(\cdot)$ be nondecreasing and that $\hat{v}(\cdot)$ satisfy (3.16) - (3.18). For lack of a better term, I call this *the principal's modified problem*. The following proposition shows that this problem is actually equivalent to the principal's problem.

Proposition 3.2 *A contract menu $(w(\cdot), y(\cdot))$, with associated indirect utility function $v(\cdot)$, solves the principal's problem, if and only if the functions $\hat{w}(\cdot), \hat{y}(\cdot)$, and $\hat{v}(\cdot)$ that are given by (3.2), (3.3), and (3.19) solve the principal's modified problem.*

Proof. The principal's problem has been shown to be equivalent to the problem of choosing $w(\cdot), y(\cdot)$, and $v(\cdot)$ so as to maximize (3.10) under the constraints that $y(\cdot)$ be nondecreasing and that $v(\cdot)$ satisfy conditions (3.11) - (3.14). This problem is equivalent to the problem of choosing functions $\hat{w}(\cdot), \hat{y}(\cdot)$, and $\hat{v}(\cdot)$ to maximize (3.15) under the constraints that $\hat{y}(\cdot)$ be nondecreasing, that $\hat{v}(\cdot)$ satisfy (3.16) - (3.18), with \hat{u} given by (3.11), and under the constraint that $\hat{w}(\cdot), \hat{y}(\cdot)$, and $\hat{v}(\cdot)$ can be represented in the form (3.19) for some functions $w(\cdot), y(\cdot)$, and $v(\cdot)$. This latter problem is the same as the principal's modified problem with the added constraint $\hat{w}(\cdot), \hat{y}(\cdot)$, and $\hat{v}(\cdot)$ take the form (3.19) for some functions $w(\cdot), y(\cdot)$, and $v(\cdot)$. To prove the proposition, it is therefore sufficient to show that the added constraint is redundant because, up to modifications on null sets, any solution $\hat{w}, \hat{y}, \hat{v}$ to the principal's modified problem satisfies (3.19) for some functions w, y, v .

For this purpose, I will show that, for almost all x_1 and x_2 , $\tau(x_1) = \tau(x_2)$ implies $\hat{w}(x_1) = \hat{w}(x_2)$, $\hat{y}(x_1) = \hat{y}(x_2)$, and $\hat{v}(x_1) = \hat{v}(x_2)$. From (3.17) and (3.11), one has

$$\hat{v}'(x) = u_t(\hat{w}(x), \hat{y}(x), \tau(x)) \tau'(x).$$

Since $\tau(x_1) = \tau(x_2)$ implies $\tau'(x) = 0$ for almost all $x \in [x_1, x_2]$, it follows that $\tau(x_1) = \tau(x_2)$ implies $\hat{v}(x_1) = \hat{v}(x_2)$. By standard arguments,¹⁵ it follows that there exists a function v such that $\hat{v}(x) = v(\tau(x))$ for all x .

Next, consider the function w^* such that

$$w^*(t) = E[\hat{w}(\tilde{x}) | \tau(\tilde{x}) = t] \quad (3.20)$$

for all t , where \tilde{x} is distributed as G . By the definition of the conditional expectation, one has

$$\int_{x_0}^{x_1} w^*(\tau(x)) dG(x) = \int_{x_0}^{x_1} \hat{w}(x) dG(x). \quad (3.21)$$

¹⁵See, e.g., Result (8), p. 43, in Hildenbrand (1974).

Given w^* , consider also the function y^* such that

$$u(w^*(t), y^*(t), t) = v(t) \quad (3.22)$$

for all t . By the definition of v , one also has

$$u(\hat{w}(x), \hat{y}(x), \tau(x)) = v(\tau(x)) \quad (3.23)$$

for all x . By the strict quasi-concavity of u , (3.22), (3.20) and (3.23) imply

$$y^*(t) \geq E[\hat{y}(\tilde{x}) | \tau(\tilde{x}) = t] \quad (3.24)$$

for all t , hence

$$\int_{x_0}^{x_1} y^*(\tau(x)) dG(x) \geq \int_{x_0}^{x_1} \hat{y}(x) dG(x). \quad (3.25)$$

Moreover, the inequality in (3.25) is strict unless one has $w^*(\tau(x)) = \hat{w}(x)$ for G -almost all x .

If the inequality in (3.25) is strict, one has

$$\int_{x_0}^{x_1} [y^*(\tau(x)) - w^*(\tau(x))] dG(x) > \int_{x_0}^{x_1} [\hat{y}(x) - \hat{w}(x)] dG(x). \quad (3.26)$$

Given that, trivially, the triple $w^* \circ \tau, y^* \circ \tau, \hat{v}$ has $y^* \circ \tau$ nondecreasing and $\hat{v} = v \circ \tau$ satisfying (3.16) - (3.18), (3.26) is incompatible with the assumption that $\hat{w}, \hat{y}, \hat{v}$ maximizes (3.15) subject to the constraints that \hat{y} be nondecreasing and that \hat{v} satisfy (3.16) - (3.18). Therefore, the inequality in (3.25) cannot be strict. It follows that $w^*(\tau(x)) = \hat{w}(x)$ and, by (3.22), $y^*(\tau(x)) = \hat{y}(x)$ for G -almost all x , as claimed in the lemma. ■

The argument is illustrated in Figure 2. If contracts are conditioned on x rather than t , the principal has room to offer a richer contract menu. In particular, if t is a mass point of the distribution F , the function $\xi(\cdot)$ is discontinuous at t , and the principal can assign different contracts to different pseudotypes $x \in (\xi(t-), \xi(t)]$. Thus, he might offer different contracts $(\hat{w}(x), \hat{y}(x))$ for $x \in (\xi(t-), \xi(t)]$, so that $\hat{w}(x)$ is uniformly distributed between w_1 and w_2 in Figure 2. However, such an arrangement cannot be optimal for him. Because all pseudotypes $x \in (\xi(t-), \xi(t)]$ correspond to the same real type $\tau(x) = t$, incentive compatibility requires that all the contracts $(\hat{w}(x), \hat{y}(x))$ for $x \in (\xi(t-), \xi(t)]$ provide the agent of type t with the same utility. Thus, in Figure 1, the contract offers $(\hat{w}(x), \hat{y}(x))$ for $x \in (\xi(t-), \xi(t)]$ all lie on the same indifference curve $I(t)$ for type t . Strict quasi-concavity of u implies that the indifference curve $I(t)$ is strictly convex. If the principal replaces the wage offers $\hat{w}(x)$ for $x \in (\xi(t-), \xi(t)]$ by their (conditional) expectation $w^* = (w_1 + w_2)/2$, he can ask for an output y^* that is strictly greater than the (conditional) expectation of $\hat{y}(x)$, $x \in (\xi(t-), \xi(t)]$. By introducing heterogeneity into the contract offers to people with the same "real" type t , the principal can only harm himself.

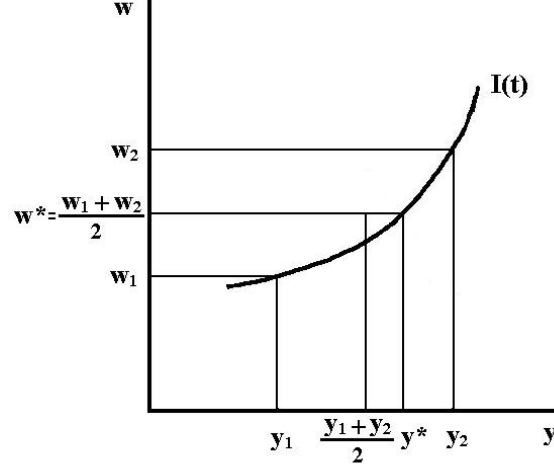


Figure 2: Multiple Contracts for Different Pseudotypes with the Same Type

4 Analysis of the Principal's Modified Problem

4.1 Preliminaries

The principal's modified problem has the same formal structure as the principal's problem itself when F has a density. However, the density $g(\cdot)$ in (3.15) is not, in general, continuous. Moreover, the functions \hat{u} , \hat{u}_w , and \hat{u}_y are not generally continuously differentiable with respect to x . From (3.16), however, one easily verifies that \hat{u} , like u , is twice continuously differentiable and strictly quasi-concave in w and y jointly, as well as increasing in w and decreasing in y . In particular, one has

$$\hat{u}_w(w, y, x) = u_w(w, y, \tau(x)) \quad (4.1)$$

and

$$\hat{u}_y(w, y, x) = u_y(w, y, \tau(x)) \quad (4.2)$$

for any w, y , and x . However, from (4.1) and (4.2), one obtains

$$\frac{\partial |\hat{u}_y(w, y, x)|}{\partial x \hat{u}_w(w, y, x)} = \frac{\partial |u_y(w, y, \tau(x))|}{\partial t u_w(w, y, \tau(x))} \tau'(x), \quad (4.3)$$

which has a discontinuity with respect to x whenever τ' and the density $g = 1 - \tau'$ have a discontinuity. By (2.2) and Lemma 3.1, (4.3) yields

$$\frac{\partial |\hat{u}_y(w, y, x)|}{\partial x \hat{u}_w(w, y, x)} \leq 0 \quad (4.4)$$

for all w, y , and x . Thus, \hat{u} inherits the weak single-crossing property from u .

4.2 Optimality Conditions

If the function $\hat{y}(\cdot)$ was known to be absolutely continuous, the principal's modified problem would be a standard control problem with $\hat{v}(\cdot)$ and $\hat{y}(\cdot)$ as state variables and with $\hat{w}(\cdot)$ and $q(\cdot) := \hat{y}'(\cdot)$, as control variables, with the constraint that $q(x) \geq 0$ for all x . The Hamiltonian function for this control problem would take the form

$$\begin{aligned} \hat{H}(\hat{v}, \hat{y}, \hat{w}, \hat{q}, \hat{\mu}, \hat{\varphi}_v, \hat{\varphi}_q, x) &= (\hat{y} - \hat{w})g(x) + \hat{\mu} (\hat{u}(\hat{w}, \hat{y}, x) - \hat{v}) \\ &\quad + \hat{\varphi}_v \hat{u}_x(\hat{w}, \hat{y}, x) + \hat{\varphi}_q q. \end{aligned} \quad (4.5)$$

However, this Hamiltonian is not, in general, continuous in x . In particular, if $\tau(x)$ is a mass point of the type distribution F and $x' < x$ implies $\tau(x') < \tau(x)$, the density $g(\cdot)$ will exhibit a discontinuity at x , and so will the Hamiltonian \hat{H} . If such a discontinuity arises, there is no reason for $\hat{y}(\cdot)$ to be absolutely continuous.

