

# CAMBRIDGE WORKING PAPERS IN ECONOMICS

## JANEWAY INSTITUTE WORKING PAPERS

### Optimal Nonlinear Savings Taxation

Charles  
Brendon  
University of  
Cambridge

#### Abstract

This paper analyses the design of optimal nonlinear savings taxation, in a multi-period consumption-savings economy where consumers face persistent, uninsurable shocks to the marginal value that they place on consuming. Its main contributions are: (a) to show that shocks of this kind generically justify positive marginal savings taxes, and (b) to characterise these taxes by reference to a limited number of sufficient statistics. The method for obtaining this characterisation is generalisable, and provides a roadmap for reconnecting ‘Mirrleesian’ and ‘sufficient statistics’ approaches to dynamic taxation. Intuitively, dynamic asymmetric information problems imply significant restrictions on intertemporal consumption elasticities. These restrictions keep sufficient statistics representations manageable, despite the multi-dimensional choice setting.

#### Reference Details

2221 2022/10	Cambridge Working Papers in Economics Janeway Institute Working Paper Series
Published	25 March 2022
Key Words	Nonlinear Taxation, Sufficient Statistics, Mirrleesian Taxation, New Dynamic Public Finance
JEL Codes	D82, E21, E61 H21, H24, H30
Websites	<a href="http://www.econ.cam.ac.uk/cwpe">www.econ.cam.ac.uk/cwpe</a> <a href="http://www.janeway.econ.cam.ac.uk/working-papers">www.janeway.econ.cam.ac.uk/working-papers</a>

# Optimal Nonlinear Savings Taxation\*

Charles Brendon<sup>†</sup>

March 25, 2022

## Abstract

This paper analyses the design of optimal nonlinear savings taxation, in a multi-period consumption-savings economy where consumers face persistent, uninsurable shocks to the marginal value that they place on consuming. Its main contributions are: (a) to show that shocks of this kind generically justify positive marginal savings taxes, and (b) to characterise these taxes by reference to a limited number of sufficient statistics. The method for obtaining this characterisation is generalisable, and provides a roadmap for reconnecting ‘Mirrleesian’ and ‘sufficient statistics’ approaches to dynamic taxation. Intuitively, dynamic asymmetric information problems imply significant restrictions on intertemporal consumption elasticities. These restrictions keep sufficient statistics representations manageable, despite the multi-dimensional choice setting.

Keywords: Nonlinear Taxation, Sufficient Statistics, Mirrleesian Taxation, New Dynamic Public Finance

JEL Codes: D82, E21, E61 H21, H24, H30

---

\*Preliminary draft: comments welcome. This work has benefited from discussions and comments from Vasco Carvalho, Lukas Freund, Andrew Hannon, Christian Hellwig, David Martimort, Ludvig Sinander and Nicolas Werquin, together with seminar participants in Cambridge and TSE. The first draft was completed during a visit to the Toulouse School of Economics: I am very grateful for the hospitality shown. Richard Nickl provided important literary motivation.

<sup>†</sup>Faculty of Economics & Queens’ College, University of Cambridge, and CEPR. Email: cfb46@cam.ac.uk

# 1 Introduction

This paper revisits the problem of optimal taxation in dynamic economies with private information. Adapting the dynamic asymmetric information problem due to Atkeson and Lucas (1992), it considers a government that would like to insure individuals against persistent, unobservable shocks over time to their marginal utility of consumption. Though this environment has been studied extensively in the social insurance literature, I derive a number of novel results that clarify the character of optimal policy within it. In doing so, I provide a more general contribution to the links between dynamic Mirrleesian and sufficient statistic analyses of optimal taxation. Specifically, I show that the assumptions embedded in the dynamic asymmetric information problem are sufficient simplifications for optimal nonlinear taxation to be characterised by a remarkably small number of conventional behavioural elasticities, despite the *a priori* complexity of an infinite-horizon, incomplete market setting.

Constrained-optimal allocations in the Atkeson-Lucas environment can be decentralised in a standard consumption-savings economy, provided savings are subject to a nonlinear tax period-by-period. The main focus of the paper is on characterising this tax schedule. I show that the marginal tax rate on savings is generically positive in all time periods, and strictly positive away from the extreme ends of the savings distribution. The revenue from this savings tax is used to fund a positive, uniform lump-sum transfer each period, which permits higher within-period consumption for those with high marginal utility, relative to what they would otherwise be able to afford. Inclusive of this lump-sum component, the expected total tax bill for individuals each period is zero *ex-ante*, and its realised value *ex-post* increases monotonically in savings.

For any given time period and shock history, I express the optimal marginal savings tax rate as a function of a small number of behavioural statistics and statistical attributes of the realised savings distribution. This ‘sufficient statistics’ representation provides a simple, intuitive statement of the mechanical, behavioural and welfare considerations that are relevant to optimal policy design. It is analogous to the well-known Saez (2001) condition for optimal labour taxation, and isomorphic to it in the special case that preference shocks are iid over time.

The derivation of this characterisation represents a promising methodological innovation for the

Mirrleesian, ‘mechanism design’ approach to dynamic taxation, of which this paper is an example.<sup>1</sup> A number of writers have expressed scepticism in recent years about the practical relevance of dynamic Mirrleesian analysis. A common complaint is that it generates implausibly complex policies, whose form is too dependent on utility functions, type distributions, and other unknowable objects, to have real-world applicability.<sup>2</sup> The results here provide a counter-point to this. The characterisation of optimal policy that we offer is not significantly more complex than a textbook Saez formula with income effects. It is written in terms of behavioural objects that are defined independently of the utility function and hidden type process, with the sole exception of social welfare weights – in which a reference to marginal utility is standard in the sufficient statistics literature.

Indeed, the most significant feature of this characterisation is precisely its simplicity. Despite the infinite-horizon setting and continuum of possible shock draws each period, at most three behavioural statistics are of relevance to an optimal marginal savings tax. These are: the compensated elasticity of savings with respect to the marginal tax rate, the marginal effect of higher income on savings, and a compensated elasticity that measures effect of a change to insurance in the present period on prior savings.

This simplification can be interpreted through the lens of the Atkinson-Stiglitz theorem. Consistent with the wider literature, we impose a Markovian structure on shocks. This means that *conditional on the next period’s type draw*, preferences across alternative insurance schemes more than one period ahead are independent of an individual’s current type. Following the logic of Atkinson-Stiglitz, this independence implies that there is no gain to distorting allocations in later time periods, in order to improve the structure of incentives more than one period previously. This is what keeps the relevant behavioural considerations for policy design to a manageable level.

The important general lesson, we argue, is that the mechanism design approach imposes structural restrictions on consumer preferences that allow optimal dynamic tax problems to become tractable. Far from complexifying, it can provide a powerful, theoretically-grounded basis for *simple* policy advice in dynamic settings.

---

<sup>1</sup>Following Mirrlees (1971), this approach focuses on the design of optimal dynamic allocations subject only to information frictions, without assuming any particular decentralisation, or *a priori* limits on the set of tax instruments.

<sup>2</sup>See, for instance, the discussions in Diamond and Saez (2011), Piketty and Saez (2013b), and Stantcheva (2020).

## 1.1 Preview of main characterisation

To substantify this discussion, we briefly preview the main characterisation result – which features as Theorem 1 in the body of the paper.<sup>3</sup> When a nonlinear savings tax decentralises the constrained-optimal allocation, we show that it satisfies the following trade-off within each period, at each contemporaneous savings level  $s'$ :

$$\mathbb{E} \left[ 1 - T'(s) \frac{ds}{dM} - g(s) \Big| s \geq s' \right] = T'(s') \epsilon^s a(s') + RT'_{-1}(s_{-1}) s_{-1} \epsilon_{-1}^s(s') \quad (1)$$

This equation can be read as comparing the costs and benefits from a cut in the marginal tax rate at  $s'$ , for a cross-section of types with a common history. The left-hand side gives the net fiscal cost of the tax cut, due to a transfer of resources to higher savers. It is made up of a mechanical unit cost, less the marginal tax revenue that is recovered through a standard income effect on savings,  $T'(s) \frac{ds}{dM}$ , less a social welfare weight,  $g(s)$ , that captures the welfare value of transferring income to an individual whose savings are  $s$ . The welfare weight – an endogenous object that evolves with individuals' wealth levels, discussed in detail below – is decreasing in savings, because higher savers have a relatively low contemporaneous marginal utility of consumption.

The right-hand side of the equation gives the fiscal benefits of substitution effects that are induced by the tax change. When taxes are cut, savings in the current period increase in proportion to the contemporaneous savings elasticity  $\epsilon^s$ . This raises revenue in proportion to the marginal tax rate  $T'(s')$ , and to the measure of types at  $s'$ , relative to those above:  $a(s')$  is the local Pareto parameter for the realised savings distribution.

Cutting taxes at  $s'$  in the current period may also change savings in the previous period, by an amount proportional to a compensated cross elasticity  $\epsilon_{-1}^s(s')$ . This raises additional income in the previous period, in proportion to that period's marginal tax rate  $T'_{-1}(s_{-1})$ , whose relative value depends on the gross real interest rate  $R$ .

The cross elasticity  $\epsilon_{-1}^s(s')$  is the least conventional of the objects in the characterisation, and is particular to consumption choice models with imperfect insurance. It is the behavioural response that arises because of a change to the state-contingent profile of returns provided by the savings instrument. Its

---

<sup>3</sup>Some notation, including time indexation, is dropped for simplicity.

relevance is intrinsically linked to type persistence in the underlying structural model: it is zero when types are iid. We discuss it in detail in the body of the paper.<sup>4</sup>

Equation (1) is also helpful for understanding the significant qualitative result, Theorem 2 in the paper, that marginal savings taxes are generically positive. Taking substitution effects – the right-hand side – in isolation, it would generally be beneficial to cut *marginal* taxes on any given agent to zero. By incentivising additional savings, this raises fiscal revenue until the last unit saved is no longer being taxed.

But against this efficiency gain is an equity loss – the left-hand side. When marginal taxes are cut at  $s'$ , income is necessarily redistributed to higher savers, whose marginal utility is lower. Reflecting the individual's own *ex-ante* insurance preferences, this is an undesirable diversion of resources. It is optimal to retain positive marginal savings taxes, as this allows more resources to be directed towards lower savers.

## 1.2 Paper outline

The rest of the paper is organised as follows. Section 2 provides an overview of related literature. Section 3 introduces the detailed setup of the dynamic information-theoretic problem that we study. Section 6 outlines how nonlinear savings taxes can be used to decentralise incentive-compatible allocations for this environment, and derives sufficient conditions on the allocation for this decentralisation to work. Like much of the optimal taxation literature dating back to Mirrlees (1971), we keep analysis tractable via a 'first-order' approach to incentive compatibility: Section 4 reminds readers of this approach, and provides a novel, intuitive increasingness condition on the allocation that guarantees its validity.

To aid understanding, the main characterisation is presented constructively, in steps. In Section 5 we use standard methods to characterise constrained-optimal allocations by reference to the costs and benefits of changing *utility* levels for a cross-section of individuals. The resulting expressions are insightful, and reveal novel features about the dynamics of consumption when types are persistent, but they rely heavily on arguments of the utility function. Section 7 explains intuitively how these utility-based expressions can be mapped to a practical characterisation of optimal tax rates, and presents interme-

---

<sup>4</sup>Section 10.

mediate results to this end. The main sufficient statistics characterisation that follows is given in Section 8. Section 9 explores the qualitative properties of optimal savings taxes, notably the result that optimal marginal tax rates are positive. Section 10 explains the limited role for intertemporal cross-elasticities in characterisation of optimal taxes, with particular reference to the Atkinson-Stiglitz theorem. Section 11 provides an illustrative quantification of marginal tax rates for top savings levels, based on the formula obtained in Section 8. Section 12 concludes.

All but the most straightforward proofs are relegated to the appendix.

## 2 Relation to literature

The basic insurance problem that we study was first popularised by Atkeson and Lucas (1992), who focused on the properties of constrained-optimal allocations in the presence of unobservable shocks to marginal utility. Their paper gave particular attention to long-run outcomes, showing that the immiseration result of Thomas and Worrall (1990) – whereby measure 1 of agents see their consumption converge to zero over time – carried over to their setting, as well as emphasising that the optimum could not be decentralised via conventional linear pricing. Technically, Atkeson and Lucas assumed an iid type distribution, with types drawn from a finite set period-by-period – a structure retained by more recent literature that explores the sensitivity of their immiseration result.<sup>5</sup> The current paper instead allows for persistent (Markovian) type draws, which has non-trivial implications for consumption dynamics relative to the iid case. I also assume types are drawn from a continuum, which proves crucial in finding a mathematical link from the mechanism design characterisation to behavioural statistics.

More broadly, the present paper is situated in the dynamic Mirrleesian public finance tradition, analysing optimal tax systems subject to the deep information frictions that necessitate departures from the Second Welfare Theorem. Most of the contributions to this literature consider the traditional Mirrlees setting of endogenous labour supply and unobservable, stochastic productivity. Seminal papers include Golosov, Kocherlakota and Tsyvinski (2003), Kockerlakota (2005) and Golosov, Tsyvinski and Werning (2006), with Kocherlakota (2010) providing an excellent overview.

Much – though not all – of this literature has focused on characterising the differences between

---

<sup>5</sup>See, for instance, Sleet and Yeltekin (2006) and Farhi and Werning (2007).

constrained-optimal allocations and laissez-faire outcomes, rather than focusing directly on tax instruments.<sup>6</sup> Emphasis in the early papers was on the well-known ‘inverse Euler equation’ – an expression that implies a distortion relative to savings behaviour under autarky, but does not directly map to any particular tax instrument.<sup>7</sup> Likewise, more recent papers by Farhi and Werning (2013) and Golosov, Troshkin and Tsyvinski (2016) have examined the properties of the ‘wedge’ between the consumption-labour marginal rate of substitution and the marginal product of labour. But in dynamic settings the link between this wedge and labour income tax rate is no longer direct. By contrast, the main characterisation in the present paper relates to the marginal savings tax rate itself – an object directly controlled by policy.

Interesting parallels to the current paper are found in Albanesi and Sleet (2006). The principal focus of their paper is the possibility of a simple market decentralisation for a specific class of dynamic Mirrleesian problems – where productivity shocks are iid, and labour and consumption separable. As in the present paper, these authors find limited intertemporal dependence in tax policy, with past choice only influencing current policy through an individual’s retained wealth level. Though they do not draw the link to Atkinson and Stiglitz (1976), their assumptions together imply that the value of real output – whether saved or consumed – is independent of one’s current type. This suggests the structural reasons for limited intertemporal dependence in policy are likely very similar to what is presented here.

The current paper follows Kapička (2013), Farhi and Werning (2013), Golosov, Troshkin and Tsyvinski (2016), Stantcheva (2017), Hellwig (2021) and Hellwig and Werquin (2022) in making use of the first-order approach to incentive compatibility. Early contributions to the dynamic Mirrlees literature were wary of the risks of neglecting global incentive compatibility, but this has faded in recent years, due both to increased understanding of the conditions for validity – to which we contribute – and the simple difficulty in making progress otherwise. Pavan, Segal and Toikka (2014) provided important new clarity on the conditions for the first-order approach to be valid. Though their main focus is on settings from the microeconomic literature with quasilinear preferences, like Hellwig (2021) the current paper adapts their methodology to a dynamic Mirrleesian setting.

---

<sup>6</sup>Kocherlakota (2005) is an important exception, though his decentralisation retains much of the spirit of a direct mechanism: agents are offered limited menus of options, with extreme punishments for behaviours inconsistent with the constrained-optimal allocation.

<sup>7</sup>The inverse Euler condition had previously been derived in different settings by Diamond and Mirrlees (1978) and Rogerson (1985).



Away from the dynamic Mirrlees literature, the results in this paper contribute to the influential movement to link policy prescriptions to observable ‘sufficient statistics’, insofar as possible. From original contributions by Diamond (1998) and Saez (2001), which re-cast the static Mirrlees (1971) model by reference to instruments rather than allocations, this approach now encompasses broad areas of macro and micro policy design.<sup>8</sup> Yet in contrast with the static literature, for dynamic tax problems ‘sufficient statistics’ and ‘mechanism design’ approaches are commonly interpreted as rivals rather than complements – that is, as mutually inconsistent methods that generate distinct policy prescriptions, rather than equivalent formulations linked by a duality relationship.

The reason for this divorce has been a desire to obtain comparably simple policy lessons for dynamic tax environments as for static, and the seeming difficulty of achieving this in a mechanism design setting. To achieve this, influential papers by Piketty and Saez (2013a) and Saez and Stantcheva (2018) have deliberately discarded information-theoretic foundations, in favour of a long-run focus: tax instruments are assumed to be time-invariant, and the effects of any changes are analysed purely by reference to their mechanical, welfare and behavioural effects *in steady state*.<sup>9</sup> This overcomes the need to consider arbitrary intertemporal cross elasticities – the response of savings in period  $t$  to taxation in  $t + 59$ , say – by asserting that all that matters is what happens in the long run.<sup>10</sup> The present paper instead shows that simple, intuitive sufficient statistics characterisations *can* arise from a mechanism design approach, attributing this to the standard preference assumptions made in these settings. Thus I highlight an alternative route to policy insight from the more radical focus on long-run outcomes alone.

By offering a novel justification for savings taxation, this paper also contributes to the large general literature on the desirability of intertemporal distortions. Work on this topic has moved on considerably from the classic Chamley (1986) and Judd (1985) zero tax results, due both to the direct assault of Straub and Werning (2020),<sup>11</sup> and the earlier findings that savings taxes could play a useful role in computa-

---

<sup>8</sup>Ferey, Lockwood and Taubinsky (2021) is a notable recent contribution, expanding the static Mirrlees problem to allow for preference heterogeneity.

<sup>9</sup>Stantcheva (2020) provides an excellent summary of the approach.

<sup>10</sup>Golosov, Tsyvinski and Werquin (2014) provide a general behavioural decomposition of the effects of tax changes, allowing for arbitrary cross-elasticities, also without direct reference to information frictions. The difficulty they encounter is the multiplicity of potential consumer substitution responses across periods and states of the world, which makes applicability a challenge.

<sup>11</sup>See Chari, Nicolini and Teles (2020) for a useful discussion of the scope of the Straub-Werning results.

tional Ramsey environments.<sup>12</sup> Yet a common theme in this literature remains that savings distortions are only introduced because of significant limitations elsewhere in the tax system – particularly credit constraints, limits on age-dependent taxation, or arbitrary tax ceilings. It provides few direct arguments *for* savings taxes. In this regard the present paper differs: if savings reveal consumption need, and the government would like to redistribute according to consumption need, then a savings tax is the most direct, appropriate intervention.

Finally, this paper has links to the growing microeconomic mechanism design literature that gives particular attention to the problems implied by type persistence.<sup>13</sup> Current work by Bloedel, Krishna and Strulovici (2020), Bloedel, Krishna and Leukhina (2020) and Makris and Pavan (2020) deploy various settings to explore the dynamics of wedges in problems without quasilinearity. The present paper provides a novel decomposition of consumption dynamics into two distinct multiplier processes – one stationary, one nonstationary – that helps shed light on the distinct roles played by type persistence and risk aversion in these settings.

### 3 Model setup

This section and the next present the information-theoretic dynamic social insurance problem that is the starting point for our analysis. The presentation is kept succinct, with a focus on the technical assumptions that are important for the subsequent arguments to work.

#### 3.1 Preliminaries

Time is discrete but infinite, indexed by the natural numbers and starting in period zero. The economy consists of a measure-1 continuum of individuals, plus a policymaker whose role is to provide some insurance mechanism against the taste shocks that consumers face period-by-period.

---

<sup>12</sup>For instance, Aiyagari (1995), İmrohoroğlu (1998), Erosa and Gervais (2002) and Conesa, Kitao and Krueger (2009).

<sup>13</sup>Pavan (2017) surveys this literature in detail.

### 3.2 Preferences and shock structure

There is an aggregate endowment  $y_t$  of real resources in each period  $t$ , which the policymaker either owns or can tax lump-sum. Each consumer values contingent consumption streams from each period  $s \geq 0$  onwards according to the criterion  $U_s$ :

$$U_s := \mathbb{E}_s \sum_{t=s}^{\infty} \beta^{t-s} \alpha_t u(c_t) \quad (2)$$

where  $c_t$  is consumption in period  $t$ ,  $\beta \in (0, 1)$  is the discount factor,  $u : \mathbb{R}^+ \rightarrow \mathbb{R}$  or  $\mathbb{R}^{++} \rightarrow \mathbb{R}$  is the period utility function, and  $\alpha_t \in [\underline{\alpha}, \bar{\alpha}] \subset \mathbb{R}^+$  is the idiosyncratic taste disturbance in  $t$ , with  $\underline{\alpha} > 0$  and  $\bar{\alpha} < \infty$ . To keep notation compact, we will refer to the interval  $[\underline{\alpha}, \bar{\alpha}]$  as  $A$ .  $\alpha^t \in A^{t+1}$  will denote a complete idiosyncratic history of taste draws up to period  $t$ , and  $\alpha_t^s \in A^{s-t+1}$  a partial sequence of draws between periods  $t$  and  $s$  (inclusive). We make the following standard assumption on the utility function:

**Assumption 1.**  *$u(\cdot)$  is twice differentiable, with  $u'(\cdot) > 0$  and  $u''(\cdot) < 0$ , and satisfies the Inada conditions.*

Type draws are assumed to be independent across individuals, so there is no aggregate risk. Since there is no other intrinsic source of uncertainty, and no policy reason to introduce one artificially, an agent's consumption in period  $t$  will be measurable with respect to their history  $\alpha^t$  alone.

The taste parameter is assumed to follow a Markov process, identical through time in all periods except the initial period 0. Conditional on drawing  $\alpha_t \in A$  in period  $t$ , the distribution of shocks in  $t+1$  is denoted  $\Pi(\alpha_{t+1}|\alpha_t)$ , with conditional density  $\pi(\alpha_{t+1}|\alpha_t)$ . The equivalent (unconditional) objects for period 0 are denoted  $\Pi(\alpha_0)$  and  $\pi(\alpha_0)$  respectively. We place the following regularity structure on the distributions:

**Assumption 2.** *Both  $\Pi(\cdot|\cdot)$  and  $\pi(\cdot|\cdot)$  are continuously differentiable on  $A^2$ , and  $\pi(\alpha_t|\alpha_{t-1}) > 0$  for all  $\alpha_{t-1} \in A$  and  $\alpha_t \in (\underline{\alpha}, \bar{\alpha})$ .  $\Pi(\cdot)$  and  $\pi(\cdot)$  are differentiable, and  $\pi(\alpha_0) > 0$  for all  $\alpha_0 \in (\underline{\alpha}, \bar{\alpha})$ .*

Notice that the density functions may approach zero at endpoints for the type distribution.

Occasionally it will be useful to make reference to the measure of type histories up to some period  $t$ . For all  $S \subseteq A^{t+1}$ ,  $\Pi_t(S)$  denotes the probability that  $\alpha^t$  will lie in  $S$ . This is induced by  $\Pi$  in the obvious way.  $\mathbb{E}_s$  denotes period- $s$  conditional expectations of a future variable under this process, given an  $\alpha^s$ .

The elasticity of expected next-period type with respect to current type features in some of the analysis that follows, where it is denoted  $\varepsilon^\alpha(\alpha_t)$ . Formally, this is defined as follows:

$$\varepsilon^\alpha(\alpha_t) := \frac{\alpha_t}{\mathbb{E}_t[\alpha_{t+1}|\alpha_t]} \frac{d\mathbb{E}[\alpha_{t+1}|\alpha_t]}{d\alpha_t} \quad (3)$$

Also important is the responsiveness of the distribution of types at  $t + 1$  with respect to type at  $t$ . This is well summarised by the statistic  $\rho(\alpha_{t+1}|\alpha_t)$ , defined as the relative responsiveness of  $\Pi(\alpha_{t+1}|\alpha_t)$  to log changes in  $\alpha_t$  by comparison with log changes in  $\alpha_{t+1}$ .<sup>14</sup>

$$\rho(\alpha_{t+1}|\alpha_t) := \frac{\alpha_t}{\alpha_{t+1}} \cdot \frac{\frac{d(1-\Pi(\alpha_{t+1}|\alpha_t))}{d\alpha_t}}{\pi(\alpha_{t+1}|\alpha_t)} \quad (4)$$

Integrating provides a link between the two preceding objects:

$$\int_{\alpha_{t+1}} \rho(\alpha_{t+1}|\alpha_t) \alpha_{t+1} \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} = \varepsilon^\alpha(\alpha_t) \mathbb{E}_t[\alpha_{t+1}|\alpha_t] \quad (5)$$

Persistence notwithstanding, higher values of  $\alpha$  are intended to imply a relative preference for current consumption. This motivates the following assumption:

**Assumption 3.**  $\rho(\alpha_{t+1}|\alpha_t) \in [0, 1)$  for all  $(\alpha_t, \alpha_{t+1}) \in A^2$ .

It is immediate from (5) that this implies  $\varepsilon^\alpha(\alpha_t) < 1$ . This implies that higher current  $\alpha$  may raise expectations about future marginal utility, but not by so much as the increase in current marginal utility. Thus preferences become more present-biased.

Some formulae will also feature the product of successive  $\rho(\alpha_{t+1}|\alpha_t)$  terms. Hence for all  $t < s$ ,  $\alpha^s \in A^{s+1}$ , we define  $D_{t,s}(\alpha^s)$ :

$$D_{t,s}(\alpha^s) := \prod_{r=t+1}^s \rho(\alpha_r|\alpha_{r-1}) \quad (6)$$

and normalise  $D_{t,t}(\alpha^t) \equiv 1$ .

---

<sup>14</sup>Note that in the lognormal case, where:

$$\log \alpha_{t+1} \sim N(\rho \log \alpha_t, \sigma^2)$$

for parameters  $\rho$  and  $\sigma$ , we have  $\rho(\alpha_{t+1}|\alpha_t) \equiv \rho$ .

Related to  $\rho$  is the elasticity of the density with respect to lagged type, denoted  $\pi^\Delta$ :

$$\pi^\Delta(\alpha_{t+1}|\alpha_t) := \frac{\alpha_t}{\pi(\alpha_{t+1}|\alpha_t)} \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} \quad (7)$$

Integration by parts allows this to be linked to  $\rho$ . For instance, for any absolutely continuous function  $f : A \rightarrow \mathbb{R}$ , we have:

$$\int_{\alpha_{t+1}} \alpha_{t+1} f'(\alpha_{t+1}) \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} = \int_{\alpha_{t+1}} f(\alpha_{t+1}) \pi^\Delta(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \quad (8)$$

Finally, we impose a standard monotone likelihood condition on  $\pi$ . This is not used in the characterisation results, but plays an important role in confirming that optimal savings taxes are positive.

**Assumption 4.** For all  $\alpha'_t, \alpha''_t$  with  $\alpha'_t < \alpha''_t$ , the ratio  $\frac{\pi(\alpha_{t+1}|\alpha''_t)}{\pi(\alpha_{t+1}|\alpha'_t)}$  is monotone increasing in  $\alpha_{t+1}$ .

An implication of this condition is that  $\pi^\Delta(\alpha_{t+1}|\alpha_t)$  is monotone increasing in  $\alpha_{t+1}$ .

