

**Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2008/4**



**A Maximum Principle for
Control Problems with
Monotonicity Constraints**

Martin Hellwig



MAX PLANCK SOCIETY



A Maximum Principle for Control Problems with Monotonicity Constraints

Martin Hellwig

August 2009

A Maximum Principle for Control Problems with Monotonicity Constraints*

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

August 12, 2009

Abstract

The paper develops a version of Pontryagin's maximum principle for optimal control problems with monotonicity constraints on control variables. Whereas the literature handles such constraints by imposing an assumption of piecewise smoothness on the control variable and treating the slope of this variable as a new control variable subject to a nonnegativity constraint, the paper obtains the maximum principle without such an additional assumption. The result is useful for studying incentive problems with hidden characteristics when the type set is a continuum and preferences satisfy a single-crossing constraint.

JEL Classification: C61

Key Words: Maximum Principle, Optimal Control, Monotonicity Constraints, Incentive Problems with Hidden Characteristics

1 Introduction

This paper develops a version of Pontryagin's maximum principle for problems of optimal control in which a control variable is required to satisfy a monotonicity constraint. Such control problems arise naturally in the analysis of incentive problems with hidden characteristics and a continuum of types. Under a standard single-crossing condition on preferences, in such

*I am grateful to Mark Hahmeier for very helpful comments on an earlier version of this paper.

problems, incentive compatibility requires that outcomes depend monotonically on the hidden characteristic.

The usual approach to handling monotonicity constraints on control variables is to recast the variable in question as a state variable, to assume that the function relating this state variable to the hidden characteristic is (piecewise) continuously differentiable, to treat its slope as a new control variable and to require that this new control variable be nonnegative.¹ This procedure has been pioneered by Brito and Oakland (1977) for an income tax problem and by Guesnerie and Laffont (1984) for a principal-agent problem. Subsequent treatments and applications can be found in Fudenberg and Tirole (1991), Ebert (1992), Brunner (1993), and Hellwig (2007 a).

Given that the control variable is endogenous, the requirement that this variable be piecewise smooth in the underlying parameter is problematic. The implications of this requirement for the exogenous data of the problem are often unclear. When they are clear, they may be unpalatable.

The paper shows that this additional condition is not actually needed. The usual set of necessary conditions for optimal control problems, including a version of the maximum principle, are obtained without any additional assumptions. If the control variables happen to be absolutely continuous anyway, the conditions given here coincide with those of Guesnerie and Laffont (1984). However, they also apply if the control variables are not continuous.

The basic idea is to exploit the monotonicity property itself. Being monotonic, the control variable in question can be treated as the sum of an absolutely continuous component and a singular component, both of which must be monotonic. The analysis of Guesnerie and Laffont can be applied to the absolutely continuous component. This yields the usual characterization of solutions in terms of costate variables and a maximum principle for the slope of the absolutely continuous part. Given this characterization, one can use the consideration that, at the optimum, a proportional decrease of the singular component of the control variable in question should not raise the value of the objective function, to show that the maximum principle also applies to the "slope" of the singular component.

The first part of the argument is based on Clarke's (1976, 1983) version of the maximum principle under minimal hypotheses. In this version, there is no need for continuity in the dependence of the various functions

¹Barron and Jensen (1980) use a dynamic-programming approach. Assuming uniform Lipschitz continuity of the functions determining the "dynamics" of the system, they show that the value function is well defined and Lipschitz continuous, and that it satisfies a Hamilton-Jacobi partial differential equation at points where it is continuously differentiable.

on the underlying parameter. Whereas traditional statements of the maximum principle assume that, e.g., the integrand in the objective function be continuous in the variable of integration, Clarke only requires measurability.

The second part of the argument involves just calculus. However, if the control problem involves explicit constraints for both endpoints, an additional assumption is needed to deal with the possibility that the variation under consideration, i.e. a proportional decrease of the singular component of the control variable in question, when taken by itself, might lead to a violation of one of the endpoint constraints. The assumption takes the form of a regularity condition ensuring that any such violation can be neutralized by a change in the initial value of one of the state variables.

In the following, Section 2 formulates the control problem with explicit constraints for one endpoint only. A paradigmatic example is given by the principal-agent problem studied in Hellwig (2009), with a hidden characteristic of the agent and a participation constraint. This example is used to explain why the assumption of absolute continuity of the control variable is problematic. Section 3 presents the main theorem, i.e. the maximum principle for the control problem with explicit constraints for one endpoint. The theorem is proved in Section 4. Section 5 extends the analysis to control problems with explicit constraints for both endpoints. Here, a paradigmatic application is provided by the optimal income tax problem of Mirrlees (1971), with an integral constraint ensuring that the government budget is balanced.

2 The Control Problem with Explicit Constraints for One Endpoint Only

I will study the problem of choosing an absolutely continuous function $x : [0, 1] \rightarrow \mathbb{R}^\ell$, and measurable functions $y : [0, 1] \rightarrow \mathbb{R}^m$, and $z : [0, 1] \rightarrow \mathbb{R}^n$ so as to maximize the functional

$$\int_0^1 f(x(t), y(t), z(t), t) dt \tag{2.1}$$

under the following constraints:

$$G_0(x(0)) \geq \bar{G}_0; \tag{2.2}$$

$$x'(t) = h(x(t), y(t), z(t), t), \tag{2.3}$$

$$y(0) \geq 0, \tag{2.4}$$

and

$$z(t) \geq 0 \tag{2.5}$$

for almost all $t \in [0, 1]$; finally,

$$y(\cdot) \text{ must be nondecreasing.} \tag{2.6}$$

In this specification, f is a function from $\mathbb{R}^\ell \times \mathbb{R}^m \times \mathbb{R}^n \times [0, 1]$ into \mathbb{R} , G_0 is a function from \mathbb{R}^ℓ into \mathbb{R}^k , where $k \leq \ell$, h is a function from $\mathbb{R}^\ell \times \mathbb{R}^m \times \mathbb{R}^n \times [0, 1]$ into \mathbb{R}^ℓ , and \bar{G}_0 is a vector in \mathbb{R}^k .

The functions f and h are assumed to be bounded, as well as measurable with respect to the product $\mathcal{B}(\mathbb{R}^\ell \times \mathbb{R}^m \times \mathbb{R}^n) \times \mathcal{L}$ of the Borel σ -algebra on $\mathbb{R}^\ell \times \mathbb{R}^m \times \mathbb{R}^n$ and the Lebesgue σ -algebra on the unit interval. In addition, for any $t \in [0, 1]$, the sections $f(\cdot, \cdot, \cdot, t)$ and $h(\cdot, \cdot, \cdot, t)$ of f and h that are determined by t are assumed to be equi-Lipschitz-continuous and continuously differentiable. The function G_0 is also continuously differentiable; its Jacobian has rank k everywhere.

Optimization problems of this type occur regularly in models of incentive provision with hidden characteristics. As an example, consider the problem studied in Hellwig (2009) of a principal who does not know the productivity of the agent to whom he proposes a contract. Given his ignorance, he offers a menu of contracts, letting the agent choose whichever contract he likes best. The menu of contracts is represented by a pair of functions that indicate for each t the output $y(t) \geq 0$ that the agent is supposed to produce and the wage $w(t)$ that the agent will receive if his productivity parameter has the value t . The principal's problem is to choose the functions $y(\cdot)$ and $w(\cdot)$ so as to maximize the expected value of his profit $y(t) - w(t)$ subject to incentive compatibility and individual-rationality conditions. In analysing this problem, it is convenient to write

$$w(t) = c(v(t), y(t), t), \tag{2.7}$$

where $v(\cdot)$ is a function that indicates, for each t , the payoff $v(t)$ that the agent receives from contracting with the principal if his productivity parameter is t and $c(v(t), y(t), t)$, the consumption level that permits a person of type t to obtain the payoff level $v(t)$ when this person is required to produce the output $y(t)$ is implicitly defined by the equation

$$v(t) = u(c(v(t), y(t), t), y(t), t). \tag{2.8}$$

If the utility function u satisfies a single-crossing condition, then, along the lines of Mirrlees (1976), one can show that incentive compatibility is equivalent to the requirement that

$$v'(t) = u_t(c(v(t), y(t), t), y(t), t) \quad (2.9)$$

for almost all $t \in [0, 1]$, and that

$$y(\cdot) \text{ must be nondecreasing.} \quad (2.10)$$

Condition (2.9) is the first-order condition for incentive compatibility, and condition (2.10) is the second-order condition for incentive compatibility. With a single-crossing condition, these conditions are sufficient as well as necessary for incentive compatibility. If u is nondecreasing in t and the utility \bar{v} that the agent gets from his outside option is independent of t , the participation constraint is satisfied for all t if and only if it is satisfied for $t = 0$, i.e., if and only if

$$v(0) \geq \bar{v}. \quad (2.11)$$

Moreover, under (2.10), the nonnegativity condition on $y(t)$ is satisfied for all t if and only if

$$y(0) \geq 0. \quad (2.12)$$

If the principal considers the distribution of the productivity parameter t to have a density γ , then his problem is to choose $v(\cdot)$ and $y(\cdot)$ so as to maximize the expected profit

$$\int_0^1 (y(t) - c(v(t), y(t), t))\gamma(t)dt \quad (2.13)$$

under the constraints (2.8) - (2.12). This is a special case of the more general problem of maximizing (2.1) under the constraints (2.2) - (2.6).

If the desired function $y(\cdot)$ were known to be absolutely continuous,² the problem of maximizing (2.1) under the constraints (2.2) - (2.6) could be formulated as a standard control problem with state variables $x(t), y(t)$ and control variables $z(t), q(t)$ where $q(\cdot)$ is the (Radon-Nikodym) derivative of $y(\cdot)$. In this formulation, the monotonicity condition would reduce to the requirement that $q(t) \geq 0$ for almost all t .

²Guesnerie and Laffont (1984) assume piecewise continuous differentiability. This allows them to use the original version of the maximum principle. Clarke's (1976, 1983) version of the maximum principle under minimal hypotheses shows that one only needs absolute continuity.

However, the presumption that $y(\cdot)$ is absolutely continuous is unattractive. The implications of this presumption for the exogenous data of the problem are unclear. Even when they are clear, they may be unduly restrictive.

To see the issue, consider the contracting example again. If it wasn't for the monotonicity constraint on $y(\cdot)$, this would be a control problem with state variable $v(t)$ and control variable $y(t)$. For any t , the control choice $y = y(t)$ should be a global maximizer of the Hamiltonian expression

$$\hat{H}(v(t), y, t; \varphi) = (y - c(v(t), y, t))\gamma(t) + \varphi u_t(c(v(t), y, t), y, t). \quad (2.14)$$

By the maximum theorem, the global maximizers of $\hat{H}(v(t), \cdot, t; \varphi)$ are continuous in t if $\hat{H}(v(t), y, t; \varphi)$ is jointly continuous in t and y , and if, for each t , the global maximizer of $\hat{H}(v(t), \cdot, t; \varphi)$ is unique. If these conditions are not satisfied, continuity cannot be taken for granted.