Even so, the principal's modified problem can be handled by control-theoretic methods. In Hellwig (2008), I formulate a maximum principle for optimal control problems with monotonicity constraints on the controls. When applied to the principal's modified problem, Theorem 3.1 in Hellwig (2008) yields:

Theorem 4.1 *If the functions $\hat{w}(\cdot)$, $\hat{y}(\cdot)$, and $\hat{v}(\cdot)$ solve the principal's modified problem, then there exists a measurable function $\hat{\mu}$ from $[x_0, x_1]$ into \mathbb{R}_+ , and there exist absolutely continuous functions $\hat{\varphi}_v$ and $\hat{\varphi}_q$ from $[x_0, x_1]$ into \mathbb{R} such that the following hold:*

(a) *For almost all $x \in [x_0, x_1]$,*

$$\hat{\varphi}'_v(x) = -\hat{H}_v(\hat{v}(x), \hat{y}(x), \hat{w}(x), \hat{q}(x), \hat{\mu}(x), \hat{\varphi}_v(x), \hat{\varphi}_q(x), x); \quad (4.6)$$

(b) *moreover,*

$$\hat{\varphi}_v(x_0) \leq 0, \quad \hat{\varphi}_v(x_0)\hat{v}(x_0) = 0, \quad (4.7)$$

and

$$\hat{\varphi}_v(x_1) = 0. \quad (4.8)$$

(c) *For almost all $x \in [x_0, x_1]$,*

$$\hat{\varphi}'_q(x) = -\hat{H}_y(\hat{v}(x), \hat{y}(x), \hat{w}(x), \hat{q}(x), \hat{\mu}(x), \hat{\varphi}_v(x), \hat{\varphi}_q(x), x); \quad (4.9)$$

(d) *moreover,*

$$\hat{\varphi}_q(x_0) \leq 0, \quad \hat{\varphi}_q(x_0) \cdot \hat{y}(x_0) = 0, \quad (4.10)$$

and

$$\hat{\varphi}_q(x_1) = 0. \quad (4.11)$$

(e) For almost every $x \in [x_0, x_1]$,

$$-g(x) + \hat{\mu}(x)\hat{u}_w(\hat{w}(x), \hat{y}(x), x) + \hat{\varphi}_v(x)\hat{u}_{xw}(\hat{w}(x), \hat{y}(x), x) = 0. \quad (4.12)$$

(f) For almost every $x \in [x_0, x_1]$,

$$\hat{\varphi}_q(x) \leq 0; \quad (4.13)$$

moreover, $\hat{\varphi}_q(x) = 0$ if $\hat{y}(\cdot)$ is strictly increasing at x .¹⁶

Proof. Reformulate the principal's modified problem by substituting for

$$\hat{w}(x) = c(\hat{v}(x), \hat{y}(x), x)$$

for all x , where $c(\cdot, \cdot, \cdot)$ be defined so that for any y and t , $c(\cdot, y, t)$ is the inverse of the section $u(\cdot, y, t)$ of u that is determined by y and t . The modified problem then satisfies the assumptions of Theorem 3.1 in Hellwig (2008). The associated Hamiltonian is

$$H = (\hat{y} - c(\hat{v}, \hat{y}, x))g(x) + \hat{\varphi}_v(x)\hat{u}_x(c(\hat{v}, \hat{y}, x), \hat{y}, x) + \hat{\varphi}_q q. \quad (4.14)$$

Theorem 3.1 in Hellwig (2008) implies the existence of $\hat{\varphi}_v, \hat{\varphi}_q$, and M such that, for almost all $x \in [x_0, x_1]$, one has

$$\hat{\varphi}'_v(x) = -H_v, \quad (4.15)$$

$$\hat{\varphi}'_q(x) = -H_y, \quad (4.16)$$

and, in addition, the transversality conditions (b) and (d), as well as statements (e) and (f) hold. I also introduce

$$\hat{\mu}(x) = (g(x) - \hat{\varphi}_v(x)\hat{u}_{xw})\frac{1}{\hat{u}_w}, \quad (4.17)$$

defined so that (4.12) holds.

From (4.14) - (4.17), one obtains

$$\begin{aligned} \hat{\varphi}'_v(x) &= \frac{\partial c}{\partial v}(\hat{v}(x), \hat{y}(x), x) - \hat{\varphi}_v(x)\hat{u}_{xw}\frac{\partial c}{\partial v}(\hat{v}(x), \hat{y}(x), x) \\ &= (g(x) - \hat{\varphi}_v(x)\hat{u}_{xw})\frac{1}{\hat{u}_w} = \hat{\mu}(x) \end{aligned}$$

and

$$\begin{aligned} \hat{\varphi}'_q(x) &= -(1 - \frac{\partial c}{\partial y}(\hat{v}(x), \hat{y}(x), x))g(x) - \hat{\varphi}_v(x)\hat{u}_{xy}\frac{\partial c}{\partial y}(\hat{v}(x), \hat{y}(x), x) \\ &= -(1 + \frac{\hat{u}_y}{\hat{u}_w})g(x) + \hat{\varphi}_v(x)\hat{u}_{xy}\frac{\hat{u}_y}{\hat{u}_w}. \\ &= -g(x) - \hat{\mu}(x)\hat{u}_y(\hat{w}(x), \hat{y}(x), x) - \hat{\varphi}_v(x)\hat{u}_{xy}(\hat{w}(x), \hat{y}(x), x) \end{aligned} \quad (4.18)$$

¹⁶A real-valued, nondecreasing function f is said to be *strictly increasing* at t if $f(t + \varepsilon) - f(t - \varepsilon) > 0$ for all $\varepsilon > 0$.

for almost all $x \in [x_0, x_1]$. Statements (a) and (c) follow immediately. ■

In the following, I will suppose that $\hat{w}(\cdot)$, $\hat{y}(\cdot)$, and $\hat{v}(\cdot)$ provide a solution to the principal's modified problem, and that $\hat{\mu}(\cdot)$, $\hat{\varphi}_v(\cdot)$, and $\hat{\varphi}_q(\cdot)$ are the associated Lagrange multiplier and costate variables.

Given the Inada condition $u_y(w, 0, t) = 0$ for all w and t , one does not have to worry about boundary solutions.

Lemma 4.2 *Any solution to the principal's modified problem satisfies $\hat{w}(x) > 0$ and $\hat{y}(x) > 0$ for all $x \in (x_0, x_1]$.*

Proof. I first show that $\hat{y}(x) > 0$ for all $x \in (x_0, x_1]$. For suppose that this claim is false. Then, because $\hat{y}(\cdot)$ is nondecreasing, there exists $\bar{x} \in (x_0, x_1]$ such that $\hat{y}(x) = 0$ for $x \in [x_0, \bar{x}]$ and $\hat{y}(x) > 0$ for $x > \bar{x}$. Because $\hat{y}(\cdot)$ is strictly increasing at \bar{x} , statement (f) of Theorem 4.1 yields

$$\hat{\varphi}_q(\bar{x}) = 0. \quad (4.19)$$

Because $u_y(w, 0, t) = 0$ for all w and t , $\hat{y}(x) = 0$ for $x \in [x_0, \bar{x}]$ also implies

$$\hat{u}_y(\hat{w}(x), \hat{y}(x), x) = u_y(\hat{w}(x), 0, \tau(x)) = 0$$

and

$$\hat{u}_{xy}(\hat{w}(x), \hat{y}(x), x) = u_{yt}(\hat{w}(x), 0, \tau(x))\tau'(x) = 0$$

for all $x \in [x_0, \bar{x}]$. By (4.18), therefore,

$$\psi'(x) = -g(x)$$

for all $x \in (x_0, \bar{x})$. By integration, using (4.19), one obtains

$$\hat{\varphi}_q(0) = \hat{\varphi}_q(\bar{x}) - \int_0^{\bar{x}} \hat{\varphi}'_q(x') dx' = G(\bar{x}). \quad (4.20)$$

Because $\bar{x} > x_0$, by the definitions of G and x_0 , one has $G(\bar{x}) = F(\tau(\bar{x})) > 0$. Therefore, (4.20) is incompatible with (4.13). The assumption that $\hat{y}(x) = 0$ for some $x \in (x_0, x_1]$ thus leads to a contradiction and must be false.

Given that $\hat{y}(x) > 0$ for $x \in (x_0, x_1]$, the individual-rationality condition $\hat{u}(\hat{w}(x), \hat{y}(x), x) \geq 0$ also yields $\hat{w}(x) > 0$ for $x \in (x_0, x_1]$. ■

Using (4.6), one can rewrite (4.12) and (4.18) as

$$\hat{\varphi}'_v(x)\hat{u}_w + \hat{\varphi}_v(x)\hat{u}_{wx} = g(x) \quad (4.21)$$

and

$$\hat{\varphi}'_q(x) = -g(x) - \hat{\varphi}'_v(x)\hat{u}_y - \hat{\varphi}_v(x)\hat{u}_{xy}, \quad (4.22)$$

where \hat{u}_w , \hat{u}_y , \hat{u}_{wx} , and \hat{u}_{yx} are all evaluated at $(\hat{w}(x), \hat{y}(x), x)$. If one uses (4.21) to substitute for $\hat{\varphi}'_v(x)$ in (4.22), one obtains:

$$\hat{\varphi}'_q(x) = -\left(1 + \frac{\hat{u}_y}{\hat{u}_w}\right)g(x) - \hat{\varphi}_v(x)\left(\hat{u}_{yx} - \frac{\hat{u}_y}{\hat{u}_w}\hat{u}_{wx}\right), \quad (4.23)$$

or, equivalently,

$$\hat{\varphi}'_q(x) = -\frac{\hat{u}_w + \hat{u}_y}{\hat{u}_w} g(x) + \hat{\varphi}_v(x) \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w}. \quad (4.24)$$

4.3 Analysis of the Optimality Conditions

Equation (4.24) is the key to assessing the efficiency properties of the contracts $(\hat{w}(x), \hat{y}(x))$, $x \in (x_0, x_1]$. A contract $(\hat{w}(x), \hat{y}(x)) \in \mathbb{R}_{++}^2$ is efficient for x if

$$\hat{u}_w(\hat{w}(x), \hat{y}(x), x) + \hat{u}_y(\hat{w}(x), \hat{y}(x), x) = 0, \quad (4.25)$$

distorted downwards from efficiency for x if

$$\hat{u}_w(\hat{w}(x), \hat{y}(x), x) + \hat{u}_y(\hat{w}(x), \hat{y}(x), x) > 0, \quad (4.26)$$

and distorted upwards from efficiency for x if

$$\hat{u}_w(\hat{w}(x), \hat{y}(x), x) + \hat{u}_y(\hat{w}(x), \hat{y}(x), x) < 0. \quad (4.27)$$

By (4.24), therefore, $(\hat{w}(x), \hat{y}(x))$ is efficient for x if

$$\hat{\varphi}'_q(x) - \hat{\varphi}_v(x) \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} = 0, \quad (4.28)$$

distorted downwards from efficiency for x if

$$\hat{\varphi}'_q(x) - \hat{\varphi}_v(x) \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} < 0, \quad (4.29)$$

and distorted upwards from efficiency for x if

$$\hat{\varphi}'_q(x) - \hat{\varphi}_v(x) \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} > 0. \quad (4.30)$$

The function $\hat{\varphi}_v(\cdot)$ is given as the solution to the differential equation (4.21) that satisfies the transversality condition $\hat{\varphi}_v(x_1) = 0$. This solution is computed as

$$\hat{\varphi}_v(x) = - \int_x^{x_1} \frac{g(x')}{\hat{u}_w(\hat{w}(x'), \hat{y}(x'), x')} \exp \left(\int_x^{x'} \frac{\hat{u}_{xw}(\hat{w}(x''), \hat{y}(x''), x'')}{\hat{u}_w(\hat{w}(x''), \hat{y}(x''), x'')} dx'' \right) dx', \quad (4.31)$$

so that one obtains:

Lemma 4.3 *The costate variable $\hat{\varphi}_v(\cdot)$ satisfies $\hat{\varphi}_v(x) < 0$ for all $x \in [x_0, x_1)$.*

At this point, a standard argument, along the lines of Mirrlees (1971, 1976) or Seade (1982), might go as follows: If there is no pooling of types, so that

$\hat{\varphi}'_q(x) = 0$, and if the single-crossing condition (4.4) is strict, i.e., if $\frac{\partial}{\partial x} \frac{|\hat{u}_y|}{\hat{u}_w} < 0$, then, for $x \in [x_0, x_1]$, Lemma 4.3 implies (4.29). Therefore, the contract $(\hat{w}(x), \hat{y}(x))$ is distorted downwards from efficiency if $x \in [x_0, x_1]$.

However, the presumptions that $\hat{\varphi}'_q(x) = 0$ and that the single-crossing condition (4.4) is strict are both not justified. Even if the original single-crossing condition (2.2) is strict, the inequality in (4.4) cannot be strict if $\tau(x)$ is a mass point of the distribution $F(\cdot)$. Moreover, if $\tau(x)$ is a mass point of the distribution $F(\cdot)$, it turns out that one must have $\hat{\varphi}'_q(x) < 0$.

Given the insufficiency of the traditional argument focussing on the sign of the costate variable $\hat{\varphi}_v(\cdot)$, the following argument focusses on the costate variable $\hat{\varphi}_q(\cdot)$ that corresponds to the monotonicity constraint. From (4.24), one finds that

$$\hat{\varphi}_q(\hat{x}) - \hat{\varphi}_q(x) = \int_x^{\hat{x}} \left[\hat{\varphi}_v(x) \hat{u}_w \frac{\partial}{\partial x} \frac{|\hat{u}_y|}{\hat{u}_w} - \frac{\hat{u}_w + \hat{u}_y}{\hat{u}_w} g(x') \right] dx' \quad (4.32)$$

for any x and \hat{x} . The following lemma relates the behaviour of $\hat{\varphi}_q(\cdot)$ to the efficiency properties of the contracts $(\hat{w}(x), \hat{y}(x))$.