### 3.3 Planner choice

The planner's aim in period 0 is to maximise a simple utilitarian sum, denoted  $W_0$ :

$$W_0 := \int_{\underline{\alpha}}^{\bar{\alpha}} U_0(\alpha_0) d\Pi(\alpha_0) \quad (9)$$

The precise utilitarian form for period 0 is not important, and easily generalised.

The planner can commit perfectly in period 0 to an allocation mechanism for all future dates. This assumption means that the revelation principle will apply, and so there is no loss in generality from initially focusing on direct revelation mechanisms in which truth-telling is optimal. Thus individuals report their type each period, and receive a consumption allocation conditional on their reports to date,  $c_t(\alpha^t)$ . A complete set of  $c_t(\alpha^t)$  functions for all  $t \geq 0$  and  $\alpha^t \in A^{t+1}$  is referred to as an **allocation**.<sup>15</sup>

The planner's choice is restricted by the resource and incentive compatibility constraints detailed

---

<sup>15</sup>The dependence of  $c_t$  on  $\alpha^t$  will be left implicit where the context allows.

below, plus a technical **interiority restriction**. This is defined by a set of scalars  $\{K_t\}_{t \geq 0}$  and the bound:

$$\left| \mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} \alpha_s u(c_s(\alpha^s)) \right| \leq K_t \quad (10)$$

for all  $\alpha^t \in A^{t+1}$ .

Condition (10) guarantees that information rents are well defined at each history node, but it does not capture meaningful economic restrictions, and cases where the constraint binds will not be our main focus. In particular, the value of  $K_t$  can be set arbitrarily large for each  $t$ . If (10) does not bind for any  $\alpha_t$  following a given  $\alpha^{t-1}$ , period- $t$  consumption levels will be called **interior** for this history. If (10) never binds, the allocation as a whole will be called interior. The focus of the remainder of the paper is exclusively on interior allocations.

### 3.4 Resources

Policy choice is subject to two main constraints: resources and incentive compatibility. Of these, the resource constraint is by far the simpler. We assume that there is an exogenous, time-invariant world real interest rate, whose gross value is  $R \leq \beta^{-1}$ . The constraint requires the net-present value of consumption to equal the net-present value of endowments:

$$\sum_{t=0}^{\infty} R^{-t} \left[ y_t - \int_{\alpha^t} c_t(\alpha^t) d\Pi_t(\alpha^t) \right] \geq 0 \quad (11)$$

This departs from the structure in Atkeson and Lucas (1992), where no savings technology exists. This is not important for the characterisation results below, as it is a simple extension to let  $R$  vary over time, and to set it period-by-period to a value that ensures  $\int_{\alpha^t} c_t(\alpha^t) d\Pi_t(\alpha^t) = y_t$  for all  $t$ .

### 3.5 Incentive compatibility

Incentive compatibility requires that truth-telling should be optimal for all types in each successive period, and after each possible history. This places a set of restrictions in  $t$  across every subset of types that share a common  $\alpha^{t-1}$ . It is helpful to characterise it by reference to continuation utilities. Let  $V_t(\alpha^{t-1}; \alpha_t)$  be the maximised value for  $U_t$  available to an individual with history of type reports  $\alpha^{t-1}$  and current

type  $\alpha_t$ . This has the recursive definition:

$$V_t(\alpha^{t-1}; \alpha_t) = \max_{\tilde{\alpha}_t} \left\{ \alpha_t u(c_t(\alpha^{t-1}, \tilde{\alpha}_t)) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha^{t-1}, \tilde{\alpha}_t; \alpha_{t+1}) d\Pi(\alpha_{t+1} | \alpha_t) \right\}$$

for all  $t \geq 0$ .<sup>16</sup>

Incentive compatibility then requires:

$$\begin{aligned} & \alpha'_t u(c_t(\alpha^{t-1}, \alpha'_t)) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha^{t-1}, \alpha'_t; \alpha_{t+1}) d\Pi(\alpha_{t+1} | \alpha'_t) \\ & \geq \alpha'_t u(c_t(\alpha^{t-1}, \alpha''_t)) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha^{t-1}, \alpha''_t; \alpha_{t+1}) d\Pi(\alpha_{t+1} | \alpha'_t) \end{aligned} \quad (12)$$

for all  $t \geq 0$ ,  $\alpha^{t-1} \in A^t$ ,  $\alpha'_t \in A$  and  $\alpha''_t \in A$ .  $\alpha'_t$  here represents the agent's true type, and  $\alpha''_t$  a candidate report.

Note that the true type  $\alpha'_t$  affects restriction (12) in two ways. Most directly, it controls the marginal utility of consumption in  $t$ . But current type also affects the distribution of future type draws,  $\Pi(\alpha_{t+1} | \alpha'_t)$ . This persistence channel complicates the link between preferences and type, relative to a canonical two-good screening problem.

An allocation that satisfies constraints (10), (11) and (12) for all histories and all time periods is described as **incentive-feasible**. The **planner's problem** is to maximise  $W_0$  on the set of incentive-feasible allocations.

## 4 First-order incentive compatibility

### 4.1 A relaxed incentive constraint

Condition (12) implies a continuum of constraints for every element of  $A$ , after every history  $\alpha^{t-1}$  – a dimensionality that is not possible to handle tractably. Since there is only one consumption level, and one continuation value, to solve for at each  $\alpha^t$ , almost all of these constraints must be redundant. In keeping with much of the literature, we thus replace them with a ‘first-order’ envelope requirement that is necessary for (12) to be true, but not sufficient. The conjecture, to be verified, is that optimal policy

---

<sup>16</sup>The Markov property of shocks implies that the value of  $V_{t+1}$  is unaffected by the truthfulness, or otherwise, of past reports.

for the simplified constraint set will also be optimal for the more complex constraint set. This is what is known as the first-order approach to mechanism design.<sup>17</sup>

The approach is easiest to define by reference to two state variables:  $\omega_t(\alpha^{t-1})$ , which corresponds to the average level of utility across agents with a common history  $\alpha^{t-1}$ , and  $\omega_t^\Delta(\alpha^{t-1})$ , which summarises information rents that arise due to the impact of current type on the distribution of future  $\alpha$  values. These objects have the following recursive definitions:

$$\omega_t(\alpha^{t-1}) := \int_{\alpha_t} \left\{ \alpha_t u(c_t(\alpha^{t-1}, \alpha_t)) + \beta \omega_{t+1}(\alpha^{t-1}, \alpha_t) \right\} d\Pi(\alpha_t | \alpha_{t-1}) \quad (13)$$

$$\omega_t^\Delta(\alpha^{t-1}) := \int_{\alpha_t} \rho(\alpha_t | \alpha_{t-1}) \cdot \left\{ \alpha_t u(c_t(\alpha^{t-1}, \alpha_t)) + \beta \omega_{t+1}^\Delta(\alpha^{t-1}, \alpha_t) \right\} d\Pi(\alpha_t | \alpha_{t-1}) \quad (14)$$

In subsequent usage the history dependence of these objects will be left implicit so long the meaning remains clear.

The following result is critical to the characterisation:

**Lemma 1.** *For all  $t \geq 0$  and any given history  $\alpha^{t-1} \in A^t$ , an incentive-feasible allocation will satisfy the following envelope condition for all current types  $\alpha'_t \in A$ :*

$$\begin{aligned} \alpha'_t u(c_t(\alpha'_t)) + \beta \omega_{t+1}(\alpha'_t) &= \underline{\alpha} u(c_t(\underline{\alpha})) + \beta \omega_{t+1}(\underline{\alpha}) \\ &+ \int_{\underline{\alpha}}^{\alpha'_t} \frac{1}{\alpha_t} \left[ \alpha_t u(c_t(\alpha_t)) + \beta \omega_{t+1}^\Delta(\alpha_t) \right] d\alpha_t \end{aligned} \quad (15)$$

We refer to equation (15) as the **relaxed incentive constraint**. For an arbitrary type  $\alpha'_t$ , it decomposes the value of lifetime utility into the value for the lowest type  $\underline{\alpha}$ , plus the sum of ‘information rents’ between  $\underline{\alpha}$  and  $\alpha'_t$  – that is, the marginal increments to utility that are needed to keep truthful reporting as a local optimum for all types. These information rents are the objects contained within the integral on the last line.

An allocation that satisfies the interiority constraint (10), resource constraint (11) and relaxed incen-

---

<sup>17</sup>The first-order approach has been used extensively in the dynamic mechanism design literature. Kapička (2013) highlighted the computational gains from a lower-dimensional state space when non-local deviations were neglected, with an application to a dynamic Mirrleesian economy. Farhi and Werning (2013) made use of similar techniques, also in a dynamic Mirrleesian setting. Pavan, Segal and Toikka (2014) provide detailed a detailed exploration of the first-order approach in problems with quasilinear preferences, and the approach in the present paper bears close resemblance to theirs.



tive constraint (15) for all periods and histories is called a **relaxed incentive-feasible allocation**. The **relaxed planner’s problem** is to maximise  $W_0$  on the set of relaxed incentive-feasible allocations.

By Lemma 1, the set of relaxed incentive-feasible allocations must contain the set of incentive-feasible allocations. If the optimal allocation from the set of relaxed incentive-feasible options is also incentive-feasible, it follows that it must be optimal for the main planner’s problem. Confirming this inclusion is the central issue in justifying the first-order approach, and the focus of the next subsection.

## 4.2 Sufficiency

When will the relaxed incentive constraint imply global incentive compatibility? In bivariate problems, this issue can be addressed by a classic Spence-Mirrlees approach. Given single crossing in preferences, an appropriate form of monotonicity in the solution is enough. This works because ‘single crossing plus monotonicity’ allows inference to be drawn about the preferences of all agents, based on the local preferences of any one.

In a multi-period setting the situation is less straightforward, because current type may influence preferences in a complex, multidimensional way. In the present environment, this occurs when types are persistent. In such a case, an increase in  $\alpha_t$  does not just make current consumption more desirable relative to future. It also changes an agent’s distribution across future draws,  $\Pi(\alpha_{t+1}|\alpha_t)$ . This means that an allocation with relatively low  $c_t$  could nonetheless be appealing to an agent with high  $\alpha_t$ , if it delivers a more advantageous distribution of future outcomes for this type.

We present two alternative criteria for confirming global incentive compatibility. The first is an ‘integral monotonicity condition’, of the type introduced in quasilinear settings by Pavan, Segal and Toikka (2014).<sup>18</sup> This has the advantage that it is both necessary and sufficient for (15) to imply global incentive compatibility, but the disadvantage that it depends on properties of the utility function rather than the nonlinear menu of options alone. In this regard it makes significantly greater informational demands than an ordinal (monotonicity) condition on the set of allocations. Thus a subsequent corollary gives a sufficient – but no longer necessary – monotonicity condition for the allocation alone to satisfy. This condition has a particularly clear interpretation by reference to the decentralisation that is introduced

---

<sup>18</sup>C.f. their Theorem 3.

later in the paper.

To simplify presentation, we adopt the convention for definite integrals that  $\int_{\alpha_t'}^{\alpha_t''} \{\cdot\} d\alpha_t$  corresponds to  $-\int_{\alpha_t''}^{\alpha_t'} \{\cdot\} d\alpha_t$  when  $\alpha_t' > \alpha_t''$ . As a necessary and sufficient condition, we have:

**Proposition 1.** *A relaxed incentive-feasible allocation is incentive-feasible if and only if for all  $t$ ,  $\alpha^{t-1} \in A^t$  and  $(\alpha_t', \alpha_t'') \in A^2$ , the following condition is true:*

$$\int_{\alpha_t'}^{\alpha_t''} \frac{1}{\alpha_t^2} \left\{ \mathbb{E}_t \left[ \sum_{s=t+1}^{\infty} \beta^{s-t} (1 - D_{t,s}(\alpha^s)) \alpha_s \left[ u \left( c_s \left( \alpha^{t-1}, \alpha_t, \alpha_{t+1}^s \right) \right) - u \left( c_s \left( \alpha^{t-1}, \alpha_t'', \alpha_{t+1}^s \right) \right) \right] \middle| \alpha_t \right] \right\} d\alpha_t \geq 0 \quad (16)$$

Since  $D_{t,s}(\alpha^s) < 1$ , the following corollary is immediate:

**Corollary 1.** *A relaxed incentive-feasible allocation is incentive-feasible if for all  $t$  and  $s$  with  $s > t$ , all  $\alpha^{t-1} \in A^t$  and all  $\alpha_{t+1}^s \in A^{s-t+1}$ , the consumption function  $c_s \left( \alpha^{t-1}, \alpha_t, \alpha_{t+1}^s \right)$  is non-increasing in  $\alpha_t$ .*

Thus global incentive compatibility is confirmed so long as future consumption is weakly decreasing in current type, along all subsequent history nodes. This condition has a particularly simple interpretation when the allocation is decentralised in the manner described in Section 6: it is equivalent to requiring that higher savings in  $t$  raise consumption along every subsequent history path, or that consumption is a normal good at every date-state. Given this, we refer to an allocation that satisfies the requirements of Corollary 1 a **normal allocation**.

Normality – and the global incentive compatibility that it implies – guarantees that period- $t$  consumption is non-decreasing in  $\alpha_t$  for types with a common history. For some purposes it is convenient to strengthen this, and neglect the possibility that multiple types bunch at the same consumption value in  $t$ . Thus, we define ‘strictly normal’ allocations as follows:

**Definition.** An allocation is called **strictly normal** if it is normal and for all  $t$  and  $\alpha^{t-1} \in A^t$  there exists  $\delta_t(\alpha^{t-1}) > 0$  such that  $\frac{c_t(\alpha^{t-1}, \alpha_t'') - c_t(\alpha^{t-1}, \alpha_t')}{\alpha_t'' - \alpha_t'} \geq \delta_t(\alpha^{t-1})$  for all  $(\alpha_t', \alpha_t'') \in A^2$ .

That is, current consumption is strictly increasing in  $\alpha_t$ , by an amount that is bounded below with respect to the change in  $\alpha_t$ .

Note from the definition that a focus on strictly normal allocations will be enough to guarantee, for any given shock history, that the inverse mapping  $\alpha_t(c)$  – the type associated with each consumption

level – will be uniquely defined, and Lipschitz continuous on all sub-intervals in  $(c_t(\alpha^{t-1}, \underline{\alpha}), c_t(\alpha^{t-1}, \bar{\alpha}))$  where a positive measure of types locate.<sup>19</sup>

## 5 Characterising direct allocations

In the present section, I provide a direct characterisation of optimal allocations for the relaxed planner’s problem. This characterisation expresses the trade-off associated with local changes to the information rents that are earned by each type at a given history node. Changes to information rents imply changes to the cross-sectional profile of utilities – and so the resulting expressions trade off the marginal costs of changing information rents with the marginal costs of providing utility to different agents. Thus the characterisation, though insightful, relies heavily on arguments of the utility function. It makes no direct reference to taxes, or the behavioural effects of changes to a system of decentralised market prices.

Yet the results here have substantial instrumental value. Sections 7 and 8 will show how the same conditions can be reformulated by reference to a consumption-savings decentralisation, introduced in Section 6. The resulting expressions are elementary manipulations of those derived here, but have interpretations in terms of observable ‘sufficient statistics’.

### 5.1 Characterisation result

By conventional techniques, we derive the following:

**Proposition 2.** *Suppose an interior allocation is optimal in the relaxed problem. For a cross-section of types in  $t$  with a common history  $\alpha^{t-1}$ , the following must hold a.e.:*

$$\mathbb{E}_{t-1} \left\{ \alpha_t \left[ 1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t | \alpha_{t-1}) \right] - \frac{\eta_t}{u'(c_t(\alpha_t))} \middle| \alpha_t > \alpha'_t \right\} = \frac{\pi(\alpha'_t | \alpha_{t-1})}{(1 - \Pi(\alpha'_t | \alpha_{t-1}))} \cdot (\alpha'_t)^2 \cdot \{ \lambda_{t+1}^\Delta(\alpha'_t) - \rho(\alpha'_t | \alpha_{t-1}) \lambda_t^\Delta \} \quad (17)$$

$$\mathbb{E}_{t-1} \left\{ \alpha_t \left[ 1 + \lambda_t + \lambda_t^\Delta \varepsilon^\alpha(\alpha_{t-1}) \right] - \frac{\eta_t}{u'(c_t(\alpha_t))} \right\} = 0 \quad (18)$$

---

<sup>19</sup>Once consumption is monotone in type, there can be at most countably many discontinuities in the function  $c_t(\alpha^{t-1}, \cdot)$ .  $\alpha_t(c)$  is not defined for values of  $c$  that lie between the left and right limits of each discontinuity in  $c_t(\alpha^{t-1}, \cdot)$ .

where  $\eta_t$  is the shadow value of resources for the planner, satisfying:

$$\eta_t = (\beta R)^{-t} \frac{\mathbb{E} [\alpha_0]}{\mathbb{E} \left[ \frac{1}{u'(c_0(\alpha_0))} \right]} \quad (19)$$

$\lambda_t$  and  $\lambda_t^\Delta$  are scalars measurable with respect to  $\alpha^{t-1}$ , satisfy  $\lambda_0 = \lambda_0^\Delta \equiv 0$ , and update according to:

$$\lambda_{t+1} = \lambda_t + \mu_t (\alpha'_t) \quad (20)$$

$$\lambda_{t+1}^\Delta = \rho (\alpha'_t | \alpha_{t-1}) \lambda_t^\Delta - \frac{1 - \Pi (\alpha'_t | \alpha_{t-1})}{\alpha_t \pi (\alpha'_t | \alpha_{t-1})} \mathbb{E}_{t-1} [\mu_t (\alpha'_t) | \alpha_t > \alpha'_t] \quad (21)$$

with  $\mu_t (\alpha'_t)$  a mean-zero object defined in the appendix. The conditional distribution and densities are replaced with their unconditional equivalents for period 0.

Expressions (17) and (18) are the main objects of interest here. (17) can be interpreted by reference to the costs and benefits of changing information rents at a particular point in the cross-sectional type distribution, for agents with a common history. This yields a direct welfare benefit, mitigated by the direct marginal resource cost of the higher utility – which together account for the objects on the left-hand side. Against this is the marginal cost of raising information rents *at* the threshold type, in order for (15) to remain true. This is captured by the object on the right-hand side: note that  $\lambda_{t+1}^\Delta (\alpha'_t) - \rho (\alpha'_t | \alpha_{t-1}) \lambda_t^\Delta$  is the shadow cost of raising  $\omega_{t+1}^\Delta (\alpha'_t)$ , holding constant  $\omega_t^\Delta$ .

If instead utility is raised uniformly for all agents with a common history prior to  $t$ , then no within-period change to information rents is required. The result is equation (18): the value of raising welfare across types must equal the resource cost of doing so.

$\lambda_t$  and  $\lambda_t^\Delta$  in these expressions are multipliers deriving from prior incentive restrictions – capturing the shadow costs of changing  $\omega_t$  and  $\omega_t^\Delta$  respectively. As highlighted by Marcat and Marimon (2019), the Pareto weights that the policymaker attaches to different agents' utility in  $t$  are updated to accommodate the shadow benefits of changes to incentives in prior periods.

## 5.2 The dynamics of consumption

A key element of the characterisation in Proposition 2 is the the inverse marginal utility of consumption – equivalently, the marginal cost of providing  $\alpha_t$  units of utility to an agent in period  $t$ . This is a widely-studied object in dynamic incentive problems, where it is commonly used to assess the long-run properties of the consumption distribution.<sup>20</sup> When shocks are iid, the inverse marginal utility is well-known to follow a quasi-martingale process, with substantial implications for long-run inequality.<sup>21</sup> There has been significant recent debate about the sensitivity of this conclusion to type persistence.<sup>22</sup> In this subsection we show, for the present model, that the dynamics of the inverse marginal utility of consumption can be described by the interaction between two multiplier processes, one stationary, and one following a martingale. In particular:

**Proposition 3.** *For all  $t$  and  $s$ ,  $s \geq t$ , and any history  $\alpha^t$ , the period- $t$  expected value of the period- $s$  inverse marginal utility of consumption satisfies:*

$$\frac{1}{\mathbb{E}_t[\alpha_s]} \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s)} \right] = \frac{1 + \lambda_{t+1}}{\eta_t} + \frac{\mathbb{E}_t [D_{t,s}(\alpha^s) \alpha_s]}{\mathbb{E}_t[\alpha_s]} \frac{\lambda_{t+1}^\Delta}{\eta_t} \quad (22)$$

According to the Proposition, the expected value of the marginal cost of utility provision evolves as a composite of two multiplier processes. For long-run expectations, what matters is the object  $\frac{1 + \lambda_{t+1}}{\eta_t}$  – the shadow value of raising lifetime utility for type  $\alpha_t$ , in periods after  $t$ . This is the only component that matters in the long run, because  $\rho(\alpha_r | \alpha_{r-1}) \in [0, 1)$ , and so  $\frac{\mathbb{E}_t [D_{t,s}(\alpha^s) \alpha_s]}{\mathbb{E}_t[\alpha_s]} \rightarrow 0$  as  $s \rightarrow \infty$ .<sup>23</sup> Since  $\lambda_{t+1}$  follows a martingale,<sup>24</sup> the implication is that shocks to this martingale process control *long-run* consumption outcomes. Intuitively, given diminishing marginal utility it is cost-efficient to spread incentives over time. If there is justification for raising the utility of type  $\alpha_t$  from  $t + 1$  on, in return for this type consuming relatively little in  $t$ , then it is cost-effective to spread this increase across all future periods.

<sup>20</sup>In particular, Rogerson (1985) first highlighted the ‘inverse Euler equation’ as a dynamic optimality condition for the marginal cost of utility provision in multi-period moral hazard settings, following its derivation in a two-period setting by Diamond and Mirrlees (1978). Thomas and Worrall (1990) showed that this condition implied almost sure immiseration in the long run, provided the discount factor was sufficiently small.

<sup>21</sup>This setting is explored in detail by Farhi and Werning (2007).

<sup>22</sup>See, for instance, Bloedel, Krishna and Strulovici (2020).

<sup>23</sup>Recall that  $D_{t,s}(\alpha^s) := \prod_{r=t+1}^s \rho(\alpha_r | \alpha_{r-1})$ , and  $\alpha_s$  is bounded.

<sup>24</sup>C.f. equation (20).

These arguments are well understood from the existing social insurance literature. But when types are persistent, there is an additional component to the short-run expected marginal cost, more in keeping with the dynamic contracting literature in quasilinear settings.<sup>25</sup> This comes from the desire to spread over time the distortions to efficient choice that are needed in order to achieve desired values for the information rents at any given  $\alpha_t$ . Type persistence implies that periods after  $t$  matter for information rents in  $t$ .<sup>26</sup> If it is desirable in  $t$  to raise rents at  $\alpha_t$  (so  $\lambda_{t+1}^\Delta > 0$ ), then it will also be desirable to pay some costs to achieve the same change in periods after  $t$ . The extent to which this incentive fades over time depends on the extent of persistence in the shock process, which controls how much outcomes in period  $s$  matter for information rents in  $t$ . In this context, note that the object  $\frac{\mathbb{E}_t [D_{t,s}(\alpha^s) \alpha_s]}{\mathbb{E}_t [\alpha_s]}$  is precisely the elasticity of  $\mathbb{E}_t [\alpha_s]$  with respect to  $\alpha_t$ .

The layering of a transitory shock component over the more conventional martingale for inverse marginal utilities complicates the derivation and interpretation of long-run results relating to inequality, but the main content of the immiseration conclusion endures. In particular, notice:

$$\frac{(1 + \lambda_{t+1})}{\eta_t} = \lim_{s \rightarrow \infty} \left\{ \frac{1}{(\beta R)^{s-t}} \frac{1}{\mathbb{E}_t [\alpha_s]} \mathbb{E}_t \left[ \frac{1}{u'(c_s)} \right] \right\} \geq 0$$

Since  $\eta_t > 0$ , from (19), it follows that  $(1 + \lambda_{t+1})$  is a bounded martingale. Thus it converges a.s. in  $t$ . So long as  $R \leq \beta^{-1}$ , this will imply convergence to zero in the *long-run expected* inverse marginal utility. In particular, note from (20) that convergence in  $t$  to a constant, positive value for  $(1 + \lambda_{t+1})$  is only possible if the allocation converges to an outcome with  $\mu_s(\alpha_s) = 0$  for all  $\alpha_s, s \geq t$ . But since  $\mu_s$  is the multiplier on the relaxed incentive constraint, this implies convergence to a first-best allocation, maximising welfare on the resource constraint alone. This violates incentive compatibility. Thus for  $R \leq \beta^{-1}$ , as  $t$  becomes large:

$$\left\{ \lim_{s \rightarrow \infty} \mathbb{E}_t \left[ \frac{1}{u'(c_s)} \right] \right\}_{\text{a.s.}} \rightarrow 0$$

That is, the long-run expectation of the marginal cost of utility provision converges almost surely to zero.

<sup>25</sup>See, for example, the discussion in Pavan (2017).

<sup>26</sup>Recall, from equation (15), that the information rent at type  $\alpha_t$  is given by:

$$\frac{1}{\alpha_t} \left[ \alpha_t u(c_t(\alpha_t)) + \beta \omega_{t+1}^\Delta(\alpha_t) \right]$$

and  $\omega_{t+1}^\Delta(\alpha_t)$  is generically non-zero once types are persistent.

## 6 A consumption-savings decentralisation

The analysis that now follows focuses on a decentralisation of the optimal direct allocation whereby agents face a simple period-by-period choice between consumption and savings, subject to a nonlinear savings tax. Specifically, individuals enter each period  $t$  with a given value of net wealth,  $M_t$ , which for convenience we normalise to include the net present value of future endowments. Wealth in  $t$  can either be allocated to period- $t$  consumption,  $c_t$ , or savings,  $s_t$ . This choice is observable, and the planner implements a non-linear tax on  $s_t$ , which may vary in the history of past savings decisions  $s^{t-1}$ . This tax is denoted  $T_t(s_t; s^{t-1})$ , or simply  $T_t(s_t)$  if context allows. The tax is normalised to equal zero on average, given  $s^{t-1}$  and the equilibrium distribution of future choices.

In  $t + 1$  the individual is allocated their residual post-tax savings, together with interest, as their new wealth level, and choice proceeds as before. The budget constraints can thus be written in sequential form as:

$$c_t + s_t = M_t \quad (23)$$

$$M_{t+1} = R \left[ s_t - T_t(s_t; s^{t-1}) \right] \quad (24)$$

Given  $M_0$ , individuals choose contingent consumption sequences to maximise  $U_0$ , subject to (23) and (24), plus a ‘no Ponzi’ constraint:

$$\lim_{t \rightarrow \infty} R^{-t} M_t \geq 0 \quad (25)$$

Notice that conditions (23) to (25) together imply a forward-looking infinite-horizon budget constraint that must be satisfied in all periods  $r \geq 0$ , for any realised consumption-savings path:

$$M_r = \sum_{t=r}^{\infty} R^{-(t-r)} \left[ c_t + T_t(s_t; s^{t-1}) \right] \quad (26)$$

That is,  $M_r$  must equal the net present value of consumption and tax payments from  $r$  onwards.

Proposition 4 provides conditions under which an incentive-feasible allocation can be decentralised by a tax scheme of this kind.