Both conditions are restrictive. Continuity of $\hat{H}(v(t), y, t; \varphi)$ in t and y requires that the density $\gamma(t)$ in (2.14) be continuous in t . If we actually think of γ as a probability density, this requirement is unproblematic. It can, however, be problematic if γ is itself derived from other data of the problem. Thus, in Hellwig (2009), I study the principal-agent problem with hidden characteristics without requiring the distribution of the productivity parameter to have a continuous density, i.e., I allow this distribution to have mass points as well as a continuous part. To make the problem amenable to control-theoretic methods, the notion of an agent's "type" t , i.e., his productivity parameter, is replaced by the notion of a "pseudo-type" x , with associated "type" $t = \tau(x)$, the idea being that a type t is a mass point of the type distribution if many pseudo-types have the same type t . The distribution G of pseudo-types has a density γ , so that, by replacing types with pseudo-types, one can formulate the principal's problem as an instance of the control problem considered here. However, if the original type distribution has mass points, the density γ of the distribution of "pseudo-types" is not continuous.³

³Let F be the distribution of productivity parameters t . If $\tau(x)$ is the productivity parameter of an agent with pseudo-type x , the value $\gamma(x)$ of the density of the pseudo-type distribution satisfies

$$\gamma(x) = \frac{F'(\tau(x))}{1 + F'(\tau(x))}$$

if, at $t = \tau(x)$, the derivative $F'(t)$ of the distribution function $F(\cdot)$ is well defined, and $\gamma(x) = 1$ otherwise. For details, see Hellwig (2009). The analysis there shows that, if the distribution $F(\cdot)$ has mass points, as well as a continuous part, the mass point will necessarily be pooled with adjacent higher and, possibly, adjacent lower types, and the

In the absence of additional assumptions about the utility function, there is also no presumption that, for any t , the function $\hat{H}(v(t), \cdot, t; \varphi(t))$ must be quasiconcave. Quasiconcavity of $u(\cdot, \cdot, t)$ implies that $c(v(t), \cdot, t)$ is a convex function and, therefore, that the first term on the right-hand side of (2.14) is concave in y . However, quasiconcavity of $u(\cdot, \cdot, t)$ is *not* sufficient to determine the curvatures of the function $y \rightarrow u_t(c(v(t), y, t), y, t)$ and of the second term on the right-hand side of (2.14).⁴ The Hamiltonian expression (2.14) may therefore have multiple maxima. In consequence, it may be impossible for a global maximizer of $\hat{H}(v(t), \cdot, t; \varphi(t))$ to depend continuously on t .

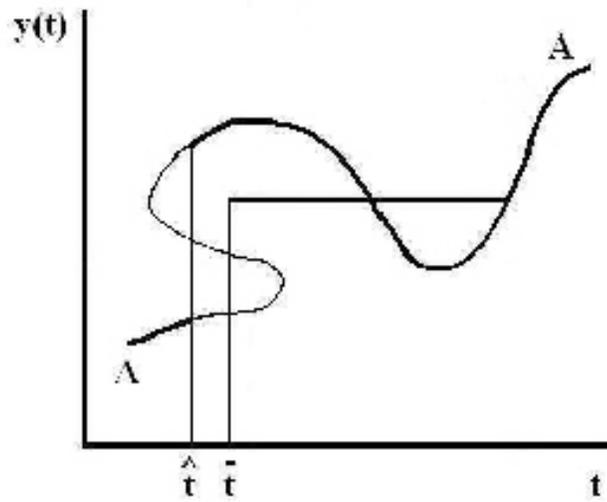
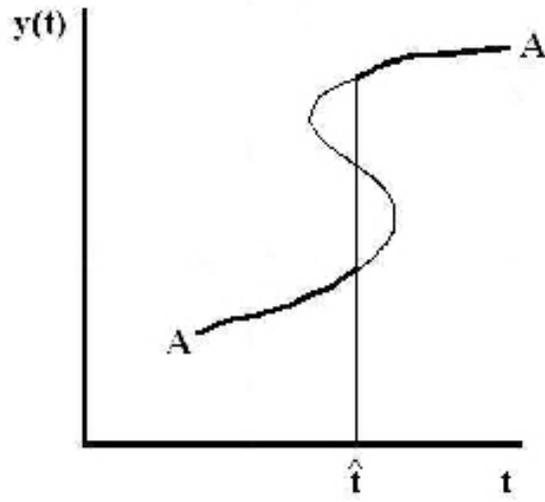
A diagrammatic illustration is given in Figures 1 and 2. The curves A-A in these figures indicate the values of y that correspond to critical points of the Hamiltonian for different t . The fat segments of these curves correspond to global maxima. As the figures are drawn, the global maximizers are *necessarily* discontinuous at some \hat{t} where the Hamiltonian has multiple maxima.

In Figure 1, the global maximizers are actually nondecreasing so that the monotonicity constraint is nonbinding. In Figure 2, by contrast, the global maximizers of the Hamiltonian violate monotonicity. The correct, nondecreasing, solution to the principal's problem therefore involves a flat segment that is given by an ironing procedure as in Guesnerie and Laffont (1984).

However, in Figure 2, there is an interdependence between the ironing procedure and the discontinuity in the optimal $y(\cdot)$: The initial point \bar{t} of the ironing interval is a discontinuity point of $y(\cdot)$. By contrast to Guesnerie and Laffont (1984), therefore, the value of $y(\cdot)$ on the ironing interval is *not* a maximizer of the Hamiltonian at the interval endpoint \bar{t} . In consequence, \bar{t} , i.e., the discontinuity point of the correct, nondecreasing solution to the principal's problem, is *not* the same as \hat{t} , the discontinuity point of the global maximizers of $\hat{H}(v(t), \cdot, t; \varphi(t))$.

output requirement $y(\cdot)$ will normally exhibit an upward jump in the transition from lower types to the pool that contains the mass point.

⁴The costate variable φ can be shown to be nonpositive. Therefore, if u is concave in c and y , quasiconcavity of \hat{H} in y is assured if the function $y \rightarrow u_t(c(v(t), y, t), y, t)$ is *convex*. For details, see Mirrlees (1986). Hellwig (2007 b) shows that convexity of the function $y \rightarrow u_t(c(v(t), y, t), y, t)$ is equivalent to the condition that u exhibit a property of nonincreasing consumption-specific risk aversion in the sense that, for each two-dimensional outcome lottery, the amount of consumption that the agent is willing to give up in order to have the lottery replaced by the vector of expected values is a nonincreasing function of t . Hellwig (2007 b) also shows that this assumption guarantees uniqueness of the solution to the optimal control problem.



In the following, I study the problem of maximizing (2.1) under the constraints (2.2) - (2.6) *without* requiring that $y(\cdot)$ be absolutely continuous. The very monotonicity condition (2.6) provides additional structure which can be used to obtain a version of the maximum principle for this problem. The resulting characterization of solutions to the maximization problem turns out to be similar to the one given by Guesnerie and Laffont (1984); however, this characterization encompasses discontinuity points of $y(\cdot)$.

3 The Main Theorem

Assuming that $y(\cdot)$ is absolutely continuous, Guesnerie and Laffont (1984) work with the Hamiltonian

$$H(x, y, z, q, t; \varphi, \chi) = f(x, y, z, t) + \varphi \cdot h(x, y, z, t) + \chi \cdot q, \quad (3.1)$$

where q corresponds to the derivative of $y(\cdot)$, and χ is the associated vector of costate variables. The monotonicity constraint on $y(\cdot)$ translates into a nonnegativity constraint on the control vector q . The maximum principle requires that, for almost every t , the chosen controls $z(t)$ and $q(t)$ should maximize the value of the Hamiltonian at t . The necessary and sufficient condition for maximization with respect to $q \geq 0$ is

$$\chi \leq 0 \text{ and } \chi \cdot q = \sum_{i=1}^m \chi_i \cdot q_i = 0. \quad (3.2)$$

If $y(\cdot)$ is not absolutely continuous, there is no presumption that, for any given t , the derivative of $y(\cdot)$ at t is well defined. Even so, for any i , one can talk heuristically about the "slope" $q_i(t)$ of $y_i(\cdot)$ at t if one allows for the possibility that q takes values in the extended nonnegative real numbers, $\mathbb{R}_+ \cup \{\infty\}$, with the proviso that, for any t , $q_i(t) = \infty$ if $\lim_{\varepsilon \downarrow 0} [y_i(t + \varepsilon) - y_i(t - \varepsilon)]/\varepsilon = \infty$; in particular, $q_i(t) = \infty$ if t is a discontinuity point of $y_i(\cdot)$. With the convention that, for any i , $\chi_i \cdot \infty = 0$ if $\chi_i = 0$, $\chi_i \cdot \infty = \infty$ if $\chi_i > 0$, and $\chi_i \cdot \infty = -\infty$ if $\chi_i < 0$, one can extend the specification of the Hamiltonian (3.1) to allow for the possibility that $q_i = \infty$ for some i ; in this extended specification, the Hamiltonian itself takes values in the extended real numbers.

Given these modifications, it still makes sense to talk about $z(t)$ and $q(t)$ maximizing the value of the Hamiltonian at t . Indeed, the necessary and sufficient condition for $q \geq 0$ to be maximizing H is the same as before. As in (3.1), one must have $\chi_i \leq 0$ for all i and $\chi_i = 0$ if $q_i > 0$; in particular, one must have $\chi_i = 0$ if $q_i = \infty$.

The following theorem shows that, with this extended notion of the "slope" of the nondecreasing function $y(\cdot)$, the Guesnerie-Laffont characterization of solutions to the control problem in terms of a maximum principle for the choice of the "slope" of $y(\cdot)$ remains valid even if $y(\cdot)$ is not absolutely continuous. To avoid the imprecision of language involved in talking about the "slope" of a function that cannot be characterized in terms of derivatives, the key condition (f) in the theorem does not actually refer to the "slope" of $y(\cdot)$. However, heuristically, this condition can be interpreted as a version of the maximum principle for this "slope".

Theorem 3.1 *Suppose that, under the specified measurability, differentiability, and equi-Lipschitz conditions on f , G_0 and h , the absolutely continuous function $x^*(\cdot)$ and the measurable functions $y^*(\cdot)$ and $z^*(\cdot)$ maximize (2.1) under the constraints (2.2) - (2.6). Let*

$$y^*(\cdot) = y_a^*(\cdot) + y_s^*(\cdot) \quad (3.3)$$

be a Lebesgue decomposition of the nondecreasing function $y^*(\cdot)$ into a nondecreasing absolutely continuous function $y_a^*(\cdot)$ and a nondecreasing singular⁵ function $y_s^*(\cdot)$, and let $q^*(\cdot)$ be the Radon-Nikodym derivative of the absolutely continuous function $y_a^*(\cdot)$. Then there exist a nonnegative vector ν_0 and absolutely continuous functions $\varphi : [0, 1] \rightarrow \mathbb{R}^\ell$ and $\chi : [0, 1] \rightarrow \mathbb{R}^m$ such that the following hold:

(a) For almost all $t \in [0, 1]$,

$$\varphi'(t) = -H_x(x^*(t), y^*(t), z^*(t), q^*(t), t; \varphi(t), \chi(t)); \quad (3.4)$$

(b) moreover,

$$\varphi(0) = -\nu_0 \cdot G_0'(x^*(0)), \quad \nu_0 \cdot G_0(x^*(0)) = 0, \quad (3.5)$$

and

$$\varphi(1) = 0. \quad (3.6)$$

(c) For almost all $t \in [0, 1]$,

$$\chi'(t) = -H_y(x^*(t), y^*(t), z^*(t), q^*(t), t; \varphi(t), \chi(t)); \quad (3.7)$$

⁵I use the term *singular* to comprise *both*, continuous singular functions and jump functions, i.e., all functions with the property that the measures which they generate are concentrated on a set of Lebesgue measure zero.

(d) moreover,

$$\chi(0) \leq 0, \quad \chi(0) \cdot y^*(0) = 0, \quad (3.8)$$

and

$$\chi(1) = 0. \quad (3.9)$$

(e) For almost every t ,

$$z^*(t) \in \arg \max_{z \geq 0} H(x^*(t), y^*(t), z, q^*(t), t; \varphi(t), \chi(t)). \quad (3.10)$$

(f) For almost every $t \in [0, 1]$, $\chi(t) \leq 0$. Moreover, for any i ,

$$\int_0^1 \chi_i(t) dy_i^*(t) = 0;$$

in particular, for almost every $t \in [0, 1]$ and $i = 1, \dots, m$, $\chi_i(t) = 0$ if $y_i^*(\cdot)$ is strictly increasing at t .⁶

To understand this result, it is useful to go back to the case where $y^*(\cdot)$ is absolutely continuous, i.e., the singular component $y_s^*(\cdot)$ in (3.3) is zero. In this case, $q^*(\cdot)$ is the (Radon-Nikodym) derivative of $y^*(\cdot)$ itself, and statement (f) is equivalent to statement

(f*) For almost every $t \in [0, 1]$, $\chi(t) \leq 0$; moreover, for almost every $t \in [0, 1]$ and $i = 1, \dots, m$, $\chi_i(t) = 0$ if $q_i^*(t) > 0$.