Lemma 4.4 *For any $x \in [x_0, x_1]$, $\hat{\varphi}_q(x) \neq 0$ implies*

$$(\hat{w}(x), \hat{y}(x)) \ll (w^*(\tau(x)), y^*(\tau(x)));$$

i.e., if the monotonicity constraint is binding, there must be a downward distortion from efficiency.

Proof. I first show that $\hat{\varphi}_q(x_0) = 0$. If $\hat{y}(x_0) > 0$, this follows from the transversality condition (4.10). If $\hat{y}(x_0) = 0$, Lemma 4.2 implies that the function $\hat{y}(\cdot)$ is strictly increasing at x_0 ; in this case, $\hat{\varphi}_q(x_0) = 0$ follows from statement (f) in Theorem 4.1.

Suppose that $\hat{\varphi}_q(\hat{x}) \neq 0$ for some \hat{x} so that $(\hat{w}(\hat{x}), \hat{y}(\hat{x}))$ is *not* distorted downwards from efficiency. Since $\hat{\varphi}_q(x_0) = 0$ and $\hat{\varphi}_q(\cdot)$ is continuous, there exists $\bar{x} \in [x_0, \hat{x})$ such that $\hat{\varphi}_q(\bar{x}) = 0$ and $\hat{\varphi}_q(x) \neq 0$ for all $x \in (\bar{x}, \hat{x}]$. By statement (f) in Theorem 4.1, one actually has $\hat{\varphi}_q(x) < 0$ and $\hat{y}(x) = \hat{y}(\hat{x})$ for all $x \in (\bar{x}, \hat{x}]$. By (3.17), one also has $\hat{w}(x) = \hat{w}(\hat{x})$ for all $x \in (\bar{x}, \hat{x}]$, i.e., all types t with $\xi(t) \in (\bar{x}, \hat{x}]$ must get the same contract.

By the single-crossing condition (4.4), it follows that, for any $x' \in (\bar{x}, \hat{x}]$, the contract $(\hat{w}(x'), \hat{y}(x')) = (\hat{w}(\hat{x}), \hat{y}(\hat{x}))$ is not distorted downwards from efficiency for x . By (4.25) - (4.27), it follows that

$$\hat{u}_w(\hat{w}(x'), \hat{y}(x'), x') + \hat{u}_y(\hat{w}(x'), \hat{y}(x'), x') \leq 0$$

for all $x' \in (\bar{x}, \hat{x}]$. From (4.32), one therefore obtains

$$\hat{\varphi}_q(\hat{x}) - \hat{\varphi}_q(\bar{x}) \geq - \int_{\bar{x}}^{\hat{x}} -\hat{\varphi}_v(x) \hat{u}_w \frac{\partial}{\partial x} \frac{|\hat{u}_y|}{\hat{u}_w} dx'. \quad (4.33)$$

By (4.4) and Lemma 4.3, (4.33) in turn yields $\hat{\varphi}_q(\hat{x}) - \hat{\varphi}_q(\bar{x}) \geq 0$, which is incompatible with the assumption that $\hat{\varphi}_q(\hat{x}) < 0$ and $\hat{\varphi}_q(\bar{x}) = 0$. The assumption

that $\hat{\varphi}_q(\hat{x}) \neq 0$ for some \hat{x} so that $(\hat{w}(\hat{x}), \hat{y}(\hat{x}))$ is *not* distorted downwards from efficiency has thus led to a contradiction and must be false. ■

Lemma 4.4 implies that, for any x for which $(\hat{w}(x), \hat{y}(x))$ is *not* downward distorted, the term $\hat{\varphi}_q(x)$ in (4.32) vanishes. Because, by (4.13), $\hat{\varphi}_q(\hat{x}) \leq 0$, it follows that, for any such x and any $\hat{x} > x$, the left-hand side of (4.32) is nonpositive, and one must have

$$\int_x^{\hat{x}} \hat{\varphi}_v(x') \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx' \leq \int_x^{\hat{x}} \frac{\hat{u}_w + \hat{u}_y}{\hat{u}_w} g(x') dx'. \quad (4.34)$$

Since $\hat{u}_w > 0$ and $\hat{u}_y \leq 0$, this is equivalent to the requirement that

$$\int_x^{\hat{x}} \hat{\varphi}_v(x') \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx' \leq \int_x^{\hat{x}} \left(1 - \frac{|\hat{u}_y|}{\hat{u}_w}\right) g(x') dx'. \quad (4.35)$$

By Lemma 4.3 and the single-crossing condition (4.4), the integrand on the left-hand side is everywhere nonnegative. What about the integrand on the right-hand side? If $(\hat{w}(x), \hat{y}(x))$ is *not* downward distorted, then, at x , the integrand on the right-hand side is zero or negative. For $x' > x$, the integrand on the right-hand side of (4.35) may change, first, because $\frac{|\hat{u}_y|}{\hat{u}_w}$ depends on x' directly, and, second, because the contract $(\hat{w}(x'), \hat{y}(x'))$ depends on x' . The following lemma exploits the monotonicity of $\hat{y}(\cdot)$ and the quasi-concavity of u and \hat{u} to provide a bound on the change that depends only on the *direct* effect of x' on the marginal rate of substitution.

Lemma 4.5 *For any $x \in [x_0, x_1]$ and any $x' \in (x, x_1]$,*

$$\frac{|\hat{u}_y(\hat{w}(x'), \hat{y}(x'), x')|}{\hat{u}_w(\hat{w}(x'), \hat{y}(x'), x')} \geq \frac{|\hat{u}_y(\hat{w}(x), \hat{y}(x), x)|}{\hat{u}_w(\hat{w}(x), \hat{y}(x), x)} + \int_x^{x'} \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx''. \quad (4.36)$$

Moreover, the inequality is strict if $\hat{y}(x') > \hat{y}(x)$.

Proof. Incentive compatibility implies that the functions $\hat{y}(\cdot)$ and $\hat{w}(\cdot)$ are co-monotonic. Indeed, by (3.16) and the incentive compatibility condition (3.17), one must have $d\hat{w}(x'') + \frac{\hat{u}_y}{\hat{u}_w} d\hat{y}(x'') = 0$ for almost all $x'' \in [x_0, x_1]$. By standard calculus, one therefore has

$$\begin{aligned} & \frac{|\hat{u}_y(\hat{w}(x'), \hat{y}(x'), x')|}{\hat{u}_w(\hat{w}(x'), \hat{y}(x'), x')} - \frac{|\hat{u}_y(\hat{w}(x), \hat{y}(x), x)|}{\hat{u}_w(\hat{w}(x), \hat{y}(x), x)} - \int_x^{x'} \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx'' \\ &= \int_x^{x'} \left[\frac{\partial |\hat{u}_y|}{\partial w \hat{u}_w} \left(-\frac{\hat{u}_y}{\hat{u}_w}\right) + \frac{\partial |\hat{u}_y|}{\partial y \hat{u}_w} \right] d\hat{y}(x'') \end{aligned} \quad (4.37)$$

for all $x \in [x_0, x_1]$ and all $x' \in (x, x_1]$. The right-hand side of (4.37) is computed as

$$\int_x^{x'} \left(-\frac{1}{\hat{u}_w^3}\right) [\hat{u}_y^2 \hat{u}_{ww} - \hat{u}_y \hat{u}_w (\hat{u}_{wy} + \hat{u}_{yw}) + \hat{u}_w^2 \hat{u}_{yy}] d\hat{y}(x''). \quad (4.38)$$

Because u and \hat{u} are quasi-concave in w and y and because $\hat{y}(\cdot)$ is nondecreasing, expression (4.38) is nonnegative. Moreover, if $\hat{y}(x') > \hat{y}(x)$, then, because the quasi-concavity of u and \hat{u} in w and y is strict, expression (4.38) is strictly positive. The left-hand side of (4.37) is therefore nonnegative. It is positive if $\hat{y}(x') > \hat{y}(x)$. ■

Upon combining (4.35) and (4.36), one finds that, for any x for which $(\hat{w}(x), \hat{y}(x))$ is *not* downward distorted, one must have

$$\int_x^{\hat{x}} \hat{\varphi}_v(x') \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x} \frac{1}{\hat{u}_w} dx' \leq \left(1 - \frac{|\hat{u}_y(\hat{w}(x), \hat{y}(x), x)|}{\hat{u}_w(\hat{w}(x), \hat{y}(x), x)} \right) \int_x^{\hat{x}} g(x') dx' - \int_x^{\hat{x}} \int_x^{x'} \frac{\partial |\hat{u}_y|}{\partial x} \frac{1}{\hat{u}_w} dx'' g(x') dx' \quad (4.39)$$

for all $\hat{x} > x$. Upon combining (4.39) with (4.4) and Lemma 4.3, one obtains

Lemma 4.6 *None of the contracts $(\hat{w}(x), \hat{y}(x))$ for $x \in [x_0, x_1]$ is distorted upwards from efficiency, i.e., all of these contracts satisfy*

$$\hat{u}_w(\hat{w}(x), \hat{y}(x), x) + \hat{u}_y(\hat{w}(x), \hat{y}(x), x) \geq 0.$$

Proof. If the lemma is false, then, by (4.27), one has

$$\frac{|\hat{u}_y(\hat{w}(x), \hat{y}(x), x)|}{\hat{u}_w(\hat{w}(x), \hat{y}(x), x)} > 1 \quad (4.40)$$

for some $x \in [x_0, x_1]$. Moreover, Lemmas 4.4 and 4.5 imply that (4.39) must hold for all $\hat{x} \in (x, x_1]$. However, (4.40) implies that, for \hat{x} sufficiently close to x , the right-hand side of (4.39) is negative. By (4.4) and Lemma 4.3, the left-hand side of (4.39) is nonnegative. The assumption that the lemma is false thus leads to a contradiction. ■

By contrast, the possibility that $(\hat{w}(x), \hat{y}(x))$ might be efficient for x cannot be entirely ruled out. The following lemma and its corollary show that, if this is the case, then for any $x' \in [x, x_1]$, the contract $(\hat{w}(x'), \hat{y}(x'))$ must also be efficient for x' . Moreover, one must have pooling of all types between $\tau(x)$ and t_1 , the top of the type set.

Lemma 4.7 *If, for some $x \in [x_0, x_1]$, the contract $(\hat{w}(x), \hat{y}(x))$ is efficient for x , then, for every $\hat{x} \in [x, x_1]$, the contract $(\hat{w}(\hat{x}), \hat{y}(\hat{x}))$ is efficient for \hat{x} .*

Proof. I will prove that

$$\frac{|\hat{u}_y(\hat{w}(x), \hat{y}(x), x)|}{\hat{u}_w(\hat{w}(x), \hat{y}(x), x)} = 1 \quad (4.41)$$

implies

$$\frac{|\hat{u}_y(\hat{w}(\hat{x}), \hat{y}(\hat{x}), \hat{x})|}{\hat{u}_w(\hat{w}(\hat{x}), \hat{y}(\hat{x}), \hat{x})} = 1 \quad (4.42)$$

for all $\hat{x} \in [x, x_1]$. If the claim is false, there exist $x \in [x_0, x_1)$ and $\hat{x} \in [x, x_1]$ such that $(\hat{w}(x), \hat{y}(x))$ is efficient and $(\hat{w}(\hat{x}), \hat{y}(\hat{x}))$ is distorted downwards from efficiency. Let $\bar{x} \geq x$ be the infimum of the set of $\hat{x} \in [x, x_1]$ for which (4.42) fails to hold. I first show that

$$\hat{u}_w(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x}) + \hat{u}_y(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x}) = 0. \quad (4.43)$$

If $\bar{x} = x$, the claim is trivial. If $\bar{x} > x$, the definition of \bar{x} implies that (4.42) holds for all $x' \in [x, \bar{x})$. By Lemma 4.5, it follows that

$$\frac{|\hat{u}_y(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x})|}{\hat{u}_w(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x})} \geq 1 + \int_{x'}^{\bar{x}} \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx''$$

for all $x' \in [x, \bar{x})$. Upon taking limits as x' converges to \bar{x} from below, one obtains $|\hat{u}_y(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x})| \geq \hat{u}_w(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x})$. By Lemma 4.6, one also has $|\hat{u}_y(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x})| \leq \hat{u}_w(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x})$. (4.43) follows immediately.