**Proposition 4.** *An incentive-feasible allocation  $\{c_t^*(\alpha^t)\}_{t, \alpha^t}$  can be decentralised by a sequence of tax functions*

$T_t(s_t; s^{t-1})$  provided:

1.  $c_t^*(\alpha^{t-1}, \alpha_t)$  is increasing in  $\alpha_t$  for all  $t$  and  $\alpha^{t-1}$ , and
2. If  $c_t^*(\alpha^{t-1}, \alpha'_t) = c_t^*(\alpha^{t-1}, \alpha''_t)$  for distinct  $\alpha'_t, \alpha''_t$ , then  $c_{t+s}^*(\alpha^{t-1}, \alpha'_t, \alpha_{t+1}^{t+s}) = c_t^*(\alpha^{t-1}, \alpha'_t, \alpha_{t+1}^{t+s})$  for all  $\alpha_{t+1}^{t+s} \in A^s, s > 0$ .

The main restriction here is to consumption allocations that are either strictly increasing in type or, if multiple types bunch at the same allocation, provide identical future allocations across ‘bunchers’. If there is strict increasingness then an agent’s consumption/saving choice in  $t$  implies a unique value for their type,  $\alpha_t$ , and so consumption choice is informationally equivalent to a direct type report. If bunching occurs, consumption choice implies a range of possible types. For the decentralisation to work, the intended allocation must not subsequently differentiate more precisely among the types that fall in this range.

A central contribution of the present paper is to provide a direct characterisation of the optimal tax schedule  $T'(s_t)$ , by reference to economic statistics that arise in this decentralisation. These statistics are, in particular, the elasticity of savings with respect to current and future marginal tax rates, the effect on savings of higher wealth,  $M_t$ , and the endogenous conditional distribution of savings at each history node.

## 7 Sufficient statistics: preliminaries

This section defines a number of important concepts and provides intermediate steps in order to map to the desired sufficient statistics representation.

### 7.1 Towards sufficient statistics: intuition and integration

Equation (17) states that the net benefits from raising the utility of high-type agents must be traded off against the costs of changing information rents in a compatible manner. But the practical value of the characterisation is weakened by the fact that its key components – the marginal costs of unit changes to utility and to information rents – are defined by reference to the utility function and the unobserved type process. These do not map easily to observables, and – in the case of utility – are only defined up to



a large class of renormalisations. It would be preferable to characterise, as far as possible, by reference to measurable objects: behavioural elasticities and observable distributions.

We achieve this by use of a novel, intuitive analytical step. The logic is as follows. Sufficient statistics characterisations in static nonlinear tax problems typically describe the costs and benefits at the margin of simple step changes in the cross-sectional profile of effective *income*, or wealth in the population. For example, a cut in the marginal tax rate *at* a certain point in the earnings (or savings) distribution raises the effective income of all types above this point, by a uniform amount.<sup>27</sup> By considering the resulting behavioural responses – a combination of standard income and substitution effects – one can arrive at an expression for the net fiscal cost of the tax cut, to be contrasted with its welfare benefits.

Proposition 2 also describes the costs and benefits of a simple step change, but in the cross-sectional profile of *utilities* rather than incomes. As already discussed, its key components are marginal costs and benefits ‘per unit change in utility’, as a result of marginal changes to information rents. But this does not prevent it from being used to discuss income changes – it just means that a conversion is needed. For any given profile of income changes, there will always be a corresponding profile of utility changes. An understanding of the former can be achieved by starting from the latter.

More substantively, a unit marginal increase in the feasible period-*t* consumption level for an agent of type  $\alpha_t$  raises their utility by  $\alpha_t u' (c_t (\alpha_t))$  at the margin. So long as the envelope condition applies, this will be true whether the additional resources are fully used on period-*t* consumption, or are partly saved. To analyse the effects of changing effective *incomes* by a uniform amount below some threshold type, it is sufficient to analyse the effects of changing *utility* by  $\alpha_t u' (c_t (\alpha_t))$  units for all agents in this range. By integrating equations (17) and (18) in appropriate proportions, this is a straightforward exercise.

To this end, we obtain the following as a corollary to Proposition 2:

**Corollary 2.** *If a strictly normal allocation is optimal in the relaxed problem, with  $c_t (\alpha_t)$  continuous at a given history node, then the following two expressions are true:*

---

<sup>27</sup>The discussion in Piketty and Saez (2013b) provides detailed treatment.

$$\begin{aligned}
& \int_{\underline{c}}^{\bar{c}} \left[ 1 - \frac{\alpha_t(c) u'(c) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c) | \alpha_{t-1})\}}{\eta_t} \right] \pi^c(c | \alpha^{t-1}) dc \\
& + \int_{\underline{c}}^{\bar{c}} (\alpha_t(c))^2 u''(c) \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c))}{\eta_t} - \rho(\alpha_t(c) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_t} \right\} \left( \frac{d\alpha_t(c)}{dc} \right)^{-1} \pi^c(c | \alpha^{t-1}) dc \\
& = 0
\end{aligned} \tag{27}$$

$$\begin{aligned}
& - \int_{\underline{c}'}^{\bar{c}'} \left[ 1 - \frac{\alpha_t(c') u'(c') \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c') | \alpha_{t-1})\}}{\eta_t} \right] \pi^c(c' | \alpha^{t-1}) dc' \\
& - \int_{\underline{c}'}^{\bar{c}'} (\alpha_t(c'))^2 u''(c') \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c'))}{\eta_t} - \rho(\alpha_t(c') | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \left( \frac{d\alpha_t(c')}{dc'} \right)^{-1} \pi^c(c' | \alpha^{t-1}) dc' \\
& + (\alpha_t(c'))^2 u'(c') \left( \frac{d\alpha_t(c')}{dc'} \right)^{-1} \pi^c(c' | \alpha^{t-1}) \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c'))}{\eta_t} - \rho(\alpha_t(c') | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \\
& = 0
\end{aligned} \tag{28}$$

the latter for almost all  $c' \in (\underline{c}', \bar{c}')$ , with  $\underline{c} = c_t(\underline{\alpha})$  and  $\bar{c} = c_t(\bar{\alpha})$ , and  $\pi^c(c_t | \alpha^{t-1})$  denoting the realised density of consumption in  $t$ , given history  $\alpha^{t-1}$ .

The restriction to continuous  $c_t(\alpha_t)$  functions is made for convenience only: the appendix shows that it is possible to incorporate jumps in consumption as type increases. Note also that the realised consumption density will equal the realised savings density, denoted  $\pi^s(s_t | \alpha^{t-1})$ , since:

$$\Pi^c(c'_t | \alpha^{t-1}) \equiv P(c_t \leq c'_t) = P(s_t \geq s'_t) \equiv 1 - \Pi^s(s'_t | \alpha^{t-1})$$

where  $s'_t := M_t(\alpha^{t-1}) - c'_t$ , so  $\frac{ds'_t}{dc'_t} = -1$ .

## 7.2 Relevant behavioural statistics

By themselves, conditions (27) and (28) appear just to be slightly more unwieldy alternatives to (17) and (18). Their value becomes evident only when their components can be equated with meaningful

behavioural statistics. The characterisation in Section 8 will do precisely this. It makes use of four distinct behavioural statistics, defined by reference to the decentralisation of Section 6. These are defined in turn here.

**The contemporaneous elasticity of savings with respect to the post-tax rate of return:** This is denoted  $\epsilon_t^s$ . For an agent whose chosen savings level is  $s_t$ , it is defined as the response to a change in the local marginal tax rate that they face:

$$\epsilon_t^s := \frac{R(1 - T_t'(s_t))}{s_t} \frac{ds_t}{dR(1 - T_t'(s_t))}$$

As for the other statistics, this value is not 'structural'. It will be endogenous to the chosen allocation, and associated tax schedule. It is a compensated elasticity, since the the total tax liability at  $s_t$ ,  $T_t(s_t)$ , remains unchanged to first order when the marginal tax rate,  $T_t'(s_t)$ , changes.

**The contemporaneous income effect on savings:** This is denoted  $\frac{ds_t}{dM_t}$ . For a given period- $t$  type, it is defined as the effect on  $s_t$  when  $M_t$  is increased at the margin, holding constant current and future savings tax schedules, and abstracting from anticipation effects prior to  $t$ .

**The compensated elasticity of lagged savings, with respect to contemporary returns:** This object captures the complementarities that may exist in economies with incomplete insurance, between the state-contingent profile of returns provided by a savings instrument, and the level of saving itself. Denoted  $\epsilon_{t-1,t}^s(s_t)$ , it measures the response of savings in  $t - 1$  to the change in the profile of insurance at  $t$  that is generated by a cut in the marginal savings tax rate in the interval  $(s_t, s_t + \Delta)$ , taking the limit as  $\Delta$  becomes small. To abstract from mechanical effects, this is normalised by (a) the relative proportion of agents in  $t$  who benefit from the tax cut,  $(1 - \Pi^s(s_t | \alpha^{t-1}))$ , where  $\Pi^s(s_t | \alpha^{t-1})$  denotes the relevant conditional distribution of savings in  $t$ , and (b) the size of the interval on which taxes are cut, denoted  $ds_t$  at

the limit. Thus, heuristically:<sup>28</sup>

$$\epsilon_{t-1,t}^s(s_t) \left(1 - \Pi^s(s_t | \alpha^{t-1})\right) ds_t := \frac{(1 - T_t'(s_t))}{s_{t-1}} \frac{ds_{t-1}}{d(1 - T_t'(s_t))} \Big|_{\text{comp}} \quad (29)$$

$\epsilon_{t-1,t}^s(s_t)$  is a compensated elasticity, viewed from the perspective of  $t - 1$ . It is calculated assuming a *uniform* compensating adjustment to lifetime utility across period- $t$  states, so that an agent with realised history  $\alpha^{t-1}$  does not experience any change to their continuation value,  $\omega_t(\alpha^{t-1})$ . Thus the behavioural change in  $t - 1$  that it captures is purely due to the re-profiling of state-by-state utility outcomes in  $t$ , and not to a first-order change in the utility value of a given quantity of savings.

How and why savings at  $t - 1$  should respond to a re-profiling of insurance in  $t$  is an issue to which we return later. Ultimately the sign and magnitude of  $\epsilon_{t-1,t}^s(s_t)$  will capture important links between tax cuts, insurance and savings – which can provide a force for additional insurance when types are persistent.

**The compensated effect of transfers on lagged savings:** Just as the insurance effects of a marginal savings tax cut in  $t$  may change savings in  $t - 1$ , so too could the insurance effects of a change in the lump-sum component of taxes. Suppose  $\underline{s}$  is the lowest realised savings level in period  $t$  after some history, and consider a marginal reduction in  $T_t(\underline{s})$ , holding constant the profile of marginal tax rates at higher savings. This tax cut will shift consumption possibilities in  $t$ , by an equal amount for all types. The marginal effect on *utility* for type  $\alpha_t$  will be  $\alpha_t u'(c_t(\alpha_t))$ , and in general this will vary in  $\alpha_t$ . This implies that the change in the lump-sum component will induce a marginal reprofiling of utility across type draws. Once again, we consider the compensated effect of this reprofiling, given a uniform adjustment to utility across period- $t$  states that leaves  $\omega_t(\alpha^{t-1})$  constant.

Since a change in the lump-sum component of taxes is equivalent to a change in  $M_t$ , we denote the compensated effect of higher period- $t$  income on  $t - 1$  savings by:

$$\frac{ds_{t-1}}{dM_t} \Big|_{\text{comp}}$$

---

<sup>28</sup>The formal definition of  $\epsilon_{t-1,t}^s(s_t)$  is complicated by the fact that marginal tax changes at a single point in period  $t$  have ‘small’ (zero measure) effects on the incentives to save in  $t - 1$ . It is given in the appendix.

### 7.3 Equivalence results

Making use of these definitions, the following Lemma provides the ingredients to link from expressions (27) and (28) to a sufficient statistics representation:

**Lemma 2.** *The following relationships hold:*

1. [Contemporaneous savings elasticity]

$$T'_t(s_t) s_t \epsilon_t^s = \frac{\lambda_{t+1}^\Delta (\alpha_t(c_t))}{\eta_t} (\alpha_t(c_t))^2 u'(c_t) \left( \frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \quad (30)$$

2. [Contemporaneous income effect]

$$T'_t(s_t) \frac{ds_t}{dM_t} = - \frac{\lambda_{t+1}^\Delta (\alpha_t(c_t))}{\eta_t} (\alpha_t(c_t))^2 u''(c_t) \left( \frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \quad (31)$$

3. [Intertemporal effect of tax cut]

$$\begin{aligned} RT'_{t-1}(s_{t-1}) s_{t-1} \epsilon_{t-1,t}^s(s'_t) &= - \rho(\alpha_t(c'_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} (\alpha_t(c'_t))^2 u'(c'_t) \left( \frac{d\alpha_t(c'_t)}{dc_t} \right)^{-1} \frac{\pi^c(c'_t | \alpha^{t-1})}{\Pi^c(c'_t | \alpha^{t-1})} \\ &+ \frac{1}{\Pi^c(c'_t | \alpha^{t-1})} \int_{\underline{c}}^{c'_t} \left\{ \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \alpha_t(c_t) (u'(c_t)) \right. \\ &\times \left. \left[ 1 + \frac{c_t u''(c_t)}{u'(c_t)} \left( \frac{c_t}{\alpha_t(c_t)} \frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \right] \pi^c(c_t | \alpha^{t-1}) \right\} dc_t \end{aligned} \quad (32)$$

4. [Intertemporal income effect]

$$\begin{aligned} RT'_{t-1}(s_{t-1}) \frac{ds_{t-1}}{dM_t} \Big|_{comp} &= \int_{\underline{c}}^{\bar{c}} \left\{ \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \alpha_t(c_t) (u'(c_t)) \right. \\ &\times \left. \left[ 1 + \frac{c_t u''(c_t)}{u'(c_t)} \left( \frac{c_t}{\alpha_t(c_t)} \frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \right] \pi^c(c_t | \alpha^{t-1}) \right\} dc_t \end{aligned} \quad (33)$$

The proof of these relationships, in the Appendix, is based on elementary manipulations. It exploits two important features of the problem. The first is the duality between welfare maximisation and cost minimisation when designing policy. This enables the multipliers  $\lambda_t^\Delta$  and  $\lambda_t$  to be linked to the marginal

cost for the policymaker of allowing additional savings at the margin. This, in turn, allows an expression for the marginal tax revenue that is raised per unit of savings.

The second feature that we exploit is the separability of consumption utility over time. This limits the dependence of contemporaneous choice on decisions in other periods, and guarantees the existence of relatively simple cross-relationships between different behavioural statistics.

## 7.4 Welfare weights

In keeping with the static literature, we make use of ‘social welfare weights’ to capture the marginal value to the policymaker of providing an extra unit of income to each type, expressed in units of current resources. In our setting these weights are defined for history  $\alpha^t$  and current consumption level  $c_t$  by:<sup>29</sup>

$$g_t(\alpha^t) := \alpha_t u'(c_t(\alpha^t)) \frac{1 + \lambda_t(\alpha^{t-1})}{\eta_t} \quad (34)$$

That is, the subjective marginal utility of consumption,  $\alpha_t u'(c_t)$ , multiplied by a term  $(1 + \lambda_t(\alpha^{t-1})) > 0$  that captures the contemporaneous value to the policymaker of providing resources to the cross-section of types with history  $\alpha^{t-1}$ , and divided by  $\eta_t$  – the shadow utility value of period- $t$  resources.

Economically, the most interesting component of the welfare weight is the object  $\lambda_t(\alpha^{t-1})$ . This updates period-by period in response to the shocks that agents receive, with mean-zero innovations:  $\mathbb{E}[\lambda_{t+1}|\alpha^{t-1}] = \lambda_t$ . In the decentralised allocation, the updating process will capture the wealth that agents accumulate along each history branch. Higher values for  $\lambda_t$  correspond to higher past savings, and therefore a higher implicit weight in period- $t$  welfare calculations. Cross-sectionally, this is equivalent to placing a higher Pareto weight on those who have accumulated a large amount of wealth, relative to those who have not.<sup>30</sup>

The link between Pareto weights and wealth in a decentralised market economy has been understood at least since Negishi (1960). The interesting feature of the present context is the fact that the evolution of the wealth distribution over time gives rise to changes in the effective welfare objective of the policymaker. A policymaker in the initial period may seek a radical utilitarian allocation, un-

<sup>29</sup>If the mapping between  $\alpha_t$  and  $s_t$  is bijective, we may sometimes write  $g_t(s_t)$  in place of  $g_t(\alpha^t)$ , leaving history implicit.

<sup>30</sup>Formally, the proof of Theorem 2 establishes that  $\lambda_t$  is decreasing in  $\alpha_{t-1}$ , for agents with a common history  $\alpha^{t-2}$ . Since higher  $\alpha_{t-1}$  corresponds to higher consumption in  $t-1$ , there is a monotone link from savings to Pareto weights.

constrained by any initial profile of asset ownership. But as time progresses, respect for the evolving pattern of wealth is a necessary counterpart to respect for past incentive constraints. An optimal plan remains cross-sectionally utilitarian, for any subset of individuals who share a common history. Across subgroups, however, substantial differentiation in treatment is likely to emerge. The time inconsistency here is evident, and provides a challenge to the plausibility of the commitment assumption.<sup>31</sup>

## 8 Sufficient statistics characterisation

### 8.1 Characterisation

Corollary 2 and Lemma 2 together deliver our main ‘sufficient statistics’ characterisation result:

**Theorem 1.** *If a strictly normal allocation is optimal in the relaxed problem, with  $c_t(\alpha_t)$  continuous for any given  $\alpha^{t-1}$ , then at  $t = 0$ , for almost all realised savings levels  $s'_0$ :*

$$\mathbb{E} \left[ 1 - T'_0(s_0) \frac{ds_0}{dM_0} - g_0(s_0) \middle| s_0 \geq s'_0 \right] = T'_0(s'_0) \varepsilon_0^s \frac{s'_0 \pi^s(s'_0)}{1 - \Pi^s(s'_0)} \quad (35)$$

and:

$$\mathbb{E} [g_0(s_0)] = \mathbb{E} \left[ 1 - T'_0(s_0) \frac{ds_0}{dM_0} \right] \quad (36)$$

Similarly, for  $t > 0$ , any given  $\alpha^{t-1}$ , and almost all realised  $s'_t$ :

$$\mathbb{E}_{t-1} \left[ 1 - T'_t(s_t) \frac{ds_t}{dM_t} - g_t(s_t) \middle| s_t \geq s'_t \right] = T'_t(s'_t) \varepsilon_t^s \frac{s'_t \pi^s(s'_t | \alpha^{t-1})}{1 - \Pi^s(s'_t | \alpha^{t-1})} + RT'_{t-1}(s_{t-1}) s_{t-1} \varepsilon_{t-1,t}^s(s'_t) \quad (37)$$

and:

$$\mathbb{E}_{t-1} [g_t(s_t)] = 1 - \mathbb{E}_{t-1} \left[ T'_t(s_t) \frac{ds_t}{dM_t} \right] - RT'_{t-1}(s_{t-1}) \frac{ds_{t-1}}{dM_t} \Big|_{comp} \quad (38)$$

*Proof.* Follows from direct substitution of the expressions in Lemma 2 into the conditions in Corollary 2, applying the definition of the social welfare weights.  $\square$

<sup>31</sup>Brendon and Ellison (2018) propose an alternative solution concept under commitment that delivers stationarity in the Pareto weights.

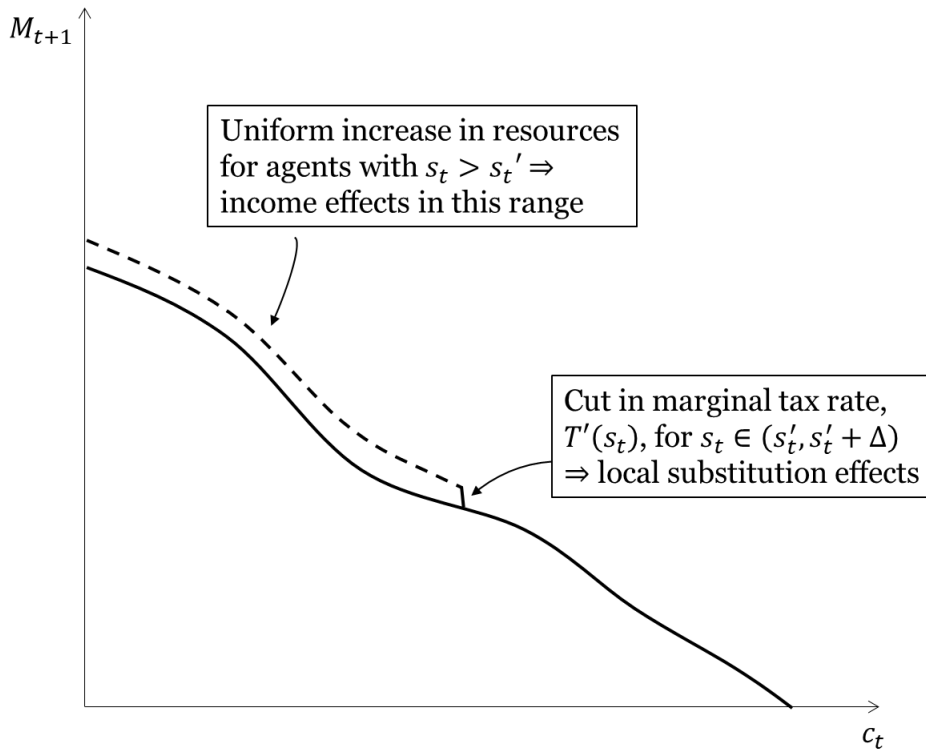


Figure 1: Effects on the budget constraint of a marginal savings tax cut at  $s'_t$

## 8.2 Intuition

As previewed, equations (35) to (38) can be understood intuitively by reference to simple changes in the intertemporal budget constraint that links consumption in one period to income in the next. For (35) and (37), the relevant exercise is a cut in the marginal tax rate at some particular savings level. As Figure 1 illustrates, the result is a rightwards shift in the budget constraint for all savings levels above the threshold. For conditions (36) and (38), the relevant exercise is a rightwards shift in the entire budget constraint, as the lump-sum component of the tax schedule is made more generous.

Condition (35) assesses the effects of the tax cut in Figure 1, when applied in the initial time period. Heuristically, the effects can be divided into those *above*  $s'_0$ , and those *at*  $s'_0$ . For those above  $s'_0$ , the tax cut serves to shift out the within-period budget constraint by a uniform amount, and the left-hand side of (35) accounts for this from the policymaker's perspective. There are three components: (1) the direct cost of the transfer, normalised to 1 per agent by construction, minus (2) the additional tax revenue that is received on whatever fraction of the additional income is saved,  $T'_0(s_0) \frac{ds_0}{dM_0}$ , minus (3) the social welfare



value of providing an additional consumption unit,  $g_0(s_0)$ . Taken together, these objects give the net fiscal cost of the transfer that high-saving agents receive.

The right-hand side of (35) relates to agents locating at  $s'_0$ . A higher post-tax rate of return – i.e., a lower savings tax rate – will induce these agents to substitute towards savings in proportion to the savings elasticity. So long as the marginal savings tax rate is positive, this is desirable to the policymaker: it generates higher tax revenue. This is captured by the object  $T'_0(s'_0) s'_0 \epsilon_0^s$ , interacted with the density of savers affected.

Condition (37) is the equivalent to (35) for  $t > 0$ . Relative to the period-0 version, it has an extra term that allows for the impact that changes to tax schedules in  $t$  have on savings in  $t - 1$ . This is the object  $RT'_{t-1}(s_{t-1}) s_{t-1} \epsilon_{t-1,t}(s'_t)$ , with the real interest rate  $R$  reflecting the value of resources raised in  $t - 1$  relative to  $t$ . Clearly this term depends critically on the sign and magnitude of the cross-elasticity  $\epsilon_{t-1,t}(s'_t)$ : do tax cuts at  $s'_t$  incentivise or deter savings at  $t - 1$ , and by how much? This will be discussed in detail in Section 10, and we defer further comment for now.

Crucially, according to Theorem 1 there is no need to keep track of arbitrary cross-elasticity statistics when working out optimal taxes – i.e., the response of savings in period  $s$  to tax changes in period  $s + r$ , for  $r > 1$ . This is an extremely helpful simplification, since the set of cross-elasticities that could potentially matter is infinite. It speaks positively for the practical applicability of dynamic Mirrleesian tax analysis – a feature that has not, to date, been considered its greatest strength.

Conditions (36) and (38) describe the consequences of shifting the entire budget constraint, rather than just an upper segment. They are essentially variants of (35) and (37) respectively, absent contemporaneous substitution effects – and when the relevant intertemporal behavioural effect is  $\left. \frac{ds_{t-1}}{dM_t} \right|_{\text{comp}}$  rather than  $\epsilon_{t-1,t}(s'_t)$ . The expressions can be read as optimal solutions for the lump-sum component of the social insurance system, discussed in more detail in the next section.

## 9 Properties of optimal taxes

### 9.1 Positive marginal rates at interior points

The characterisation can be used to analyse the qualitative properties of an optimal savings tax schedule in the decentralised allocation. The most general result is the following:

**Theorem 2.** *Suppose the optimal allocation is strictly normal. Then for all time periods and shock histories, marginal savings taxes are strictly positive at all interior points in the type distribution.*

This provides a very direct qualitative description of the optimal social insurance scheme. Recall from Section 6 that the average value of  $T_t(s_t)$  is constructed to be zero, given the history  $s^{t-1}$ . Since the marginal rate is positive, an optimal social insurance scheme must therefore provide a positive transfer (negative  $T$ ) at the lowest savings level, which is then taxed away as savings increase.

Intuitively, this is consistent with the basic problem that the social insurance scheme seeks to address: how to distribute income to those with a high consumption need in period  $t$ , given that need is unobservable? The solution is to exploit the relative preference of high-need consumers for current rather than future consumption. A universal transfer is made available to all, financed by those who choose to save. The act of saving signals a relatively low consumption need, and thus attracts a high net fiscal contribution. Optimal policy faces the familiar trade-off between redistributing towards those with a higher social welfare weight, as revealed by their low savings, and the distortion of savings decisions that is implied by this.

### 9.2 Limiting outcomes

The general finding of strictly positive marginal savings tax rates need not extend to endpoints of the type distribution, where limiting rates may instead reach zero. A critical role is played by the limiting properties of the distribution and density of savings. Zero tax results generally follow if the density remains positive at endpoints, or converges to zero relatively slowly, since this implies that the local efficiency costs of taxation become large relative to any redistributive benefit. These properties may be inherited from primitive assumptions on the type distribution – for instance, it is straightforward to

show that zero marginal taxes are optimal at endpoints if the conditional type density  $\pi(\cdot|\cdot)$  is everywhere bounded above zero.<sup>32</sup>

More generally, for optimal top marginal tax rates the Pareto statistic for savings provides the relevant weighting between efficiency redistributive effects. We denote this  $a_t(s_t|\alpha^{t-1})$ :

$$a_t(s_t|\alpha^{t-1}) := \frac{s_t \pi^s(s_t|\alpha^{t-1})}{1 - \Pi^s(s_t|\alpha^{t-1})}$$

If  $\pi^s(\bar{s}|\alpha^{t-1}) > 0$ , where  $\bar{s}$  is the conditional upper bound for savings, then  $\lim_{s_t \rightarrow \bar{s}} [a_t(s_t|\alpha^{t-1})] = \infty$ . This is true when  $\pi(\underline{\alpha}|\alpha_{t-1}) > 0$ . More generally, however, it is quite possible for  $\lim_{s_t \rightarrow \bar{s}} [a_t(s_t|\alpha^{t-1})] < \infty$ , and the primitive assumptions that we have placed on the problem admit either of these outcomes.

