UNIVERSITY OF CAMBRIDGE
Faculty of Economics

JANEWAY INSTITUTE CAMBRIDGE

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	Cambridge Working Papers in Economics
2210	Janeway Institute Working Paper Series
Published	25 March 2022
Revised	3 February 2023
Key Words	Nonlinear Taxation, Sufficient Statistics, Mirrleesian Taxation, New Dynamic Public Finance
JEL Codes	D82, E21, E61 H21, H24, H30
Websites	www.econ.cam.ac.uk/cwpe www.janeway.econ.cam.ac.uk/working-papers

Optimal Nonlinear Savings Taxation*

Charles Brendon†

February 3, 2023

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.

*First draft: October 2021. This work has benefited from discussions and comments from Vasco Carvalho, Paweł Doligalski, Lukas Freund, Andrew Hannon, Christian Hellwig, David Martimort, Ludvig Sinander and Nicolas Werquin, together with seminar participants in Bristol, Cambridge, EUI and TSE. The first draft was completed when I was a visitor at TSE: I am very grateful for the hospitality shown. Richard Nickl provided important literary motivation.

†Faculty of Economics & Queens' College, University of Cambridge, Toulouse School of Economics, and CEPR. Email: cfb46@cam.ac.uk

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

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

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 provide sufficient simplifications for optimal nonlinear taxation to be characterised by a small number of conventional behavioural elasticities, despite the *a priori* complexity of an infinite-horizon, incomplete market setting.

As a first contribution, I show that a large class of incentive-feasible allocations in the Atkeson-Lucas setting can be decentralised in a simple market economy, in which individuals make period-by-period choices between consumption and savings. Relative to a conventional market economy, the only additional component is a nonlinear tax on each period's savings. The decentralisation is possible for any allocation that gives greater future consumption to individuals who choose lower current consumption. This amounts to saying that I focus on allocations where future consumption is increasing in current savings.

The main focus of the paper is on the optimal determination of the savings tax schedule. I show that it is generically optimal to set a positive marginal tax rate on savings in all time periods and for all shock histories – 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 be able to afford under a laissez-faire system.

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.[1] A number of writers have expressed scepticism in recent years about the practical relevence 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.[2] The results here provide a counter-point to this. The characterisation of optimal policy that I derive 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, I 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 keeps the relevant behavioural considerations for policy design to a manageable level.

The important general lesson, I argue, is that a mechanism design approach can imply structural

[1]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.

[2]See, for instance, the discussions in Diamond and Saez (2011), Piketty and Saez (2013b), and Stantcheva (2020).

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.

1.1 Preview of main characterisation

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

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

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, $T'\left(s\right)\frac{ds}{dM}$, less a social welfare weight, $g\left(s\right)$, that captures the welfare value of transfering income to an individual whose savings are s. The welfare weight – an endogenous object that evolves with individuals' wealth levels – varies across types with a common history in proportion to their marginal utility of consumption alone. Thus the policymaker is conditionally utilitarian. It 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 ε^{s}. This raises revenue in proportion to the marginal tax rate $T'\left(s'\right)$, and to the measure of types at s', relative to those above: $a\left(s'\right)$ 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 $\epsilon^{s}_{-1}\left(s'\right)$. This raises additional income in the previous

[3] Some notation, including time indexation, is dropped for simplicity.

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 $\epsilon^{s}_{-1}(s')$ 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 relevance is intrinsically linked to type persistence in the underlying structural model. It is zero when types are iid.

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 I 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), I 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 I 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 intermediate 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 I 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 most of the literature that explores the sensitivity of their immiseration result.[4] 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

[4] See, for instance, Sleet and Yeltekin (2006) and Farhi and Werning (2007). Recent work by Bloedel, Krishna and Leukhina (2021) considers persistent shocks in the original Thomas-Worral environment, again retaining a finite type set.

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 constrained-optimal allocations and laissez-faire outcomes, rather than focusing directly on tax instruments.[5] 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.[6] 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

[5] 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.

[6] The inverse Euler condition had previously been derived in different settings by Diamond and Mirrlees (1978) and Rogerson (1985).

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.

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.[7] 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*.[8] 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.[9] 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.

[7] See, for instance, Scheuer and Werning (2017) and Ferey, Lockwood and Taubinsky (2021). Kleven (2021) provides a broad discussion of the sufficient statistics approach.

[8] Stantcheva (2020) gives an excellent summary of the approach.

[9] 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.

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 assult of Straub and Werning (2020),[10] and the earlier findings that savings taxes could play a useful role in computational Ramsey environments.[11] 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.[12] Current work by Bloedel, Krishna and Strulovici (2020), Bloedel, Krishna and Leukhina (2021) 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 the analysis.

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 policmaker whose role is to provide some

[10]Chari, Nicolini and Teles (2020) and Greulich, Laczo and Marcet (2022) explore the limitations of the Straub-Werning results.