Because the contract $(\hat{w}(\bar{x}), \hat{y}(\bar{x}))$ satisfies (4.43), Lemmas 4.4 and 4.5 imply that (4.39) must hold for $x = \bar{x}$ and any $\hat{x} \in (\bar{x}, x_1]$; moreover, by (4.43), the first term on the right-hand side of (4.39) is zero. Thus, one must have

$$\int_{\bar{x}}^{\hat{x}} \hat{\varphi}_v(x') \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx' \leq - \int_{\bar{x}}^{\hat{x}} \int_{\bar{x}}^{x'} \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx'' g(x') dx' \quad (4.44)$$

for all $\hat{x} \in (\bar{x}, x_1]$. By (4.4) and (3.8), the right-hand side of (4.44) is no greater than $- \int_{\bar{x}}^{\hat{x}} \int_{\bar{x}}^{x'} \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx'' dx'$. Therefore, (4.44) implies that

$$\int_{\bar{x}}^{\hat{x}} \hat{\varphi}_v(x') \hat{u}_w \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx' \leq -(\hat{x} - \bar{x}) \int_{\bar{x}}^{\hat{x}} \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx'',$$

or, equivalently, that

$$\int_{\bar{x}}^{\hat{x}} [\hat{\varphi}_v(x') \hat{u}_w + \hat{x} - x] \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} dx' \leq 0 \quad (4.45)$$

for all $\hat{x} \in (\bar{x}, x_1]$. By Lemma 4.3, there exists $A > 0$ such that $\hat{\varphi}_v(x') \hat{u}_w \leq -A$ if x' is sufficiently close to \bar{x} . If $\hat{x} - \bar{x} < A$, the integrand in (4.45) is everywhere nonnegative; moreover, it is strictly positive if $\frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} < 0$. For (4.45) to hold, one must therefore have $\frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} = 0$ for all $x' \in [\bar{x}, \hat{x}]$. By Lemma 4.5, it follows that

$$\frac{|\hat{u}_y(\hat{w}(x'), \hat{y}(x'), x')|}{\hat{u}_w(\hat{w}(x'), \hat{y}(x'), x')} \geq \frac{|\hat{u}_y(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x})|}{\hat{u}_w(\hat{w}(\bar{x}), \hat{y}(\bar{x}), \bar{x})} \quad (4.46)$$

for all $x' \in [\bar{x}, \hat{x}]$. By (4.43), this implies $|\hat{u}_y(\hat{w}(x'), \hat{y}(x'), x')| \geq \hat{u}_w(\hat{w}(x'), \hat{y}(x'), x')$. By Lemma 4.6, one also has $|\hat{u}_y(\hat{w}(x'), \hat{y}(x'), x')| \leq \hat{u}_w(\hat{w}(x'), \hat{y}(x'), x')$. Therefore, (4.42) holds for all $x' \in [\bar{x}, \hat{x}]$, contrary to the assumption that $(\hat{w}(\hat{x}), \hat{y}(\hat{x}))$ is distorted downwards from efficiency. The assumption that one can have $x \in [x_0, x_1)$ and $\hat{x} \in [x, x_1]$ such that $(\hat{w}(x), \hat{y}(x))$ is efficient and $(\hat{w}(\hat{x}), \hat{y}(\hat{x}))$ is distorted downwards from efficiency has thus led to a contradiction. ■

Lemma 4.8 *If, for some $x \in [x_0, x_1]$, the contract $(\hat{w}(x), \hat{y}(x))$ is efficient for x , then, for every $\hat{x} \in [x, x_1]$,*

$$(\hat{w}(\hat{x}), \hat{y}(\hat{x})) = (\hat{w}(x), \hat{y}(x)) \quad (4.47)$$

$$\frac{\partial}{\partial x'} \frac{|\hat{u}_y(\hat{w}(x), \hat{y}(x), x')|}{\hat{u}_w(\hat{w}(x), \hat{y}(x), x')}(\hat{x}) = 0. \quad (4.48)$$

Proof. If, for some $x \in [x_0, x_1]$, the contract $(\hat{w}(x), \hat{y}(x))$ is efficient for x , then, by Lemmas 4.7 and 4.4, one must have $\hat{\varphi}_q(\hat{x}) = 0$ for all $\hat{x} \in [x, x_1]$. Hence, also, $\hat{\varphi}'_q(\hat{x}) = 0$ for all $\hat{x} \in [x, x_1]$. By the optimality condition (4.24), it follows that

$$\hat{\varphi}_v(\hat{x})\hat{u}_w \frac{\partial}{\partial x} \frac{|\hat{u}_y|}{\hat{u}_w} = 0$$

for all $\hat{x} \in [x, x_1]$. By Lemma 4.3, therefore,

$$\frac{\partial}{\partial x'} \frac{|\hat{u}_y(\hat{w}(\hat{x}), \hat{y}(\hat{x}), x')|}{\hat{u}_w(\hat{w}(\hat{x}), \hat{y}(\hat{x}), x')}(\hat{x}) = 0 \quad (4.49)$$

for all $\hat{x} \in [x, x_1]$. By Lemma 4.5, it follows that

$$\frac{|\hat{u}_y(\hat{w}(\hat{x}), \hat{y}(\hat{x}), \hat{x})|}{\hat{u}_w(\hat{w}(\hat{x}), \hat{y}(\hat{x}), \hat{x})} \geq \frac{|\hat{u}_y(\hat{w}(x), \hat{y}(x), x)|}{\hat{u}_w(\hat{w}(x), \hat{y}(x), x)} \quad (4.50)$$

for all $\hat{x} \in [x, x_1]$, and the inequality is strict if $\hat{y}(\hat{x}) > \hat{y}(x)$. Because Lemma 4.7 implies that the two sides of (4.50) are both equal to one, it follows that $\hat{y}(\hat{x}) \not> \hat{y}(x)$. Hence $\hat{y}(\hat{x}) = \hat{y}(x)$, and, by incentive compatibility, $\hat{w}(\hat{x}) = \hat{w}(x)$. This establishes (4.47). (4.48) follows from (4.49) and (4.47). ■

As a last step on the way towards proving Theorem 2.1, I show that, even if there is no $x \in [x_0, x_1]$ for which the contract $(\hat{w}(x), \hat{y}(x))$ is efficient, yet there is no distortion at the top.

Lemma 4.9 *For any sequence $\{x^k\}$ that converges to x_1 from below, one has*

$$\lim_{k \rightarrow \infty} [\hat{u}_w(\hat{w}(x^k), \hat{y}(x^k), x^k) + \hat{u}_y(\hat{w}(x^k), \hat{y}(x^k), x^k)] = 0. \quad (4.51)$$

Proof. Let $\{x^k\}$ be any sequence that converges to x_1 from below. Without loss of generality, one may assume that the sequence is nondecreasing. The contract sequence $\{(\hat{w}(x^k), \hat{y}(x^k))\}$ is then also nondecreasing. Because this sequence is bounded by $(\hat{w}(x_1), \hat{y}(x_1))$, it must have a limit (\bar{w}, \bar{y}) .

Recall that, from the transversality condition (4.11) and the optimality condition (4.13), one has $\hat{\varphi}_q(x_1) = 0$ and $\hat{\varphi}_q(x) \leq 0$. For any k , therefore, condition (4.32) yields

$$\frac{1}{1 - G(x^k)} \int_{x^k}^{x_1} \hat{\varphi}_v(x)\hat{u}_w \frac{\partial}{\partial x} \frac{|\hat{u}_y|}{\hat{u}_w} dx \geq \frac{1}{1 - G(x^k)} \int_{x^k}^{x_1} \frac{\hat{u}_w + \hat{u}_y}{\hat{u}_w} g(x) dx. \quad (4.52)$$

I claim that, as $k \rightarrow \infty$, the left-hand side of this inequality converges to zero. To establish this claim, I first observe that, by (4.31) and (4.4), the left-hand side of (4.52) is nonnegative. Also, by (4.31) and (4.4), the left-hand side of (4.52) can be written in the form

$$\frac{1}{1 - G(x^k)} \int_{x^k}^{x_1} \int_x^{x_1} h(x, x') g(x') dx' \hat{u}_w \left| \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} \right| dx, \quad (4.53)$$

where, for any x and x' ,

$$h(x, x') := \frac{1}{\hat{u}_w(\hat{w}(x'), \hat{y}(x'), x')} \exp \left(\int_x^{x'} \frac{\hat{u}_{xw}(\hat{w}(x''), \hat{y}(x''), x'')}{\hat{u}_w(\hat{w}(x''), \hat{y}(x''), x'')} dx'' \right). \quad (4.54)$$

For $x \in [x^k, x_1]$, one has $1 - G(x^k) \geq 1 - G(x)$. Expression (4.53) is therefore bounded above by

$$\int_{x^k}^{x_1} \frac{\int_x^{x_1} h(x, x') dG(x')}{1 - G(x)} \hat{u}_w \left| \frac{\partial |\hat{u}_y|}{\partial x \hat{u}_w} \right| dx, \quad (4.55)$$

which converges to zero as $k \rightarrow \infty$ and x^k converges to x_1 from below. Given that the left-hand side of (4.52) is bounded between zero and (4.53), it must also converge to zero as $k \rightarrow \infty$.

By Lemma 4.6, the right-hand side of (4.52) is nonnegative for all k . Because the left-hand side of (4.52) converges to zero, it follows that the right-hand side also converges to zero as $k \rightarrow \infty$ and x^k converges to x_1 from below. Therefore, one must have

$$\frac{\hat{u}_w(\bar{w}, \bar{y}, x_1) + \hat{u}_y(\bar{w}, \bar{y}, x_1)}{\hat{u}_w(\bar{w}, \bar{y}, x_1)} = 0. \quad (4.56)$$

The lemma follows immediately. ■

4.4 Proofs of Theorems 2.1 and 2.3

Proof of Theorem 2.1. Given the optimal contract menu $(w(\cdot), y(\cdot))$, let $(\bar{w}(\cdot), \bar{y}(\cdot))$ be the associated strongly equivalent contract menu that is given by Lemmas 2.9 and 2.8, and let $v(\cdot)$ be the associated indirect utility function. Further, let $\hat{w}(\cdot)$, $\hat{y}(\cdot)$, and $\hat{v}(\cdot)$ be given by (3.2), (3.3), and (3.19). By Proposition 3.2, these functions solve the principal's modified problem.

If, for all $x \in (x_0, x_1]$, the contract $(\hat{w}(x), \hat{y}(x))$ is efficient for x , set $\hat{t} = t_0$. Statement (a.1) of the theorem is then trivially true, and statements (a.2) and (b) are moot.

Suppose therefore that the set of $x \in [x_0, x_1]$ for which the contract $(\hat{w}(x), \hat{y}(x))$ is distorted downward from efficiency is nonempty, let $\hat{x} > x_0$ be the supremum of this set, and let $\hat{t} = \tau(\hat{x})$. Lemma 4.7 implies that, for $x \in [x_0, \hat{x})$, the contract $(\hat{w}(x), \hat{y}(x))$ is distorted downward from efficiency. By (3.19), it follows that, for $t \in [t_0, \hat{t})$, the contract $(\bar{w}(\cdot), \bar{y}(\cdot))$ is distorted downward from efficiency. This

confirms statement (b) of the theorem. Statement (a.1) of the theorem follows from Lemma 4.8, statement (a.2) from Lemma 4.9. ■