As a direct corollary of Theorem 1, we can write an expression for the optimal marginal tax rate at the top of the savings distribution:

**Corollary 3.** *Given  $\alpha^{t-1}$ , the optimal marginal tax rate at the upper limit of the savings distribution,  $\bar{s}$ , satisfies:*

$$T'_t(\bar{s}) = \frac{1 - g_t(\bar{s}) - RT'_{t-1}(s_{t-1}) \frac{s_{t-1}}{\bar{s}} \epsilon_{t-1,t}^s(\bar{s}) a_t(\bar{s}|\alpha^{t-1})}{\left. \frac{ds_t}{dM_t} \right|_{\bar{s}} + \epsilon_t^s a_t(\bar{s}|\alpha^{t-1})} \quad (39)$$

In keeping with the static literature on optimal tax design, it is possible to use this equation to obtain approximate figures for upper marginal tax rates, given a shock history. Section 11 provides an indicative exercise to this end.

### 9.3 Optimal transfers

Conditions (36) and (38) describe the optimal determination of the lump-sum component to the tax system after each history. In the initial period, this is a straightforward trade-off between the welfare benefits of transferring an extra unit of income across all agents, captured by  $\mathbb{E}[g_0(s_0)]$ , and the net cost of doing so,  $\mathbb{E}\left[1 - T'(s_0) \frac{ds_0}{dM_0}\right]$ . So long as contemporaneous income effects on savings are positive, it is optimal to increase transfers even beyond the level where the average welfare weight is unity – the usual benchmark in the labour supply literature with quasilinear preferences – because the net cost of

<sup>32</sup>This follows from equation (65) in the appendix, and the fact that  $T'_t(s_t(\alpha_t)) = 0$  when  $\lambda_{t+1}^\Delta(\alpha_t) = 0$ , shown in the proof of Theorem 2.

the transfer is mitigated by tax revenue on the additional savings it induces.

Outcomes in periods after 0 are additionally influenced by the complementarity of insurance and past savings. A compensated increase in the lump-sum component of the tax system in period  $t$  raises the insurance value of savings, since the marginal utility of this additional income is increasing in  $\alpha_t$ . With type persistence, this raises savings at the margin in  $t - 1$  – for reasons discussed in Section 10 below. This implies it is optimal to set transfers above the value that equates the average welfare weight in period  $t$  with the within-period net cost of the transfer:

$$\mathbb{E}_{t-1} [g_t (s_t)] < \mathbb{E}_{t-1} \left[ 1 - T' (s_t) \frac{ds_t}{dM_t} \right] \quad (40)$$

The general message is that type persistence motivates a more generous insurance scheme, because insurance acts as a complement to past savings. The reasons for this are explored in more detail in the next Section.

## 10 Limited intertemporal elasticities: Atkinson-Stiglitz revisited

The most significant theoretical insight from Theorem 1 is the limited extent to which the conventional Saez (2001) condition needs to change when moving from a simple two-good screening problem to an infinite-horizon, persistent-type setting. Indeed, the only statistic preventing the within-period characterisation from being isomorphic to a textbook Saez formula is the elasticity of lagged savings,  $\epsilon_{t-1,t}^s (s'_t)$ . The purpose of this Section is to discuss its role, and relate it to more familiar intuition.

So long as marginal savings taxes are positive in period  $t - 1$ , the marginal social value of additional savings in that period necessarily exceeds the marginal private value. This means that there are social benefits to inducing more savings, and the policymaker should be willing to pay some costs at the margin to achieve this. In particular, following the well-known logic of Atkinson and Stiglitz (1976), the tax system should favour any goods that are complements to the main behaviour being taxed – in the Atkinson-Stiglitz setting labour supply; here saving.

When types are persistent, there is one relevant complement to savings in  $t - 1$ : the level of insurance in period  $t$ . To see why, suppose individuals are ordered by their savings levels in  $t - 1$ . Given the link

between type and behaviour, those with lower  $s_{t-1}$  necessarily have higher values for  $\alpha_{t-1}$ . Type persistence means that those with higher  $\alpha_{t-1}$  place relatively more weight on the likelihood that they will find themselves with a high consumption need in period  $t$ . This means that they have a relative preference for greater cross-sectional insurance at  $t$ , by comparison with those whose savings are marginally higher. A policy that increases the level of insurance that is provided in  $t$  for each level of  $t-1$  savings, holding constant the expected utility of those who keep their savings unchanged, will raise the marginal attractiveness of savings in  $t-1$  for all types.

Thus the presence of  $\epsilon_{t-1,t}^s(s'_t)$  in equation (37) is precisely to capture the effect of the tax cut on insurance in  $t$ , and, through this, on savings in  $t-1$ . It represents an additional distortion to outcomes from  $t$  onwards, relative to an optimal plan from the perspective of  $t$  alone. This distortion is justified by the consequent reduction in under-saving prior to  $t$ . It implies a second, more prosaic source of time inconsistency in the setting, distinct from the more fundamental societal challenge of a widening wealth distribution. A policymaker re-optimising in  $t$  would have no incentive to consider the effect of their choices on savings in  $t-1$ , which would by now already be determined.

Atkinson and Stiglitz (1976) showed that when consumption goods were independent of labour supply, there were no gains to differential consumption taxation. The counterpart to this result in our setting is provided by the case of iid types. There, the level of  $s_{t-1}$  is independent of preferences across period- $t$  outcomes. Any change to the profile of utilities at  $t$  will be viewed identically by all types in  $t-1$ . *Ex post* insurance is neither a complement nor substitute to savings. This means that it is not optimal for the distribution of outcomes in  $t$  to be influenced by concerns relating to  $t-1$  or earlier:  $\epsilon_{t-1,t}^s(s_t) \equiv 0$ , and only contemporaneous elasticities matter.

More generally, this line of reasoning also indicates that the Markov property of shocks is crucial to ensuring that just two elasticities feature in (37). Markovian shocks imply that the preferences of two distinct  $\alpha_{t-1}$  types across alternative allocations from  $t+1$  on are identical, conditional on drawing a particular  $\alpha_t$ . Compensated distortions to  $t+1$  allocations, that hold constant expected lifetime utility for each period- $t$  type draw, can do nothing to induce additional saving in  $t-1$  or earlier.

## 10.1 Intertemporal elasticities: a source of progressivity

The main lesson from the previous subsection is that it may be desirable to provide more additional insurance in  $t$ , beyond what is contemporaneously optimal, because this helps to increase savings in  $t - 1$ . Additional insurance is often associated with additional progressivity in the tax system, and in this subsection we confirm this link. The intertemporal elasticity  $\epsilon_{t-1,t}^s(s'_t)$  exhibits systematic cross-sectional variation, in a manner that contributes to greater progressivity in the marginal tax rate in  $t$ , and thus a greater degree of insurance *ex-post*. Formally, we have the following result:

**Proposition 5.** *The statistic  $s_{t-1}\epsilon_{t-1,t}^s(s'_t)$  is monotonically decreasing in  $s'_t$ . It is positive for sufficiently low  $s'_t$ , and negative for sufficiently high  $s'_t$ .*

The implications of this Proposition can be seen by comparing policies that satisfy condition (37) with those that neglect intertemporal cross-elasticities – as would be optimal for a policymaker re-optimising in  $t$ . At our optimum, we have:

$$\mathbb{E}_{t-1} \left[ 1 - T'_t(s_t) \frac{ds_t}{dM_t} - g_t(s_t) \middle| s_t \geq s'_t \right] > T'_t(s'_t) s'_t e_t^s a_t (s'_t | \alpha^{t-1}) \quad (41)$$

for all  $s'_t$  below a threshold and

$$\mathbb{E}_{t-1} \left[ 1 - T'_t(s_t) \frac{ds_t}{dM_t} - g_t(s_t) \middle| s_t \geq s'_t \right] < T'_t(s'_t) s'_t e_t^s a_t (s'_t | \alpha^{t-1}) \quad (42)$$

for all  $s'_t$  above the same threshold. By contrast, the re-optimising policymaker would set the two sides of these expressions equal at all  $s'_t$ . The left-hand side represents the marginal redistributive cost of cutting taxes, and the right-hand side the marginal revenue gain due to substitution effects. At least locally, therefore, the re-optimising policymaker would prefer to raise marginal tax rates at low  $s'_t$ , and cut them at high  $s'_t$ . The optimum has ‘too much’ progressivity, viewed *ex-post*.

*Why* does the elasticity change sign in the manner described in Proposition 5? The explanation derives from the varying importance of two contrasting effects on insurance as the value of  $s'_t$  is changed. A tax cut at  $s'_t$  has two consequences for insurance. The first is to redistribute resources to high savers, relative to the population as a whole. Since high savers have a low marginal consumption utility, this

worsens the insurance value of savings, and pushes the value of  $\epsilon_{t-1,t}^s(s'_t)$  downwards. The second effect is to provide a uniform increase in income within the sub-group of savers above the threshold  $s'_t$ . Uniform income provision implies non-uniform effects on *utility* within this sub-group. Lower savers have higher marginal utility, and so benefit more. This improves the insurance profile of savings, raising the value of  $\epsilon_{t-1,t}^s(s'_t)$ .

When taxes are cut at a low level of savings, the within-group insurance gains are large relative to cross-group effects, as the within-group dispersion of marginal utilities is large. As the threshold  $s'_t$  increases, cross-group effects come to be relatively more significant, worsening the insurance properties of the change and ensuring a lower value for  $\epsilon_{t-1,t}^s(s'_t)$ .

## 11 An indicative quantification

To give our results clearer focus, we provide an indicative quantification of the optimal top marginal savings tax, based on the formula (39). The main aim is to highlight the important sources of uncertainty, as much as to provide a precise figure.

The exercise is more straightforward if the term in the lagged marginal tax rate is assumed to be small in magnitude relative to the remaining  $1 - g_t(\bar{s})$  in the denominator – an assumption that is exactly satisfied if  $t = 0$ , or type draws are independent over time. Given Proposition 5, this will tend to bias the resulting figure downwards. It implies that the optimality formula simplifies to:

$$T'_t(\bar{s}) = \frac{1 - g_t(\bar{s})}{\left. \frac{ds_t}{dM_t} \right|_{\bar{s}} + \epsilon_t^s a_t(\bar{s}|\alpha^{t-1})}$$

If consumption approaches zero at the upper limit for savings, an empirical value of  $\left. \frac{ds_t}{dM_t} \right|_{\bar{s}}$  equal to 1 is an obvious approximation. The value of  $\epsilon_t^s a_t(\bar{s}|\alpha^{t-1})$  is equal to the Pareto parameter for the lower tail of the conditional *consumption* distribution, multiplied by the compensated intertemporal elasticity of consumption. Toda and Walsh (2015) suggest a value of approximately 4 for the lower consumption Pareto parameter, based on cross-sectional US data. Unlike  $a_t$  in our formulation, this is based on a sample of the entire population, rather than a conditional cross section with common wealth, and is likely to be a significant under-estimate of the relevant statistic.

The compensated intertemporal elasticity of consumption is strictly less than the Frisch, for which a value of around 0.5 is standard.<sup>33</sup> Since the difference between the two becomes small as the share of current consumption in marginal expenditure becomes small, this value is a reasonable starting point.

Taken together, these assumptions would imply a value for the top marginal tax rate equal to  $\frac{1-g_t(\bar{s})}{3}$ , where  $g_t(\bar{s})$  is the value of the social welfare weight at the highest savings level. The social welfare weight is a harder object to quantify. Unlike static Mirrleesian environments, it does not make sense to assume that it approaches zero for the least-favoured types. To see why, recall that  $g_t(s_t)$  is directly proportional to the marginal utility of consumption  $\alpha_t u'(c_t)$ , which is optimally set equal to the marginal value of savings. Even individuals whose draw for  $\alpha_t$  is arbitrarily close to zero will have returns to saving that are bounded above zero, and will optimally reduce their period- $t$  consumption until a high value for  $u'(c_t)$  offsets low  $\alpha_t$ .

Equation (38) shows that the average value for  $g_t$  is 1 minus the value of tax recovered through income effects at the margin, both in  $t$  and  $t - 1$ . Assuming a generous ceiling for total income effects on tax revenue of 0.1, the top marginal rate goes to  $\frac{0.1+(\bar{g}_t-g_t(\bar{s}))}{3}$  where  $\bar{g}_t$  is the cross-sectional average value of  $g_t$ . An individual with type draw  $\underline{\alpha}$  in  $t$  is assumed to have minimal consumption needs in period  $t$ , and can devote all resources to saving. Even if this raises lifetime consumption by 5 per cent relative to the mean, with time-separable utility and an EIS of 0.5 it can at best imply a marginal utility reduction of around 10 per cent through the effect of higher consumption alone. Persistence in  $\alpha$  provides a direct channel for lower period- $t$  types to value the future less, but since insurance is incomplete this effect will be limited: the marginal value of saving will be dominated by the risk of high type draws.

The main lesson is that the top marginal tax rate on savings could perhaps reach 10 per cent, but only with some very favourable assumptions, particularly on the value of the Pareto parameter. If – as conjectured by Toda and Walsh (2015) – Pareto tails for the overall consumption distribution only emerge as a consequence of demographic evolution, and are absent within cohorts, then  $a_t(\bar{s}) = \infty$  would be a more reasonable assumption, and we have zero distortion at the top.

---

<sup>33</sup>See, for instance, Attanasio and Weber (2010).



## 12 Conclusion

This paper contains two main messages. The first, from a policy perspective, is that a widely-used model of social insurance under imperfect information implies a novel justification for taxing savings. Faced with a population whose consumption needs are heterogeneous and unobserved, it is best for the policymaker to provide a uniform lump-sum resource transfer to all agents period-by-period, and to tax the savings of those whose very decision to save reveals that their need is low.

The second main message of the paper is of relevance to the wider dynamic tax literature. It is that – contrary to widespread perceptions – the ‘mechanism design’ approach to dynamic optimal taxation can give rise to simple, intuitive ‘sufficient statistics’ representations of optimal taxes. Indeed, it is precisely the assumptions of the mechanism design approach – additively-separable utility over time, and Markovian shock processes – that appear to simplify behavioural responses in a way that keeps them tractable. In a multi-period world, tax design must inevitably make some simplifying assumptions, to avoid being overwhelmed by the multitude of possible cross-period behavioural responses. One option, pursued in the literature already, is to focus exclusively on steady-state outcomes. Though defensible, this is a significant departure from conventional approaches, both positive and normative. This paper suggests that mechanism design offers a theory-guided alternative route.

## References

- [1] Aiyagari, S.R. (1995), ‘Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting’, *Journal of Political Economy*, 103(6): 1158–75.
- [2] Albanesi, S., and C. Sleet (2006), ‘Dynamic Optimal Taxation with Private Information’, *Review of Economic Studies*, 73(1): 1–30.
- [3] Atkeson, A., and R.E. Lucas (1992), ‘On Efficient Distribution With Private Information’, *Review of Economic Studies*, 59(3): 427–453.
- [4] Atkinson, A.B., and J.E. Stiglitz (1976) , ‘The Design of Tax Structure: Direct versus Indirect Taxation’, *Journal of Public Economics*, 6: 55–75.

- [5] Attanasio, O.P., and G. Weber (2010), 'Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy' *Journal of Economic Literature*, 48 (3): 693–751.
- [6] Bloedel, A.W., R.V. Krishna and O. Leukhina (2020), 'Insurance and Inequality with Persistent Private Information', Working Paper, Stanford University and Federal Reserve Bank of St. Louis.
- [7] Bloedel, A.W., R.V. Krishna and B. Strulovici (2020), 'Persistent Private Information Revisited', Working Paper.
- [8] Brendon, C.F., and M. Ellison (2018), 'Time-Consistently Undominated Policies', Working Paper.
- [9] Chamley, C. (1986), 'Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives', *Econometrica*, 54(3): 607–22.
- [10] Chari, V.V., J.P. Nicolini and P. Teles (2020), 'Optimal Capital Taxation Revisited', *Journal of Monetary Economics*, (116): 147–165.
- [11] Conesa, J.C., S. Kitao and D. Krueger (2009), 'Taxing Capital? Not a Bad Idea After All!', *American Economic Review*, 99(1): 25– 48.
- [12] Diamond, P. (1998), 'Optimal Taxation: An Example with a U-shaped Pattern of Optimal Marginal Tax Rates', *American Economic Review*, 88: 83–95.
- [13] Diamond, P., and J.A. Mirrlees (1978), 'A Model of Social Insurance with Variable Retirement', *Journal of Public Economics*, 10: 295-336.
- [14] Diamond, P., and E. Saez (2011), 'The Case for a Progressive Tax: From Basic Research to Policy Recommendations', *Journal of Economic Perspectives*, 25(4): 165–90.
- [15] Erosa, A., and M. Gervais (2002), 'Optimal Taxation in Life-Cycle Economies', *Journal of Economic Theory*, 105(2): 338–69.
- [16] Farhi, E., and I. Werning (2007), 'Inequality and Social Discounting', *Journal of Political Economy*, 115(3): 365–402.

- [17] Farhi, E., and I. Werning (2013), ‘Insurance and Taxation over the Life Cycle’, *Review of Economic Studies*, 810(2): 596–635.
- [18] Ferey, A., B. Lockwood and D. Taubinsky (2021), ‘Sufficient Statistics for Nonlinear Tax Systems with Preference Heterogeneity’, NBER Working Paper 29582
- [19] Golosov, M., N.R. Kocherlakota and A. Tsyvinski (2003), ‘Optimal Indirect and Capital Taxation’, *Review of Economic Studies*, 70(3): 569–87.
- [20] Golosov, M., M. Troshkin and A. Tsyvinski (2016), ‘Redistribution and Social Insurance’, *American Economic Review*, 106(2): 359–386.
- [21] Golosov, M., A. Tsyvinski and I. Werning (2006), ‘New Dynamic Public Finance: A User’s Guide’, in *NBER Macroeconomics Annual 2006*, Vol. 21, edited by D. Acemoglu, K. Rogoff, and M. Woodford, 317–63. Cambridge, MA: MIT Press.
- [22] Golosov, M., A. Tsyvinski and N. Werquin (2014), ‘A Variational Approach to the Analysis of Tax Systems’, NBER Working Paper 20780.
- [23] Hellwig, C. (2021), ‘Static and Dynamic Mirrleesian Taxation with Non-separable Preferences: A Unified Approach’, TSE Working Paper 1224
- [24] Hellwig, C., and N. Werquin (2022), ‘A Fair Day’s Pay for a Fair Day’s Work: Optimal Tax Design as Redistributive Arbitrage’, CEPR Working Paper
- [25] İmrohoroğlu, S. (1998), ‘A Quantitative Analysis of Capital Income Taxation’, *International Economic Review*, 39(2): 307–28.
- [26] Judd K.L. (1985), ‘Redistributive Taxation in a Simple Perfect Foresight Model’, *Journal of Public Economics*, 28(1): 59–83.
- [27] Kapička, M. (2013), ‘Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: a First-order Approach’, *Review of Economic Studies*, 80: 1027–54
- [28] Kockerlakota, N.R. (2005), ‘Zero Expected Wealth Taxes: a Mirrlees Approach to Dynamic Optimal Taxation’, *Econometrica*, 73: 1587–621

- [29] Kocherlakota, N.R. (2010), *The New Dynamic Public Finance*, Princeton University Press.
- [30] Makris, M., and A. Pavan (2020), 'Wedge Dynamics with Evolving Private Information', Working Paper.
- [31] Marcet, A., and R. Marimon (2019), 'Recursive Contracts', *Econometrica*, 87(5): 1589–1631.
- [32] Milgrom, P., and I. Segal (2002), 'Envelope Theorems for Arbitrary Choice Sets', *Econometrica*, 70(2): 583–601.
- [33] Mirrlees, J.A. (1971), 'An Exploration in the Theory of Optimum Income Taxation', *Review of Economic Studies*, 38: 175–208.
- [34] Negishi, T. (1960), 'Welfare Economics and Existence of an Equilibrium for a Competitive Economy', *Metroeconomica*, 12: 92–97.
- [35] Pavan, A. (2017), 'Dynamic Mechanism Design: Robustness and Endogenous Types', in *Advances in Economics and Econometrics: Proceeding of the 2015 World Congress of the Econometric Society*.
- [36] Pavan, A., I. Segal and J. Toikka (2014), 'Dynamic Mechanism Design: A Myersonian Approach', *Econometrica*, 82(2): 601–653.
- [37] Piketty, T., and E. Saez (2013a), 'A Theory of Optimal Inheritance Taxation', *Econometrica*, 81(5): 1851–1886
- [38] Piketty, T., and E. Saez (2013b), 'Optimal Labour Income Taxation', Ch 7 in A.J. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.), *Handbook of Public Economics*, Vol 5: 391–474.
- [39] Rogerson, W.P. (1985), 'The First-order Approach to Principal-Agent Problems', *Econometrica*, 53(6), 1357–67.
- [40] Saez, E. (2001), 'Using Elasticities to Derive Optimal Income Tax Rates', *Review of Economic Studies*, 68: 205–229.
- [41] Saez, E., and S. Stantcheva (2018), 'A Simpler Theory of Optimal Capital Taxation', *Journal of Public Economics*, 162: 120–142.

- [42] Sleet, C., and S. Yeltekin (2006), ‘Credibility and Endogenous Societal Discounting’, *Review of Economic Dynamics*, 9: 410–437.
- [43] Stantcheva, S. (2017), ‘Optimal Taxation and Human Capital Policies over the Life Cycle’, *Journal of Political Economy*, 125: 1931–90.
- [44] Stantcheva, S. (2020), ‘Dynamic Taxation’, *Annual Review of Economics*, 12: 801–31
- [45] Straub, L., and I. Werning (2020), ‘Positive Long-Run Capital Taxation: Chamley-Judd Revisited’, *American Economic Review*, 110(1): 86–119.
- [46] Thomas, J., and T. Worrall (1990), ‘Income Fluctuations and Asymmetric Information: An Example of the Repeated Principal-Agent Problem’, *Journal of Economic Theory*, 51: 367–390.
- [47] Toda, A.A., and K. Walsh (2015), ‘The Double Power Law in Consumption and Implications for Testing Euler Equations’, *Journal of Political Economy*, 123(5): 1177–1200.

## A Appendix

### A.1 Proof of Lemma 1

This result is an application of Theorem 2 in Milgrom and Segal (2002), plus elementary manipulations.

First, note that the utility of type  $\alpha'_t$  from arbitrary type report  $\alpha''_t$  in period  $t$  can be written in the form:

$$\alpha'_t u \left( c_t \left( \alpha^{t-1}, \alpha''_t \right) \right) + \beta \int_{\alpha_{t+1}} V \left( \alpha^{t-1}, \alpha''_t, \alpha_{t+1} \right) d\Pi \left( \alpha_{t+1} | \alpha'_t \right) \quad (43)$$

The boundedness of lifetime utility (constraint (10)) and the differentiability of the conditional density  $\pi \left( \alpha_{t+1} | \alpha'_t \right)$  in  $\alpha'_t$  (Assumption 2) together imply that this expression is differentiable in  $\alpha'_t$  for  $\alpha'_t \in (\underline{\alpha}, \bar{\alpha})$ . Its derivative with respect to  $\alpha'_t$  is:

$$u \left( c_t \left( \alpha^{t-1}, \alpha''_t \right) \right) + \beta \int_{\alpha_{t+1}} V \left( \alpha^{t-1}, \alpha''_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+1} | \alpha'_t \right)}{d\alpha'_t} d\alpha_{t+1} \quad (44)$$

Constraint (10) implies that  $u \left( c_t \left( \alpha^{t-1}, \alpha''_t \right) \right)$  and  $V \left( \alpha^{t-1}, \alpha''_t, \alpha_{t+1} \right)$  are bounded for all  $\alpha''_t$  and  $\alpha_{t+1}$ .  $\pi \left( \alpha_{t+1} | \alpha'_t \right)$  is continuously differentiable in  $\alpha'_t$  by assumption, and  $\alpha'_t$  inhabits a compact interval, so  $\frac{d\pi \left( \alpha_{t+1} | \alpha'_t \right)}{d\alpha'_t}$  is also bounded by construction. Taken together this implies that the object in (44) is bounded in absolute value, uniformly across type reports  $\alpha''_t$ . Since the allocation satisfies the general incentive compatibility restriction (12) under the condition

of the Lemma, the set of optimal choices for all types must, trivially, be nonempty. This establishes the conditions required for the Milgrom and Segal's Theorem 2 to be applied.