Statement (f*) in turn is equivalent to statement

(f**) For almost every $t \in [0, 1]$,

$$q^*(t) \in \arg \max_{q \geq 0} H(x^*(t), y^*(t), z^*(t), q, t; \varphi(t), \chi(t)).$$

This is just the maximum principle for the choice of the slope $q^*(t)$ of $y^*(\cdot)$ at t . Thus, if $y^*(\cdot)$ is absolutely continuous, Theorem 3.1 can be read as a version of the maximum principle in the formulation of Clarke (1976, 1983), which does without most of the regularity conditions in the more traditional formulations of the principle.

⁶ A real-valued, nondecreasing function f is said to be *strictly increasing* at t if $f(t + \varepsilon) - f(t - \varepsilon) > 0$ for all $\varepsilon > 0$.

However, Theorem 3.1 is more general than that. It does not actually presume that $y^*(\cdot)$ is absolutely continuous. In the general case, statement (f) stipulates that $\chi_i(t) = 0$ for almost every t at which $y_i^*(\cdot)$ is strictly increasing. In particular, $\chi_i(t) = 0$ if the slope $q_i^*(t)$ of the absolutely continuous component $y_{ia}^*(\cdot)$ of $y_i^*(\cdot)$ at t is strictly positive. In addition, however, $\chi_i(t) = 0$ if the singular component $y_{is}^*(\cdot)$ of $y_i^*(\cdot)$ is strictly increasing at t , e.g., if $y_{is}^*(\cdot)$ has a jump discontinuity at t . As discussed above, this requirement can also be interpreted as an instance of a maximum principle for the "slope" of $y_i^*(\cdot)$.

In Hellwig (2009), Theorem 3.1 provides the basis for a unified analysis of the contracting problem with unidimensional hidden characteristics, without any special assumptions about the distribution of the hidden characteristic and without any regularity assumption about the dependence of outcomes on types. In addition to the familiar properties of no distortion at the top and downward distortions below the top of the type distribution, the analysis actually shows that any mass point of the type distribution "below the top" must be bunched with immediately adjacent higher types. Moreover, unless the mass point is also bunched with lower types, the optimal incentive scheme is likely to be discontinuous in the transition from immediately adjacent lower types to the mass point.⁷ The discontinuity is due to the fact that, in (2.13), the density $\gamma(t)$ is not continuous in t if this density is merely a mathematical artefact that is introduced to make the incentive problem with an arbitrary type distribution amenable to control-theoretic methods.

4 Proof of the Theorem

Before turning to the proof of the theorem, I give two general results about the solutions to the constraint (2.3). Detailed proofs of these results are given in the Appendix. They are based on routine arguments from the theory of ordinary differential equations,⁸ adapted so as to take account of the fact that h need not be continuous in the variable of integration.

I begin with a statement concerning existence and uniqueness of solu-

⁷More precisely, the paper gives a simple set of sufficient conditions for such a discontinuity to occur. It also explains why such a discontinuity should be taken to be the rule rather than the exception.

⁸See, e.g., Luenberger (1969), pp. 276 f., or Shilov (1974), pp. 163 ff.

tions. The constraint (2.3) is equivalent to the integral equation

$$x(t) = x(0) + \int_0^t h(x(\tau), y(\tau), z(\tau), \tau) d\tau \quad (4.1)$$

holding for all t .

Lemma 4.1 *For any $x(0), y(\cdot), z(\cdot)$, there is a unique function $x(\cdot|x(0), y(\cdot), z(\cdot))$ that satisfies (4.1) for almost all t . The function $x(\cdot|x(0), y(\cdot), z(\cdot))$ is continuous.*

The following result concerns the differentiability properties of the relation between the exogenous data and the solution of the integral equation. Rather than studying this function as such, I fix a point $(x(0), y(\cdot), z(\cdot))$ and consider the function

$$(\Delta x, \Delta y(\cdot), \Delta z(\cdot)) \rightarrow x_\Delta(\cdot) = x(\cdot|x(0)+\Delta x, y(\cdot)+\Delta y(\cdot), z(\cdot)+\Delta z(\cdot)), \quad (4.2)$$

i.e., I study how a *change* in the initial state and in the time paths of the control variables affect the solution to (4.1). In this analysis, the image $x_\Delta(\cdot)$ is treated as an element of the space $C^\ell([0, 1])$ of continuous functions from $[0, 1]$ into \mathbb{R}^ℓ , endowed with the usual sup norm. The triple $\Delta = (\Delta x, \Delta y(\cdot), \Delta z(\cdot))$ is treated as an element of the product space $\mathbb{R}^\ell \times L_1^m \times L_1^n$ where \mathbb{R}^ℓ has the usual Euclidean norm, and L_1^m, L_1^n have the L_1 norms for integrable functions from $[0, 1]$ into \mathbb{R}^m , respectively \mathbb{R}^n .⁹ The product space $\mathbb{R}^\ell \times L_1^m \times L_1^n$ is given the norm $\|\Delta\| = \|\Delta x\| + \int \|\Delta y(\tau)\| d\tau + \int \|\Delta z(\tau)\| d\tau$. With this specification of a norm, one obtains:

Lemma 4.2 *For any $(x(0), y(\cdot), z(\cdot))$ and any t , the function (4.2) is Fréchet differentiable. At $\Delta = (\Delta x, \Delta y(\cdot), \Delta z(\cdot))$, the Fréchet differential of this function with increment $\delta = (\delta x, \delta y(\cdot), \delta z(\cdot))$ is given by the solution to the linear integral equation*

$$\xi(t) = \delta x + \int_0^t [h_x \cdot \xi(\tau) + h_y \cdot \delta y(\tau) + h_z \cdot \delta z(\tau)] d\tau, \quad (4.3)$$

where h_x, h_y, h_z are evaluated at $(x_\Delta(\tau), y(\tau) + \Delta y(\tau), z(\tau) + \Delta z(\tau))$.

⁹The desire to use the L_1 norm is the reason for looking at the dependence of $x(\cdot)$ on the increments $\Delta x, \Delta y(\cdot), \Delta z(\cdot)$, rather than $x(0), y(\cdot), z(\cdot)$. I do not want to assume that the function $y(\cdot)$ itself is integrable.

I now turn to the proof of Theorem 3.1. There is no loss of generality in assuming that the Lebesgue decomposition (3.3) satisfies $y_a^*(0) = y^*(0)$ and $y_s^*(0) = 0$. Monotonicity of $y_a^*(\cdot)$ implies that $q^*(t) \geq 0$ for all t .

Lemma 4.3 *For given $y_s^*(\cdot)$, the functions $x^*(\cdot), y_a^*(\cdot), z^*(\cdot), q^*(\cdot)$ maximize the integral*

$$\int_0^1 f(x(t), y_s^*(t) + y_a(t), z(t), t) dt \quad (4.4)$$

subject to the constraints that $x(\cdot)$ and $y_a(\cdot)$ be absolutely continuous and $z(\cdot)$ and $q(\cdot)$ be measurable, that

$$G_0(x(0)) \geq \bar{G}_0 \text{ and } y_s^*(0) + y_a(0) \geq 0 \quad (4.5)$$

and that, for almost all $t \in [0, 1]$,

$$x'(t) = h(x(t), y_s^*(t) + y_a(t), z(t), t), \quad (4.6)$$

$$y_a'(t) = q(t), \quad (4.7)$$

$$y_s^*(t) + y_a(t) \geq 0, \quad (4.8)$$

$$z(t) \geq 0, \quad (4.9)$$

and, finally,

$$q(t) \geq 0. \quad (4.10)$$

Proof. Because the functions $x^*(\cdot), y^*(\cdot)$, and $z^*(\cdot)$ satisfy the constraints (2.2) - (2.6), for given $y_s^*(\cdot)$, the functions $x^*(\cdot), y_a^*(\cdot), z^*(\cdot)$, and $q^*(\cdot)$ also satisfy the constraints (4.5) - (4.10). Further, if any functions $x(\cdot), y_a(\cdot), z(\cdot)$, and $q(\cdot)$ satisfy the constraints (4.5) - (4.10), then the functions $x(\cdot) = x^*(\cdot), y(\cdot) = y_s^*(\cdot) + y_a(\cdot)$, and $z(\cdot)$, satisfy the constraints (2.2) - (2.6) of the original maximization problem. Because $x^*(\cdot), y^*(\cdot)$, and $z^*(\cdot)$ maximize (2.1) under the constraints (2.2) - (2.6), it follows that, for any $x(\cdot), y_a(\cdot), z(\cdot)$, and $q(\cdot)$ which satisfy (4.5) - (4.10), one has

$$\begin{aligned} \int_0^1 f(x(t), y_s^*(t) + y_a(t), z(t), t) dt &\leq \int_0^1 f(x^*(t), y^*(t), z^*(t), t) dt \\ &= \int_0^1 f(x^*(t), y_s^*(t) + y_a^*(t), z^*(t), t) dt. \end{aligned}$$

The lemma follows immediately. ■

Lemma 4.4 *There exists a multiplier $\lambda \in \{0, 1\}$, a nonnegative vector ν_0 , and absolutely continuous functions $\varphi : [0, 1] \rightarrow \mathbb{R}^\ell$ and $\chi : [0, 1] \rightarrow \mathbb{R}^m$, such that the following hold:*

(a) *For almost all $t \in [0, 1]$,*

$$\varphi'(t) = -\lambda f_x(x^*(t), y^*(t), z^*(t), t) - \varphi(t) \cdot h_x(x^*(t), y^*(t), z^*(t), t); \quad (4.11)$$

(b) *moreover,*

$$\varphi(0) = -\nu_0 \cdot G_0'(x^*(0)), \quad \nu_0 \cdot G_0(x^*(0)) = 0, \quad (4.12)$$

and

$$\varphi(1) = 0. \quad (4.13)$$

(c) *For almost all $t \in [0, 1]$,*

$$\chi'(t) = -\lambda f_y(x^*(t), y^*(t), z^*(t), t) - \varphi(t) \cdot h_y(x^*(t), y^*(t), z^*(t), t); \quad (4.14)$$

(d) *moreover,*

$$\chi(0) \leq 0, \quad \chi(0) \cdot y^*(0) = 0, \quad (4.15)$$

and

$$\chi(1) = 0. \quad (4.16)$$

(e) *For almost all t ,*

$$\begin{aligned} (z^*(t), q^*(t)) \in \arg \max_{z \geq 0, q \geq 0} & [\lambda f(x^*(t), y^*(t), z, t) + \varphi(t) \cdot h(x^*(t), y^*(t), z, t) \\ & + \left(\chi(t) - \int_0^t \eta(\tau) d\mu(\tau) \right) \cdot q]. \end{aligned} \quad (4.17)$$

(f) $\lambda + \|\varphi\| + \|\chi\| > 0$.