Proof of Theorem 2.3. Let $(\bar{w}(\cdot), \bar{y}(\cdot))$, \hat{t} , and $t \in [t_0, \hat{t}]$ be as specified in the theorem, and let $\hat{w}(\cdot), \hat{y}(\cdot)$ be given by (3.2), (3.3), and (3.19). Let $x = \xi(t)$ and $\underline{x} = \sup_{t' < t} \xi(t')$. Then $\tau(\underline{x}) = \tau(x) = t$. For any $x' \in [\underline{x}, x]$, one has $(\hat{w}(x'), \hat{y}(x')) = (\bar{w}(t), \bar{y}(t))$, and, since $t < \hat{t}$, the contract $(\hat{w}(x'), \hat{y}(x'))$ is distorted downward from efficiency for x' . For any $x' \in [\underline{x}, x]$, one also has $\tau'(x') = 0$. By (4.3) and (3.8), it follows that

$$\frac{\partial |\hat{u}_y(\hat{w}(x'), \hat{y}(x'), x')|}{\partial x \hat{u}_w(\hat{w}(x'), \hat{y}(x'), x')} = 0$$

and for all $x' \in [\underline{x}, x]$. By (4.24) and the fact that $(\hat{w}(x'), \hat{y}(x')) = (\bar{w}(t), \bar{y}(t))$ is distorted downward from efficiency, it follows that $\hat{\varphi}'_q(x') < 0$. Because $\hat{\varphi}_q(\underline{x}) \leq 0$ and $\hat{\varphi}'_q(x') < 0$ for all $x' \in (\underline{x}, x)$, it follows that $\hat{\varphi}_q(x) < 0$ and, hence, that $\hat{\varphi}_q(x'') < 0$ for any $x'' > x$ that is sufficiently close to x . Therefore, there exists $\bar{x} > x$ such that $(\hat{w}(x'), \hat{y}(x')) = (\hat{w}(x), \hat{y}(x)) = (\bar{w}(t), \bar{y}(t))$ for all $x' \in (\underline{x}, \bar{x})$. Set $\bar{t} = \tau(\bar{x})$. Because $\bar{x} > x$, one must have $\bar{t} > t$. Since $(\hat{w}(x'), \hat{y}(x')) = (\bar{w}(t), \bar{y}(t))$ is distorted downward from efficiency for all $x' \in (\underline{x}, \bar{x})$, one must also have $\bar{t} \leq \hat{t}$. ■

4.5 Proofs of Propositions 2.4, 2.5, 2.6

Proof of Proposition 2.4. Let t and \underline{t} be as specified in the proposition. By Theorem 2.3, there exists $\bar{t} \in (t, \hat{t})$ such that the functions $\bar{w}(\cdot), \bar{y}(\cdot)$ are constant on (t, \bar{t}) . If $\underline{t} = t_0$, there is nothing to prove. Suppose therefore that $\underline{t} > t_0$ and that the contract menu $(\bar{w}(\cdot), \bar{y}(\cdot))$ is continuous at \underline{t} . Let $(\hat{w}(\cdot), \hat{y}(\cdot))$ be the associated solution to the principal's modified problem, and let $\underline{x} := \xi(\underline{t})$ and $\bar{x} := \xi(\bar{t})$. The constancy of $(\bar{w}(\cdot), \bar{y}(\cdot))$ on (\underline{t}, \bar{t}) implies that $(\hat{w}(\cdot), \hat{y}(\cdot))$ is constant on (\underline{x}, \bar{x}) . The continuity of $(\bar{w}(\cdot), \bar{y}(\cdot))$ at \underline{t} implies that $(\hat{w}(\cdot), \hat{y}(\cdot))$ is continuous at \underline{x} . If (\hat{w}, \hat{y}) is the common value of $(\hat{w}(\cdot), \hat{y}(\cdot))$ on (\underline{x}, \bar{x}) , then, by the definitions of \underline{x} and \underline{t} , one has $(\hat{w}(x), \hat{y}(x)) \ll (\hat{w}, \hat{y})$ for $x < \underline{x}$.

By statement (f) in Theorem 4.1, it follows that $\hat{\varphi}_q(\underline{x}) = 0$. One must also have $\hat{\varphi}_q(x^k) = 0$ for all k along some sequence $\{x^k\}$ that converges to \underline{x} from below. For suppose that $\hat{\varphi}_q(x) < 0$ for all x below \underline{x} and sufficiently close to \underline{x} . Then $(\hat{w}(\cdot), \hat{y}(\cdot))$ must be constant on an interval which has \underline{x} as its supremum. Given that $(\hat{w}(x), \hat{y}(x)) \ll (\hat{w}, \hat{y})$ for $x < \hat{x}$, this would contradict the continuity of $(\hat{w}(\cdot), \hat{y}(\cdot))$ at \hat{x} .

Given that $\hat{\varphi}_q(x^k) = 0$ for all k along some sequence $\{x^k\}$ that converges to \underline{x} from below, one must also have $\hat{\varphi}'_q(x^k) \leq 0$ for all k along some sequence $\{\chi^k\}$ that converges to \underline{x} from below. By (4.24) and the definition (3.11) of the function \hat{u} , one then has

$$-\left(1 + \frac{u_y^k}{u_w^k}\right) g(\chi^k) + \hat{\varphi}_v(\chi^k) u_w^k \frac{\partial |u_y^k|}{\partial t u_w^k} \cdot \tau'(\chi^k) \leq 0 \quad (4.57)$$

for all k , where u_y^k, u_w^k , etc. are all evaluated at $(\hat{w}(\chi^k), \hat{y}(\chi^k), \tau(\chi^k))$. Without loss of generality, one may suppose that, for any k , at $\tau(\chi^k)$, the distribution function F has a derivative. The value of this derivative is $f_a(\tau(\chi^k))$. By Lemma 3.1, one then has

$$\tau'(\chi^k) = \frac{1}{1 + f_a(\tau(\chi^k))} \quad \text{and} \quad g(\chi^k) = \frac{f_a(\tau(\chi^k))}{1 + f_a(\tau(\chi^k))}$$

for all k . Thus, (4.57) implies that

$$\frac{\hat{\varphi}_v(\chi^k) u_w^k}{f_a(\tau(\chi^k))} \frac{\partial |u_y^k|}{\partial t u_w^k} \leq \left(1 + \frac{u_y^k}{u_w^k} \right) \quad (4.58)$$

for all k . Upon taking limits as k goes out of bounds, using the presumed continuity of $(\hat{w}(\cdot), \hat{y}(\cdot))$ at \underline{x} , one infers that

$$\frac{\hat{\varphi}_v(\underline{x}) u_w(\hat{w}, \hat{y}, \tau(\underline{x}))}{\bar{f}_a} \frac{\partial |u_y|}{\partial t u_w}(\hat{w}, \hat{y}, \tau(\underline{x})) \leq \left(1 + \frac{u_y(\hat{w}, \hat{y}, \tau(\underline{x}))}{u_w(\hat{w}, \hat{y}, \tau(\underline{x}))} \right), \quad (4.59)$$

where $\bar{f}_a := \lim_{k \rightarrow \infty} f_a(\tau(\chi^k))$; this limit exists because $f_a(\cdot)$ is a nondecreasing function.

Without loss of generality, one may also suppose that \bar{x} is the supremum of the set on which $(\hat{w}(\cdot), \hat{y}(\cdot))$ takes the value (\hat{w}, \hat{y}) . Thus, $(\hat{w}(\cdot), \hat{y}(\cdot))$ is strictly increasing at \bar{x} . Therefore, $\hat{\varphi}_q(\bar{x}) = 0$. By statement (f) in Theorem 4.1, it follows that $\hat{\varphi}_q$ is maximal at \bar{x} . Therefore, there exists (a nonnull set of) $\zeta < \bar{x}$, close to \bar{x} such that $\hat{\varphi}_q'(\zeta) \geq 0$. Because $t < \hat{t}$, Theorem 2.1 implies that $u_w(\hat{w}, \hat{y}, \tau(\underline{x})) + u_y(\hat{w}, \hat{y}, \tau(\underline{x})) > 0$. By the single-crossing condition, it follows that $u_w(\hat{w}, \hat{y}, \tau(\zeta)) + u_y(\hat{w}, \hat{y}, \tau(\zeta)) > 0$ for ζ close to \bar{x} . By (4.24), therefore, $\hat{\varphi}_q'(\zeta) \geq 0$ implies $\tau'(\zeta) > 0$ so that, at $\tau(\zeta)$, the type distribution F again has a density, with the value $f_a(\tau(\zeta))$. Upon using (4.24), (3.11), and Lemma 3.1, as before, one infers that

$$\frac{\hat{\varphi}_v(\zeta) u_w(\hat{w}, \hat{y}, \tau(\zeta))}{f_a(\tau(\zeta))} \frac{\partial |u_y|}{\partial t u_w}(\hat{w}, \hat{y}, \tau(\zeta)) \geq \left(1 + \frac{u_y(\hat{w}, \hat{y}, \tau(\zeta))}{u_w(\hat{w}, \hat{y}, \tau(\zeta))} \right). \quad (4.60)$$

By (4.21) and the fact that $(\hat{w}(x), \hat{y}(x)) = (\hat{w}, \hat{y})$ for all $x \in (\underline{x}, \zeta)$, one also has $\hat{\varphi}_v(\zeta) u_w(\hat{w}, \hat{y}, \tau(\zeta)) > \hat{\varphi}_v(\underline{x}) u_w(\hat{w}, \hat{y}, \tau(\underline{x}))$. Since $\tau(\zeta) > \tau(\chi^k)$ for all k and f_a is a nondecreasing function, one also has $f_a(\tau(\zeta)) \geq f_a(\tau(\chi^k))$ for all k and, therefore, $f_a(\tau(\zeta)) \geq \bar{f}_a$. Because, by Lemma 4.3, $\hat{\varphi}_v(\zeta) < 0$ and $\hat{\varphi}_v(\underline{x}) < 0$, it follows that

$$\frac{\hat{\varphi}_v(\zeta) u_w(\hat{w}, \hat{y}, \tau(\zeta))}{f_a(\tau(\zeta))} > \frac{\hat{\varphi}_v(\underline{x}) u_w(\hat{w}, \hat{y}, \tau(\underline{x}))}{\bar{f}_a}. \quad (4.61)$$

By the log-convexity of $\frac{|u_y(w, y, t)|}{u_w(w, y, t)}$ in t , one also has

$$\frac{u_w(\hat{w}, \hat{y}, \tau(\zeta))}{|u_y(\hat{w}, \hat{y}, \tau(\zeta))|} \frac{\partial |u_y|}{\partial t u_w}(\hat{w}, \hat{y}, \tau(\zeta)) \geq \frac{u_w(\hat{w}, \hat{y}, \tau(\underline{x}))}{|u_y(\hat{w}, \hat{y}, \tau(\underline{x}))|} \frac{\partial |u_y|}{\partial t u_w}(\hat{w}, \hat{y}, \tau(\underline{x})). \quad (4.62)$$

Upon combining (4.59) - (4.62), using the fact that, by Lemma 4.3 and the single-crossing condition, $\hat{\varphi}_v$ and $\frac{\partial |u_y|}{\partial t} \frac{|u_y|}{u_w}$ are negative, one obtains

$$\begin{aligned}
& \frac{u_w(\hat{w}, \hat{y}, \tau(\zeta))}{|u_y(\hat{w}, \hat{y}, \tau(\zeta))|} \left(1 + \frac{u_y(\hat{w}, \hat{y}, \tau(\zeta))}{u_w(\hat{w}, \hat{y}, \tau(\zeta))} \right) \\
\leq & \frac{\hat{\varphi}_v(\zeta) u_w(\hat{w}, \hat{y}, \tau(\zeta))}{f_a(\tau(\zeta))} \frac{u_w(\hat{w}, \hat{y}, \tau(\zeta))}{|u_y(\hat{w}, \hat{y}, \tau(\zeta))|} \frac{\partial |u_y|}{\partial t} \frac{|u_y|}{u_w}(\hat{w}, \hat{y}, \tau(\zeta)) \\
< & \frac{\hat{\varphi}_v(\underline{x}) u_w(\hat{w}, \hat{y}, \tau(\underline{x}))}{f_a} \frac{u_w(\hat{w}, \hat{y}, \tau(\underline{x}))}{|u_y(\hat{w}, \hat{y}, \tau(\underline{x}))|} \frac{\partial |u_y|}{\partial t} \frac{|u_y|}{u_w}(\hat{w}, \hat{y}, \tau(\underline{x})) \\
\leq & \frac{u_w(\hat{w}, \hat{y}, \tau(\underline{x}))}{|u_y(\hat{w}, \hat{y}, \tau(\underline{x}))|} \left(1 + \frac{u_y(\hat{w}, \hat{y}, \tau(\underline{x}))}{u_w(\hat{w}, \hat{y}, \tau(\underline{x}))} \right).
\end{aligned}$$

Thus,

$$\frac{u_w(\hat{w}, \hat{y}, \tau(\zeta))}{|u_y(\hat{w}, \hat{y}, \tau(\zeta))|} - 1 < \frac{u_w(\hat{w}, \hat{y}, \tau(\underline{x}))}{|u_y(\hat{w}, \hat{y}, \tau(\underline{x}))|} - 1, \quad (4.63)$$

which is incompatible with the single-crossing condition. The assumption that $\hat{t} > t_0$ and that the contract menu $(w(\cdot), y(\cdot))$ is continuous at \hat{t} has thus led to a contradiction and must be false. This completes the proof of Proposition 2.4. ■