A direct application gives that  $V_t(\alpha^{t-1}; \alpha_t)$  is absolutely continuous in  $\alpha_t$  for all  $t$  and  $\alpha^{t-1}$ , with:

$$\begin{aligned} & \alpha'_t u \left( c_t \left( \alpha^{t-1}, \alpha'_t \right) \right) + \beta \int_{\alpha_{t+1}} V_{t+1} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) d\Pi \left( \alpha_{t+1} | \alpha'_t \right) \\ = & \underline{\alpha} u \left( c_t \left( \alpha^{t-1}, \underline{\alpha} \right) \right) + \beta \int_{\alpha_{t+1}} V_{t+1} \left( \alpha^{t-1}, \underline{\alpha}, \alpha_{t+1} \right) d\Pi \left( \alpha_{t+1} | \underline{\alpha} \right) \\ & + \int_{\underline{\alpha}}^{\alpha'_t} \frac{1}{\alpha_t} \left\{ \alpha_t u \left( c_t \left( \alpha^{t-1}, \alpha_t \right) \right) + \beta \alpha_t \int_{\alpha_{t+1}} V_{t+1} \left( \alpha^{t-1}, \alpha_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+1} | \alpha_t \right)}{d\alpha_t} d\alpha_{t+1} \right\} d\alpha_t \end{aligned} \quad (45)$$

To obtain the representation in the main text, we then make use of the following sub-Lemma:

**Lemma 3.** For all  $(\alpha_t, \alpha'_t) \in A^2$  and  $\alpha^{t-1} \in A^t$ :

$$\beta \alpha_t \int_{\alpha_{t+1}} V_{t+1} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+1} | \alpha_t \right)}{d\alpha_t} d\alpha_{t+1} = \mathbb{E}_t \left[ \sum_{s=t+1}^{\infty} \beta^{s-t} D_{t,s} \left( \alpha^s \right) \alpha_s u \left( c_s \left( \alpha^{t-1}, \alpha'_t, \dots, \alpha_s \right) \right) \right] \Big|_{\alpha_t}$$

*Proof.* Given the absolute continuity of the value function, the object:

$$\beta \alpha_t \int_{\alpha_{t+1}} V_{t+1} \left( \alpha^{t-1}, \alpha_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+1} | \alpha_t \right)}{d\alpha_t} d\alpha_{t+1}$$

can be integrated by parts, giving:

$$\begin{aligned} & \beta \alpha_t \int_{\alpha_{t+1}} V_{t+1} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+1} | \alpha_t \right)}{d\alpha_t} d\alpha_{t+1} \\ = & \beta \alpha_t \int_{\alpha_{t+1}} \left[ u_{t+1} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) + \beta \int_{\alpha_{t+2}} V_{t+2} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+2} | \alpha_{t+1} \right)}{d\alpha_{t+1}} d\alpha_{t+2} \right] \frac{d \left( 1 - \Pi \left( \alpha_{t+1} | \alpha_t \right) \right)}{d\alpha_t} d\alpha_{t+1} \\ = & \beta \int_{\alpha_{t+1}} \frac{\alpha_t \frac{d \left( 1 - \Pi \left( \alpha_{t+1} | \alpha_t \right) \right)}{d\alpha_t}}{\alpha_{t+1} \pi \left( \alpha_{t+1} | \alpha_t \right)} \left[ \alpha_{t+1} u_{t+1} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) + \beta \alpha_{t+1} \int_{\alpha_{t+2}} V_{t+2} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+2} | \alpha_{t+1} \right)}{d\alpha_{t+1}} d\alpha_{t+2} \right] d\Pi \left( \alpha_{t+1} | \alpha_t \right) \\ = & \int_{\alpha_{t+1}} \beta \rho \left( \alpha_{t+1} | \alpha_t \right) \left[ \alpha_{t+1} u_{t+1} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) + \beta \alpha_{t+1} \int_{\alpha_{t+2}} V_{t+2} \left( \alpha^{t-1}, \alpha'_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+2} | \alpha_{t+1} \right)}{d\alpha_{t+1}} d\alpha_{t+2} \right] d\Pi \left( \alpha_{t+1} | \alpha_t \right) \end{aligned}$$

where  $u_{t+1}(\alpha^{t-1}, \alpha'_t, \alpha_{t+1})$  is used as shorthand for  $u(c_{t+1}(\alpha^{t-1}, \alpha'_t, \alpha_{t+1}))$ . Applying this result recursively, together with the assumption that  $\rho(\alpha_{t+1} | \alpha_t) \in (0, 1)$  (Assumption 3), and the boundedness of value in  $t$ , the result follows.  $\square$

Using the definition of  $\omega_{t+1}^\Delta(\alpha^{t-1}, \alpha_t)$ , setting  $\alpha'_t = \alpha_t$  gives:

$$\alpha_t \int_{\alpha_{t+1}} V_{t+1} \left( \alpha^{t-1}, \alpha_t, \alpha_{t+1} \right) \frac{d\pi \left( \alpha_{t+1} | \alpha_t \right)}{d\alpha_t} d\alpha_{t+1} = \omega_{t+1}^\Delta \left( \alpha^{t-1}, \alpha_t \right)$$

Using this and the definition of  $\omega_{t+1}(\alpha^{t-1}, \alpha_t)$ , (45) collapses to (15).

## A.2 Proof of Proposition 1

### A.2.1 'If'

Suppose that global incentive compatibility fails. By the time separability of preferences, for some  $t$  and history  $\alpha^{t-1}$  there must exist  $\alpha'_t, \alpha''_t$  such that:

$$\alpha'_t u_t(\alpha''_t) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha''_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha'_t) d\alpha_{t+1} > \alpha'_t u_t(\alpha'_t) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha'_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha'_t) d\alpha_{t+1}$$

or equivalently:

$$u_t(\alpha''_t) + \frac{\beta}{\alpha'_t} \int_{\alpha_{t+1}} V_{t+1}(\alpha''_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha'_t) d\alpha_{t+1} > u_t(\alpha'_t) + \frac{\beta}{\alpha'_t} \int_{\alpha_{t+1}} V_{t+1}(\alpha'_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha'_t) d\alpha_{t+1} \quad (46)$$

where  $u_t(\alpha_t)$  is used as shorthand for  $u(c_t(\alpha^{t-1}, \alpha_t))$ , and dependence of  $V_{t+1}$  on  $\alpha^{t-1}$  is similarly suppressed. By the absolute continuity of lifetime utility in type:

$$\begin{aligned} & u_t(\alpha''_t) + \frac{\beta}{\alpha'_t} \int_{\alpha_{t+1}} V_{t+1}(\alpha''_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha'_t) d\alpha_{t+1} \\ & - u_t(\alpha'_t) + \frac{\beta}{\alpha'_t} \int_{\alpha_{t+1}} V_{t+1}(\alpha'_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha'_t) d\alpha_{t+1} \\ & = \int_{\alpha'_t}^{\alpha''_t} \frac{d}{d\alpha_t} \left\{ \frac{1}{\alpha_t} \left[ \alpha_t u_t(\alpha_t) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} \right] \right\} d\alpha_t \\ & = \int_{\alpha'_t}^{\alpha''_t} \left\{ \begin{array}{l} -\frac{1}{\alpha_t^2} \left[ \alpha_t u_t(\alpha_t) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} \right] \\ + \frac{1}{\alpha_t} \frac{d}{d\alpha_t} \left[ \alpha_t u_t(\alpha_t) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha_t, \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} \right] \end{array} \right\} d\alpha_t \\ & = \int_{\alpha'_t}^{\alpha''_t} \left\{ \begin{array}{l} -\frac{1}{\alpha_t^2} \left[ \alpha_t u_t(\alpha_t) + \beta \omega_{t+1}(\alpha^{t-1}, \alpha_t) \right] \\ + \frac{1}{\alpha_t^2} \left[ \alpha_t u_t(\alpha_t) + \beta \omega_{t+1}^\Delta(\alpha^{t-1}, \alpha_t) \right] \end{array} \right\} d\alpha_t \\ & = -\beta \int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t^2} \left( \omega_{t+1}(\alpha^{t-1}, \alpha_t) - \omega_{t+1}^\Delta(\alpha^{t-1}, \alpha_t) \right) d\alpha_t \end{aligned}$$

where the penultimate line has made use of the relaxed incentive constraint (15).

Applying this result in (46) yields:

$$u_t(\alpha'_t) + \frac{\beta}{\alpha'_t} \int_{\alpha_{t+1}} V_{t+1}(\alpha'_t, \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha'_t) d\alpha_{t+1} > u_t(\alpha''_t) + \frac{\beta}{\alpha''_t} \int_{\alpha_{t+1}} V_{t+1}(\alpha''_t, \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha''_t) d\alpha_{t+1} \\ + \beta \int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t^2} \left[ \omega_{t+1}(\alpha^{t-1}, \alpha_t) + \omega_{t+1}^\Delta(\alpha^{t-1}, \alpha_t) \right] d\alpha_t$$

Or:

$$\beta \int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t^2} \int_{\alpha_{t+1}} (V_{t+1}(\alpha'_t, \alpha_{t+1}) - V_{t+1}(\alpha_t, \alpha_{t+1})) \left( \pi(\alpha_{t+1}|\alpha_t) - \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} \right) d\alpha_{t+1} d\alpha_t > 0$$

Applying Lemma 3 and the definition of  $V_{t+1}$ , this is equivalent to:

$$\int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t^2} \left\{ \mathbb{E}_t \left[ \sum_{s=t+1}^{\infty} \beta^{s-t} (1 - D_{t,s}(\alpha^s)) \alpha_s \left[ u_s(\alpha^{t-1}, \alpha'_t, \dots, \alpha_s) - u_s(\alpha^s) \right] \middle| \alpha_t \right] \right\} d\alpha_t > 0$$

But this directly contradicts the integral monotonicity condition given in the Proposition.

### A.2.2 'Only if'

Suppose integral monotonicity fails for some  $(\alpha'_t, \alpha''_t)$ , i.e.:

$$\int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t^2} \left\{ \mathbb{E}_t \left[ \sum_{s=t+1}^{\infty} \beta^{s-t} (1 - D_{t,s}(\alpha^s)) \alpha_s \left[ u_s(\alpha^{t-1}, \alpha''_t, \dots, \alpha_s) - u_s(\alpha^s) \right] \middle| \alpha_t \right] \right\} d\alpha_t > 0$$

Applying the steps for the previous subsection in reverse, this is equivalent the inequality:

$$\alpha'_t u_t(\alpha''_t) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha'_t, \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha'_t) d\alpha_{t+1} > \alpha'_t u_t(\alpha'_t) + \beta \int_{\alpha_{t+1}} V_{t+1}(\alpha'_t, \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha'_t) d\alpha_{t+1}$$

Thus global incentive compatibility must be violated for type  $\alpha'_t$ .

## A.3 Proof of Proposition 2

The key first-order conditions are constructed by studying differential changes to the allocation that remain within the constraint space. The derivations are taken by reference to Lagrange multipliers, which are assumed to take a standard algebraic form. It is simple to show that these multipliers can be eliminated by taking linear combinations of the resulting expressions. Thus by construction, the results could equivalently be derived through a calculus of variations approach, and do not depend on the assumed form for the multipliers.



The relaxed planner's problem is to solve:

$$\max_{\{c_t(\alpha^t)\}_{\alpha^t}} \sum_{t=0}^{\infty} \int_{\alpha^t} \alpha_t u(c_t(\alpha^t)) d\Pi_t(\alpha^t)$$

subject to the resource constraint:

$$\sum_{t=0}^{\infty} R^{-t} \left[ y_t - \int_{\alpha^t} c_t(\alpha^t) d\Pi_t(\alpha^t) \right] \geq 0 \quad (47)$$

and the relaxed incentive constraint:

$$\begin{aligned} \alpha'_t u(c_t(\alpha^{t-1}, \alpha'_t)) + \beta \omega_{t+1}(\alpha^{t-1}, \alpha'_t) &= \underline{\alpha} u(c_t(\alpha^{t-1}, \underline{\alpha})) + \beta \omega_{t+1}(\alpha^{t-1}, \underline{\alpha}) \\ &+ \int_{\underline{\alpha}}^{\alpha'_t} \frac{1}{\alpha_t} \left[ \alpha_t u(c_t(\alpha^{t-1}, \alpha_t)) + \beta \omega_{t+1}^{\Delta}(\alpha^{t-1}, \alpha_t) \right] d\alpha_t \end{aligned} \quad (48)$$

with, for all  $t \geq 0$ :

$$\omega_{t+1}(\alpha^t) := \int_{\alpha_{t+1}} \{ \alpha_{t+1} u(c_{t+1}(\alpha^t, \alpha_{t+1})) + \beta \omega_{t+2}(\alpha^t, \alpha_{t+1}) \} d\Pi(\alpha_{t+1} | \alpha_t) \quad (49)$$

$$\omega_{t+1}^{\Delta}(\alpha^t) := \int_{\alpha_{t+1}} \rho(\alpha_{t+1} | \alpha_t) \cdot \{ \alpha_t u(c_{t+1}(\alpha^t, \alpha_{t+1})) + \beta \omega_{t+2}^{\Delta}(\alpha^t, \alpha_{t+1}) \} d\Pi(\alpha_{t+1} | \alpha_t) \quad (50)$$

plus the interiority restriction, which is assumed not to bind for the Proposition. We place multiplier  $\eta$  on (47),  $\beta^t \mu_t(\alpha^{t-1}, \alpha_t) d\Pi_t(\alpha^{t-1}, \alpha_t)$  on (48),  $\beta^{t+1} \lambda_{t+1}(\alpha^t) d\Pi_t(\alpha^t)$  on (49) and  $\beta^{t+1} \lambda_{t+1}^{\Delta}(\alpha^t) d\Pi_t(\alpha^t)$  on (50). Necessary first-order optimality conditions are:

- With respect to  $c_t(\alpha^t)$ , a.e.:

$$0 = \alpha_t u'(c_t(\alpha^t)) \cdot \left[ \begin{aligned} &1 + \lambda_t(\alpha^{t-1}) + \lambda_t^{\Delta}(\alpha^{t-1}) \rho(\alpha_t | \alpha_{t-1}) + \mu_t(\alpha^t) \\ & - \frac{1}{\alpha_t \pi(\alpha_t | \alpha_{t-1})} \int_{\alpha_t}^{\tilde{\alpha}} \mu_t(\alpha^{t-1}, \tilde{\alpha}_t) \pi(\tilde{\alpha}_t | \alpha_{t-1}) d\tilde{\alpha}_t \end{aligned} \right] - (\beta R)^{-t} \eta \quad (51)$$

- With respect to  $\omega_{t+1}(\alpha^t)$ , a.e.:

$$0 = -\lambda_{t+1}(\alpha^t) + \lambda_t(\alpha^{t-1}) + \mu_t(\alpha^{t-1}, \alpha_t) \quad (52)$$

- With respect to  $\omega_{t+1}^{\Delta}(\alpha^t)$ , a.e.:

$$\begin{aligned} 0 = & -\lambda_{t+1}^{\Delta}(\alpha^t) + \lambda_t^{\Delta}(\alpha^{t-1}) \rho(\alpha_t | \alpha_{t-1}) \\ & - \frac{1}{\alpha_t \pi(\alpha_t | \alpha_{t-1})} \int_{\alpha_t}^{\tilde{\alpha}} \mu_t(\alpha^{t-1}, \tilde{\alpha}_t) \pi(\tilde{\alpha}_t | \alpha_{t-1}) d\tilde{\alpha}_t \end{aligned} \quad (53)$$

- With respect to  $c_t(\alpha^{t-1}, \underline{\alpha})$ :

$$0 = \int_{\underline{\alpha}}^{\bar{\alpha}} \mu_t(\alpha^{t-1}, \alpha_t) \pi(\alpha_t | \alpha_{t-1}) d\alpha_t \quad (54)$$

Throughout here, we normalise  $\lambda_0 = \lambda_0^\Delta \equiv 0$ , and let  $\pi(\alpha_t | \alpha_{t-1})$  be replaced with  $\pi(\alpha_0)$  when  $t = 0$ .

Using (52) and (53) in (51) gives:

$$\frac{(\beta R)^{-t} \eta}{\alpha_t u'(c_t(\alpha^t))} = 1 + \lambda_{t+1}(\alpha^t) + \lambda_{t+1}^\Delta(\alpha^t) \quad (55)$$

Condition (51) can be rearranged to:

$$\begin{aligned} & \frac{(\beta R)^{-t} \eta}{u'(c_t(\alpha^t))} - \alpha_t \left[ 1 + \lambda_t(\alpha^{t-1}) + \lambda_t^\Delta(\alpha^{t-1}) \rho(\alpha_t | \alpha_{t-1}) \right] \\ &= \alpha_t \mu_t(\alpha^{t-1}, \alpha_t) - \frac{1}{\pi(\alpha_t | \alpha_{t-1})} \int_{\alpha_t}^{\bar{\alpha}} \mu_t(\alpha^{t-1}, \tilde{\alpha}_t) \pi(\tilde{\alpha}_t | \alpha_{t-1}) d\tilde{\alpha}_t \end{aligned} \quad (56)$$

This can be integrated across all  $\alpha_t$ :

$$\begin{aligned} & \int_{\underline{\alpha}}^{\bar{\alpha}} \left\{ \frac{(\beta R)^{-t} \eta}{u'(c_t(\alpha^{t-1}, \alpha_t))} - \alpha_t \left[ 1 + \lambda_t(\alpha^{t-1}) + \lambda_t^\Delta(\alpha^{t-1}) \rho(\alpha_t | \alpha_{t-1}) \right] \right\} \pi(\alpha_t | \alpha_{t-1}) d\alpha_t \\ &= \int_{\underline{\alpha}}^{\bar{\alpha}} \left\{ \alpha_t \mu_t(\alpha^{t-1}, \alpha_t) \pi(\alpha_t | \alpha_{t-1}) - \int_{\alpha_t}^{\bar{\alpha}} \mu_t(\alpha^{t-1}, \tilde{\alpha}_t) \pi(\tilde{\alpha}_t | \alpha_{t-1}) d\tilde{\alpha}_t \right\} d\alpha_t \end{aligned} \quad (57)$$

Integrating by parts, making use of (54):

$$\int_{\underline{\alpha}}^{\bar{\alpha}} \int_{\alpha_t}^{\bar{\alpha}} \mu_t(\alpha^{t-1}, \tilde{\alpha}_t) \pi(\tilde{\alpha}_t | \alpha_{t-1}) d\tilde{\alpha}_t d\alpha_t = \int_{\underline{\alpha}}^{\bar{\alpha}} \alpha_t \mu_t(\alpha^{t-1}, \alpha_t) \pi(\alpha_t | \alpha_{t-1}) d\alpha_t$$

and so:

$$\int_{\underline{\alpha}}^{\bar{\alpha}} \left\{ \frac{(\beta R)^{-t} \eta}{u'(c_t(\alpha^{t-1}, \alpha_t))} - \alpha_t \left[ 1 + \lambda_t(\alpha^{t-1}) + \lambda_t^\Delta(\alpha^{t-1}) \rho(\alpha_t | \alpha_{t-1}) \right] \right\} \pi(\alpha_t | \alpha_{t-1}) d\alpha_t = 0 \quad (58)$$

or, using (5):

$$\frac{1}{\mathbb{E}[\alpha_t | \alpha_{t-1}]} \mathbb{E} \left[ \frac{(\beta R)^{-t} \eta}{u'(c_t(\alpha^t))} \middle| \alpha^{t-1} \right] = \left[ 1 + \lambda_t(\alpha^{t-1}) + \lambda_t^\Delta(\alpha^{t-1}) \varepsilon^\alpha(\alpha_{t-1}) \right] \quad (59)$$

Rearranging (59) for period 0 gives an expression for  $\eta$ :

$$\eta = \frac{\mathbb{E}[\alpha_0]}{\mathbb{E} \left[ \frac{1}{u'(c_0(\alpha_0))} \right]} \quad (60)$$

Combining (55) and (59) gives expressions for the objects  $1 + \lambda_{t+1}(\alpha^t)$  and  $\lambda_{t+1}^\Delta(\alpha^t)$ :

$$1 + \lambda_{t+1}(\alpha^t) = (\beta R)^{-t} \eta \frac{1}{1 - \varepsilon^\alpha(\alpha_t)} \left\{ \frac{1}{\mathbb{E}[\alpha_{t+1}|\alpha_t]} \mathbb{E} \left[ \frac{1}{\beta R u'(c_{t+1}(\alpha^{t+1}))} \middle| \alpha^t \right] - \frac{\varepsilon^\alpha(\alpha_t)}{\alpha_t u'(c_t(\alpha^t))} \right\} \quad (61)$$

$$\lambda_{t+1}^\Delta(\alpha^t) = (\beta R)^{-t} \eta \frac{1}{1 - \varepsilon^\alpha(\alpha_t)} \left\{ \frac{1}{\alpha_t u'(c_t(\alpha^t))} - \frac{1}{\mathbb{E}[\alpha_{t+1}|\alpha_t]} \mathbb{E} \left[ \frac{1}{\beta R u'(c_{t+1}(\alpha^{t+1}))} \middle| \alpha^t \right] \right\} \quad (62)$$

Integrating (56) above any given  $\alpha'_t$  gives:

$$\begin{aligned} & \int_{\alpha'_t}^{\bar{\alpha}} \left\{ \frac{(\beta R)^{-t} \eta}{u'(c_t(\alpha^{t-1}, \alpha_t))} - \alpha_t \left[ 1 + \lambda_t(\alpha^{t-1}) + \lambda_t^\Delta(\alpha^{t-1}) \rho(\alpha_t|\alpha_{t-1}) \right] \right\} \pi(\alpha_t|\alpha_{t-1}) d\alpha_t \\ &= \int_{\alpha'_t}^{\bar{\alpha}} \left\{ \alpha_t \mu_t(\alpha^{t-1}, \alpha_t) \pi(\alpha_t|\alpha_{t-1}) - \int_{\alpha_t}^{\bar{\alpha}} \mu_t(\alpha^{t-1}, \tilde{\alpha}_t) \pi(\tilde{\alpha}_t|\alpha_{t-1}) d\tilde{\alpha}_t \right\} d\alpha_t \end{aligned} \quad (63)$$

and integrating by parts, we have:

$$\begin{aligned} & \int_{\alpha'_t}^{\bar{\alpha}} \int_{\alpha_t}^{\bar{\alpha}} \mu_t(\alpha^{t-1}, \tilde{\alpha}_t) \pi(\tilde{\alpha}_t|\alpha_{t-1}) d\tilde{\alpha}_t d\alpha_t \\ &= -\alpha'_t \int_{\alpha'_t}^{\bar{\alpha}} \mu_t(\alpha^{t-1}, \alpha_t) \pi(\alpha_t|\alpha_{t-1}) d\alpha_t + \int_{\alpha'_t}^{\bar{\alpha}} \alpha_t \mu_t(\alpha^{t-1}, \alpha_t) \pi(\alpha_t|\alpha_{t-1}) d\alpha_t \end{aligned} \quad (64)$$

so:

$$\begin{aligned} & \int_{\alpha'_t}^{\bar{\alpha}} \left\{ \frac{(\beta R)^{-t} \eta}{u'(c_t(\alpha^{t-1}, \alpha_t))} - \alpha_t \left[ 1 + \lambda_t(\alpha^{t-1}) + \lambda_t^\Delta(\alpha^{t-1}) \rho(\alpha_t|\alpha_{t-1}) \right] \right\} \pi(\alpha_t|\alpha_{t-1}) d\alpha_t \\ &= \alpha'_t \int_{\alpha'_t}^{\bar{\alpha}} \mu_t(\alpha^{t-1}, \alpha_t) \pi(\alpha_t|\alpha_{t-1}) d\alpha_t \\ &= -\pi(\alpha'_t|\alpha_{t-1}) (\alpha'_t)^2 \left\{ \lambda_{t+1}^\Delta(\alpha^t) - \rho(\alpha'_t|\alpha_{t-1}) \lambda_t^\Delta(\alpha^{t-1}) \right\} \end{aligned} \quad (65)$$

Applying the definition of a conditional expectation, and letting  $\eta_t := (\beta R)^{-t} \eta$ , this immediately gives the main condition in the Proposition.

#### A.4 Proof of Proposition 3

The proof of Proposition 2 has already established the result for  $s = t$  and  $s = t + 1$ . From (61) and (62) we immediately have:

$$\frac{1}{\alpha_t u'(c_t)} = \frac{1 + \lambda_{t+1} + \lambda_{t+1}^\Delta}{\eta_t} \quad (66)$$

$$\frac{1}{\mathbb{E}_t[\alpha_{t+1}]} \mathbb{E}_t \left[ \frac{1}{\beta R u'(c_{t+1})} \right] = \frac{1 + \lambda_{t+1} + \lambda_{t+1}^\Delta \varepsilon^\alpha(\alpha_t)}{\eta_t} \quad (67)$$

and note that  $\varepsilon^\alpha(\alpha_t) = \frac{\mathbb{E}_t[D_{t,t+1}(\alpha^{t+1})\alpha_{t+1}]}{\mathbb{E}_t[\alpha_{t+1}]}$ .

The proof then works recursively. Suppose that, for  $r < s$ :

$$\mathbb{E}_r \left[ \frac{\eta_r}{(\beta R)^{s-r} u'(c_s)} \right] = [1 + \lambda_{r+1}] \mathbb{E}_r [\alpha_s] + \lambda_{r+1}^\Delta \mathbb{E}_r [D_{r,s}(\alpha^s) \alpha_s] \quad (68)$$

Then:

$$\begin{aligned} & \mathbb{E}_{r-1} \left[ \frac{\eta_{r-1}}{(\beta R)^{s-r+1} u'(c_s)} \right] \quad (69) \\ &= \mathbb{E}_{r-1} \left\{ [1 + \lambda_{r+1}] \mathbb{E}_r [\alpha_s] + \lambda_{r+1}^\Delta \mathbb{E}_r [D_{r,s}(\alpha^s) \alpha_s] \right\} \\ &= \int_{\alpha_r} \left\{ \left[ \rho(\alpha_r | \alpha_{r-1}) \lambda_r^\Delta - \frac{1}{\alpha_r \pi(\alpha_r | \alpha_{r-1})} \int_{\alpha_r}^{\bar{\alpha}} \mu_r(\tilde{\alpha}_r) \pi(\tilde{\alpha}_r | \alpha_{r-1}) d\tilde{\alpha}_r \right] \cdot \mathbb{E}_r [D_{r,s}(\alpha^s) \alpha_s] \right. \\ & \quad \left. + [1 + \lambda_r + \mu_r(\alpha_r)] \cdot \mathbb{E}_r [\alpha_s] \right\} \pi(\alpha_r | \alpha_{r-1}) d\alpha_r \quad (70) \end{aligned}$$

By an identical argument to Lemma 3, we have:

$$\frac{d}{d\alpha_r} [\mathbb{E}_r [\alpha_s]] = \frac{1}{\alpha_r} \mathbb{E}_r [D_{r,s}(\alpha^s) \alpha_s]$$

Integrating by parts, we therefore have:

$$\int_{\alpha_r} \frac{1}{\alpha_r} \mathbb{E}_r [D_{r,s}(\alpha^s) \alpha_s] \left[ \int_{\alpha_r}^{\bar{\alpha}} \mu_r(\tilde{\alpha}_r) \pi(\tilde{\alpha}_r | \alpha_{r-1}) d\tilde{\alpha}_r \right] d\alpha_r = \int_{\alpha_r} \mathbb{E}_r [\alpha_s] \mu_r(\tilde{\alpha}_r) \pi(\alpha_r | \alpha_{r-1}) d\alpha_r \quad (71)$$

where we have used condition (54). Using this in the preceding expression, the terms in  $\mu_r$  cancel:

$$\begin{aligned} \mathbb{E}_{r-1} \left[ \frac{\eta_{r-1}}{(\beta R)^{s-r+1} u'(c_s)} \right] &= \int_{\alpha_r} \left\{ [1 + \lambda_r] \cdot \mathbb{E}_r [\alpha_s] + \rho(\alpha_r | \alpha_{r-1}) \lambda_r^\Delta \cdot \mathbb{E}_r [D_{r,s}(\alpha^s) \alpha_s] \right\} \pi(\alpha_r | \alpha_{r-1}) d\alpha_r \quad (72) \\ &= [1 + \lambda_r] \cdot \mathbb{E}_{r-1} [\alpha_s] + \lambda_r^\Delta \cdot \mathbb{E}_{r-1} [D_{r-1,s}(\alpha^s) \alpha_s] \end{aligned}$$

which makes use of the definition of  $D_{r-1,s}(\alpha^s)$ . Thus we have iterated expectations backwards a period from condition (68). Now, for any  $s > 0$ , condition (67) implies:

$$\mathbb{E}_{s-1} \left[ \frac{\eta_{s-1}}{\beta R u'(c_s)} \right] = [1 + \lambda_s] \mathbb{E}_{s-1} [\alpha_s] + \lambda_s^\Delta \mathbb{E}_{s-1} [D_{s-1,s}(\alpha^s) \alpha_s] \quad (73)$$

The preceding arguments allow this to be iterated back to  $t$ , as required.

## A.5 Proof of Proposition 4

We proceed constructively, showing how to map from the allocation  $\{c_t^*(\alpha^t)\}_{t,\alpha^t}$  to tax functions  $T_t(s^{t-1}, s_t)$ , with the property every budget-feasible sequence of savings choices over time implies a consumption sequence that is part of the target incentive-feasible allocation  $\{c_t^*(\alpha^t)\}_{t,\alpha^t}$ , and every consumption sequence from the target allocation can be chosen via a feasible sequence of savings decisions. This implies that the menu of choices at every history node under the decentralised allocation is the same as under the direct mechanism, and so the decentralised scheme must implement the target allocation.