Proof. If one writes

$$\hat{f}(x(t), y_a, z, t) := f(x, y_s^*(t) + y_a, z, t), \quad (4.18)$$

and

$$\hat{h}(x(t), y, z, t) := h(x, y_s^*(t) + y, z, t), \quad (4.19)$$

one sees that, for given $y_s^*(\cdot)$, the maximization problem specified in Lemma 4.3 is a standard control problem with absolutely continuous state variables

$x(\cdot)$ and $y_a(\cdot)$ and with measurable control variables $z(\cdot)$ and $q(\cdot)$. If the constraint (4.8) on the state variable $y_a(\cdot)$ is rewritten in the form,

$$-\min_i [y_{ai}(t) + y_{si}^*(t)] \leq 0, \quad (4.20)$$

this control problem takes the form assumed in Theorem 5.2.1, p. 211, of Clarke (1983). The Hamiltonian function is given by the formula

$$\hat{H}(x, y_a, z, q, t; \lambda, \varphi, \chi) = \lambda \hat{f}(x, y_a, z, t) + \varphi \cdot \hat{h}(x, y_a, z, t) + \chi \cdot q. \quad (4.21)$$

By Theorem 5.2.1, p. 211, in Clarke (1983),¹⁰ it follows that there exist $\lambda, \varphi, \chi, \eta, \mu$, as specified in the lemma, that satisfy statements (a), (c), (e), and (f), as well as the following:

The pair $(\varphi(0), \chi(0))$ belongs to the Clarke normal cone at $(x^*(0), y_a^*(0))$ to the set of (x, y_a) satisfying $G_0(x) \geq \bar{G}_0$ and $y_a \geq -y_s^*(0)$. By the Corollary to Theorem 2.4.5, p 54, and Corollary 2 to Theorem 2.4.7, p. 56 f., in Clarke (1983), it follows that $\varphi(0)$ and $\chi(0)$ satisfy (4.12) and (4.15). Moreover, in the absence of any explicit constraints¹¹ on the the pair $(x(1), y_a(1))$, the pair $(-\varphi(1), -\chi(1))$ belongs to the Clarke normal cone at $(x^*(1), y_a^*(1))$ to the set \mathbb{R} . This yields (4.13) and (4.16). ■

Lemma 4.5 *The multiplier λ in Lemma 4.4 is nonzero, i.e., $\lambda = 1$.*

Proof. If the multiplier λ in Lemma 4.4 were equal to zero, equations (4.11) and (4.14) would take the form

$$\varphi'(t) = -\varphi(t) \cdot h_x(x^*(t), y^*(t), z^*(t), t) \quad (4.22)$$

and

$$\chi'(t) = -\varphi(t) \cdot h_y(x^*(t), y^*(t), z^*(t), t). \quad (4.23)$$

The validity of (4.22) and (4.23) for almost all t is equivalent to the integral equations

$$\varphi(t) = \varphi(1) - \int_t^1 \varphi(\tau) \cdot h_x(x^*(\tau), y^*(\tau), z^*(\tau), \tau) d\tau \quad (4.24)$$

¹⁰In applying Clarke's theorem, due allowance must be made for the fact that he deals with a minimization, rather than a maximization problem. Also, the function η that I use is the negative of Clarke's γ .

¹¹The constraint $y_a(1) \geq -y_s^*(1)$ is automatically satisfied if $y_a(0) \geq -y_s^*(0)$, $q(t) \geq 0$ for all t , and $y_s^*(1) \geq y_s^*(0)$.

and

$$\chi(t) = \chi(1) - \int_t^1 \varphi(\tau) \cdot h_y(x^*(\tau), y^*(\tau), z^*(\tau), \tau) d\tau \quad (4.25)$$

holding for all t . For the boundary conditions (4.13) and (4.16), this pair of integral equations has a solution satisfying

$$\varphi(t) = 0 \text{ and } \chi(t) = 0 \quad (4.26)$$

for all t . By Lemma A.2 in the Appendix, the solution is unique.

Thus, $\lambda = 0$ implies $\lambda + \|\varphi\| + \|\chi\| = 0$, which contradicts statement (f) of Lemma 4.4. ■

Given that $\lambda = 1$, statements (a) - (d) of Lemma 4.4 are the same to statements (a) - (d) of Theorem 3.1. Moreover, statement (e) of the lemma is equivalent to the requirement that, for almost all t , one has

$$z^*(t) \in \arg \max_{z \geq 0} [\lambda f(x^*(t), y^*(t), z, t) + \varphi(t) h(x^*(t), y^*(t), z, t)] \quad (4.27)$$

and

$$\chi(t) \leq 0 \text{ and } \chi(t) \cdot q^*(t) = 0. \quad (4.28)$$

Statement (e) of the lemma thus implies statement (e) and the first part of statement (f) of the theorem.

Statement (e) of the lemma also implies that the second part of statement (f) of the theorem holds for almost every t at which the slope q_i^* of the absolutely continuous part of y_i^* is strictly positive. Indeed, upon integrating the equation in (4.28), one obtains

$$\int_0^1 \chi(t) \cdot dy_a^*(t) = 0, \quad (4.29)$$

which shows that, for any i and almost every t , $\chi_i(t) = 0$ if t is a point of increase of the absolutely continuous part $y_{ai}^*(t)$ of $y_i^*(t)$.

To complete the proof of Theorem 3.1, it remains to be shown that (4.29) continues to hold if y_a^* is replaced by y^* itself. This is the point of the following lemma.

Lemma 4.6 *The functions χ and y^* satisfy the equation*

$$\int_0^1 \chi(t) \cdot dy^*(t) = 0. \quad (4.30)$$

Proof. Given (4.29), it suffices to prove that

$$\int_0^1 \chi(t) \cdot dy_s^*(t) = 0. \quad (4.31)$$

For this purpose, fix some (large) vector \bar{y} in \mathbb{R}_+^m and consider the effects of replacing the function $y^*(\cdot)$ by the function

$$y(\cdot|\alpha, \bar{y}) := y^*(\cdot) + \alpha y_s^*(\cdot|\bar{y}), \quad (4.32)$$

where α is close to zero and $y_s^*(\cdot|\bar{y})$ is given by the formula

$$y_s^*(t|\bar{y}) = \min(y_s^*(t), \bar{y}). \quad (4.33)$$

Since, trivially, $y^*(\cdot) = y(\cdot|0, \bar{y})$, this amounts to considering the effects of changes in α away from zero. If $y(\cdot|\alpha, \bar{y})$ replaces $y(\cdot|0, \bar{y})$, the path of the state variable is given as

$$x(\cdot|\alpha, \bar{y}) = x(\cdot|x^*(0), y(\cdot|\alpha, \bar{y}), z^*(\cdot)) = x(\cdot|x^*(0), y^*(\cdot) + \alpha y_s^*(\cdot|\bar{y}), z^*(\cdot)). \quad (4.34)$$

Because the function $y_s^*(\cdot|\bar{y})$ is bounded, it belongs to L_1^m . By Lemma 4.2, therefore, the function $\alpha \rightarrow x(\cdot|\alpha, \bar{y})$ is continuously differentiable. Its derivative is given by the solution to the integral equation

$$\frac{\partial x(t|\alpha, \bar{y})}{\partial \alpha} = \int_0^t \left[h_x \frac{\partial x(\tau|\alpha, \bar{y})}{\partial \alpha} + h_y \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha} \right] d\tau. \quad (4.35)$$

By (4.32), one also has

$$\frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha} = y_s^*(t|\bar{y}) \quad (4.36)$$

for all t .

If $\alpha \geq -1$, the function $y(\cdot|\alpha, \bar{y})$ is nonnegative-valued and nondecreasing. For such α , therefore, the triple $(x(\cdot|\alpha, \bar{y}), y(\cdot|\alpha, \bar{y}), z^*(\cdot))$ satisfies the constraints (2.2) - (2.6) of the original maximization problem. Because $(x^*(\cdot), y^*(\cdot), z^*(\cdot)) = (x(\cdot|0, \bar{y}), y(\cdot|0, \bar{y}), z^*(\cdot))$ is a solution to this problem, it follows that

$$\int_0^1 f(x(t|0, \bar{y}), y(t|0, \bar{y}), z^*(t), t) dt \geq \int_0^1 f(x(t|\alpha, \bar{y}), y(t|\alpha, \bar{y}), z^*(t), t) dt. \quad (4.37)$$

for all $\alpha \geq -1$. Because the functions $(x, y, z) \rightarrow f(x, y, z, t)$ and $\alpha \rightarrow (x(t|\alpha, \bar{y}), y(t|\alpha, \bar{y}))$ are continuously differentiable, the function

$$\alpha \rightarrow f(x(t|\alpha, \bar{y}), y(t|\alpha, \bar{y}), z^*(t), t)$$

is also continuously differentiable, with derivative

$$\frac{\partial}{\partial \alpha} f(x(t|\alpha, \bar{y}), y(t|\alpha, \bar{y}), z^*(t), t) = f_x \cdot \frac{\partial x(t|\alpha, \bar{y})}{\partial \alpha} + f_y \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha}.$$

Because the functions $(x, y, z) \rightarrow f(x, y, z, t)$, $t \in [0, 1]$, are equi-Lipschitzian, because the solution to the integral equation (4.35) is a Fréchet derivative of the function $\alpha \rightarrow x(\cdot|\alpha, \bar{y})$, and because the function $y_s^*(\cdot|\bar{y})$ belongs to L_1^m , condition (4.37) in fact yields the first-order condition

$$\int_0^1 \left[f_x \cdot \frac{\partial x(t|\alpha, \bar{y})}{\partial \alpha} + f_y \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha} \right] dt = 0 \quad (4.38)$$

at $\alpha = 0$, where, for each t , f_x and f_y are evaluated at the point $(x^*(t), y^*(t), z^*(t), t)$.

From (4.11), (4.14), and the fact that $\lambda = 1$, it follows that

$$\int_0^1 [-\varphi'(t) - \varphi(t) \cdot h_x] \cdot \frac{\partial x(t|\alpha, \bar{y})}{\partial \alpha} dt + \int_0^1 [-\chi'(t) - \varphi(t) \cdot h_y] \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha} dt = 0. \quad (4.39)$$

where for each t , h_x and h_y are also evaluated at $(x^*(t), y^*(t), z^*(t), t)$.

By (4.35), one also has

$$\int_0^1 \varphi'(t) \cdot \frac{\partial x(t|\alpha, \bar{y})}{\partial \alpha} dt = \int_0^1 \varphi'(t) \cdot \int_0^t \left[h_x \cdot \frac{\partial x(\tau|\alpha, \bar{y})}{\partial \alpha} + h_y \cdot \frac{\partial y(\tau|\alpha, \bar{y})}{\partial \alpha} \right] d\tau dt. \quad (4.40)$$

The right-hand side of (4.40) can be integrated by parts. Because (4.13) yields $\varphi(1) = 0$, it follows that

$$\int_0^1 \varphi'(t) \cdot \frac{\partial x(t|\alpha, \bar{y})}{\partial \alpha} dt = - \int_0^1 \varphi(t) \cdot \left[h_x \cdot \frac{\partial x(t|\alpha, \bar{y})}{\partial \alpha} + h_y \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha} \right] dt. \quad (4.41)$$

Equation (4.39) therefore implies

$$- \int_0^1 \chi'(t) \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha} dt = 0. \quad (4.42)$$

Upon using (4.36) to substitute for $\frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha}$ and integrating (4.42) by parts, one obtains the equation

$$- [\chi(t) \cdot y_s^*(t|\bar{y})]_0^1 + \int_0^1 \chi(t) \cdot dy_s^*(t|\bar{y}) = 0 \quad (4.43)$$

Because $y_s^*(0|\bar{y}) = 0$, it follows that

$$\int_0^1 \chi(t) \cdot dy_s^*(t|\bar{y}) - \chi(1) \cdot y_s^*(1|\bar{y}) = 0. \quad (4.44)$$

By (4.16), therefore,

$$\int_0^1 \chi(t) \cdot dy_s^*(t|\bar{y}) = 0.$$

Upon taking limits as \bar{y}_i goes out of bounds for all i , one obtains (4.31). ■

Corollary 4.7 *For almost every t and any i , $\chi_i(t) = 0$ if $y_i^*(\cdot)$ is increasing at t .*

Proof. For any i , let Y_i be the measure on $[0, 1]$ that is induced by the nondecreasing function $y_i^*(\cdot)$. Because $\chi_i(t)$ is nonpositive for all i and almost all t , (4.30) implies that $\chi_i(t) = 0$ for Y_i -almost all t . If $y_i^*(\cdot)$ is increasing at t_0 , one has $Y_i((t_0 - \varepsilon, t_0 + \varepsilon]) > 0$ for all $\varepsilon > 0$. Hence there exists a sequence $\{t^k\}$ converging to t_0 such that $\chi_i(t^k) = 0$ for all k . Because $\chi_i(\cdot)$ is absolutely continuous, it follows that $\chi_i(t_0) = 0$. ■

The proof of Theorem 3.1 is now complete.

5 Control Problems with Integral Constraints or Explicit Constraints on Both Endpoints

None of the constraints (2.2) - (2.6) is of the form

$$\int_0^1 g(x(t), y(t), t) dt \geq 0. \quad (5.1)$$

However, such constraints occur frequently in economic applications. As an example, consider the optimal income tax problem of Mirrlees (1971). This problem involves the same utility specification and the same incentive compatibility conditions as the agency problem in Section 2. However, the indirect utility function $v(\cdot)$ and output requirement function $y(\cdot)$ are chosen to maximize the welfare objective

$$\int_0^1 v(t) dt \quad (5.2)$$

subject to the feasibility constraint

$$\int_0^1 (y(t) - c(v(t), y(t), t)) \gamma(t) dt \geq 0, \quad (5.3)$$

and the incentive compatibility conditions (2.9) and (2.10), where, as before, $c(v(t), y(t), t)$ is implicitly defined by the equation $v(t) = u(c(v(t), y(t), t), y(t), t)$.