Proof of Proposition 2.5 (Sketch). The argument is similar to the one in the proof of Proposition 2.4: If there were a pooling interval $[\underline{t}, \bar{t}]$, the costate variable $\hat{\varphi}_q(\cdot)$ would have to be nonincreasing just above $\underline{x} = \xi(\underline{t})$ and nondecreasing just below $\bar{x} = \xi(\bar{t})$. If the type distribution has no singular component, one has $\tau'(x) = 1/(1 + f(\tau(x)))$ for all x . By the same reasoning as before, therefore, $\hat{\varphi}'_q(\underline{x}+) \leq 0$ and $\hat{\varphi}'_q(\bar{x}-) \geq 0$ would yield (4.59) and (4.60), with \bar{f}_a and $f_a(\tau(\zeta))$ replaced by $f(\underline{t}) = f(\tau(\underline{x}))$ and $f(\bar{t}) = f(\tau(\bar{x}))$. With strict positivity and monotonicity of f and with log-convexity of $\frac{|u_y(w, y, t)|}{u_w(w, y, t)}$ in t , one then again obtains (4.63). Because this is incompatible with the single-crossing condition, the assumption that there is a pooling interval must be false. ■

Proof of Proposition 2.6. Proceeding indirectly, suppose that the proposition is false. Then there is an optimal contract menu $(w(\cdot), y(\cdot))$ that exhibits a discontinuity at some $\bar{t} \in [t_0, t_1]$. If $(\hat{w}(\cdot), \hat{y}(\cdot))$ is the associated solution to the principal's modified problem, then $(\hat{w}(\cdot), \hat{y}(\cdot))$ exhibits a discontinuity at $\bar{x} = \xi(\bar{t})$. Then $(\hat{w}(\cdot), \hat{y}(\cdot))$ is strictly increasing at \bar{x} . This implies $\hat{\varphi}_q(\bar{x}) = 0$. By statement (f) in Theorem 4.1, it follows that $\hat{\varphi}_q$ is maximal at \bar{x} . Therefore there exist sequences $\{\zeta_k\}$ converging to \bar{x} from below and $\{\chi^\ell\}$ converging to \bar{x} from above such that $\hat{\varphi}'_q(\zeta_k) \geq 0$ and $\hat{\varphi}'_q(\chi^\ell) \leq 0$ for all k and ℓ . By (4.24), one then has

$$\hat{\varphi}_v(\zeta_k) u_w^k \frac{\partial |u_y^k|}{\partial t} \frac{|u_y^k|}{u_w^k} \cdot \tau'(\zeta_k) \geq \left(1 + \frac{u_y^k}{u_w^k} \right) g(\zeta_k) \quad (4.64)$$

and

$$\hat{\varphi}_v(\chi^\ell) u_w^\ell \frac{\partial |u_y^\ell|}{\partial t u_w^\ell} \cdot \tau'(\chi^\ell) \leq \left(1 + \frac{u_y^\ell}{u_w^\ell}\right) g(\chi^\ell) \quad (4.65)$$

for all k and ℓ , where u_y^k, u_w^k are evaluated at $(\hat{w}(\zeta_k), \hat{y}(\zeta_k), \tau(\zeta_k))$ and u_y^ℓ, u_w^ℓ , etc. are evaluated at $(\hat{w}(\chi^\ell), \hat{y}(\chi^\ell), \tau(\chi^\ell))$.

Because the type distribution has a continuous density, Lemma 3.1, yields

$$\tau'(x) = \frac{1}{1 + f(\tau(x))} \quad \text{and} \quad g(x) = \frac{f(\tau(x))}{1 + f(\tau(x))}$$

for all x . Conditions (4.64) and (4.65) can therefore be rewritten as

$$\frac{\hat{\varphi}_v(\zeta_k)}{f(\tau(\zeta_k))} u_w^k \frac{\partial |u_y^k|}{\partial t u_w^k} \geq \left(1 + \frac{u_y^k}{u_w^k}\right) \quad (4.66)$$

and

$$\frac{\hat{\varphi}_v(\chi^\ell)}{f(\tau(\chi^\ell))} u_w^\ell \frac{\partial |u_y^\ell|}{\partial t u_w^\ell} \leq \left(1 + \frac{u_y^\ell}{u_w^\ell}\right). \quad (4.67)$$

Upon taking limits as k and ℓ become large, using the continuity of f and τ and the monotonicity of the contract menu, one obtains

$$\begin{aligned} \frac{\hat{\varphi}_v(\bar{x})}{f(\tau(\bar{x}))} u_w(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x})) \frac{\partial |u_y|}{\partial t u_w}(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x})) \\ \geq \left(1 + \frac{u_y(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x}))}{u_w(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x}))}\right) \end{aligned} \quad (4.68)$$

and

$$\begin{aligned} \frac{\hat{\varphi}_v(\bar{x})}{f(\tau(\bar{x}))} u_w(\hat{w}(\bar{x}+), \hat{y}(\bar{x}+), \tau(\bar{x})) \frac{\partial |u_y|}{\partial t u_w}(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x})) \\ \leq \left(1 + \frac{u_y(\hat{w}(\bar{x}+), \hat{y}(\bar{x}+), \tau(\bar{x}))}{u_w(\hat{w}(\bar{x}+), \hat{y}(\bar{x}+), \tau(\bar{x}))}\right). \end{aligned} \quad (4.69)$$

Because u is strictly quasi-concave in w and y , $(\hat{w}(\bar{x}+), \hat{y}(\bar{x}+)) \gg (\hat{w}(\bar{x}-), \hat{y}(\bar{x}-))$ implies that the right-hand side of (4.68) is strictly greater than the right-hand side of (4.69). Because $\frac{\hat{\varphi}_v(\bar{x})}{f(\tau(\bar{x}))} < 0$ and $|u_y| = -u_y$, it follows that

$$\begin{aligned} u_w(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x})) \frac{\partial u_y}{\partial t u_w}(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x})) \\ > u_w(\hat{w}(\bar{x}+), \hat{y}(\bar{x}+), \tau(\bar{x})) \frac{\partial u_y}{\partial t u_w}(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x})), \end{aligned}$$

or, equivalently,

$$u_{yt}(-) - \frac{u_y(-)}{u_w(-)} u_{wt}(-) > u_{yt}(+) - \frac{u_y(+)}{u_w(+)} u_{wt}(+), \quad (4.70)$$

where "−" and "+" in the arguments stand in for evaluation at $(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x}))$ and $(\hat{w}(\bar{x}+), \hat{y}(\bar{x}+), \tau(\bar{x}))$. However, since IC implies

$$u(\hat{w}(\bar{x}-), \hat{y}(\bar{x}-), \tau(\bar{x})) = u(\hat{w}(\bar{x}+), \hat{y}(\bar{x}+), \tau(\bar{x})),$$

one easily verifies that (4.70) is incompatible with the convexity of the function $y \rightarrow u_t(c(v, y, \tau(\bar{x})), y, \tau(\bar{x}))$. The assumption that the proposition is false has thus led to a contradiction. ■

References

- [1] Brett, C., and J.A. Weymark (2003), Financing Education Using Optimal Redistributive Taxation, *Journal of Public Economics* 87, 2549 - 2569.
- [2] Clarke, F.H. (1976), The Maximum Principle under Minimal Hypotheses, *SIAM Journal of Control and Optimization* 14, 1078 - 1091.
- [3] Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis*, Wiley, New York.
- [4] Crémer, J., F. Khalil, and J.-C. Rochet (1998 a), Strategic Information Gathering before a Contract is Offered, *Journal of Economic Theory* 81, 163 - 200.
- [5] Crémer, J., F. Khalil, and J.-C. Rochet (1998 b), Contracts and Productive Information Gathering, *Games and Economic Behavior* 25, 174 - 193.
- [6] Diamond, P.A. (1998), Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Income Tax Rates, *American Economic Review* 88, 83 - 95.
- [7] Fudenberg, D., and J. Tirole (1991), *Game Theory*, MIT Press, Cambridge, Massachusetts.
- [8] Guesnerie, R., and J.-J. Laffont (1984), A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm, *Journal of Public Economics* 25, 329 - 369.
- [9] Hellwig, M.F. (2004), Risk Aversion in the Small and in the Large When Outcomes Are Multidimensional, Discussion Paper No. 04-06, Max Planck Institute for Research on Collective Goods, Bonn, Germany, http://www.coll.mpg.de/pdf_dat/2004_06online.pdf.
- [10] Hellwig, M.F. (2004/2008), Optimal Income Taxation, Public-Goods Provision and Public-Sector Pricing: A Contribution to the Foundations of Public Economics, 2008 revision of Preprint 14/2004, Max Planck Institute for Research on Collective Goods, Bonn, Germany, http://www.coll.mpg.de/pdf_dat/2004_14online.pdf.
- [11] Hellwig, M.F. (2007 a), A Contribution to the Theory of Optimal Utilitarian Income Taxation, *Journal of Public Economics* 91 (2007), 1449 - 1477; http://www.coll.mpg.de/pdf_dat/2007_02online.pdf.
- [12] Hellwig, M.F. (2007 b), The Undesirability of Randomized Income Taxation under Decreasing Risk Aversion, *Journal of Public Economics* 91 (2007), 791 - 816.

- [13] Hellwig, M.F. (2008), A Maximum Principle for Control Problems with Monotonicity Constraints, Preprint 04/2008, Max Planck Institute for Research on Collective Goods, Bonn, Germany, http://www.coll.mpg.de/pdf_dat/2008_04online.pdf.
- [14] Hildenbrand, W. (1974), *Core and Equilibria of a Large Economy*, Princeton University Press, Princeton, N.J.
- [15] Laffont, J.-J., and D. Martimort (2001), *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press, Princeton, New Jersey.
- [16] Mas-Colell, A., M.D. Whinston, and J.R. Green (1995), *Microeconomic Theory*, Oxford University Press, Oxford.
- [17] Mirrlees, J.M. (1971), An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies* 38, 175 - 208.
- [18] Mirrlees, J.M. (1976), Optimal Tax Theory: A Synthesis, *Journal of Public Economics* 6, 327 - 358.
- [19] Mirrlees, J.M. (1986), The Theory of Optimal Taxation, in: K.J. Arrow and M.D. Intriligator (eds.), *Handbook of Mathematical Economics*, North Holland, Elsevier Publishers, Amsterdam *et al.*, Vol. III, Ch. 24, 1198 - 1249.
- [20] Roberts, K. (2000), A reconsideration of the Optimal Income Tax, in: P.J. Hammond and G.D. Myles (eds.), *Incentives, Organization and Public Economics: Papers in Honour of Sir James Mirrlees*, Oxford University Press, Oxford, 171 - 189.
- [21] Saez, E. (2001), Using Elasticities to Derive Optimal Income Tax Rates, *Review of Economic Studies* 68, 205 - 229.
- [22] Seade, J. (1977), On the Shape of Optimal Tax Schedules, *Journal of Public Economics* 7, 203 - 236.
- [23] Seade, J. (1982), On the Sign of the Optimum Marginal Income Tax, *Review of Economic Studies* 49, 637 - 643.
- [24] Szalay, D., (2005), The Economics of Extreme Options and Clear Advice, *Review of Economic Studies* 72, 1173-1198.
- [25] Werning, I. (2007), Pareto Efficient Income Taxation, MIT Working Paper, <http://econ-www.mit.edu/files/1281>.

Online Appendix to:
Incentive Problems with Unidimensional Hidden
Characteristics: A Unified Approach

Martin F. Hellwig
Max Planck Institute for Research on Collective Goods
Kurt-Schumacher-Str. 10, D - 53113 Bonn, Germany
hellwig@coll.mpg.de

April 7, 2010

This Appendix provides formal proofs of Lemmas 2.7 - 2.9 in the text of the paper.

1 Proof of Lemma 2.7

Lemma 1 (Lemma 2.7) *A contract menu $(w(\cdot), y(\cdot))$ that is defined on $T \subset [t_0, t_1]$ is incentive-compatible and individually rational on T if and only if there exists an extension of $(w(\cdot), y(\cdot))$ to the interval $[t_0, t_1]$ that is incentive-compatible and individually rational on $[t_0, t_1]$.*

Before proving this lemma, I note some implications of the weak single-crossing condition, i.e., the requirement that

$$\frac{\partial |u_y(w, y, t)|}{\partial t u_w(w, y, t)} \leq 0 \tag{1}$$

for all w, y , and t .