First, set  $M_0$  equal the net-present value of resources per capita in period zero:

$$M_0 := \sum_{t=0}^{\infty} R^{-t} y_t$$

For all  $\alpha_0$ , let the savings level  $s_0(\alpha_0)$  then be defined by:

$$s_0(\alpha_0) := M_0 - c_0^*(\alpha_0)$$

and denote the range of  $s_0$  values across  $\alpha_0$  by  $S_0$ :

$$S_0 := \{s_0(\alpha_0)\}_{\alpha_0 \in A}$$

Since consumption is increasing, the minimum value for savings is  $s_0(\bar{\alpha})$  and its maximum is  $s_0(\underline{\alpha})$ , and so  $S_0 \subseteq [s_0(\bar{\alpha}), s_0(\underline{\alpha})]$ . We denote by  $S_0^c$  the complement of  $S_0$  in  $\mathbb{R}$ .

For all  $\tilde{s}_0 \in S_0^c$ , let  $T_0(\tilde{s}_0) = \tilde{s}_0$ , so that  $M_1(\tilde{s}_0) = 0$ , and for all  $t > 0$  and subsequent savings choices  $\{s_1, \dots, s_t\}$ , let  $T_t(\tilde{s}_0, s_1, \dots, s_t) > \varepsilon$  for some  $\varepsilon > 0$ . Combining the budget constraints (23) and (24), we have, along all future consumption paths:

$$0 = M_1(\tilde{s}_0) = \sum_{t=0}^T R^{-t} [c_{t+1} + T_{t+1}(\tilde{s}_0, \dots, s_{t+1})] + R^{-T-2} M_{T+2}$$

and so:

$$\sum_{t=0}^T R^{-t} c_{t+1} = - \sum_{t=0}^T R^{-t} T_{t+1}(\tilde{s}_0, \dots, s_{t+1}) - R^{-T-2} M_{T+2}$$

for all  $T \geq 0$ . By the ‘no Ponzi’, the final term on the right-hand side satisfies  $\lim_{T \rightarrow \infty} R^{-T-2} M_{T+2} \geq 0$ , and so the positive bound on taxes implies that the right-hand side must be negative for large enough  $T$ . But this implies negative consumption in at least one period, which is not possible. It follows that  $\tilde{s}_0$  will not be chosen.

For all  $s_0 \in S_0$ , let  $M_1(s_0)$  be given by:

$$M_1(s_0) := \mathbb{E} \left[ \sum_{t=1}^{\infty} R^{1-t} c_t^*(\alpha^t) \middle| \alpha_0 \in \alpha_0(s_0) \right]$$

where  $\alpha_0(s_0) : S_0 \rightarrow \mathbb{R}$  is the inverse of  $s_0(\alpha_0)$ . Where  $c_0^*(\alpha_0)$  is strictly increasing,  $\alpha_0(s_0)$  is singleton-valued, and expectations are with respect to the evolution of types subsequent to period 0 given this  $\alpha_0$ . More generally  $\alpha_0(s_0)$  may take values from a convex interval in  $A$ , and in this case expectations satisfy Bayes's rule in the obvious way. Condition 2 in the Proposition guarantees that the continuation allocation is identical across types within any such set, and so they do not need to be separated in their market treatment.

Given  $M_1(s_0)$ , we then define  $T_0(s_0)$  by:

$$T_0(s_0) := s_0 - R^{-1}M_1(s_0)$$

The logic can then proceed recursively for on-equilibrium choices. Fix  $t > 0$ . Suppose that a mapping from type history  $\alpha^{t-1} \in A^t$  to savings history  $s^{t-1} \in \mathbb{R}^t$  is known, denoted  $s^{t-1}(\alpha^{t-1})$ , with range  $S^{t-1}$ :

$$S^{t-1} := \left\{ s^{t-1}(\alpha^{t-1}) \right\}_{\alpha^{t-1} \in A^t}$$

and that this mapping has an inverse correspondence  $\alpha^{t-1}(s^{t-1})$ , with  $\alpha^{t-1} : S^{t-1} \rightarrow A^t$ . For  $t = 1, S^{t-1} = S_0$ . Suppose further that there is a known wealth level  $M_t(s^{t-1}(\alpha^{t-1}))$  corresponding to each  $\alpha^{t-1} \in A^t$ . For all  $\alpha_t \in A$ , let  $s_t(\alpha^{t-1}, \alpha_t)$  be given by:

$$s_t(\alpha^{t-1}, \alpha_t) := M_t(s^{t-1}(\alpha^{t-1})) - c_t^*(\alpha^{t-1}, \alpha_t)$$

By the assumed increasingness of  $c_t^*$ ,  $s_t(\alpha^t)$  is decreasing in  $\alpha_t$ , with minimum  $s_t(\alpha^{t-1}, \bar{\alpha})$  and maximum  $s_t(\alpha^{t-1}, \underline{\alpha})$ . Denote its range  $S_t(\alpha^{t-1})$ :

$$S_t(\alpha^{t-1}) := \left\{ s_t(\alpha^{t-1}, \alpha_t) \right\}_{\alpha_t \in A}$$

Given  $\alpha^{t-1}$ , the inverse function  $\alpha_t(s_t; \alpha^{t-1})$  gives the convex subset of types corresponding to savings choice  $s_t$ , for any  $s_t \in S_t(\alpha^{t-1})$ . The mapping  $s^t(\alpha^t)$  is then given by extending  $s^{t-1}(\alpha^{t-1})$ :

$$s^t(\alpha^t) := \left\{ s^{t-1}(\alpha^{t-1}), s_t(\alpha^t) \right\}$$

and  $\alpha^t(s^t)$  by:

$$\alpha^t(s^t) := \left\{ \alpha^{t-1}(s^{t-1}), \alpha_t(s_t; \alpha^{t-1}(s^{t-1})) \right\}$$

For all  $s_t \in S_t(\alpha^{t-1})$ ,  $M_{t+1}(s^{t-1}(\alpha^{t-1}), s_t)$  can then be given by:

$$M_{t+1}(s^{t-1}(\alpha^{t-1}), s_t) := \mathbb{E} \left[ \sum_{r=t+1}^{\infty} R^{t+1-r} c_r(\alpha^r) \middle| \alpha^t \in \alpha^t(s^t) \right]$$

with expectations again taken with respect to the evolution of types, applying Bayes's rule if  $\alpha_t$  is not uniquely identified. This leaves the tax function  $T_t(s^{t-1}, s_t)$  to be given by:

$$T_t(s^{t-1}, s_t) := s_t - R^{-1} M_{t+1}(s^{t-1}, s_t)$$

for all  $s_t \in S_t(\alpha^{t-1}(s^{t-1}))$ .

As in period-zero, we need to rule out allocation choices that do not feature under the direct mechanism. Denote by  $S_t^c(\alpha^{t-1})$  the complement of  $S_t(\alpha^{t-1})$  in  $\mathbb{R}$ , and for all  $\tilde{s}_t \in S_t^c(\alpha^{t-1}(s^{t-1}))$ , set  $T_t(s^{t-1}, \tilde{s}_t)$  equal to  $\tilde{s}_t$ . For all  $r > t$ , set  $T_t(s^{t-1}, \tilde{s}_t, \dots, s_r) > \varepsilon$  for some  $\varepsilon > 0$ . Again, this implies that choosing  $\tilde{s}_t$  is inconsistent with satisfying the no-Ponzi condition.

## A.6 Corollary 2: details

Monotonicity of  $c_t(\alpha_t)$  implies that  $c_t(\alpha_t)$  is continuous a.e.. Suppose it is continuous on some open interval  $(\alpha', \alpha'') \subset [\underline{\alpha}, \bar{\alpha}]$ , and write  $c' = \lim_{\alpha \searrow \alpha'} (c_t(\alpha))$  and  $c'' = \lim_{\alpha \nearrow \alpha''} (c_t(\alpha))$  (i.e. limits as  $\alpha$  approaches from above and below respectively). Integrating (17) across this range gives:

$$\begin{aligned} & \int_{c'}^{c''} \left\{ \int_{\alpha_t(c_t)}^{\bar{\alpha}} \left[ \frac{1}{u'(c_t(\alpha))} - \frac{\alpha \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha | \alpha_{t-1})\}}{\eta_t} \right] \pi(\alpha | \alpha_{t-1}) d\alpha \right\} \frac{du'(c_t)}{dc_t} dc_t \\ &= - \int_{c'}^{c''} (\alpha_t(c_t))^2 \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c_t))}{\eta_t} - \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \pi(\alpha_t(c_t) | \alpha_{t-1}) \frac{du'(c_t)}{dc_t} dc_t \end{aligned} \quad (74)$$

Integrating the left-hand side by parts, we have:

$$\begin{aligned}
& \int_{c'}^{c''} \left\{ \int_{\alpha_t(c_t)}^{\bar{\alpha}} \left[ \frac{1}{u'(c_t(\alpha))} - \frac{\alpha \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha | \alpha_{t-1})\}}{\eta_t} \right] \pi(\alpha | \alpha_{t-1}) d\alpha \right\} \frac{du'(c_t)}{dc_t} dc_t \\
&= \left[ \left\{ \int_{\alpha_t(c_t)}^{\bar{\alpha}} \left[ \frac{1}{u'(c_t(\alpha))} - \frac{\alpha \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha | \alpha_{t-1})\}}{\eta_t} \right] \pi(\alpha | \alpha_{t-1}) d\alpha \right\} \left\{ \int_{c'}^{c_t} \frac{du'(c)}{dc} dc \right\} \right]_{c_t=c'}^{c''} \\
&+ \int_{c'}^{c''} \left[ \frac{1}{u'(c_t)} - \frac{\alpha_t(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \left[ \int_{c'}^{c_t} \frac{du'(c)}{dc} dc \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t \\
&= \int_{\alpha''}^{\bar{\alpha}} \left[ \frac{1}{u'(c_t(\alpha))} - \frac{\alpha \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha | \alpha_{t-1})\}}{\eta_t} \right] \pi(\alpha | \alpha_{t-1}) d\alpha [u'(c'') - u'(c')] \\
&+ \int_{c'}^{c''} \left[ \frac{1}{u'(c_t)} - \frac{\alpha_t(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] [u'(c_t) - u'(c')] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t \\
&= \int_{c''}^{\bar{c}} \left[ \frac{1}{u'(c_t)} - \frac{\alpha_t(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t (u'(c'') - u'(c')) \\
&+ \int_{c'}^{c''} \left[ \frac{1}{u'(c_t)} - \frac{\alpha_t(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] [u'(c_t) - u'(c')] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t \\
&= \int_{c'}^{c''} \left[ 1 - \frac{\alpha_t(c_t) u'(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t \\
&+ u'(c'') \int_{c''}^{\bar{c}} \left[ \frac{1}{u'(c_t)} - \frac{\alpha_t(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t \\
&- u'(c') \int_{c'}^{\bar{c}} \left[ \frac{1}{u'(c_t)} - \frac{\alpha_t(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t
\end{aligned}$$

where the strict normality assumption guarantees that  $\frac{d\alpha_t(c_t)}{dc_t}$  is defined a.e.. Note also that:

$$\pi^c(c_t | \alpha^{t-1}) := \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1})$$

is a measure of the empirical density of consumption at  $c_t$ , across types with the given history, defined a.e..



The main condition thus becomes:

$$\begin{aligned}
& \int_{c'}^{c''} \left[ 1 - \frac{\alpha_t(c_t) u'(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t \\
& + u'(c'') \int_{c'}^{\bar{c}} \left[ \frac{1}{u'(c_t)} - \frac{\alpha_t(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t \\
& - u'(c') \int_{c'}^{\bar{c}} \left[ \frac{1}{u'(c_t)} - \frac{\alpha_t(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t \\
& = - \int_{c'}^{c''} (\alpha_t(c_t))^2 u''(c_t) \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c_t))}{\eta_t} - \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \pi(\alpha_t(c_t) | \alpha_{t-1}) dc_t
\end{aligned} \tag{75}$$

Suppose first that there are no discontinuities in  $c_t(\alpha_t)$ . Making use of (58), over the entire range (75) gives:

$$\begin{aligned}
& \int_{\underline{c}}^{\bar{c}} \left[ 1 - \frac{\alpha_t(c_t) u'(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \pi^c(c_t | \alpha^{t-1}) dc_t \\
& + \int_{\underline{c}}^{\bar{c}} (\alpha_t(c_t))^2 u''(c_t) \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c_t))}{\eta_t} - \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \left( \frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \pi^c(c_t | \alpha^{t-1}) dc_t \\
& = 0
\end{aligned} \tag{76}$$

And for each  $c' \in (\underline{c}, \bar{c})$ , making use of (65):

$$\begin{aligned}
& \int_{c'}^{\bar{c}} \left[ 1 - \frac{\alpha_t(c_t) u'(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \pi^c(c_t | \alpha^{t-1}) dc_t \\
& + (\alpha_t(c'))^2 u'(c') \pi^c(c' | \alpha^{t-1}) \left( \frac{d\alpha_t(c')}{dc'} \right)^{-1} \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c'))}{\eta_t} - \rho(\alpha_t(c') | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \\
& + \int_{c'}^{\bar{c}} (\alpha_t(c_t))^2 u''(c_t) \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c_t))}{\eta_t} - \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \left( \frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \pi^c(c_t | \alpha^{t-1}) dc_t \\
& = 0
\end{aligned} \tag{77}$$

as given in the text.

### A.6.1 Allowing discontinuities

Discontinuities in  $c_t(\alpha_t)$  can be included with minimal additional manipulation. Since  $c_t(\alpha_t)$  is monotone, the set of  $\alpha_t$  values at which it is discontinuous is at most countable. Denote this set  $\mathcal{A} \subset A$ , and for all  $\alpha_t \in \mathcal{A}$  let  $c^u(\alpha_t) = \lim_{\alpha \searrow \alpha_t} c_t(\alpha)$  and  $c^l(\alpha_t) = \lim_{\alpha \nearrow \alpha_t} c_t(\alpha)$  denote the upper and lower limits for consumption

respectively. Summing (75) across intervals, over the entire range we have:

$$\begin{aligned}
& \int_{\underline{c}}^{\bar{c}} \left[ 1 - \frac{\alpha_t(c_t) u'(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \pi^c(c_t | \alpha^{t-1}) dc_t \\
& + \sum_{\alpha_t \in \mathcal{A}} \left( u'(c^u(\alpha_t)) - u'(c^l(\alpha_t)) \right) \pi(\alpha_t | \alpha_{t-1}) (\alpha_t)^2 \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t)}{\eta_t} - \rho(\alpha_t | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \\
& + \int_{c'}^{c''} (\alpha_t(c_t))^2 u''(c_t) \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c_t))}{\eta_t} - \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \pi^c(c_t | \alpha^{t-1}) dc_t \\
& = 0
\end{aligned} \tag{78}$$

And for each  $c' \in (\underline{c}, \bar{c})$ :

$$\begin{aligned}
& \int_{c'}^{\bar{c}} \left[ 1 - \frac{\alpha_t(c_t) u'(c_t) \{1 + \lambda_t + \lambda_t^\Delta \rho(\alpha_t(c_t) | \alpha_{t-1})\}}{\eta_t} \right] \pi^c(c_t | \alpha^{t-1}) dc_t \\
& + \sum_{\alpha_t \in \mathcal{A} \cap (\alpha_t(c'), \bar{\alpha})} \left( u'(c^u(\alpha_t)) - u'(c^l(\alpha_t)) \right) \pi(\alpha_t | \alpha_{t-1}) (\alpha_t)^2 \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t)}{\eta_t} - \rho(\alpha_t | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \\
& + (\alpha_t(c'))^2 u'(c') \pi^c(c' | \alpha^{t-1}) \left( \frac{d\alpha_t(c')}{dc'} \right)^{-1} \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c'))}{\eta_t} - \rho(\alpha_t(c') | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \\
& + \int_{c'}^{\bar{c}} (\alpha_t(c_t))^2 u''(c_t) \left\{ \frac{\lambda_{t+1}^\Delta(\alpha_t(c_t))}{\eta_t} - \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \right\} \left( \frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \pi^c(c_t | \alpha^{t-1}) dc_t \\
& = 0
\end{aligned} \tag{79}$$

### A.7 Definitions of $\epsilon_{t-1,t}^s(s_t)$ and $\left. \frac{ds_{t-1}}{dM_t} \right|_{\text{comp}}$

$\epsilon_{t-1,t}^s(s_t)$  is defined formally by reference to the Fréchet derivative of  $s_{t-1}$ , with respect to an arbitrary profile of changes to the tax schedule in  $t$ . These perturbations are defined pointwise at each value for  $M_{t+1}$ , so that the change in total taxes at a given  $M_{t+1}$  corresponds directly to the available change in period- $t$  consumption if  $M_{t+1}$  is held constant – i.e., to the size of the period- $t$  income effect for agents at the chosen  $M_{t+1}$ .

I first present the general approach to perturbing the tax schedule. Let  $s(M_{t+1})$  be the level of period- $t$  savings corresponding to  $t+1$  wealth level  $M_{t+1}$  under the chosen allocation, uniquely defined if the assumptions of Proposition 4, required for the decentralisation to work, are satisfied. This function is defined implicitly by:

$$M_{t+1} = R [s(M_{t+1}) - T(s(M_{t+1}))] \tag{80}$$

Notice that:

$$s'(M_{t+1}) = \frac{1}{R(1 - T'(s(M_{t+1})))} \quad (81)$$

That is, the derivative of  $s(M_{t+1})$  gives the localised period- $t$  cost of an extra unit of  $t + 1$  resources.

Suppose that for each choice of  $M_{t+1}$ , the tax schedule is perturbed to:<sup>34</sup>

$$T_t(s(M_{t+1})) - \Gamma f(s(M_{t+1})) \quad (82)$$

where  $\Gamma \in \mathbb{R}$  and  $f(\cdot)$  is an arbitrary bounded, differentiable function on the interval of realised savings. Using the intertemporal budget constraint (24), continued choice of  $M_{t+1}$  implies a savings level  $s_t$  that satisfies:

$$M_{t+1} = R[s_t - T_t(s(M_{t+1})) + \Gamma f(s(M_{t+1}))] \quad (83)$$

Combining with (80), this implies that realising  $M_{t+1}$  requires savings equal to:

$$s_t = s(M_{t+1}) - \Gamma f(s(M_{t+1})) \quad (84)$$

and thus consumption equal to:

$$\begin{aligned} c_t &= M_t - s(M_{t+1}) + \Gamma f(s(M_{t+1})) \\ &= c(M_{t+1}) + \Gamma f(s(M_{t+1})) \end{aligned} \quad (85)$$

where  $c(M_{t+1})$  is the consumption corresponding to  $M_{t+1}$  in the unperturbed allocation. Thus, holding constant  $M_{t+1}$ , the perturbation allows an increase in  $c_t$  by  $f(s(M_{t+1}))$  units per unit increase in  $\Gamma$ .

The derivative of savings with respect to  $M_{t+1}$  is:

$$\frac{ds_t}{dM_{t+1}} = s'(M_{t+1})(1 - \Gamma f'(s(M_{t+1}))) = \frac{1 - \Gamma f'(s(M_{t+1}))}{R(1 - T'(s(M_{t+1})))} \quad (86)$$

The inverse of this is the rate of return on savings when  $M_{t+1}$  is the realised wealth level in  $t + 1$ :

$$\left[ \frac{ds_t}{dM_{t+1}} \right]^{-1} = \frac{R(1 - T'(s(M_{t+1})))}{1 - \Gamma f'(s(M_{t+1}))} \quad (87)$$

---

<sup>34</sup>Defining the perturbations as functions of  $M_{t+1}$  rather than  $s_t$  ensures that  $f(\cdot)$  defines the 'rightward' shift in the budget constraint linking  $c_t$  to  $M_{t+1}$ , and so  $f(\cdot)$  corresponds to the magnitude of the income effect, in units of period- $t$  income. This allows for a simple characterisation of  $\epsilon_{t-1,t}^s(s_t)$  by reference to income and substitution effects at  $t$ , and the impact these have on information rents.

The derivative of this with respect to  $\Gamma$  is:

$$\frac{d}{d\Gamma} \left\{ \left[ \frac{ds_t}{dM_{t+1}} \right]^{-1} \right\} = \frac{f'(s(M_{t+1})) R(1 - T'(s(M_{t+1})))}{[1 - \Gamma f'(s(M_{t+1}))]^2} \quad (88)$$

and locally when  $\Gamma = 0$ , this gives:

$$\frac{d}{d\Gamma} \left\{ \left[ \frac{ds_t}{dM_{t+1}} \right]^{-1} \right\} \Big|_{\Gamma=0} = f'(s(M_{t+1})) R(1 - T'(s(M_{t+1})))$$

That is, a marginal increase in  $\Gamma$  from zero changes the post-tax rate of return at  $M_{t+1}$  by  $f'(s(M_{t+1})) R(1 - T'(s(M_{t+1})))$ . Thus the value of  $f'(s(M_{t+1}))$  corresponds to the proportional change in the slope of the intertemporal budget constraint at  $M_{t+1}$ .

The proof of Lemma 2, below, establishes that the compensated derivative of  $s_{t-1}$  with respect to  $\Gamma$ , evaluated at  $\Gamma = 0$ , will be linear in  $f'(s_t)$ .<sup>35</sup> That is:

$$\frac{ds_{t-1}}{d\Gamma} \Big|_{\Gamma=0, \text{comp}} = f(\underline{s}) H + \int_{s_t} f'(s_t) h(s_t) ds_t \quad (89)$$

with the function  $h(s_t)$  and scalar  $H$  independent of the choice of  $f(\cdot)$ . Since  $f'(s_t)$  is precisely the proportional increase in the rate of return on  $s_t$  per unit change in  $\Gamma$ , it suffices to normalise:

$$h(s_t) := s_{t-1} \epsilon_{t-1,t}^s(s_t) \left( 1 - \Pi^s(s_t | s^{t-1}) \right) \quad (90)$$

where  $\Pi^s(s_t | s^{t-1})$  is the induced measure of savings in  $t$ , given history of savings  $s^{t-1}$ . This implicitly defines  $\epsilon_{t-1,t}^s(s_t)$ . Thus, the aggregate change in  $s_{t-1}$  induced by an arbitrary change to marginal tax rates at  $t$  is, by definition, the integral of the changes induced piecewise by tax cuts at each point, and  $\epsilon_{t-1,t}^s(s_t)$  is constructed to capture the response at each point.

The compensated income effect  $\frac{ds_{t-1}}{dM_t} \Big|_{\text{comp}}$  is defined as the derivative  $\frac{ds_{t-1}}{d\Gamma} \Big|_{\Gamma=0, \text{comp}}$  in the simpler case where the perturbation schedule satisfies  $f(s(M_{t+1})) = 1$  for all realised  $M_{t+1}$  values. It is thus equated to  $H$  in (89).

## A.8 Proof of Lemma 2

### Conditions 1 and 2

We start with two Lemmata:

<sup>35</sup>This is established formally in the proof of Lemma 2: c.f. equation (122) below.

**Lemma 4.** *The marginal tax rate satisfies:*

$$T'_t(s(\alpha_t)) = \frac{\lambda_{t+1}^\Delta(\alpha_t)}{\eta_t} \left\{ \alpha_t u'(c_t) - \beta R(1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1} \right\}$$

where  $V_M(M_{t+1}; \alpha_{t+1})$  denotes the marginal increase in lifetime utility in  $t + 1$  when  $M_{t+1}$  is increased at the margin, given type draw  $\alpha_{t+1}$ .