Integral constraints of the form (5.1) can in principle be handled by introducing additional state variables. Thus, condition (5.1) is equivalent to the requirement that the additional state variable η that is defined by the requirements that

$$\eta(0) = 0 \tag{5.4}$$

and that

$$\eta'(t) = g(x(t), y(t), t) \tag{5.5}$$

for all t should also satisfy the endpoint condition

$$\eta(1) \geq 0. \tag{5.6}$$

If the vector of state variables is written as $(x(\cdot), \eta(\cdot))$, one can replace the integral constraint (5.1) by conditions (5.4) - (5.6). Given that (5.4) is equivalent to the twin constraints $\eta(0) \geq 0$ and $-\eta(0) \geq 0$, conditions (5.4) and (5.5) are in the format of conditions (2.2) and (2.3). However, (5.6) requires the addition of an endpoint condition for $t = 1$. The question of how to deal with integral constraints can thus be seen as a special case of the more general question of how to deal with endpoint conditions that are imposed for $t = 1$ in addition to the endpoint condition for $t = 0$.

The additional endpoint condition raises the possibility that, in a neighbourhood of the optimal $(x^*(\cdot), y^*(\cdot), z^*(\cdot))$, the constraint set might be degenerate, in which case the multiplier λ in Lemma 4.4 might be equal to zero. The additional endpoint condition also raises the possibility that the variation $y(\cdot|\alpha, \bar{y})$ of the control $y^*(\cdot)$ that is considered in the proof of Lemma 4.6 might not be feasible because, for given $x^*(0)$ and $z^*(\cdot)$, the resulting $x(1|\delta, \bar{y})$ violates the endpoint condition for $t = 1$. To avoid these difficulties, one needs an additional assumption about the data of the problem.

Regularity Condition The state variable can be written in the form

$$x(t) = (\zeta(t), \eta(t))$$

so that

(i) the endpoint conditions take the form

$$G_0(\eta(0)) \geq \bar{G}_0, \tag{5.7}$$

$$G_1(\eta(1)) \geq \bar{G}_1, \tag{5.8}$$

i.e., they are independent of $\zeta(0)$ and $\zeta(1)$; the functions G_0 and G_1 are continuously differentiable functions, with Jacobians that have maximal rank;

(ii) moreover, for any $x(0) = (\zeta(0), \eta(0)), y(\cdot), z(\cdot)$, there exists $\Delta\zeta$ such that, at the point $\alpha = 0$, the derivative of the function

$$\alpha \rightarrow G_1(\eta(1|(\zeta(0) + \alpha\Delta\zeta, \eta(0)), y(\cdot), z(\cdot))) \quad (5.9)$$

is strictly positive; i.e., a suitable change in $\zeta(0)$, leaving $\eta(0)$, $y(\cdot)$, and $z(\cdot)$ unchanged, is sufficient to move the component $\eta(1)$ of the state

$$\begin{aligned} & x(1|(\zeta(0) + \alpha\Delta\zeta, \eta(0)), y(\cdot), z(\cdot)) \\ &= (\zeta(1|(\zeta(0) + \alpha\Delta\zeta, \eta(0)), y(\cdot), z(\cdot)), \eta(1|(\zeta(0) + \alpha\Delta\zeta, \eta(0)), y(\cdot), z(\cdot))) \end{aligned}$$

at $t = 1$ in such a way that the left-hand side of the endpoint condition is increased.

In the absence of participation constraints, the *optimal income tax problem* satisfies the Regularity Condition with $\zeta = v$ and

$$\eta(t) = \int_0^t (y(\tau) - c(v(\tau), y(\tau), \tau))\gamma(\tau) d\tau$$

for all t . When there are no participation constraints, the state variable v is not subject to any endpoint constraints. Moreover, under the incentive compatibility condition (2.9), a reduction in $v(0)$, keeping $y(\cdot)$ unchanged, induces a decrease in $c(v(t), y(t), t)$ for all t and thereby an increase in the surplus $\int_0^1 (y(t) - c(v(t), y(t), t))\gamma(t) dt$.

Even so, the Regularity Condition is quite restrictive. It should, however, be regarded as a prototype, rather than the most general condition that is available. Its simplicity facilitates the presentation of the argument. This makes it easy to see what aspects of the condition matter and what scope there would be extending the analysis. I will return to this issue at the end of the paper.

Theorem 5.1 *Assume that, in addition to the specified measurability, differentiability, and equi-Lipschitz conditions, the Regularity Condition holds. If $x^*(\cdot) = (\zeta^*(\cdot), \eta^*(\cdot)), y^*(\cdot), z^*(\cdot)$ maximizes (2.1) subject to the constraints (5.7), (5.8), and (2.3) - (2.6), there exists a function $q^* : [0, 1] \rightarrow \mathbb{R}_+^m$, and*

there exist nonnegative vectors ν_0 and ν_1 , and absolutely continuous functions $\varphi = (\varphi_\zeta, \varphi_\eta) : [0, 1] \rightarrow \mathbb{R}^\ell$, $\chi : [0, 1] \rightarrow \mathbb{R}^m$, such that $x^*(\cdot), y^*(\cdot), z^*(\cdot), q^*(\cdot), \varphi(\cdot)$, and $\chi(\cdot)$ satisfy statements (a) and (c) - (f) of Theorem 3.1, and, in addition, the transversality condition

(b*)

$$\varphi_\zeta(0) = \varphi_\zeta(1) = 0, \quad (5.10)$$

$$\varphi_\eta(0) = -\nu_0 \cdot G'_0(\eta^*(0)) \text{ and } \nu_0 \cdot G_0(\eta^*(0)) = 0, \quad (5.11)$$

and

$$\varphi_\eta(1) = \nu_1 \cdot G'_1(\eta^*(1)) \text{ and } \nu_1 \cdot G_1(\eta^*(1)) = 0. \quad (5.12)$$

According to this result, optimality conditions for the control problem with endpoint conditions (5.7), (5.8) are the same as the optimality conditions in Theorem 3.1, except that the transversality conditions (3.5) - (3.6) are replaced by (5.10) - (5.12). The change reflects the difference in endpoint conditions.

The proof of Theorem 5.1 is, by and large, the same as the proof of Theorem 3.1. Therefore, I will merely provide a sketch, focussing on those arguments that are different. I first note that the *analogues of Lemmas 4.3 and 4.4*, with duly modified endpoint and transversality conditions, can be established by the same arguments as the earlier lemmas. The details are left to the reader. I next turn to the *analogue of Lemma 4.5*. For this, one needs the Regularity Condition.

Lemma 5.2 *The multiplier λ in the analogue of Lemma 4.4 is nonzero, i.e., $\lambda = 1$.*

Proof. If one had $\lambda = 0$, the costate variables would again satisfy (4.22) and (4.23). If (4.22) is postmultiplied by a vector $\xi(t)$, one obtains

$$\varphi'(t) \cdot \xi(t) = -\varphi(t) \cdot h_x(x^*(t), y^*(t), z^*(t), t) \cdot \xi(t). \quad (5.13)$$

If the vectors $\xi(t)$, $t \in [0, 1]$, are chosen to satisfy the integral equation

$$\xi(t) = \delta x + \int_0^t h_x(x^*(\tau), y^*(\tau), z^*(\tau), \tau) \cdot \xi(\tau) d\tau \quad (5.14)$$

for some $\delta x = (\delta\zeta, \delta\eta) \in \mathbb{R}^\ell$, then it follows that $\varphi(t) \cdot \xi(t)$ is independent of t , and one has

$$\varphi(1) \cdot \xi(1) = \varphi(0) \cdot \xi(0). \quad (5.15)$$

By (5.14) and (5.10), it follows that

$$\varphi(1) \cdot \xi(1) = \varphi(0) \cdot \delta x = \varphi_\zeta(0) \cdot \delta \zeta + \varphi_\eta(0) \cdot \delta \eta = \varphi_\eta(0) \cdot \delta \eta.$$

Therefore,

$$\varphi(1) \cdot \xi(1) = 0 \tag{5.16}$$

if $\delta \eta = 0$. Equation (5.16) can also be written in the form

$$\varphi_\zeta(1) \cdot \xi_\zeta(1) + \varphi_\eta(1) \cdot \xi_\eta(1) = 0, \tag{5.17}$$

where $\xi(1) = (\xi_\zeta(1), \xi_\eta(1))$. By (5.10) and (5.12), it follows that

$$\nu_1 \cdot G'_1(\eta^*(1)) \cdot \xi_\eta(1) = 0. \tag{5.18}$$

By Lemma 4.2, $\xi(1)$ is the derivative, evaluated at $\alpha = 0$, of the function

$$\alpha \rightarrow x(1|x^*(0) + \alpha \delta x, y^*(\cdot), z^*(\cdot)) \tag{5.19}$$

and $\xi_\eta(1)$ is the derivative, evaluated at $\alpha = 0$, of the function

$$\alpha \rightarrow \eta(1|x^*(0) + \alpha \delta x, y^*(\cdot), z^*(\cdot)). \tag{5.20}$$

By the Regularity Condition, it follows that

$$G'_1(\eta^*(1)) \cdot \xi_\eta(1) \gg 0 \tag{5.21}$$

if $(\delta \zeta, \delta \eta) = (\Delta \zeta, 0)$. From (5.18) and (5.21) together, one infers that $\nu_1 = 0$. By (5.12) and (5.10), it follows that $\varphi(1) = 0$ and, by (4.22), that $\varphi(t) = 0$ for all t . The rest of the argument proceeds as before: By (4.23) and (4.16), $\varphi(t) = 0$ for all t implies $\chi(t) = \chi(1) = 0$ for all t . Thus, $\lambda = 0$ implies $\lambda + \|\varphi\| + \|\chi\| = 0$, which contradicts the analogue of statement (f) in Lemma 4.4. ■

The Regularity Condition is also critical for the *analogue of Lemma 4.6*. Two issues arise. First, if one replaces the function $y^*(\cdot)$ by the function

$$y(\cdot|\alpha, \bar{y}) := y^*(\cdot) + \alpha y_s^*(\cdot|\bar{y}), \tag{5.22}$$

the change may induce a violation of the endpoint condition (5.8). Second, the derivation of the equation (4.42),

$$-\int_0^1 \chi'(t) \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha}(0) dt = 0, \tag{5.23}$$

relied on the transversality condition (4.13), $\varphi(1) = 0$, which, in the present context, has been replaced by the conditions $\varphi_\zeta(1) = 0$ and $\varphi_\eta(1) = \nu_1 \cdot G'_1(\eta^*(1))$. The following lemma shows that the Regularity Condition yields $\varphi(1) = 0$, as in the absence of the endpoint constraint for $t = 1$.