Lemma 2 *Let t, \bar{t} , and $(w, y), (\bar{w}, \bar{y})$ be such that $\bar{t} > t$, $\bar{y} < y$, and*

$$u(w, y, t) > u(\bar{w}, \bar{y}, t). \tag{2}$$

Then

$$u(w, y, \bar{t}) > u(\bar{w}, \bar{y}, \bar{t}). \tag{3}$$

The argument is practically identical to the argument for the analogous implication of strict single-crossing in Milgrom and Shannon (1994). It is therefore left to the reader.

Given this result, the following lemma shows that, under the weak single-crossing condition (1), incentive-compatible contract menus are monotone, *unless* all the relevant types are indifferent between the contracts in question.

Lemma 3 Let t, \bar{t} , and $(w, y), (\bar{w}, \bar{y})$ be such that $\bar{t} > t, \bar{y} < y$,

$$u(w, y, t) \geq u(\bar{w}, \bar{y}, t) \quad (4)$$

and

$$u(\bar{w}, \bar{y}, \bar{t}) \geq u(w, y, \bar{t}). \quad (5)$$

Then

$$u(\bar{w}, \bar{y}, t') = u(w, y, t') \quad (6)$$

for all $t' \in [t, \bar{t}]$, i.e., all types between t and \bar{t} are indifferent between the contracts (w, y) and (\bar{w}, \bar{y}) .

Proof. By Lemma 2, $y > \bar{y}$ and (4) imply $u(w + \varepsilon, y, t') > u(\bar{w}, \bar{y}, t')$ for all $t' \in [t, \bar{t}]$ and all $\varepsilon > 0$, hence,

$$u(w, y, t') \geq u(\bar{w}, \bar{y}, t') \quad (7)$$

for all $t' \in [t, \bar{t}]$. Similarly, $y > \bar{y}$ and (5) imply

$$u(\bar{w}, \bar{y}, t') \geq u(w, y, t') \quad (8)$$

for all $t' \in [t, \bar{t}]$. (6) follows immediately. ■

Lemma 4 A contract menu $(w(\cdot), y(\cdot))$ that is defined on a set $X \subset [t_0, t_1]$ is incentive-compatible on X if and only if, for any $\underline{t} \in X$ and $\bar{t} \in X$ such that $(\underline{t}, \bar{t}) \cap X = \emptyset$, there exists an extension of $(w(\cdot), y(\cdot))$ to $X \cup (\underline{t}, \bar{t})$ that is incentive-compatible on $X \cup (\underline{t}, \bar{t})$.

Proof. The "if"-part of the lemma is trivial. To prove the "only if"-part, let $X \subset [t_0, t_1]$ and suppose that $(w(\cdot), y(\cdot))$ is incentive-compatible on X . Let $\underline{t} \in X$ and $\bar{t} \in X$ be such that $(\underline{t}, \bar{t}) \cap X = \emptyset$. Incentive compatibility of $(w(\cdot), y(\cdot))$ on X implies that

$$u(w(\underline{t}), y(\underline{t}), \underline{t}) \geq u(w(\bar{t}), y(\bar{t}), \underline{t}) \quad (9)$$

and

$$u(w(\bar{t}), y(\bar{t}), \bar{t}) \geq u(w(\underline{t}), y(\underline{t}), \bar{t}). \quad (10)$$

Because u is continuous, there exists $\hat{t} \in [\underline{t}, \bar{t}]$ such that

$$u(w(\underline{t}), y(\underline{t}), \hat{t}) = u(w(\bar{t}), y(\bar{t}), \hat{t}). \quad (11)$$

Extend $w(\cdot)$ and $y(\cdot)$ to the interval (\underline{t}, \bar{t}) by setting

$$(w(t), y(t)) = (w(\underline{t}), y(\underline{t})) \text{ for } t \in (\underline{t}, \hat{t}), \quad (12)$$

$$(w(t), y(t)) = (w(\bar{t}), y(\bar{t})) \text{ for } t \in (\hat{t}, \bar{t}), \quad (13)$$

and

$$(w(\hat{t}), y(\hat{t})) = (w(\bar{t}), y(\bar{t})) \text{ if } \hat{t} > \underline{t}. \quad (14)$$

To prove that the extended contract menu is incentive-compatible, I first note that, if $y(\bar{t}) \geq y(\underline{t})$, then Lemma 2 and (11) imply

$$u(w(\underline{t}), y(\underline{t}), t) \geq u(w(\bar{t}), y(\bar{t}), t) \quad (15)$$

for all $t \in (\underline{t}, \hat{t}]$ and

$$u(w(\bar{t}), y(\bar{t}), t) \geq u(w(\underline{t}), y(\underline{t}), t) \quad (16)$$

for all $t \in [\hat{t}, \bar{t}]$. If $y(\bar{t}) < y(\underline{t})$, these same inequalities follow from Lemma 3; indeed, in this case, (15) and (16) must hold as equations for all $t \in [\underline{t}, \bar{t}]$.

Consider the incentive compatibility condition

$$u(w(t), y(t), t) \geq u(w(t'), y(t'), t) \quad (17)$$

for $t \in [\underline{t}, \bar{t}]$ and $t' \in X$ satisfying $y(t') \leq y(\underline{t})$. Incentive compatibility of $(w(\cdot), y(\cdot))$ on X implies that

$$u(w(\underline{t}), y(\underline{t}), \underline{t}) \geq u(w(t'), y(t'), \underline{t}) \quad (18)$$

for all $t' \in X \cap [t_0, \underline{t}]$. By Lemma 2, it follows that

$$u(w(\underline{t}), y(\underline{t}), t) \geq u(w(t'), y(t'), t) \quad (19)$$

for all $t \in [\underline{t}, \bar{t}]$. By (12) - (14) and (15), it follows that

$$u(w(t), y(t), t) \geq u(w(t'), y(t'), t). \quad (20)$$

A precisely symmetric argument shows that (20) must also hold for $t \in [\underline{t}, \bar{t}]$ and $t' \in X$ satisfying $y(t') \leq y(\bar{t})$.

For $t' \in X$ satisfying $y(\underline{t}) < y(t') < y(\bar{t})$, Lemma 3 implies that

$$u(w(\underline{t}), y(\underline{t}), \underline{t}) = u(w(t'), y(t'), \underline{t}) \quad (21)$$

if $t' < \underline{t}$, and

$$u(w(\bar{t}), y(\bar{t}), \bar{t}) = u(w(\underline{t}), y(\underline{t}), \bar{t}) \quad (22)$$

if $t' > \bar{t}$. In either case, one again obtains (20) for all $t \in [\underline{t}, \bar{t}]$.

For $t \in X$ and $t' \in X \cup (\underline{t}, \bar{t})$, the validity of (17) follows trivially from the incentive compatibility of $(w(\cdot), y(\cdot))$ on X and the observation that the extension of the domain of the contract menu to $X \cup (\underline{t}, \bar{t})$ has not changed its range. ■

Proof of Lemma 1. The "if" part of the lemma is trivial. To prove the "only if" part, observe that the set $[t_0, t_1] \setminus T$ can be represented as a countable union of open intervals I_1, I_2, \dots . If one applies Lemma 4 successively, with $X_1 = T$, $(\underline{t}_1, \bar{t}_1) = I_1$, $X_2 = T \cup I_1$, $(\underline{t}_2, \bar{t}_2) = I_2$, etc., then, in the limit, one obtains an extension of $(w(\cdot), y(\cdot))$ that is incentive compatible on $[t_0, t_1]$.

To prove that this extended contract menu is also individually rational, one notes that, because $t_0 \in T$, the contract $(w(t_0), y(t_0))$ that is assigned to the lowest type has not been changed. By the individual rationality of the original contract menu, $u(w(t_0), y(t_0), t_0) \geq 0 = u(0, 0, t_0)$. By Lemma 2, it follows that $u(w(t_0), y(t_0), t) \geq u(0, 0, t) = 0$ for any $t > t_0$. By incentive compatibility, one also has $u(w(t), y(t), t) \geq u(w(t_0), y(t_0), t)$, hence, $u(w(t), y(t), t) \geq 0$. ■

2 Proof of Lemma 2.8

Lemma 5 (Lemma 2.8) *A nondecreasing contract menu $(w(\cdot), y(\cdot))$ is incentive-compatible and individually rational on $[t_0, t_1]$ if and only if the induced indirect utility function $v(\cdot)$ satisfies the integral equation*

$$v(t) = v(t_0) + \int_{t_0}^t u_t(w(\tau), y(\tau), \tau) d\tau \quad (23)$$

for $t \in [t_0, t_1]$ and, moreover,

$$v(t_0) \geq 0. \quad (24)$$

As mentioned in the text, Lemma 5 is little more than a slight generalization of the characterization result in Mirrlees (1976).

Lemma 6 *If a contract menu $(w(\cdot), y(\cdot))$ is nondecreasing and incentive-compatible on $[t_0, t_1]$, then the induced indirect utility function $v(\cdot)$ satisfies*

$$v(t) = v(t_0) + \int_{t_0}^t u_t(w(\tau), y(\tau), \tau) d\tau \quad (25)$$

for all $t \in [t_0, t_1]$.

Proof. The argument follows Baron and Myerson (1982). For any t and \bar{t} , incentive compatibility implies

$$v(t) = u(w(t), y(t), t) \geq u(w(\bar{t}), y(\bar{t}), t) \quad (26)$$

and

$$v(\bar{t}) = u(w(\bar{t}), y(\bar{t}), \bar{t}) \geq u(w(t), y(t), \bar{t}). \quad (27)$$

From (26) and (27), one obtains

$$u(w(t), y(t), t) - u(w(t), y(t), \bar{t}) \geq v(t) - v(\bar{t}) \geq u(w(\bar{t}), y(\bar{t}), t) - u(w(\bar{t}), y(\bar{t}), \bar{t}), \quad (28)$$

hence

$$\frac{\int_{\bar{t}}^t u_t(w(t), y(t), \tau) d\tau}{t - \bar{t}} \geq \frac{v(t) - v(\bar{t})}{t - \bar{t}} \geq \frac{\int_{\bar{t}}^t u_t(w(\bar{t}), y(\bar{t}), \tau) d\tau}{t - \bar{t}}. \quad (29)$$

Because $u_t(\cdot, \cdot, \cdot)$ is continuous and, by monotonicity, the triples $(w(t), y(t), \tau)$ and $(w(\bar{t}), y(\bar{t}), \tau)$ belong to the compact set $[0, w(t_1)] \times [0, y(t_1)] \times [t_0, t_1]$, the integrands on both sides of (29) are uniformly bounded. The function $v(\cdot)$ is therefore Lipschitz continuous, hence absolutely continuous on $[t_0, t_1]$.