*Proof.* Recall that the decentralisation in Proposition 4 sets the value of  $M_t(\alpha^{t-1})$  equal to the expected present value of consumption from  $t$  onwards, for agents with history  $\alpha^{t-1}$ :

$$M_t(\alpha^{t-1}) = \mathbb{E} \left[ \sum_{r=t}^{\infty} R^{t-r} c_r(\alpha^r) \middle| \alpha^{t-1} \right]$$

By definition, the marginal tax rate on savings is the net revenue raised by the policymaker, per unit, when savings are increased by a unit at the margin. Since the agent is optimising, a marginal change to savings relative to the optimum leaves them indifferent. Thus the marginal tax rate can be obtained from the optimal direct allocation as the difference between the marginal cost to the policymaker of providing resources in  $t$ , and the (discounted) shadow marginal resource cost of providing the utility increase implied by a unit increase in savings. By construction, savings raise period- $t + 1$  lifetime utility  $\omega_{t+1}$  at the margin by the amount:

$$R(1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1}$$

and raise  $\omega_{t+1}^\Delta$  by the amount:

$$R(1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1}$$

The marginal resource cost of increasing  $\omega_{t+1}$  by a unit at the margin will be the relevant shadow cost from the cost-minimisation dual. By standard arguments, an expression for this is obtained by dividing the marginal value of an increase to  $\omega_{t+1}$  in the main problem by the resource multiplier:

$$\frac{\beta^{t+1} (1 + \lambda_{t+1}(\alpha^t))}{\eta}$$

Similarly, the marginal resource cost of increasing  $\omega_{t+1}^\Delta$  by a unit is:

$$\frac{\beta^{t+1} \lambda_{t+1}^\Delta(\alpha^t)}{\eta}$$

The direct marginal resource gain from a unit increase in savings in period  $t$  is  $R^{-t}$ , and this is also the relative value of a unit of tax revenue from that period. Combining, we thus have:

$$R^{-t}T'_t(s_t) = R^{-t} - \frac{\beta^{t+1}(1 + \lambda_{t+1}(\alpha^t))}{\eta} R(1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \\ - \frac{\beta^{t+1}\lambda_{t+1}^\Delta(\alpha^t)}{\eta} R(1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1}$$

Or:

$$T'_t(s_t) = 1 - \frac{(1 + \lambda_{t+1}(\alpha^t))}{\eta(\beta R)^{-t-1}} (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \\ - \frac{\lambda_{t+1}^\Delta(\alpha^t)}{\eta(\beta R)^{-t-1}} (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1}$$

Expressions for  $\frac{(1 + \lambda_{t+1}(\alpha^t))}{\eta_t}$  and  $\frac{\lambda_{t+1}^\Delta(\alpha^t)}{\eta_t}$ , with  $\eta_t = \eta(\beta R)^{-t}$ , follow from (61) and (62) respectively:

$$\frac{1 + \lambda_{t+1}(\alpha^t)}{\eta_t} = \frac{1}{1 - \varepsilon^\alpha(\alpha_t)} \left\{ \frac{1}{\mathbb{E}[\alpha_{t+1}|\alpha_t]} \mathbb{E} \left[ \frac{1}{\beta R u'(c_{t+1}(\alpha^{t+1}))} \middle| \alpha^t \right] - \frac{\varepsilon^\alpha(\alpha_t)}{\alpha_t u'(c_t(\alpha^t))} \right\} \quad (91)$$

$$\frac{\lambda_{t+1}^\Delta(\alpha^t)}{\eta_t} = \frac{1}{1 - \varepsilon^\alpha(\alpha_t)} \left\{ \frac{1}{\alpha_t u'(c_t(\alpha^t))} - \frac{1}{\mathbb{E}[\alpha_{t+1}|\alpha_t]} \mathbb{E} \left[ \frac{1}{\beta R u'(c_{t+1}(\alpha^{t+1}))} \middle| \alpha^t \right] \right\} \quad (92)$$

Substituting in (91) gives:

$$T'_t(s_t) = 1 - \frac{1}{1 - \varepsilon^\alpha(\alpha_t)} \left\{ \frac{1}{\mathbb{E}[\alpha_{t+1}|\alpha_t]} \mathbb{E} \left[ \frac{1}{\beta R u'(c_{t+1}(\alpha^{t+1}))} \middle| \alpha^t \right] - \frac{\varepsilon^\alpha(\alpha_t)}{\alpha_t u'(c_t(\alpha^t))} \right\} \cdot \alpha_t u'(c_t(\alpha^t)) \\ - \frac{\lambda_{t+1}^\Delta(\alpha^t)}{\eta_t} \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1}$$

where we have used the consumer optimality condition:

$$\alpha_t u'(c_t(\alpha^t)) = \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1}$$

Rearranging the first line:

$$\begin{aligned}
& 1 - \frac{1}{1 - \varepsilon^\alpha(\alpha_t)} \left\{ \frac{1}{\mathbb{E}[\alpha_{t+1}|\alpha_t]} \mathbb{E} \left[ \frac{1}{\beta R u'(c_{t+1}(\alpha^{t+1}))} \middle| \alpha^t \right] - \frac{\varepsilon^\alpha(\alpha_t)}{\alpha_t u'(c_t(\alpha^t))} \right\} \cdot \alpha_t u'(c_t(\alpha^t)) \\
&= - \frac{1}{1 - \varepsilon^\alpha(\alpha_t)} \left\{ \frac{1}{\mathbb{E}[\alpha_{t+1}|\alpha_t]} \mathbb{E} \left[ \frac{1}{\beta R u'(c_{t+1}(\alpha^{t+1}))} \middle| \alpha^t \right] - \frac{\varepsilon^\alpha(\alpha_t)}{\alpha_t u'(c_t(\alpha^t))} - \frac{(1 - \varepsilon^\alpha(\alpha_t))}{\alpha_t u'(c_t(\alpha^t))} \right\} \cdot \alpha_t u'(c_t(\alpha^t)) \\
&= \frac{\lambda_{t+1}^\Delta(\alpha^t)}{\eta_t} \cdot \alpha_t u'(c_t(\alpha^t))
\end{aligned}$$

So:

$$T'_t(s(\alpha_t)) = \frac{\lambda_{t+1}^\Delta(\alpha_t)}{\eta_t} \left\{ \alpha_t u'(c_t) - \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1} \right\}$$

as stated.  $\square$

**Lemma 5.** *The contemporaneous income effect and labour supply elasticity satisfy, respectively:*

$$\left\{ \alpha_t u'(c_t) - \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1} \right\}^{-1} = - \frac{\frac{ds_t}{dM_t}}{\alpha_t^2 u''(c_t) \frac{dc_t}{d\alpha_t}} \quad (93)$$

$$\left\{ \alpha_t u'(c_t) - \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1} \right\}^{-1} = \frac{s_t \varepsilon_t^s}{\alpha_t^2 u'(c_t) \frac{dc_t}{d\alpha_t}} \quad (94)$$

*Proof.* Start with (93). Given the decentralised scheme, and holding constant actions prior to  $t$ , consider a joint marginal change to  $\alpha_t$  and  $M_t$  would leave  $s_t$  constant for an optimising individual. From the budget constraint, this implies setting a value for  $\frac{dM_t}{d\alpha_t}$  such that:

$$\frac{dc_t}{d\alpha_t} + \frac{dc_t}{dM_t} \frac{dM_t}{d\alpha_t} = \frac{dM_t}{d\alpha_t} \quad (95)$$

where  $\frac{dc_t}{d\alpha_t}$  and  $\frac{dc_t}{dM_t}$  denote optimal responses. So long as  $\frac{dc_t}{dM_t} \neq 1$ , this is possible. But since the consumer optimality condition is:

$$\alpha_t u'(c_t) = \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \quad (96)$$

then so long as the right-hand side is defined, we could only have  $\frac{dc_t}{dM_t} = 1$  (implying  $\frac{ds_t}{dM_t} = 0$ ) in the quasilinear case  $u''(c_t) = 0$ , which has been ruled out by primitive assumptions.

Differentiating (96) with respect to  $\alpha_t$ , given constant savings, yields:

$$\begin{aligned}
u'(c_t) - \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1} + \alpha_t u''(c_t) \left[ \frac{dc_t}{d\alpha_t} + \frac{dc_t}{dM_t} \frac{dM_t}{d\alpha_t} \right] &= 0 \\
u'(c_t) - \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1} + \alpha_t u''(c_t) \frac{dM_t}{d\alpha_t} &= 0
\end{aligned}$$

Rearranging (95):

$$\begin{aligned}\frac{dM_t}{d\alpha_t} &= \frac{\frac{dc_t}{d\alpha_t}}{1 - \frac{dc_t}{dM_t}} \\ &= \frac{\frac{dc_t}{d\alpha_t}}{\frac{ds_t}{dM_t}}\end{aligned}$$

Plugging this into the previous expression, trivial manipulations give (93).

Reasoning in a similar way for (94), consider the effect of a change to  $(1 - T'_t(s_t))$  at the margin, for an agent saving at  $s_t$ , coupled with a change to  $M_t$  that holds constant period- $t$  consumption. That is, set  $\frac{dM_t}{d(1-T'_t(s_t))}$  to solve:

$$\frac{dc_t}{d(1-T'_t(s_t))} + \frac{dc_t}{dM_t} \frac{dM_t}{d(1-T'_t(s_t))} = \frac{dM_t}{d(1-T'_t(s_t))} \quad (97)$$

Differentiating (96) with respect to  $(1 - T'_t(s_t))$  under this joint change gives:

$$\begin{aligned}-\beta R \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} + \alpha_t u''(c_t) \left[ \frac{dc_t}{d(1-T'_t(s_t))} + \frac{dc_t}{dM_t} \frac{dM_t}{d(1-T'_t(s_t))} \right] &= 0 \\ -\alpha_t u'(c_t) \frac{1}{1-T'_t(s_t)} + \alpha_t u''(c_t) \frac{\frac{dc_t}{d(1-T'_t(s_t))}}{\frac{ds_t}{dM_t}} &= 0\end{aligned}$$

Rearranging, and noting  $\frac{dc_t}{d(1-T'_t(s_t))} = -\frac{ds_t}{d(1-T'_t(s_t))}$ :

$$\frac{s_t \varepsilon_t^s}{u'(c_t)} = -\frac{\frac{ds_t}{dM_t}}{u''(c_t)}$$

and so (94) follows, given (93). □

Combining the results in these two sub-Lemmata immediately delivers the first two statements in the main Lemma.

### Conditions 3 and 4

The third statement relates the change in savings at  $t$  to compensated changes in the profile of insurance at  $t + 1$ . It is obtained by constructing offsetting perturbations to the marginal value of saving, based on two expressions for this object that are true in any decentralised allocation. First, from the consumer optimality condition:

$$\alpha_t u'(c_t) = \beta R (1 - T'_t(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} \quad (98)$$



Second, differentiating the relaxed incentive constraint:

$$\alpha_t u'(c_t) \frac{dc_t(\alpha_t)}{d\alpha_t} + u(c_t) + \beta \frac{d\omega_{t+1}(\alpha_t)}{d\alpha_t} = u(c_t) + \beta \frac{1}{\alpha_t} \omega_{t+1}^\Delta(\alpha_t)$$

or:

$$\alpha_t u'(c_t) = \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}} \beta \left\{ \frac{1}{\alpha_t} \omega_{t+1}^\Delta(\alpha_t) - \frac{d\omega_{t+1}(\alpha_t)}{d\alpha_t} \right\} \quad (99)$$

The right-hand sides of (98) and (99) thus give alternative expressions for the shadow value of savings at the chosen allocation. Denote this object  $\gamma_t(\alpha_t)$ , i.e.:

$$\gamma_t(\alpha_t) := R(1 - T_t'(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} \quad (100)$$

$$= \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}} \beta \left\{ \frac{1}{\alpha_t} \omega_{t+1}^\Delta(\alpha_t) - \frac{d\omega_{t+1}(\alpha_t)}{d\alpha_t} \right\} \quad (101)$$

Suppose we are interested in the response of  $c_t(\alpha_t)$  to a generic change in the consumer's constraint set  $\Delta_i$ . We have:

$$\alpha_t u''(c_t(\alpha_t)) \frac{dc_t(\alpha_t)}{d\Delta_i} = \frac{d\gamma_t(\alpha_t)}{d\Delta_i} \quad (102)$$

Consider two such changes,  $\Delta_i$  and  $\Delta_j$ , plus a scalar  $\Gamma$ , with the property:

$$\frac{d\gamma_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{d\gamma_t(\alpha_t)}{d\Delta_j} = 0 \quad (103)$$

Using (103) in (102):

$$\frac{dc_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{dc_t(\alpha_t)}{d\Delta_j} = 0 \quad (104)$$

So long as the perturbations  $\Delta_i$  and  $\Delta_j$  are constructed to leave  $M_t$  unaffected, this last result in turn implies:

$$\frac{ds_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{ds_t(\alpha_t)}{d\Delta_j} = 0 \quad (105)$$

and so

$$\frac{d\gamma_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{d\gamma_t(\alpha_t)}{d\Delta_j} = \left. \frac{d\gamma_t(\alpha_t)}{d\Delta_i} \right|_{s_t, c_t} + \Gamma \left. \frac{d\gamma_t(\alpha_t)}{d\Delta_j} \right|_{s_t, c_t} = 0 \quad (106)$$

– the notation on the right-hand side denoting that the derivative can be taken under the greatly simplifying assumption of fixed savings and consumption in  $t$ . For any pair of differential changes, (102) gives:

$$\begin{aligned}\frac{dc_t(\alpha_t)}{d\Delta_i} &= -\Gamma \frac{dc_t(\alpha_t)}{d\Delta_j} \\ &= \frac{\left. \frac{dy_t(\alpha_t)}{d\Delta_i} \right|_{s_t}}{\left. \frac{dy_t(\alpha_t)}{d\Delta_j} \right|_{s_t}} dc_t(\alpha_t)\end{aligned}\quad (107)$$

We now take the derivatives of  $y_t(\alpha_t)$  for two changes to the consumer's budget constraint, as viewed in  $t$ . The first is a simple change to contemporaneous post-tax returns,  $(1 - T'_t(s_t))$ . From (100):

$$\begin{aligned}\left. \frac{dy_t(\alpha_t)}{d(1 - T'_t(s_t))} \right|_{s_t} &= \frac{1}{1 - T'_t(s_t)} y_t(\alpha_t) \\ &= \frac{1}{1 - T'_t(s_t)} \alpha_t u'(c_t)\end{aligned}\quad (108)$$

The second change is a more general perturbation to the nonlinear budget constraint in  $t + 1$ , compensated so that  $\omega_{t+1}$  is left unaffected. This budget constraint can be rewritten as follows:

$$c_{t+1} = M_{t+1} - s(M_{t+2}) \quad (109)$$

where  $s(M_{t+2})$  is defined implicitly for all realised  $M_{t+2}$  values by:

$$M_{t+2} \equiv R[s(M_{t+2}) - T_{t+1}(s(M_{t+2}))] \quad (110)$$

We will focus on perturbations of the form:

$$c_{t+1} = M_{t+1} - s(M_{t+2}) + \Gamma f(s(M_{t+2})) \quad (111)$$

for an arbitrary bounded, a.e. differentiable function  $f$  and scalar  $\Gamma$ . Thus  $\Gamma f(s(M_{t+2}))$  gives the increase in period- $t + 1$  consumption that is made possible by the perturbation when period- $t + 2$  wealth is held constant at  $M_{t+2}$ . The focus of interest will be differential movements in  $\Gamma$  away from zero. Taking the derivative from (101), since  $c_t$  and  $\omega_{t+1}$  are being held constant we can write:

$$\left. \frac{dy_t(\alpha_t)}{d\Gamma} \right|_{s_t, c_t} = \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}} \beta \frac{1}{\alpha_t} \frac{d\omega_{t+1}^\Delta(\alpha_t)}{d\Gamma} \quad (112)$$

Thus the critical object to evaluate is  $\frac{d\omega_{t+1}^\Delta(\alpha_t)}{d\Gamma}$ . The algebraic steps for this are consigned to a Lemma:

**Lemma 6.**  $\frac{d\omega_{t+1}^\Delta(\alpha_t)}{d\Gamma}$  satisfies the following expression:

$$\begin{aligned} \frac{d\omega_{t+1}^\Delta(\alpha_t)}{d\Gamma} = & \int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \left\{ \alpha_{t+1}^2 u'(c_{t+1}) \frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) \right. \\ & \left. - \frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} \int_{\underline{\alpha}}^{\alpha_{t+1}} \left[ \tilde{\alpha}_{t+1} (u'(c_{t+1})) + \tilde{\alpha}_{t+1}^2 u''(c_{t+1}) \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \rho(\tilde{\alpha}_{t+1}|\alpha_t) \pi(\tilde{\alpha}_{t+1}|\alpha_t) d\tilde{\alpha}_{t+1} \right\} d\alpha_{t+1} \\ & + f(s_{t+1}(\bar{\alpha})) \int_{\alpha_{t+1}} \left\{ \alpha_{t+1} (u'(c_{t+1})) + \alpha_{t+1}^2 u''(c_{t+1}) \frac{dc_{t+1}}{d\alpha_{t+1}} \right\} \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \end{aligned} \quad (113)$$

*Proof.* We obtain the result by combining the income and substitution effects of the perturbation. The size of the income effect will be proportional to the increase in  $c_{t+1}$  at each  $M_{t+2}$  along the budget constraint, given by:

$$\left. \frac{dc_{t+1}}{d\Gamma} \right|_{M_{t+2}} = f(s(M_{t+2})) \quad (114)$$

The size of the substitution effect will be proportional to the change in the slope of the budget constraint for each  $M_{t+2}$ . This slope is given by:

$$\frac{dc_{t+1}}{dM_{t+2}} = -s'(M_{t+2}) (1 - \Gamma f'(s(M_{t+2}))) \quad (115)$$

$$= -\frac{1 - \Gamma f'(s(M_{t+2}))}{R [1 - T'_{t+1}(s(M_{t+2}))]} \quad (116)$$

The effect on this as  $\Gamma$  changes is:

$$\frac{d}{d\Gamma} \left[ \frac{dc_{t+1}}{dM_{t+2}} \right] = \frac{f'(s(M_{t+2}))}{R [1 - T'_{t+1}(s(M_{t+2}))]} \quad (117)$$

Notice that this is equal to:

$$f'(s(M_{t+2})) (1 - T'_{t+1}(s(M_{t+2}))) \frac{d}{d(1 - T'_{t+1}(s(M_{t+2})))} \left[ \frac{dc_{t+1}}{dM_{t+2}} \right] \Big|_{\Gamma=0} \quad (118)$$

i.e. the perturbation has an equivalent effect on the slope of the budget constraint at  $M_{t+2}$  to an increase in the post-tax rate of return by the proportional amount  $f'(s(M_{t+2}))$ .

Substitution effects have an impact on  $\omega_{t+1}^\Delta$  to the extent that consumption is deferred:

$$\begin{aligned}
\left. \frac{d\omega_{t+1}^\Delta}{d\Gamma} \right|_{\text{sub}} &= \int_{\alpha_{t+1}} \left\{ -\alpha_{t+1} u'(c_{t+1}) + \beta R (1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))) \int_{\alpha_{t+1}} V_M(\alpha_{t+1}) \frac{d\pi(\alpha_{t+2}|\alpha_{t+1})}{d\alpha_{t+1}} d\alpha_{t+2} \right\} \\
&\quad \times \frac{ds_{t+1}(\alpha_{t+1})}{d(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1})))} (1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))) f'(s_{t+1}(\alpha_{t+1})) \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \\
&= \int_{\alpha_{t+1}} \left\{ -\alpha_{t+1}^2 u'(c_{t+1}) \frac{dc_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} \frac{1}{\varepsilon_{t+1}^s s_{t+1}(\alpha_{t+1})} \right\} \\
&\quad \times \frac{ds_{t+1}(\alpha_{t+1})}{d(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1})))} (1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))) f'(s_{t+1}(\alpha_{t+1})) \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \\
&= - \int_{\alpha_{t+1}} \alpha_{t+1}^2 u'(c_{t+1}) \frac{dc_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \tag{119}
\end{aligned}$$

where the intermediate line makes use of Lemma 5, and we have used the fact that the specified perturbation is the equivalent of a change in  $(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1})))$  by  $(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))) f'(s_{t+1}(\alpha_{t+1}))$  units.

Similarly, the income effect on  $\omega_{t+1}^\Delta$  will be:

$$\begin{aligned}
\left. \frac{d\omega_{t+1}^\Delta}{d\Gamma} \right|_{\text{inc}} &= \int_{\alpha_{t+1}} f(s_{t+1}(\alpha_{t+1})) \rho(\alpha_{t+1}|\alpha_t) \left\{ \alpha_{t+1} (u'(c_{t+1})) + \frac{ds_{t+1}(\alpha_{t+1})}{dM_{t+1}} [-\alpha_{t+1} u'(c_{t+1}) \right. \\
&\quad \left. + \beta R (1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))) \int_{\alpha_{t+1}} V_M(\alpha_{t+1}) \frac{d\pi(\alpha_{t+2}|\alpha_{t+1})}{d\alpha_{t+1}} d\alpha_{t+2} \right\} \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \\
&= \int_{\alpha_{t+1}} f(s_{t+1}(\alpha_{t+1})) \left\{ \alpha_{t+1} (u'(c_{t+1})) + \alpha_{t+1}^2 u''(c_{t+1}) \frac{dc_{t+1}}{d\alpha_{t+1}} \right\} \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \\
&= f(s_{t+1}(\bar{\alpha})) \int_{\alpha_{t+1}} \left\{ \alpha_{t+1} (u'(c_{t+1})) + \alpha_{t+1}^2 u''(c_{t+1}) \frac{dc_{t+1}}{d\alpha_{t+1}} \right\} \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \tag{120} \\
&\quad - \int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} \\
&\quad \times \left\{ \int_{\underline{\alpha}}^{\alpha_{t+1}} \left[ \tilde{\alpha}_{t+1} (u'(c_{t+1})) + \tilde{\alpha}_{t+1}^2 u''(c_{t+1}) \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \rho(\tilde{\alpha}_{t+1}|\alpha_t) \pi(\tilde{\alpha}_{t+1}|\alpha_t) d\tilde{\alpha}_{t+1} \right\} d\alpha_{t+1}
\end{aligned}$$

where the second equality makes use of (93). Taking substitution and income effects, (119) and (120), together, we obtain the result.  $\square$

Applying (107), (108) and (112), we have:

$$\frac{dc_t}{d\Gamma} = \left[ \frac{\alpha_t u'(c_t)}{(1 - T'_t(s_t))} \right]^{-1} \frac{dc_t}{d(1 - T'_t(s_t))} \frac{1}{\frac{dc_t}{d\alpha_t}} \beta \frac{1}{\alpha_t} \frac{d\omega_{t+1}^\Delta}{d\Gamma} \tag{121}$$

So:

$$\begin{aligned}
\frac{1}{s_t} \frac{ds_t}{d\Gamma} &= \left[ \frac{\alpha_t u'(c_t)}{(1 - T'_t(s_t))} \right]^{-1} \frac{1}{s_t} \frac{ds_t}{d(1 - T'_t(s_t))} \frac{1}{\frac{dc_t}{d\alpha_t}} \beta \frac{1}{\alpha_t} \frac{d\omega_{t+1}^\Delta}{d\Gamma} \\
&= \int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \varepsilon_t^s \frac{\alpha_{t+1} \frac{ds_{t+1}}{d\alpha_{t+1}}}{\alpha_t \frac{ds_t}{d\alpha_t}} \left\{ -\frac{\beta \alpha_{t+1} u'(c_{t+1})}{\alpha_t u'(c_t)} \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) \right. \\
&\quad \left. + \frac{1}{\alpha_{t+1}} \int_{\underline{\alpha}}^{\alpha_{t+1}} \frac{\beta \tilde{\alpha}_{t+1} (u'(c_{t+1}))}{\alpha_t u'(c_t)} \left[ 1 + \frac{c_{t+1} u''(c_{t+1})}{u'(c_{t+1})} \frac{\tilde{\alpha}_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \rho(\tilde{\alpha}_{t+1}|\alpha_t) \pi(\tilde{\alpha}_{t+1}|\alpha_t) d\tilde{\alpha}_{t+1} \right\} d\alpha_{t+1} \\
&\quad + f(s_{t+1}(\bar{\alpha})) \varepsilon_t^s \frac{1}{\alpha_t \frac{dc_t}{d\alpha_t}} \int_{\alpha_{t+1}} \frac{\beta \alpha_{t+1} u'(c_{t+1})}{\alpha_t u'(c_t)} \left[ 1 + \frac{c_{t+1} u''(c_{t+1})}{u'(c_{t+1})} \frac{\alpha_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\alpha_{t+1}} \right] \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1}
\end{aligned} \tag{122}$$

A unit change in  $\Gamma$  changes the slope of the  $t+1$  budget constraint at  $s_{t+1}$  by  $(1 - T'_{t+1}(s_{t+1})) f'(s_{t+1})$  units, and shifts it uniformly for all higher savings levels by the same amount. Consistent with the definition in Appendix A.7, we have:

$$\begin{aligned}
\frac{1}{s_t} \frac{ds_t}{d\Gamma} &\equiv \int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \varepsilon_{t,t+1}(s_{t+1}(\alpha_{t+1})) \Pi(\alpha_{t+1}|\alpha_t) \left( -\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} \right) d\alpha_{t+1} \\
&\quad + f(s_{t+1}(\bar{\alpha})) \frac{1}{s_t} \frac{ds_t}{dM_{t+1}} \Big|_{\text{comp}}
\end{aligned} \tag{123}$$

where  $\frac{ds_t}{dM_{t+1}} \Big|_{\text{comp}}$  denotes the effect on  $s_t$  of a compensated, uniform income increase at  $t+1$ . Since  $\varepsilon_{t,t+1}$  is independent of the choice of  $f'$ , we have:

$$\varepsilon_{t,t+1}(s_{t+1}(\alpha_{t+1})) = -\varepsilon_t^s \frac{\alpha_{t+1}}{\Pi(\alpha_{t+1}|\alpha_t) \alpha_t \frac{ds_t}{d\alpha_t}} \left\{ -\frac{\beta \alpha_{t+1} u'(c_{t+1})}{\alpha_t u'(c_t)} \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) \right. \tag{124}$$

$$\begin{aligned}
&\quad \left. + \frac{1}{\alpha_{t+1}} \int_{\underline{\alpha}}^{\alpha_{t+1}} \frac{\beta \tilde{\alpha}_{t+1} (u'(c_{t+1}))}{\alpha_t u'(c_t)} \left[ 1 + \frac{c_{t+1} u''(c_{t+1})}{u'(c_{t+1})} \frac{\tilde{\alpha}_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \right. \\
&\quad \left. \times \rho(\tilde{\alpha}_{t+1}|\alpha_t) \pi(\tilde{\alpha}_{t+1}|\alpha_t) d\tilde{\alpha}_{t+1} \right\}
\end{aligned} \tag{125}$$

So:

$$RT'_t (s_t) s_t \epsilon_{t,t+1} (s_{t+1} (\alpha_{t+1})) \quad (126)$$

$$= -RT'_t (s_t) s_t \epsilon_t^s \frac{\alpha_{t+1}}{\Pi (\alpha_{t+1} | \alpha_t) \alpha_t \frac{ds_t}{d\alpha_t}} \left\{ -\frac{\beta \alpha_{t+1} u' (c_{t+1})}{\alpha_t u' (c_t)} \rho (\alpha_{t+1} | \alpha_t) \pi (\alpha_{t+1} | \alpha_t) \right. \quad (127)$$

$$\left. + \frac{1}{\alpha_{t+1}} \int_{\underline{\alpha}}^{\alpha_{t+1}} \frac{\beta \tilde{\alpha}_{t+1} (u' (c_{t+1}))}{\alpha_t u' (c_t)} \left[ 1 + \frac{c_{t+1} u'' (c_{t+1})}{u' (c_{t+1})} \frac{\tilde{\alpha}_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \rho (\tilde{\alpha}_{t+1} | \alpha_t) \pi (\tilde{\alpha}_{t+1} | \alpha_t) d\tilde{\alpha}_{t+1} \right\}$$

$$= \frac{\lambda_{t+1}^\Delta}{\eta_t} \alpha_t u' (c_t) \left\{ -\frac{\beta R \alpha_{t+1} u' (c_{t+1})}{\alpha_t u' (c_t)} \rho (\alpha_{t+1} | \alpha_t) \frac{\alpha_{t+1} \pi (\alpha_{t+1} | \alpha_t)}{\Pi (\alpha_{t+1} | \alpha_t)} \right. \quad (128)$$

$$\left. + \frac{1}{\Pi (\alpha_{t+1} | \alpha_t)} \int_{\underline{\alpha}}^{\alpha_{t+1}} \frac{\beta R \tilde{\alpha}_{t+1} (u' (c_{t+1}))}{\alpha_t u' (c_t)} \left[ 1 + \frac{c_{t+1} u'' (c_{t+1})}{u' (c_{t+1})} \frac{\tilde{\alpha}_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \rho (\tilde{\alpha}_{t+1} | \alpha_t) \pi (\tilde{\alpha}_{t+1} | \alpha_t) d\tilde{\alpha}_{t+1} \right\}$$

$$= -\beta R \alpha_{t+1} u' (c_{t+1}) \rho (\alpha_{t+1} | \alpha_t) \frac{\lambda_{t+1}^\Delta}{\eta_t} \frac{\alpha_{t+1} \pi (\alpha_{t+1} | \alpha_t)}{\Pi (\alpha_{t+1} | \alpha_t)} \quad (129)$$

$$+ \frac{1}{\Pi (\alpha_{t+1} | \alpha_t)} \frac{\lambda_{t+1}^\Delta}{\eta_t} \int_{\underline{\alpha}}^{\alpha_{t+1}} \beta R \tilde{\alpha}_{t+1} (u' (c_{t+1})) \left[ 1 + \frac{c_{t+1} u'' (c_{t+1})}{u' (c_{t+1})} \frac{\tilde{\alpha}_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \rho (\tilde{\alpha}_{t+1} | \alpha_t) \pi (\tilde{\alpha}_{t+1} | \alpha_t) d\tilde{\alpha}_{t+1}$$

Changing the unit of integration and using  $\pi (\alpha_{t+1} | \alpha_t) \frac{d\alpha_{t+1}(c_{t+1})}{dc_{t+1}} = \pi^c (c_{t+1} | \alpha^t)$  gives condition 3.