Lemma 5.3 *Under the Regularity Condition, $\nu_1 = 0$ and $\varphi(1) = 0$.*

Proof. Consider the function

$$\alpha \rightarrow \hat{x}(\cdot|\alpha) = x(\cdot|x^*(0) + \alpha(\Delta\zeta, 0), y^*(\cdot), z^*(\cdot)),$$

where $\Delta\zeta$ is the vector provided by the Regularity Condition. For any α , the triple $\hat{x}(\cdot|\alpha), y^*(\cdot), z^*(\cdot)$ satisfies the constraints (5.7), (5.8), and (2.3) - (2.6). By the same argument as in the proof of Lemma 4.6, one must therefore have

$$\int_0^1 f_x \cdot \frac{\partial \hat{x}(t|\alpha)}{\partial \alpha}(0) dt = 0. \quad (5.24)$$

By the analogues of (4.11) in Lemma 4.4 and Lemma 4.5, it follows that

$$\int_0^1 (\varphi'(t) + h_x \cdot \varphi(t)) \cdot \frac{\partial \hat{x}(t|\alpha)}{\partial \alpha}(0) dt = 0. \quad (5.25)$$

The derivative $\frac{\partial \hat{x}(t|\alpha)}{\partial \alpha}$ satisfies the integral equation

$$\frac{\partial \hat{x}(t|\alpha)}{\partial \alpha} = (\Delta\zeta, 0) + \int_0^t h_x \cdot \frac{\partial \hat{x}(\tau|\alpha)}{\partial \alpha}(0) d\tau. \quad (5.26)$$

By the same procedure as in the proof of Lemma 4.6, one can therefore compute the integral on the left-hand side of (5.25). This yields the equation

$$\varphi(1) \int_0^1 h_x \cdot \frac{\partial \hat{x}(\tau|\alpha)}{\partial \alpha}(0) d\tau = 0, \quad (5.27)$$

or

$$\varphi(1) \cdot \frac{\partial \hat{x}(1|\alpha)}{\partial \alpha} - \varphi_\zeta(1) \cdot \Delta\zeta - \varphi_\eta(1) \cdot 0 = 0. \quad (5.28)$$

Trivially, the last term vanishes. By (5.10), the second term also vanishes, and the first term is equal to $\varphi(1) \cdot \frac{\partial \hat{\eta}(1|\alpha)}{\partial \alpha}$, where

$$\hat{\eta}(\cdot|\alpha) := \eta(\cdot|x^*(0) + \alpha(\Delta\zeta, 0), y^*(\cdot), z^*(\cdot)).$$

By (5.12), therefore, (5.28) implies

$$\nu_1 \cdot G'_1(\eta^*(1)) \cdot \frac{\partial \hat{\eta}(1|\alpha)}{\partial \alpha} = 0. \quad (5.29)$$

Because $\nu_1 \leq 0$ and the Regularity Condition implies that $G'_1(\eta^*(1)) \cdot \frac{\partial \hat{\eta}(1|\alpha)}{\partial \alpha} \gg 0$, it follows that $\nu_1 = 0$. By (5.10) and (5.12), therefore, $\varphi(1) = 0$. ■

Lemma 5.4 *The functions χ , M , and y^* satisfy equation*

$$\int_0^1 (\chi(t) - M(t)) \cdot dy^*(t) = 0. \quad (5.30)$$

Proof. As in the proof of Lemma 4.6, I consider the effects of replacing the function $y^*(\cdot)$ by the function

$$y(\cdot|\alpha, \bar{y}) := y^*(\cdot) + \alpha y_s^*(\cdot|\bar{y}). \quad (5.31)$$

To neutralize any effects that this may have on the endpoint condition (5.8), I will add a suitable scalar multiple ε of the vector $(\Delta\zeta, 0)$ to the initial state $x^*(0)$. The path of the state variable is then given as

$$x(\cdot|\alpha, \varepsilon, \bar{y}) = x(\cdot|x^*(0) + \varepsilon(\Delta\zeta, 0), y^*(\cdot) + \alpha y_s^*(\cdot|\bar{y}), z^*(\cdot)).$$

The function $(\alpha, \varepsilon) \rightarrow x(\cdot|\alpha, \varepsilon, \bar{y}) = (\zeta(\cdot|\alpha, \varepsilon, \bar{y}), \eta(\cdot|\alpha, \varepsilon, \bar{y}))$ is continuously differentiable. If one sets $\varepsilon = -c\alpha$, then, at $\alpha = 0$, the derivative of the function

$$\alpha \rightarrow G_1(\eta(1|\alpha, -c\alpha, \bar{y}))$$

is equal to

$$G_1'(\eta^*(1)) \left[\frac{\partial \eta(1|\alpha, \varepsilon, \bar{y})}{\partial \alpha}(0, 0) - c \frac{\partial \eta(1|\alpha, \varepsilon, \bar{y})}{\partial \varepsilon}(0, 0) \right] \quad (5.32)$$

The Regularity Condition implies that $G_1'(\eta^*(1)) \frac{\partial \eta(1|\alpha, \varepsilon, \bar{y})}{\partial \varepsilon}(0, 0) \gg 0$. By choosing c large enough, one can ensure that the derivative (5.32) is strictly negative. This implies that, for $\alpha < 0$, sufficiently close to zero, the triple $x(\cdot|\alpha, -c\alpha, \bar{y}), y(\cdot|\alpha, \bar{y}), z^*(\cdot)$ satisfies the endpoint constraint (5.8), as well as (5.7) and (2.3) - (2.6). Because $(x^*(\cdot), y^*(\cdot), z^*(\cdot)) = (x(\cdot|0, 0, \bar{y}), y(\cdot|0, \bar{y}), z^*(\cdot))$ maximizes (2.2) under these constraints, it follows that

$$\int_0^1 f(x(t|0, 0, \bar{y}), y(t|0, \bar{y}), z^*(t), t) dt \geq \int_0^1 f(x(t|\alpha, -c\alpha, \bar{y}), y(t|\alpha, \bar{y}), z^*(t), t) dt. \quad (5.33)$$

for all negative α that are close to zero. By the same arguments as in the proof of Lemma 4.6, this yields the first-order condition

$$\int_0^1 \left[f_x \cdot \frac{dx(t|\alpha, -c\alpha, \bar{y})}{d\alpha}(0-) + f_y \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha}(0-) \right] dt \geq 0. \quad (5.34)$$

and, from this first-order condition, the inequality

$$\int_0^1 [-\varphi'(t) - \varphi(t) \cdot h_x] \cdot \frac{dx(t|\alpha, -c\alpha, \bar{y})}{d\alpha} dt + \int_0^1 [-\chi'(t) - \varphi(t) \cdot h_y] \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha} dt \geq 0. \quad (5.35)$$

The derivative $\frac{\partial x(t|\alpha, -c\alpha, \bar{y})}{\partial \alpha}$ is given by the integral equation

$$\frac{dx(t|\alpha, -c\alpha, \bar{y})}{d\alpha} = -c(\Delta\zeta, 0) + \int_0^t \left[h_x \cdot \frac{dx(\tau|\alpha, -c\alpha, \bar{y})}{d\alpha} + h_y \cdot \frac{\partial y(\tau|\alpha, \bar{y})}{\partial \alpha} \right] d\tau. \quad (5.36)$$

Therefore, one computes

$$\begin{aligned} & \int_0^1 \varphi'(t) \cdot \frac{dx(t|\alpha, -c\alpha, \bar{y})}{d\alpha} dt \\ = & -c[\varphi_\zeta(1) - \varphi_\zeta(0)] \cdot \Delta\zeta \\ & + \int_0^1 \varphi'(t) \cdot \int_0^t \left[h_x \cdot \frac{dx(\tau|\alpha, -c\alpha, \bar{y})}{d\alpha} + h_y \cdot \frac{\partial y(\tau|\alpha, \bar{y})}{\partial \alpha} \right] d\tau dt \end{aligned}$$

By (5.10), the first term on the right-hand side is equal to zero. If one integrates the second term by parts, using the fact that, by Lemma 5.3, one has $\varphi(1) = 0$, one obtains

$$\int_0^1 \varphi'(t) \cdot \frac{dx(t|\alpha, -c\alpha, \bar{y})}{d\alpha} dt = - \int_0^1 \varphi(t) \cdot \left[h_x \cdot \frac{dx(\tau|\alpha, -c\alpha, \bar{y})}{d\alpha} + h_y \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha} \right] dt,$$

as in the proof of Lemma 4.6. By the same calculation as before, one then obtains

$$- \int_0^1 \chi'(t) \cdot \frac{\partial y(t|\alpha, \bar{y})}{\partial \alpha}(0) dt \leq 0, \quad (5.37)$$

and

$$\int_0^1 \chi(t) \cdot dy_s^*(t) \geq 0. \quad (5.38)$$

By the same argument as in the proof of Corollary 4.7, it follows that, for almost every t and $i = 1, \dots, m$, $\chi_i(t) = 0$ if $y_i^*(\cdot)$ is increasing at t , and, therefore, that $\int_0^1 \chi(t) \cdot dy_s^*(t) = 0$. Given that, by the analogue of Lemma 4.4, $\chi(t) \cdot q^*(t) = 0$ for all t , the lemma follows immediately. ■

To conclude the paper, I briefly discuss the role of the Regularity Condition. This condition is used to establish that $\lambda = 1$ and to establish the inequality (5.38). In the proof that $\lambda = 1$, the Regularity Condition could be replaced by any other condition ensuring that, in a neighbourhood of the

optimal $(x^*(\cdot), y^*(\cdot), z^*(\cdot))$, the constraint is not degenerate. To establish (5.38), however, something more specific is needed.

In this latter context, two aspects of the Regularity Condition matter. First, if the variation $y(\cdot|\alpha, \bar{y})$ of $y^*(\cdot)$ that is considered in the proof of Lemma 4.6 induces a violation of the endpoint condition for $t = 1$, it must be possible to find a countervailing variation in some other variable. Here, this variable is the component $\zeta(0)$ of the initial state $x(0)$. Second, at the level of first-order conditions, the countervailing variation in the other variable does not affect the evaluation of the overall variation because, at the optimum, one is indifferent to small changes in this variable. Under the Regularity Condition, a small change in $\zeta(0)$ is a matter of indifference because this component of the initial state $x(0)$ is not subject to any endpoint condition, and the associated costate variable $\varphi_\zeta(0)$ is zero.

One easily verifies that the conclusions of Theorem 5.1 remain valid whenever these two conditions hold for *some* variable of the optimization. Suppose, for example, that a uniform increase $\Delta y \geq 0$ in the level of $y(t)$ for all t can be used to raise $G_1(x(1))$ and, thereby, to counteract any effects that the replacement of $y^*(\cdot)$ by $y(\cdot|\alpha, \bar{y})$ might have on the endpoint constraint. Suppose also that the nonnegativity constraint on $y(\cdot)$ is nowhere binding, i.e., that the function $M(\cdot)$ in the preceding analysis is identically equal to zero. Then a minor variation of the argument given here shows that the conclusions of Theorem 5.1 are still valid.

However, the last part of the argument breaks down if the nonnegativity constraints on $y(\cdot)$ are binding somewhere, so that a small uniform increase in the level of $y(t)$ for all t is not a matter of indifference. In this case, the argument yields the inequality

$$\int_0^1 \chi(t) \cdot dy_s^*(t|\bar{y}) \geq -\hat{c}(\chi(1) - \chi(0))\Delta y = \hat{c}\chi(0)\Delta y \leq 0, \quad (5.39)$$

where \hat{c} is a constant indicating the relation between the multiple of Δy by which $y(t)$ is raised for all t and the amount by which the parameter α lies below zero. By contrast to (5.38), this condition is not suitable for signing the left-hand side.

The same arguments apply if a suitable change in the other control variables, i.e., the function $z(\cdot)$, can be used to counteract the effects of the replacement of $y^*(\cdot)$ by $y(\cdot|\alpha, \bar{y})$ on the endpoint constraint. If the optimal $z^*(\cdot)$ is characterized by first-order conditions holding as equations, the conclusions of Theorem 5.1 are still valid. If the optimal $z^*(\cdot)$ is characterized by first-order conditions holding as strict inequalities, one ends up with a condition like (5.39), which does not permit any further inference. The

question then is whether this invalidates the conclusions of the theorem or whether these conclusions could still be obtained by some other argument. For now, unfortunately, I have to leave this question open.

A Appendix

Proof of Lemma 4.1. As mentioned in the text, the argument is by and large the same as in the standard treatment of existence, uniqueness and continuity properties of solutions to a system of ordinary differential equations. Recalling the fact that the functions $h(\cdot, \cdot, \cdot, \tau)$, $\tau \in [0, 1]$, are equi-Lipschitzian, let K be the common Lipschitz constant. For $r = 0, 1, \dots, 4K - 2$, let $t_r = \frac{r}{4K}$, and consider the integral equation

$$x(t) = x(t_r) + \int_{t_r}^t h(x(\tau), y(\tau), z(\tau), \tau) d\tau \quad (\text{A.1})$$

on the interval $[t_r, t_r + \frac{1}{2K}]$. I claim that, for given $x(t_r)$, $y(\cdot)$, and $z(\cdot)$, this equation has a unique solution $x(\cdot|t_r, x(t_r), y(\cdot), z(\cdot))$.