If t is a continuity point of the contract menu $(w(\cdot), y(\cdot))$ and if \bar{t} is close to t , then, by standard arguments, the right-hand side and the left-hand side of (29) are both approximately equal to $u_t(w(t), y(t), t)$. In this case, the (ordinary) derivative of $v(\cdot)$ at t exists and is equal to $u_t(w(t), y(t), t)$. Because the nondecreasing function $t \rightarrow (w(t), y(t))$ has at most countably many points

of discontinuity, it follows that the function $t \rightarrow u_t(w(t), y(t), t)$ is a Radon-Nikodym derivative for the absolutely continuous function $v(\cdot)$. The validity of (25) follows immediately. ■

Lemma 7 *If a contract menu $(w(\cdot), y(\cdot))$ on an interval $[t_0, t_1]$ is nondecreasing and the induced indirect utility function $v(\cdot)$ satisfies the integral equation (25), then $(w(\cdot), y(\cdot))$ is incentive-compatible on $[t_0, t_1]$.*

Proof. The argument follows Mirrlees (1976); see also Appendix B in the online version of Hellwig (2007). For any t and \bar{t} , (25) implies

$$u(w(t), y(t), t) - u(w(\bar{t}), y(\bar{t}), \bar{t}) = \int_{\bar{t}}^t u_t(w(\tau), y(\tau), \tau) d\tau. \quad (30)$$

Because the left-hand side of (30) is equal to

$$\int_{\bar{t}}^t u_w(w(\tau), y(\tau), \tau) dw(\tau) + \int_{\bar{t}}^t u_y(w(\tau), y(\tau), \tau) dy(\tau) + \int_{\bar{t}}^t u_t(w(\tau), y(\tau), \tau) d\tau,$$

it follows that

$$\int_{\bar{t}}^t u_w(w(\tau), y(\tau), \tau) dw(\tau) + \int_{\bar{t}}^t u_y(w(\tau), y(\tau), \tau) dy(\tau) = 0 \quad (31)$$

for all t and \bar{t} . By standard arguments, it follows that, for every measurable function f , one has

$$\int_{t_0}^{t_1} f(\tau) u_w(w(\tau), y(\tau), \tau) dw(\tau) + \int_{t_0}^{t_1} f(\tau) u_y(w(\tau), y(\tau), \tau) dy(\tau) = 0. \quad (32)$$

In particular, therefore,

$$\int_{\bar{t}}^t \chi(\tau, t) u_w(w(\tau), y(\tau), \tau) dw(\tau) + \int_{\bar{t}}^t \chi(\tau, t) u_y(w(\tau), y(\tau), \tau) dy(\tau) = 0 \quad (33)$$

for every \bar{t}, t , and every measurable function χ . If one sets $\chi(\tau, t) = \frac{u_w(w(\tau), y(\tau), t)}{u_w(w(\tau), y(\tau), \tau)}$, one infers that

$$\int_{\bar{t}}^t u_w(w(\tau), y(\tau), t) dw(\tau) + \int_{\bar{t}}^t \frac{u_w(w(\tau), y(\tau), t)}{u_w(w(\tau), y(\tau), \tau)} u_y(w(\tau), y(\tau), \tau) dy(\tau) = 0 \quad (34)$$

for all \bar{t}, t , and τ between \bar{t} and t .

The single-crossing condition implies that

$$\frac{u_w(w(\tau), y(\tau), \tau)}{|u_y(w(\tau), y(\tau), \tau)|} \leq \frac{u_w(w(\tau), y(\tau), t)}{|u_y(w(\tau), y(\tau), t)|} \text{ as } \tau \leq t,$$

hence, since u_y takes negative values,

$$\frac{u_y(w(\tau), y(\tau), \tau)}{u_w(w(\tau), y(\tau), \tau)} \leq \frac{u_y(w(\tau), y(\tau), t)}{u_w(w(\tau), y(\tau), t)} \text{ as } \tau \leq t. \quad (35)$$

Because u_w takes positive values and $y(\cdot)$ is nondecreasing, (34) and (35) imply that

$$\int_{\bar{t}}^t u_w(w(\tau), y(\tau), t) dw(\tau) + \int_{\bar{t}}^t u_y(w(\tau), y(\tau), t) dy(\tau) \geq 0 \quad (36)$$

for all t and \bar{t} . Therefore,

$$u(w(t), y(t), t) - u(w(\bar{t}), y(\bar{t}), t) \geq 0 \quad (37)$$

for all t and \bar{t} . ■

Lemma 8 *A contract menu $(w(\cdot), y(\cdot))$ that is incentive-compatible on an interval $[t_0, t_1]$ has induced utility satisfying $v(t) \geq 0$ if and only if $v(t_0) \geq 0$.*

■ **Proof.** It suffices to observe that, by (25), $v(\cdot)$ is a nondecreasing function.

Lemma 5 follows from Lemmas 6 - 8.

3 Proof of Lemma 2.9

Lemma 9 (Lemma 2.9) *For any incentive-compatible contract menu $(w(\cdot), y(\cdot))$, there exists a nondecreasing incentive-compatible contract menu $(\bar{w}(\cdot), \bar{y}(\cdot))$ that provides the agent with the same payoff $v(t) = u(w(t), y(t), t)$ for all t and that satisfies*

$$\int [\bar{y}(t) - \bar{w}(t)] dF(t) \geq \int [y(t) - w(t)] dF(t); \quad (38)$$

moreover, the inequality in (38) is strict unless the contract menus $(w(\cdot), y(\cdot))$ and $(\bar{w}(\cdot), \bar{y}(\cdot))$ are equivalent.

To prove this lemma, I need a stronger version of Lemma 3. The following result implies that, if an incentive-compatible contract menu violates monotonicity, then, in the relevant part of their domains, the indifference curves of the relevant types must coincide.

Lemma 10 *If any two types t and \bar{t} are both indifferent between two contracts (w, y) and $(\bar{w}, \bar{y}) \ll (w, y)$, then, between these two contracts, their indifference curves coincide, i.e., for any (w', y') with $(\bar{w}, \bar{y}) \leq (w', y') \leq (w, y)$,*

$$u(w', y', t) = u(w, y, t) \text{ if and only if } u(w', y', \bar{t}) = u(\bar{w}, \bar{y}, \bar{t}). \quad (39)$$

Proof. Without loss of generality, suppose that $t \leq \bar{t}$. If $(\bar{w}, \bar{y}) \leq (w', y') \leq (w, y)$, then, by Lemma 2,

$$u(w', y', t) = u(w, y, t) \quad (40)$$

implies $u(w', y' + \varepsilon, \bar{t}) < u(w, y, \bar{t})$ for all $\varepsilon > 0$, hence,

$$u(w', y', \bar{t}) \leq u(w, y, \bar{t}). \quad (41)$$

Because type t is indifferent between (w, y) and (\bar{w}, \bar{y}) and because $(\bar{w}, \bar{y}) \leq (w', y')$, by Lemma 2, (40) also implies $u(w' + \varepsilon, y', \bar{t}) > u(\bar{w}, \bar{y}, \bar{t})$ for all $\varepsilon > 0$, hence,

$$u(w', y', \bar{t}) \geq u(\bar{w}, \bar{y}, \bar{t}). \quad (42)$$

Because type \bar{t} is also indifferent between (w, y) and (\bar{w}, \bar{y}) , (41) and (42) imply

$$u(w', y', \bar{t}) = u(\bar{w}, \bar{y}, \bar{t}). \quad (43)$$

Thus, (40) implies (43). By a precisely symmetric argument, one also finds that (43) implies (40). ■

To proceed with the proof of Lemma 9 itself, I need some additional notation. Given an incentive-compatible contract menu $(w(\cdot), y(\cdot))$ with associated indirect utility function $v(\cdot)$, for any t , let

$$I(t) = \{\bar{t} \in T \mid u(w(\bar{t}), y(\bar{t}), t) = v(t) \text{ and } u(w(t), y(t), \bar{t}) = v(\bar{t})\} \quad (44)$$

be the set of types \bar{t} such that both t and \bar{t} are indifferent between the pairs $(w(t), y(t))$ and $(w(\bar{t}), y(\bar{t}))$. By Lemma 3, we know that, for any t , the set $I(t)$ contains any $\bar{t} > t$ for which $y(\bar{t}) < y(t)$; indeed, if such \bar{t} exists, the set $I(t)$ has the entire interval $[t, \bar{t}]$ as a subset. By Lemmas 3 and 10 jointly, in this case, the set $I(t)$ also contains any $t' < \bar{t}$ for which $y(t') \geq y(\bar{t})$; indeed, it has the entire interval $[t', \bar{t}]$ as a subset.

Given the set $I(t)$, let

$$\chi(t) := \{(w, y) \mid u(w, y, t) = v(t) \text{ and } y(t') \leq y \leq y(t'') \text{ for some } t', t'' \in I(t)\} \quad (45)$$

be the segment of type t 's indifference curve through $(w(t), y(t))$ that lies "between" the contracts assigned to types in $I(t)$, and let $\bar{\chi}(t)$ be the closure of $\chi(t)$. Any contract (w, y) in $\bar{\chi}(t)$ provides type t with the same utility $v(t)$ as the contract $(w(t), y(t))$. It is therefore of interest to ask which of these contracts is most profitable for the principal.

Lemma 11 *If $(w(\cdot), y(\cdot))$ is an incentive-compatible contract menu, then, for any $t \in [t_0, t_1]$, the problem*

$$\max_{(w, y) \in \bar{\chi}(t)} [y - w] \quad (46)$$

has a unique solution $(\bar{w}(t), \bar{y}(t))$. The contract menu $(\bar{w}(\cdot), \bar{y}(\cdot))$ is nondecreasing and incentive-compatible.

Proof. Uniqueness of the solution to the problem $\max_{(w,y) \in \bar{\chi}(t)} [y-w]$ follows from the strict quasi-concavity of u in w and y . To prove weak monotonicity, suppose that $t < \bar{t}$ and $\bar{y}(t) > \bar{y}(\bar{t})$. By Lemma 3, one has $\bar{t} \in I(t)$. By the definition of $I(\cdot)$, it follows that $I(t) = I(\bar{t})$. By Lemma 10, one then also has $\chi(t) = \chi(\bar{t})$, hence $\bar{\chi}(t) = \bar{\chi}(\bar{t})$. Because the solution to problem (46) is unique and depends on t only through the constraint set $\bar{\chi}(t)$, it follows that $(\bar{w}(t), \bar{y}(t)) = (\bar{w}(\bar{t}), \bar{y}(\bar{t}))$. The assumption that $t < \bar{t}$ and $\bar{y}(t) > \bar{y}(\bar{t})$ has thus led to a contradiction and must be false. Thus $t < \bar{t}$ implies $\bar{y}(t) \geq \bar{y}(\bar{t})$. Weak monotonicity of $\bar{w}(\cdot)$ then follows from incentive compatibility.

By construction,

$$u(\bar{w}(t), \bar{y}(t), t) = v(t) \quad (47)$$

for all t . To prove incentive compatibility, it therefore suffices to show that

$$v(t) \geq u(\bar{w}(\bar{t}), \bar{y}(\bar{t}), t) \quad (48)$$

for all t and all \bar{t} . Since $(\bar{w}(\bar{t}), \bar{y}(\bar{t})) \in \bar{\chi}(\bar{t})$, there exists a sequence $\{(w^k(\bar{t}), y^k(\bar{t}))\}$ of elements of $\chi(\bar{t})$ that converges to $(\bar{w}(\bar{t}), \bar{y}(\bar{t}))$. To prove (48), it therefore suffices to show that

$$v(t) \geq u(w^k(\bar{t}), y^k(\bar{t}), t) \quad (49)$$

for all k .

By the definition of $\chi(\bar{t})$, there exist sequences $\{\bar{t}'_k\}, \{\bar{t}''_k\}$ of elements of $I(\bar{t})$ such that, for any k , one has

$$y(\bar{t}'_k) \leq y^k(\bar{t}) \leq y(\bar{t}''_k); \quad (50)$$

moreover,

$$u(w^k(\bar{t}), y^k(\bar{t}), \bar{t}) = u(w(\bar{t}'_k), y(\bar{t}'_k), \bar{t}) = u(w(\bar{t}''_k), y(\bar{t}''_k), \bar{t}). \quad (51)$$

If $t < \bar{t}$, (50) and (51) in combination with Lemma 2 imply

$$u(w^k(\bar{t}), y^k(\bar{t}), t) \leq u(w(\bar{t}'_k), y(\bar{t}'_k), t). \quad (52)$$

Because incentive compatibility requires

$$v(t) \geq u(w(\bar{t}'_k), y(\bar{t}'_k), t), \quad (53)$$

(49) follows immediately. If $t > \bar{t}$, one similarly obtains

$$u(w^k(\bar{t}), y^k(\bar{t}), t) \leq u(w(\bar{t}''_k), y(\bar{t}''_k), t) \leq v(t), \quad (54)$$

which also yields (49). ■

To establish Lemma 9, it now suffices to observe that, by construction, one has

$$\bar{y}(t) - \bar{w}(t) \geq y(t) - w(t) \quad (55)$$

for all t , and that the inequality in (55) is strict unless $(w(t), y(t)) = (\bar{w}(t), \bar{y}(t))$.

References

- [1] Baron, D.P., and R.B. Myerson (1982), Regulating a Monopolist with Unknown Costs, *Econometrica* 50, 911 - 930.
- [2] Hellwig, M.F. (2007), A Contribution to the Theory of Optimal Utilitarian Income Taxation, *Journal of Public Economics* 91 (2007), 1449 - 1477; http://www.coll.mpg.de/pdf_dat/2007_02online.pdf.
- [3] Milgrom, P., and C. Shannon (1994), Monotone Comparative Statics, *Econometrica* 62, 157-80.
- [4] Mirrlees, J.M. (1976), Optimal Tax Theory: A Synthesis, *Journal of Public Economics* 6, 327 - 358.