For condition 4, we have established:

$$\frac{1}{s_t} \frac{ds_t}{dM_{t+1}} \Big|_{\text{comp}} = \epsilon_t^s \frac{1}{\alpha_t \frac{dc_t}{d\alpha_t}} \int_{\alpha_{t+1}} \frac{\beta \alpha_{t+1} u' (c_{t+1})}{\alpha_t u' (c_t)} \left[ 1 + \frac{c_{t+1} u'' (c_{t+1})}{u' (c_{t+1})} \frac{\alpha_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\alpha_{t+1}} \right] \rho (\alpha_{t+1} | \alpha_t) \pi (\alpha_{t+1} | \alpha_t) d\alpha_{t+1} \quad (130)$$

So:

$$RT'_t (s_t) \frac{ds_t}{dM_{t+1}} \Big|_{\text{comp}} = \frac{\lambda_{t+1}^\Delta}{\eta_t} \alpha_t u' (c_t) \int_{\alpha_{t+1}} \frac{\beta R \alpha_{t+1} u' (c_{t+1})}{\alpha_t u' (c_t)} \left[ 1 + \frac{c_{t+1} u'' (c_{t+1})}{u' (c_{t+1})} \frac{\alpha_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\alpha_{t+1}} \right] \rho (\alpha_{t+1} | \alpha_t) \pi (\alpha_{t+1} | \alpha_t) d\alpha_{t+1} \quad (131)$$

$$= \frac{\lambda_{t+1}^\Delta}{\eta_t} \int_{\alpha_{t+1}} \beta R \alpha_{t+1} u' (c_{t+1}) \left[ 1 + \frac{c_{t+1} u'' (c_{t+1})}{u' (c_{t+1})} \frac{\alpha_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\alpha_{t+1}} \right] \rho (\alpha_{t+1} | \alpha_t) \pi (\alpha_{t+1} | \alpha_t) d\alpha_{t+1} \quad (132)$$

Again, a change to the unit of integration gives the result.

## A.9 Proof of Theorem 2

Equation (31), above, gives:

$$T'_t (s_t) \frac{ds_t}{dM_t} = -\frac{\lambda_{t+1}^\Delta (\alpha_t)}{\eta_t} (\alpha_t)^2 u'' (c_t) \left( \frac{dc_t}{d\alpha_t} \right)$$

The utility function is time-separable and concave in consumption at each date-state, which together straightforwardly imply  $\frac{ds_t}{dM_t} > 0$ . Concavity further gives  $u'' (c_t) < 0$ , the strict increasingness assumption implies  $\frac{dc_t}{d\alpha_t} > 0$ ,

and  $\eta > 0$  from (19). It follows that  $T'_t(s_t)$  has the same sign as  $\lambda_{t+1}^\Delta(\alpha_t(c_t))$ , and so we focus on signing the latter object.

To demonstrate positive taxes at interior points we start with the following Lemma:

**Lemma 7.** *For all  $t$  and  $\alpha^{t-1}$ , the long-run expectation of variation in the inverse, discounted marginal utility of consumption is bounded, i.e.:*

$$\lim_{s \rightarrow \infty} \left\{ \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \bar{\alpha}))} - \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \underline{\alpha}))} \right] \right\} < \infty$$

*Proof.* We have just established that  $\lambda_{s+1}^\Delta = 0$  for  $\alpha_s = \bar{\alpha}$  and  $\alpha_s = \underline{\alpha}$ , and so:

$$\begin{aligned} & \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \bar{\alpha}))} - \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \underline{\alpha}))} \right] \\ &= \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} \mathbb{E}_s[\alpha_{s+1}|\bar{\alpha}]} \mathbb{E}_s \left[ \frac{1}{\beta R u'(c_{s+1}(\alpha'_t, \dots, \bar{\alpha}, \alpha_{s+1}))} \middle| \bar{\alpha} \right] \right. \\ & \quad \left. - \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} \mathbb{E}_s[\alpha_{s+1}|\underline{\alpha}]} \mathbb{E}_s \left[ \frac{1}{\beta R u'(c_{s+1}(\alpha'_t, \dots, \underline{\alpha}, \alpha_{s+1}))} \middle| \underline{\alpha} \right] \right] \right] \end{aligned} \quad (133)$$

If the right-hand side of (133) is unbounded in  $s$  then so is the expression:

$$\begin{aligned} & \frac{1}{(\beta R)^{s-t}} \mathbb{E}_t \left[ \mathbb{E}_s \left[ \frac{1}{\beta R u'(c_{s+1}(\alpha'_t, \dots, \bar{\alpha}, \alpha_{s+1}))} \middle| \bar{\alpha} \right] - \frac{\alpha \mathbb{E}_s[\alpha_{s+1}|\bar{\alpha}]}{\bar{\alpha} \mathbb{E}_s[\alpha_{s+1}|\underline{\alpha}]} \mathbb{E}_s \left[ \frac{1}{\beta R u'(c_{s+1}(\alpha'_t, \dots, \underline{\alpha}, \alpha_{s+1}))} \middle| \underline{\alpha} \right] \right] \\ &= \frac{1}{(\beta R)^{s+1-t}} \mathbb{E}_t \left[ \int_{\alpha_{s+1}} \left[ \frac{1}{u'(c_{s+1}(\alpha'_t, \dots, \bar{\alpha}, \alpha_{s+1}))} \pi(\alpha_{s+1}|\bar{\alpha}) - \frac{1}{u'(c_{s+1}(\alpha'_t, \dots, \underline{\alpha}, \alpha_{s+1}))} \frac{\alpha \mathbb{E}_s[\alpha_{s+1}|\bar{\alpha}]}{\bar{\alpha} \mathbb{E}_s[\alpha_{s+1}|\underline{\alpha}]} \pi(\alpha_{s+1}|\underline{\alpha}) \right] d\alpha_{s+1} \right] \end{aligned} \quad (134)$$

By normality, if this is unbounded in  $s$  then so too is the object:

$$\frac{1}{(\beta R)^{s+1-t}} \mathbb{E}_t \left[ \int_{\alpha_{s+1}} \frac{1}{u'(c_{s+1}(\alpha'_t, \dots, \underline{\alpha}, \alpha_{s+1}))} \left[ \frac{\pi(\alpha_{s+1}|\bar{\alpha}) - \pi(\alpha_{s+1}|\underline{\alpha}) \frac{\alpha \mathbb{E}_s[\alpha_{s+1}|\bar{\alpha}]}{\bar{\alpha} \mathbb{E}_s[\alpha_{s+1}|\underline{\alpha}]} }{\pi(\alpha_{s+1}|\underline{\alpha})} \right] \pi(\alpha_{s+1}|\underline{\alpha}) d\alpha_{s+1} \right] \quad (135)$$

Moreover, continuity of the density implies that the object:

$$\left[ \frac{\pi(\alpha_{s+1}|\bar{\alpha}) - \pi(\alpha_{s+1}|\underline{\alpha}) \frac{\alpha \mathbb{E}_s[\alpha_{s+1}|\bar{\alpha}]}{\bar{\alpha} \mathbb{E}_s[\alpha_{s+1}|\underline{\alpha}]} }{\pi(\alpha_{s+1}|\underline{\alpha})} \right]$$

is bounded in  $\alpha_{s+1}$ . Thus unboundedness of (135) in  $s$  implies:

$$\lim_{s \rightarrow \infty} \left\{ \frac{1}{(\beta R)^{s+1-t}} \mathbb{E}_t \left[ \int_{\alpha_{s+1}} \frac{1}{u'(c_{s+1}(\alpha'_t, \dots, \underline{\alpha}, \alpha_{s+1}))} \pi(\alpha_{s+1}|\underline{\alpha}) d\alpha_{s+1} \right] \right\} = \infty$$

This implies:

$$\lim_{s \rightarrow \infty} \left\{ \frac{1}{(\beta R)^{s-t}} \mathbb{E}_t \left[ \frac{\underline{\alpha}}{\mathbb{E}_s [\alpha_{s+1} | \underline{\alpha}]} \mathbb{E}_s \left[ \frac{1}{\beta R u' (c_{s+1} (\alpha'_t, \dots, \underline{\alpha}, \alpha_{s+1}))} \middle| \underline{\alpha} \right] \right] \right\} = \infty$$

and thus, since  $\lambda_{s+1}^\Delta (\underline{\alpha}) = 0$ :

$$\lim_{s \rightarrow \infty} \left\{ \frac{1}{(\beta R)^{s-t}} \mathbb{E}_t \left[ \frac{1}{u' (c_s (\alpha'_t, \dots, \underline{\alpha}))} \right] \right\} = \infty$$

But since consumption is increasing in  $\alpha_s$ , we have:

$$\frac{1}{(\beta R)^{s-t}} \mathbb{E}_t \left[ \frac{1}{u' (c_s (\alpha'_t, \dots, \underline{\alpha}))} \right] \leq \frac{1}{(\beta R)^{s-t}} \mathbb{E}_t \left[ \frac{1}{u' (c_s (\alpha'_t, \dots, \alpha_s))} \right]$$

Thus:

$$\lim_{s \rightarrow \infty} \left\{ \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u' (c_s (\alpha'_t, \dots, \tilde{\alpha}))} - \frac{1}{(\beta R)^{s-t} u' (c_s (\alpha'_t, \dots, \underline{\alpha}))} \right] \right\} = \infty$$

must imply

$$\lim_{s \rightarrow \infty} \left\{ \frac{1}{\mathbb{E}_t [\alpha_s]} \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u' (c_s)} \right] \right\} = \infty$$

But:

$$\begin{aligned} \lim_{s \rightarrow \infty} \left\{ \frac{1}{\mathbb{E}_t [\alpha_s]} \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u' (c_s)} \right] \right\} &= \frac{1 + \lambda_{t+1} (\alpha_t)}{\eta_t} \\ &= \frac{1}{1 - \varepsilon^\alpha (\alpha_t)} \left\{ \frac{1}{\mathbb{E}_t [\alpha_{t+1}]} \mathbb{E}_t \left[ \frac{1}{\beta R u' (c_{t+1})} \right] - \frac{\varepsilon^\alpha (\alpha_t)}{\alpha_t u' (c_t)} \right\} \end{aligned} \quad (136)$$

which is finite, since the resource constraint rules out infinite expected consumption. Thus we have a contradiction.  $\square$

Turning to the main argument, a combination of (52) and (53) gives:

$$\lambda_{t+1}^\Delta (\alpha^t) - \lambda_t^\Delta (\alpha^{t-1}) \rho (\alpha_t | \alpha_{t-1}) = - \frac{1}{\alpha_t \pi (\alpha_t | \alpha_{t-1})} \int_{\alpha_t}^{\tilde{\alpha}} [(1 + \lambda_{t+1} (\tilde{\alpha}_t)) - (1 + \lambda_t)] \pi (\tilde{\alpha}_t | \alpha_{t-1}) d\tilde{\alpha}_t$$

Or, using the definitions of  $\rho$  and  $\pi^\Delta$ :

$$\lambda_{t+1}^\Delta (\alpha_t) \alpha_t \pi (\alpha_t | \alpha_{t-1}) = \int_{\alpha_t}^{\tilde{\alpha}} \left[ (1 + \lambda_t + \lambda_t^\Delta \pi^\Delta (\tilde{\alpha}_t | \alpha_{t-1})) - (1 + \lambda_{t+1} (\tilde{\alpha}_t)) \right] \pi (\tilde{\alpha}_t | \alpha_{t-1}) d\tilde{\alpha}_t \quad (137)$$

with:

$$\int_{\underline{\alpha}}^{\tilde{\alpha}} \left[ (1 + \lambda_t + \lambda_t^\Delta \pi^\Delta (\alpha_t | \alpha_{t-1})) - (1 + \lambda_{t+1} (\alpha_t)) \right] \pi (\alpha_t | \alpha_{t-1}) d\alpha_t = 0 \quad (138)$$

Since  $\pi^\Delta (\alpha_t | \alpha_{t-1})$  is monotone increasing in  $\alpha_t$ , and  $\lambda_0^\Delta = 0$ , a sufficient condition for the right-hand side of (137) to



be weakly positive for all  $t$  and all histories is that  $(1 + \lambda_{t+1}(\alpha_t))$  should be non-increasing in  $\alpha_t$ . Lemma 7 enables this to be established. We have:

$$\begin{aligned} \frac{1 + \lambda_{t+1}(\alpha_t)}{\eta_t} &= \lim_{s \rightarrow \infty} \left\{ \frac{1 + \lambda_{t+1}(\alpha_t)}{\eta_t} + \frac{\mathbb{E}_t [D_{t,s}(\alpha^s) \alpha_s]}{\mathbb{E}_t [\alpha_s]} \frac{1 + \lambda_{t+1}^\Delta(\alpha_t)}{\eta_t} \right\} \\ &= \lim_{s \rightarrow \infty} \left\{ \frac{1}{\mathbb{E}_t [\alpha_s]} \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s)} \right] \right\} \end{aligned}$$

where convergence is uniform across  $\alpha_t$ , from the first line:  $\frac{\mathbb{E}_t [D_{t,s}(\alpha^s) \alpha_s]}{\mathbb{E}_t [\alpha_s]} \leq \bar{\rho}^{s-t}$  where  $\bar{\rho} = \sup_{\alpha, \alpha'} [\rho(\alpha'|\alpha)] < 1$ .

Thus we wish to show non-increasingness in the object:

$$\lim_{s \rightarrow \infty} \left\{ \frac{1}{\mathbb{E}_t [\alpha_s]} \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s)} \right] \right\}$$

Clearly  $\mathbb{E}_t [\alpha_s]$  is weakly increasing in  $\alpha_t$ , so a sufficient condition is that  $\mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s)} \right]$  is non-increasing at the limit.

Consider the difference:

$$\mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha_t'', \dots, \alpha_s))} \middle| \alpha_t'' \right] - \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha_t', \dots, \alpha_s))} \middle| \alpha_t' \right] \quad (139)$$

for  $\alpha_t'' > \alpha_t'$ . We wish to show that this is weakly negative. We have:

$$\begin{aligned} &\mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha_t'', \dots, \alpha_s))} \middle| \alpha_t'' \right] - \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha_t', \dots, \alpha_s))} \middle| \alpha_t' \right] \\ = &\mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha_t'', \dots, \alpha_s))} \middle| \alpha_t'' \right] - \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha_t', \dots, \alpha_s))} \middle| \alpha_t'' \right] \\ &+ \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha_t', \dots, \alpha_s))} \middle| \alpha_t'' \right] - \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha_t', \dots, \alpha_s))} \middle| \alpha_t' \right] \end{aligned}$$

The first term is weakly negative, by normality. The second can be rewritten:

$$\begin{aligned}
& \int_{\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \alpha_s))} \right] \pi(\alpha_s | \alpha_{s-1}) d\alpha_s \dots \pi(\alpha_{t+2} | \alpha_{t+1}) d\alpha_{t+2} \right\} [\pi(\alpha_{t+1} | \alpha'_t) - \pi(\alpha_{t+1} | \alpha_t)] d\alpha_{t+1} \\
&= \int_{\alpha'_t}^{\alpha''_t} \int_{\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \alpha_s))} \right] \pi(\alpha_s | \alpha_{s-1}) d\alpha_s \dots \pi(\alpha_{t+2} | \alpha_{t+1}) d\alpha_{t+2} \right\} \frac{d\pi(\alpha_{t+1} | \alpha_t)}{d\alpha_t} d\alpha_{t+1} d\alpha_t \\
&= \int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t} \int_{\alpha_{t+1}} \alpha_{t+1} \frac{d}{d\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \alpha_s))} \right] \pi(\alpha_s | \alpha_{s-1}) d\alpha_s \dots \pi(\alpha_{t+2} | \alpha_{t+1}) d\alpha_{t+2} \right\} \\
&\quad \times \rho(\alpha_{t+1} | \alpha_t) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} d\alpha_t \\
&= \int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t} \int_{\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \alpha_{t+1} \frac{d}{d\alpha_{t+1}} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \alpha_s))} \right] \pi(\alpha_s | \alpha_{s-1}) d\alpha_s \dots \pi(\alpha_{t+2} | \alpha_{t+1}) d\alpha_{t+2} \right\} \\
&\quad \times \rho(\alpha_{t+1} | \alpha_t) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} d\alpha_t \\
&+ \int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t} \int_{\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \alpha_s))} \right] \pi(\alpha_s | \alpha_{s-1}) d\alpha_s \dots \alpha_{t+1} \frac{d\pi(\alpha_{t+2} | \alpha_{t+1})}{d\alpha_{t+1}} d\alpha_{t+2} \right\} \\
&\quad \times \rho(\alpha_{t+1} | \alpha_t) \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1} d\alpha_t \\
&= \dots \\
&= \int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t} \left\{ \sum_{r=t+1}^s \mathbb{E}_t \left[ D_{t,r}(\alpha^r) \alpha_r \frac{d}{d\alpha_r} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \alpha_r, \dots, \alpha_s))} \right] \Big| \alpha_t \right] \right\} d\alpha_t
\end{aligned}$$

By normality, this object has negative terms except for the period- $s$  entry:

$$\begin{aligned}
& \mathbb{E}_t \left[ D_{t,s}(\alpha^s) \alpha_s \frac{d}{d\alpha_s} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \alpha_s))} \right] \Big| \alpha_t \right] \\
&\leq \int_{\alpha_{t+1}} \cdots \int_{\alpha_s} \bar{\rho}^{s-t} \bar{\alpha} \frac{d}{d\alpha_s} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \alpha_s))} \right] \bar{\pi} d\alpha_s \dots \pi(\alpha_{t+1} | \alpha_t) d\alpha_{t+1}
\end{aligned}$$

where  $\bar{\pi}$  is an upper bound on  $\pi$  (which exists, by continuity and the compactness of  $A$ ). The right hand side of this inequality is equal to:

$$\bar{\rho}^{s-t} \mathbb{E}_t \left[ \bar{\alpha} \bar{\pi} \left[ \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \bar{\alpha}))} - \frac{1}{(\beta R)^{s-t} u'(c_s(\alpha'_t, \dots, \underline{\alpha}))} \right] \Big| \alpha_t \right]$$

Lemma 7 implies that the expectation is finite as  $s \rightarrow \infty$ , and so this object converges to zero as  $s$  becomes large. Since all other components of the difference (139) are weakly negative, uniform convergence guarantees that  $\frac{1+\lambda_{t+1}(\alpha'_t)}{\eta_t} - \frac{1+\lambda_{t+1}(\alpha_t)}{\eta_t} \leq 0$ .

Thus  $1 + \lambda_{t+1}(\alpha_t)$  is weakly decreasing in  $\alpha_t$ , for all  $t$  and  $\alpha^{t-1}$ . This leaves two options:

1.  $\lambda_{t+1}(\alpha'_t) < \lambda_{t+1}(\alpha_t)$  for some  $\alpha'_t < \alpha_t$

2.  $\lambda_{t+1}(\alpha_t)$  is constant in  $\alpha_t$

The first case implies  $\lambda_{t+1}^\Delta(\alpha_t) > 0$  everywhere except endpoints, from (137) and the fact that the integral in (137) is zero over the full range.

It remains to rule out that  $1 + \lambda_{t+1}(\alpha_t)$  is constant in  $\alpha_t$ . If this were true but  $\lambda_t^\Delta > 0$ , we would still have positive taxes except at endpoints, so the case is only problematic for  $\lambda_t^\Delta = 0$  (or  $t = 0$ ), in which case it would imply  $\lambda_{t+1}^\Delta = 0$  everywhere – and so zero taxes for interior values of  $\alpha_t$  at the given node. Suppose this were true. From the definition of  $\lambda_{t+1}^\Delta$ , the implication is:

$$\frac{1}{\alpha_t u'(c_t)} = \frac{1}{\mathbb{E}_t[\alpha_{t+1}]} \mathbb{E}_t \left[ \frac{1}{\beta R u'(c_{t+1})} \right] \quad (140)$$

with both objects constant in  $\alpha_t$ . Suppose for now that types are persistent ( $\rho(\alpha|\alpha') > 0$  for all type pairs). For the right-hand side of (140) to be constant in  $\alpha_t$ , and given normality, a necessary requirement is that the partial derivatives due to persistence are weakly positive for all  $\alpha_t$ :

$$\frac{\int_{\alpha_{t+1}} \frac{1}{u'(c_{t+1})} \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1}}{\mathbb{E}_t \left[ \frac{1}{u'(c_{t+1})} \right]} - \frac{\int_{\alpha_{t+1}} \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1}}{\mathbb{E}_t[\alpha_{t+1}]} \geq 0$$

or:

$$\int_{\alpha_{t+1}} \left\{ \frac{1}{u'(c_{t+1})} - \frac{\alpha_{t+1}}{\mathbb{E}_t[\alpha_{t+1}]} \mathbb{E}_t \left[ \frac{1}{u'(c_{t+1})} \right] \right\} \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \geq 0$$

But since  $\lambda_{t+1}^\Delta = 0$ , and we have already established  $\lambda_s^\Delta \geq 0$  for all  $s$ , condition (65) implies:

$$\int_{\alpha'_{t+1}} \left\{ \frac{1}{u'(c_{t+1})} - \alpha_{t+1} \frac{1 + \lambda_{t+1}}{\eta_t} \right\} \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \leq 0$$

for all  $\alpha'_{t+1}$ , with:

$$\frac{1 + \lambda_{t+1}}{\eta_t} = \frac{1}{\mathbb{E}_t[\alpha_{t+1}]} \mathbb{E}_t \left[ \frac{1}{u'(c_{t+1})} \right]$$

By MLRP,  $\frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t}$  is a strictly increasing function, and so:

$$\int_{\alpha_{t+1}} \left\{ \frac{1}{u'(c_{t+1})} - \frac{\alpha_{t+1}}{\mathbb{E}_t[\alpha_{t+1}]} \mathbb{E}_t \left[ \frac{1}{u'(c_{t+1})} \right] \right\} \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \leq 0$$

with the inequality strict (a contradiction) unless the object in curly brackets is zero everywhere. This in turn would imply that  $\lambda_{t+2}^\Delta = 0$  everywhere. Repeating the argument, this would imply  $\lambda_{t+3}^\Delta = 0$  at all successor nodes, and so on. Thus the only possibility consistent with  $\lambda_{t+1}^\Delta = 0$  at interior points is that  $\lambda^\Delta$  is zero from  $t$  onwards at

all successor nodes. But this implies a first-best allocation, with  $\alpha_t u'(c_t)$  constant over time and histories. This is clearly not incentive-compatible.

It remains to provide equivalent arguments when types are iid. In this case  $\lambda_{t+1}^\Delta(\alpha_t) = 0$  for all  $\alpha_t$  implies:

$$\frac{1}{\alpha_t u'(c_t)} = \frac{1}{\mathbb{E}_t[\alpha_s]} \mathbb{E}_t \left[ \frac{1}{(\beta R)^{s-t} u'(c_s)} \right]$$

for all  $s > t$ , with both sides constant in  $\alpha_t$ . But if the right-hand side is constant in  $\alpha_t$  then future consumption must be constant a.e. at all horizons, which is inconsistent with incentive compatibility, given that period- $t$  consumption must increase strictly in  $\alpha_t$  to keep the left-hand side constant.

## A.10 Proof of Proposition 5

From Lemma 2,  $RT'_{t-1}(s_{t-1}) s_{t-1} \epsilon_{t-1,t}(s'_t)$  is equal to:

$$\begin{aligned} & -\rho(\alpha_t(c'_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} (\alpha_t(c'_t))^2 u'(c'_t) \left( \frac{d\alpha_t(c'_t)}{dc_t} \right)^{-1} \frac{\pi^c(c'_t | \alpha^{t-1})}{\Pi^c(s'_t | \alpha^{t-1})} \\ & + \frac{1}{\Pi^c(s'_t | \alpha^{t-1})} \int_{\underline{c}}^{c'_t} \rho(\alpha_t(c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \left[ \alpha_t(c_t) (u'(c_t)) + (\alpha_t(c_t))^2 u''(c_t) \left( \frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \right] \pi^c(c_t | \alpha^{t-1}) dc_t \end{aligned} \quad (141)$$

Switching to express arguments in terms of  $\alpha_t$ :

$$\begin{aligned} & -\rho(\alpha_t | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} (\alpha_t)^2 u'(c_t(\alpha'_t)) \frac{\pi(\alpha'_t | \alpha_{t-1})}{\Pi(\alpha'_t | \alpha_{t-1})} \\ & + \frac{1}{\Pi(\alpha'_t | \alpha_{t-1})} \int_{\underline{\alpha}}^{\alpha'_t} \rho(\alpha_t | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \left[ \alpha_t (u'(c_t(\alpha_t))) + \alpha_t^2 u''(c_t(\alpha_t)) \frac{dc_t(\alpha_t)}{d\alpha_t} \right] \pi(\alpha_t | \alpha_{t-1}) d\alpha_t \end{aligned} \quad (142)$$

Integration by parts gives the following relationship:

$$\begin{aligned} & \int_{\underline{\alpha}}^{\alpha'_t} \rho(\alpha_t | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \left[ \alpha_t (u'(c_t(\alpha_t))) + \alpha_t^2 u''(c_t(\alpha_t)) \frac{dc_t(\alpha_t)}{d\alpha_t} \right] \pi(\alpha_t | \alpha_{t-1}) d\alpha_t \\ & - \rho(\alpha_t | \alpha_{t-1}) \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} (\alpha_t)^2 u'(c_t(\alpha'_t)) \pi(\alpha'_t | \alpha_{t-1}) \\ & = \beta R \frac{\lambda_t^\Delta}{\eta_{t-1}} \int_{\underline{\alpha}}^{\alpha'_t} \alpha_t u'(c_t(\alpha_t)) \frac{\alpha_{t-1} \frac{d\pi(\alpha_t | \alpha_{t-1})}{d\alpha_{t-1}}}{\pi(\alpha_t | \alpha_{t-1})} \pi(\alpha_t | \alpha_{t-1}) d\alpha_t \end{aligned}$$

So:

$$s_{t-1} \epsilon_{t-1,t}(s'_t) = \beta \frac{\left( \frac{\lambda_t^\Delta}{\eta_{t-1}} \right)}{T'_{t-1}(s_{t-1})} \mathbb{E}_{t-1} \left[ \alpha_t u'(c_t(\alpha_t)) \frac{\alpha_{t-1} \frac{d\pi(\alpha_t | \alpha_{t-1})}{d\alpha_{t-1}}}{\pi(\alpha_t | \alpha_{t-1})} \Big| \alpha_t \leq \alpha'_t \right]$$

The expectation term contains two objects that are monotone increasing in  $\alpha_t$ , under the maintained assumptions (including MLRP). The term  $\frac{d\pi(\alpha_t|\alpha_{t-1})}{\pi(\alpha_t|\alpha_{t-1})}$  is zero in expectation, whilst  $\alpha_t u'(c_t(\alpha_t))$  is strictly positive. Thus for sufficiently low  $\alpha'_t$  (corresponding to high  $s'_t$ ) both sides of the expression must be negative, whilst positive correlation between the components implies it is positive for sufficiently high  $\alpha'_t$  (low  $s'_t$ ). From Lemma 4:

$$\frac{\left(\frac{\lambda'_t}{\eta_{t-1}}\right)}{T'_{t-1}(s_{t-1})} = \left\{ \alpha_t u'(c_t) - \beta R (1 - T'_{t-1}(s_t)) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1} \right\}^{-1} > 0$$

where the final inequality follows a step used in the proof of Lemma 7. The result follows.