To establish this claim, let $C^\ell([t_r, t_r + \frac{1}{2K}])$ be the space of bounded continuous functions from $[t_r, t_r + \frac{1}{2K}]$ into \mathbb{R}^ℓ , endowed with the uniform metric. For any $x_r, y(\cdot), z(\cdot)$, consider the mapping $T_{r, x_r, y(\cdot), z(\cdot)}$ from $C^\ell([t_r, t_r + \frac{1}{2K}])$ into itself that is defined by the formula

$$T_{r, x_r, y(\cdot), z(\cdot)} x(\cdot)(t) = x_r + \int_{t_r}^t h(x(\tau), y(\tau), z(\tau), \tau) d\tau. \quad (\text{A.2})$$

For any two elements $x(\cdot)$ and $\bar{x}(\cdot)$ of $C^\ell([t_r, t_r + \frac{1}{2K}])$ and any $t \in [t_r, t_r + \frac{1}{2K}]$, one has

$$\begin{aligned} & \|T_{r, x_r, y(\cdot), z(\cdot)} x(\cdot)(t) - T_{r, x_r, y(\cdot), z(\cdot)} \bar{x}(\cdot)(t)\| \\ & \leq \int_{t_r}^t |h(x(\tau), y(\tau), z(\tau), \tau) - h(\bar{x}(\tau), y(\tau), z(\tau), \tau)| d\tau \\ & \leq K \int_{t_r}^t \|x(\tau) - \bar{x}(\tau)\| d\tau \\ & \leq K \frac{1}{2K} \max_{\tau} \|x(\tau) - \bar{x}(\tau)\| \leq \frac{1}{2} \max_{\tau} \|x(\tau) - \bar{x}(\tau)\|. \end{aligned}$$

This shows that the mapping $T_{r, x_r, y(\cdot), z(\cdot)}$ is a contraction mapping with modulus $\frac{1}{2}$. As such, it has a unique fixed point $x_r(\cdot|x_r, y(\cdot), z(\cdot)) \in C^\ell([t_r, t_r + \frac{1}{2K}])$. This fixed point satisfies (A.1) for $t \in [t_r, t_r + \frac{1}{2K}]$.

I claim that, if

$$x_{r+1} = x_r(t_{r+1}|x_r, y(\cdot), z(\cdot)), \quad (\text{A.3})$$

then, for any $t \in [t_{r+1}, t_r + \frac{1}{2K}]$, one must have

$$x_{r+1}(t|x_{r+1}, y(\cdot), z(\cdot)) = x_r(t|x_r, y(\cdot), z(\cdot)), \quad (\text{A.4})$$

i.e., the different fixed points coincide on the intersections over their domains. To establish this claim, I first note that, with a slight abuse of notation, $T_{r+1, x_{r+1}, y(\cdot), z(\cdot)}$ can also be treated as a mapping from $C^\ell([t_{r+1}, t_r + \frac{1}{2K}])$ into itself. By the same arguments as before, this mapping is a contraction. Its unique fixed point coincides with the restriction of $x_{r+1}(\cdot|x_{r+1}, y(\cdot), z(\cdot))$ to the interval $[t_{r+1}, t_r + \frac{1}{2K}]$. Further, if (A.3) holds, then, for any $t \in [t_{r+1}, t_r + \frac{1}{2K}]$, (A.2) yields

$$\begin{aligned} & x_r(t|x_r, y(\cdot), z(\cdot)) \\ &= T_{r, x_r, y(\cdot), z(\cdot)} x_r(\cdot|x_r, y(\cdot), z(\cdot))(t) \\ &= x_r + \int_{t_r}^t h(x_r(\tau|x_r, y(\cdot), z(\cdot)), y(\tau), z(\tau), \tau) d\tau \\ &= T_{r, x_r, y(\cdot), z(\cdot)} x_r(\cdot|x_r, y(\cdot), z(\cdot))(t_{r+1}) + \int_{t_r}^t h(x_r(\tau|x_r, y(\cdot), z(\cdot)), y(\tau), z(\tau), \tau) d\tau \\ &= x_r(t_{r+1}|x_r, y(\cdot), z(\cdot)) + \int_{t_{r+1}}^t h(x_r(\tau|x_r, y(\cdot), z(\cdot)), y(\tau), z(\tau), \tau) d\tau \\ &= x_{r+1} + \int_{t_{r+1}}^t h(x_r(\tau|x_r, y(\cdot), z(\cdot)), y(\tau), z(\tau), \tau) d\tau \\ &= T_{r+1, x_{r+1}, y(\cdot), z(\cdot)} x_r(\cdot|x_r, y(\cdot), z(\cdot))(t) \end{aligned}$$

so that the restriction of $x_r(\cdot|x_r, y(\cdot), z(\cdot))$ to the interval $[t_{r+1}, t_r + \frac{1}{2K}]$ is in fact a fixed point of the mapping $T_{r+1, x_{r+1}, y(\cdot), z(\cdot)}$ when treated as a mapping from $C^\ell([t_{r+1}, t_r + \frac{1}{2K}])$ into itself. Because the fixed point of $T_{r+1, x_{r+1}, y(\cdot), z(\cdot)}$ is unique, it follows that the restrictions of $x_r(\cdot|x_r, y(\cdot), z(\cdot))$ and of $x_{r+1}(\cdot|x_{r+1}, y(\cdot), z(\cdot))$ to the interval $[t_{r+1}, t_r + \frac{1}{2K}]$ are equal.

Now let $x(0)$ be given, and define $x(\cdot)$ so that, for any r ,

$$x(t) = x_r(t|x_r, y(\cdot), z(\cdot)) \text{ if } t \in [t_r, t_r + \frac{1}{2K}], \quad (\text{A.5})$$

where $x_0 = x(0)$ and x_1, x_2, \dots are defined inductively via (A.3). Because the functions x_0, x_1, x_2, \dots coincide on the intersections of their domains, the function $x(\cdot)$ is well defined. Moreover, one easily sees that it satisfies the integral equation (4.1).

To prove uniqueness, I note that, if $\bar{x}(\cdot)$ is any (other) solution to the integral equation (4.1), the restriction of $\bar{x}(\cdot)$ to any interval $[t_r, t_r + \frac{1}{2K}]$ is a fixed point of the mapping $T_{r, \bar{x}(t_r), y(\cdot), z(\cdot)}$. Moreover, $\bar{x}(0) = x(0)$. The restriction of $\bar{x}(\cdot)$ to the interval $[0, \frac{1}{2K}]$ must therefore coincide with the restriction of $x(\cdot)$ to this interval. This in turn implies that $\bar{x}(t_1) = x(t_1)$ and, therefore, that the restriction of $\bar{x}(\cdot)$ to the interval $[t_1, t_1 + \frac{1}{2K}]$ must coincide with the restriction of $x(\cdot)$ to this interval. By a straightforward induction argument, it follows that $\bar{x}(t_r) = x(t_r)$ for all r , and that the restriction of $\bar{x}(\cdot)$ to any one of the intervals $[t_r, t_r + \frac{1}{2K}]$ must coincide with the restriction of $x(\cdot)$ to this interval. Therefore, $\bar{x}(\cdot) = x(\cdot)$. ■

Lemma A.1 *For any $(x(0), y(\cdot), z(\cdot))$, the map $(\Delta x, \Delta y, \Delta z) \rightarrow x(\cdot|x(0) + \Delta x, y(\cdot) + \Delta y, z(\cdot) + \Delta z)$ from $\mathbb{R}^\ell \times L_1^m \times L_1^n$ into $C^\ell([0, 1])$ is Lipschitz continuous.*

Proof. I will compare the functions $x(\cdot) = x(\cdot|x(0), y(\cdot), z(\cdot))$ and $x_\Delta(\cdot) = x(\cdot|x(0) + \Delta x, y(\cdot) + \Delta y, z(\cdot) + \Delta z)$. As in the proof of Lemma 4.1, define $t_r = \frac{r}{4K}$ for $r = 0, 1, \dots, 4K - 2$, and note that, for $t \in [t_r, t_r + \frac{1}{2K}]$, (A.5) implies

$$x(t) = x_r(t|x(t_r), y(\cdot), z(\cdot))$$

and

$$x_\Delta(t) = x_r(t|x_\Delta(t_r), y(\cdot) + \Delta y, z(\cdot) + \Delta z),$$

where $x_r(\cdot|x(t_r), y(\cdot), z(\cdot))$ and $x_r(\cdot|x_\Delta(t_r), y(\cdot) + \Delta y, z(\cdot) + \Delta z)$ are the fixed points of the contraction mappings $T_{r, x(t_r), y(\cdot), z(\cdot)}$ and $T_{r, x_\Delta(t_r), y(\cdot) + \Delta y, z(\cdot) + \Delta z}$ on $C^\ell([t_r, t_r + \frac{1}{2K}])$. Because the iterates of a contraction mapping converge uniformly to the unique fixed point, one then has

$$\begin{aligned} & x_\Delta(t) - x(t) \\ &= \lim_{n \rightarrow \infty} T_{r, x_\Delta(t_r), y(\cdot) + \Delta y, z(\cdot) + \Delta z}^n x(\cdot)(t) - x(t) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n [T_{r, x_\Delta(t_r), y(\cdot) + \Delta y, z(\cdot) + \Delta z}^i x(\cdot)(t) - T_{r, x_\Delta(t_r), y(\cdot) + \Delta y, z(\cdot) + \Delta z}^{i-1} x(\cdot)(t)], \end{aligned}$$

where I have written $T_{r, x_\Delta(t_r), y(\cdot) + \Delta y, z(\cdot) + \Delta z}^0 x(\cdot)(t)$ instead of $x(t)$. Because

$T_{r,x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z}$ is a contraction mapping of modulus $\frac{1}{2}$, it follows that

$$\begin{aligned}
& \|x_\Delta(t) - x(t)\| \\
& \leq \lim_{n \rightarrow \infty} \sum_{i=1}^n \left\| T_{r,x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z}^i x(\cdot)(t) - T_{r,x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z}^{i-1} x(\cdot)(t) \right\| \\
& \leq \lim_{n \rightarrow \infty} \sum_{i=1}^n \left(\frac{1}{2}\right)^i \left\| T_{r,x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z}^1 x(\cdot)(t) - T_{r,x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z}^0 x(\cdot)(t) \right\| \\
& = 2 \left\| T_{r,x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z} x(\cdot)(t) - x(t) \right\|.
\end{aligned}$$

Because $x_r(\cdot|x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z)$ is a fixed point of the mapping $T_{r,\bar{x}(t_r),\bar{y}(\cdot),\bar{z}(\cdot)}$, this inequality in turn implies that

$$\|x_\Delta(t) - x(t)\| \leq 2 \left\| T_{r,x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z} x(\cdot)(t) - T_{r,x(t_r),y(\cdot),z(\cdot)} x(\cdot)(t) \right\| \quad (\text{A.6})$$

for all $t \in [t_r, t_r + \frac{1}{2K}]$.

By (A.2) and the equi-Lipschitz property of h , one also has

$$\begin{aligned}
& \left\| T_{r,x_\Delta(t_r),y(\cdot)+\Delta y,z(\cdot)+\Delta z} x(\cdot)(t) - T_{r,x(t_r),y(\cdot),z(\cdot)} x(\cdot)(t) \right\| \\
& \leq \|x_\Delta(t_r) - x(t_r)\| \\
& \quad + \int_{t_r}^t |h(x(\tau), y(\tau) + \Delta y, z(\tau) + \Delta z, \tau) - h(x(\tau), y(\tau), z(\tau), \tau)| \, d\tau \quad (\text{A.7}) \\
& \leq \|x_\Delta(t_r) - x(t_r)\| + K \int_{t_r}^t \|(\Delta y(\tau), \Delta z(\tau))\| \, d\tau \\
& \leq \|x_\Delta(t_r) - x(t_r)\| + K(\|\Delta y\| + \|\Delta z\|). \quad (\text{A.8})
\end{aligned}$$

Upon combining this inequality with (A.6), one infers that, for all r and all $t \in [t_r, t_r + \frac{1}{2K}]$, one has

$$\|x_\Delta(t) - x(t)\| \leq 2(\|x_\Delta(t_r) - x(t_r)\| + K(\|\Delta y\| + \|\Delta z\|)). \quad (\text{A.9})$$

Because, for any r , one has $t_{r+1} = \frac{r+1}{4K} = t_r + \frac{1}{4K}$, it follows, in particular, that

$$\|x_\Delta(t_{r+1}) - x(t_{r+1})\| \leq 2(\|x_\Delta(t_r) - x(t_r)\| + K(\|\Delta y\| + \|\Delta z\|))$$

and, hence,

$$\begin{aligned}
\|x_\Delta(t_{r+1}) - x(t_{r+1})\| & \leq 2^{r+1} \|\Delta x\| + (2^{r+1} - 1)K(\|\Delta y\| + \|\Delta z\|) \\
& \leq 2^{4K} \max(K, 1)(\|\Delta x\| + \|\Delta y\| + \|\Delta z\|).
\end{aligned}$$

Upon combining this inequality with (A.9), one obtains the Lipschitz continuity of the map $(\Delta x, \Delta y, \Delta z) \rightarrow x(\cdot|x(0) + \Delta x, y(\cdot) + \Delta y, z(\cdot) + \Delta z)$.

■

Lemma A.2 *For any $(x(\cdot), y(\cdot), z(\cdot))$ and any $\delta = (\delta x, \delta y, \delta z) \in \mathbb{R}^\ell \times L_1^m \times L_1^n$, the integral equation*

$$\xi_\delta(t|x(\cdot), y(\cdot), z(\cdot)) = \delta x + \int_0^t [h_x \cdot \xi_\delta(\tau|x(\cdot), y(\cdot), z(\cdot)) + h_y \cdot \delta y(\tau) + h_z \cdot \delta z(\tau)] d\tau, \quad (\text{A.10})$$

with h_x, h_y, h_z are evaluated at $(x(\tau), y(\tau), z(\tau))$ for all τ , has a unique solution $\xi_\delta(\cdot)$. For any $(x(\cdot), y(\cdot), z(\cdot))$, the map $\delta \rightarrow \xi_\delta(\cdot)$ from $\mathbb{R}^\ell \times L_1^m \times L_1^n$ into $C^\ell([0, 1])$ is Lipschitz continuous.

Proof. It suffices to note that, for given $(x(\cdot), y(\cdot), z(\cdot))$, the functions

$$(\xi(\tau), \delta y(\tau), \delta z(\tau)) \rightarrow [h_x \cdot \xi(\tau) + h_y \cdot \delta y(\tau) + h_z \cdot \delta z(\tau)],$$

$\tau \in [0, 1]$, exhibit the same differentiability and equi-Lipschitz continuity properties as the functions $h(\cdot, \cdot, \cdot, \tau)$. Therefore, the first statement of the lemma is an instance of Lemma 4.1, the second statement an instance of Lemma A.1. ■

Lemma A.3 *Fix $(x(\cdot), y(\cdot), z(\cdot))$. For any $\Delta = (\Delta x, \Delta y, \Delta z) \in \mathbb{R}^\ell \times L_1^m \times L_1^n$, the map*

$$\delta \rightarrow \xi_\delta(\cdot|x_\Delta(\cdot), y(\cdot) + \Delta y(\cdot), z(\cdot) + \Delta z(\cdot)) \quad (\text{A.11})$$

that is given by Lemma A.2 with $(x(\cdot), y(\cdot), z(\cdot))$ replaced by $(x_\Delta(\cdot), y(\cdot) + \Delta y(\cdot), z(\cdot) + \Delta z(\cdot))$, with $x_\Delta(\cdot) = x(\cdot|x(0) + \Delta x, y(\cdot) + \Delta y(\cdot), z(\cdot) + \Delta z(\cdot))$, defines a continuous linear operator A_Δ from $\mathbb{R}^\ell \times L_1^m \times L_1^n$ into $C^\ell([0, 1])$.

Proof. Continuity follows from Lemma A.2. To show that the operator A_Δ is linear, consider any $\delta = (\delta x, \delta y, \delta z)$ and $d = (dx, dy, dz)$ in $\mathbb{R}^\ell \times L_1^m \times L_1^n$ and any two scalars α, β . Because ξ_δ and ξ_d satisfy the integral equations

$$\xi_\delta(t) = \delta x + \int_0^t [h_x \cdot \xi_\delta(\tau) + h_y \cdot \delta y(\tau) + h_z \cdot \delta z(\tau)] d\tau \quad (\text{A.12})$$

and

$$\xi_d(t) = dx + \int_0^t [h_x \cdot \xi_d(\tau) + h_y \cdot dy(\tau) + h_z \cdot dz(\tau)] d\tau, \quad (\text{A.13})$$

their linear combination $\alpha\xi_\delta + \beta\xi_d$ satisfies the integral equation

$$\begin{aligned} \alpha\xi_\delta(t) + \beta\xi_d(t) &= \alpha\delta x + \beta dx + \int_0^t h_x \cdot (\alpha\xi_\delta(\tau) + \beta\xi_d(\tau))d\tau \\ &+ \int_0^t h_y \cdot a\delta y(\tau) + \beta dy(\tau) + h_z \cdot (\alpha\delta z(\tau) + \beta dz(\tau))d\tau. \end{aligned} \quad (\text{A.14})$$

By definition, $\xi_{\alpha\delta+\beta d}$ is also a solution to this integral equation. Because, by Lemma A.2, the solution to the integral equation (A.14) is unique, it follows that $\xi_{\alpha\delta+\beta d} = \alpha\xi_\delta + \beta\xi_d$, which proves that the map $\delta \rightarrow \xi_\delta(\cdot)$ is linear. ■

Proof of Lemma 4.2. Fix $(x(0), y(\cdot), z(\cdot))$ and $\Delta = (\Delta x, \Delta y, \Delta z) \in \mathbb{R}^\ell \times L_1^m \times L_1^n$, and let A_Δ be the linear operator from $\mathbb{R}^\ell \times L_1^m \times L_1^n$ to $C^\ell([0, 1])$ that is given by Lemma A.3. From (4.1) and Lemma A.1, one has

$$x_{\Delta+\delta}(t) - x_\Delta(t) = \delta x + \int_0^t [h_x \cdot (x_{\Delta+\delta}(\tau) - x_\Delta(\tau)) + Q(\tau)] d\tau + o(\delta) \quad (\text{A.15})$$

for all t , where $Q(\tau) := h_y\delta y(\tau) + h_z\delta z(\tau)$ and $o(\delta)$ is a term such that $\frac{\|o(\delta)\|}{\|\delta\|}$ goes to zero if $\|\delta\|$ goes to zero. From (A.15) and Lemma A.3, one then obtains

$$\begin{aligned} x_{\Delta+\delta}(t) - x_\Delta(t) - (A_\Delta\delta)(t) &= x_{\Delta+\delta}(\hat{t}) - x_\Delta(\hat{t}) - (A_\Delta\delta)(\hat{t}) \\ &+ \int_{\hat{t}}^t h_x(x_{\Delta+\delta}(\tau) - x_\Delta(\tau) - (A_\Delta\delta)(\tau))d\tau \\ &+ \int_{\hat{t}}^t Q(\tau)d\tau - \xi_\delta(t) + \xi_\delta(\hat{t}) + \\ &\int_{\hat{t}}^t h_x \cdot \xi_\delta(\tau)d\tau + o(\delta) \end{aligned} \quad (\text{A.16})$$

for all t and $\hat{t} < t$. From (A.10), one has

$$\xi_\delta(t) = \xi_\delta(\hat{t}) + \int_{\hat{t}}^t h_x \cdot \xi_\delta(\tau)d\tau + \int_{\hat{t}}^t Q(\tau)d\tau$$

for all t and $\hat{t} < t$. Therefore, (A.16) can be rewritten as

$$\begin{aligned} x_{\Delta+\delta}(t) - x_\Delta(t) - (A_\Delta\delta)(t) &= x_{\Delta+\delta}(\hat{t}) - x_\Delta(\hat{t}) - (A_\Delta\delta)(\hat{t}) \\ &+ \int_0^t h_x(x_{\Delta+\delta}(\tau) - x_\Delta(\tau) - (A_\Delta\delta)(\tau))d\tau \\ &+ o(\delta). \end{aligned} \quad (\text{A.17})$$

If one defines

$$\eta(\tau) := \lim_{\varepsilon \rightarrow 0} \sup_{\|\delta\| \leq \varepsilon} \frac{\|x_{\Delta+\delta}(\tau) - x_{\Delta}(\tau) - (A_{\Delta}\delta)(\tau)\|}{\|\delta\|} \quad (\text{A.18})$$

for any τ , then (A.17) yields

$$\eta(t) \leq \eta(\hat{t}) + K \int_{\hat{t}}^t \eta(\tau) d\tau \quad (\text{A.19})$$

for all t and $\hat{t} < t$. In particular, if one sets $t_r = \frac{r}{2K}$, $r = 0, 1, \dots, 4K - 2$, as before, one has

$$\begin{aligned} \eta(t_{r+1}) &\leq \sup_{t \in [t_r, t_{r+1}]} \eta(t) \leq \eta(t_r) + K \int_{t_r}^{t_{r+1}} \sup_{t \in [t_r, t_{r+1}]} \eta(t) d\tau \\ &= \eta(t_r) + \frac{1}{4} \sup_{t \in [t_r, t_{r+1}]} \eta(t), \end{aligned} \quad (\text{A.20})$$

hence

$$\eta(t_{r+1}) \leq \frac{4}{3} \eta(t_r) \quad (\text{A.21})$$

for all r . By (4.1) and (A.10), one also has $\eta(0) = 0$. Therefore, (A.20) and (A.21) imply $\eta(t) = 0$ for all t . It follows that, for any $t \in [0, 1]$, the map $\Delta \rightarrow A_{\Delta}$ is the Fréchet derivative of the map $\Delta \rightarrow x(\cdot|x(0) + \Delta x, y(\cdot) + \Delta y, z(\cdot) + \Delta z)$. ■

References

- [1] Barron, E.N., and R. Jensen (1980), Optimal Control Problems with No Turning Back, *Journal of Differential Equations* 36, 223 - 248.
- [2] Brito, D.L., and W.H. Oakland (1977), Some Properties of the Optimal Income Tax, *International Economic Review* 18, 407 - 423.
- [3] Brunner, J.K. (1993), A Note on the Optimal Income Tax, *Journal of Public Economics* 50, 445 - 451.
- [4] Clarke, F.H. (1976), The Maximum Principle under Minimal Hypotheses, *SIAM Journal of Control and Optimization* 14, 1078 - 1091.
- [5] Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis*, Wiley, New York.
- [6] Ebert, U. (1992), A Reexamination of the Optimal Nonlinear Income Tax, *Journal of Public Economics* 49, 47 - 73.
- [7] Fudenberg, D., and J. Tirole (1991), *Game Theory*, MIT Press,
- [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. (2007 a), A Contribution to the Theory of Optimal Income Taxation, *Journal of Public Economics* 91, 1449 - 1477.
- [10] Hellwig, M.F. (2007 b), The Undesirability of Randomized Income Taxation under Decreasing Risk Aversion, *Journal of Public Economics* 91, 791 - 816.
- [11] Hellwig, M.F. (2009), Incentive Problems with Hidden Characteristics: A Unified Approach, Preprint 06-26, Max Planck Institute for Research on Collective Goods, Bonn, 2006, http://www.coll.mpg.de/pdf_dat/2006_26online.pdf, revised 2009.
- [12] Luenberger, D.G. (1969), *Optimization by Vector Space Methods*, Wiley, New York.
- [13] Mirrlees, J.M. (1971), An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies* 38, 175 - 208.

- [14] Mirrlees, J.M. (1976), Optimal Tax Theory: A Synthesis, *Journal of Public Economics* 6, 327 - 358.
- [15] 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.
- [16] Shilov, G.E. (1974), *Elementary Functional Analysis*, MIT Press, Cambridge, Massachusetts.