[11]For instance, Aiyagari (1995), İmrohoroğlu (1998), Erosa and Gervais (2002) and Conesa, Kitao and Krueger (2009).

[12]Pavan (2017) surveys this literature in detail.

insurance mechanism against the taste shocks that consumers face period-by-period.

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:

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

where c_t is consumption in period t, $\beta \in (0,1)$ is the discount factor, $u : \mathbb{R}^+ \to \mathbb{R}$ or $\mathbb{R}^{++} \to \mathbb{R}$ is the period utility function, and $\alpha_t \in \left[\underline{\alpha}, \bar{\alpha}\right] \subset \mathbb{R}^+$ 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 $\left[\underline{\alpha}, \bar{\alpha}\right]$ 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). I 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 α^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 $\Pi\left(\alpha_{t+1}|\alpha_t\right)$, with conditional density $\pi\left(\alpha_{t+1}|\alpha_t\right)$. The equivalent (unconditional) objects for period 0 are denoted $\Pi\left(\alpha_0\right)$ and $\pi\left(\alpha_0\right)$ respectively. We place the following reguarity structure on the distributions:

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

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 α^t will lie in S. This is induced by Π in the obvious way. $\mathbb{E}_s$ denotes period-s conditional expectations of a future variable under this process, given an α^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:

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

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 α_t by comparison with log changes in α_{t+1}:[13]

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

Integrating provides a link between the two preceding objects:

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

Persistence notwithstanding, higher values of α 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 α 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 $D_{t,s}(\alpha^s)$:

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

[13]Note that in the lognormal case, where:

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

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

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

Related to ρ is the elasticity of the density with respect to lagged type, denoted π^{Δ}:

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

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

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

Finally, I impose a standard monotone likelihood condition on π. 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 α_t', α_t'' with $\alpha_t' < \alpha_t''$, the ratio $\frac{\pi\left(\alpha_{t+1}|\alpha_t''\right)}{\pi\left(\alpha_{t+1}|\alpha_t'\right)}$ is monotone increasing in α_{t+1}.*

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

3.3 Planner choice

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

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

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, $c_t\left(\alpha^t\right)$. A complete set of $c_t\left(\alpha^t\right)$ functions for all $t \geq 0$ and $\alpha^t \in A^{t+1}$ is referred to as an **allocation**.[14]

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

[14]The dependence of c_t on α^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:

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

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 α_t following a given α^{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. Taking resources first:. I 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:

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

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 $\int_{\alpha^t} c_t\left(\alpha^t\right) d\Pi_t\left(\alpha^t\right) = y_t$ 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 α^{t-1}. It is helpful to characterise it by reference to continuation utilities. Let $V_t\left(\alpha^{t-1};\alpha_t\right)$ be the maximised value for U_t available to an individual with history of type reports α^{t-1} and current type α_t. This has the recursive definition:

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

for all $t \geq 0$. [15]

Incentive compatibility then requires:

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

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

Note that the true type α'_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, $\Pi\left(\alpha_{t+1}|\alpha'_t\right)$. 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 α^{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 α^t, almost all of these constraints must be redundant. In keeping with much of the literature, I 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 for

[15] The Markov property of shocks implies that the value of V_{t+1} is unaffected by the truthfulness, or otherwise, of past reports.

the simplified constraint set will also be optimal for the more complex constraint set. This is known as the first-order approach to mechanism design.[16]

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

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

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

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 $\alpha_t' \in A$:*

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

I refer to equation (15) as the **relaxed incentive constraint**. For an arbitrary type α_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 α_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-

[16]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 α_t does not just make current consumption more desirable relative to future. It also changes an agent's distribution across future draws, $\Pi\left(\alpha_{t+1}|\alpha_t\right)$. This means that an allocation with relatively low c_t could nonetheless be appealing to an agent with high α_t, if it delivers a more advantageous distribution of future outcomes for this type.

I 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).[17] 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 not necessary – monotonicity condition on the allocation alone. This condition has a particularly clear interpretation by reference to the decentralisation that is introduced later in the paper.

[17] C.f. their Theorem 3.

In the appendix I show:[18]

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:*

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

Since $D_{t,s}\left(\alpha^s\right) < 1$, 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* α_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. Formally:

Definition. An allocation is called **normal** if $c_s\left(\alpha^{t-1}, \alpha_t, \alpha_{t+1}^s\right)$ is non-increasing in α_t for all t and s with $s > t$, all $\alpha^{t-1} \in A^t$ and Π-almost all $\alpha_{t+1}^s \in A^{s-t+1}$.

Normality – and the global incentive compatibility that it implies – guarantees that period-t consumption is non-decreasing in α_t for types with a common history. It has particular relevance because, as I show in Section6, normality is also sufficient for a simple market decentralisation with nonlinear taxes to be possible.

For some purposes it is convenient to strengthen the condition, and neglect the possibility that multiple types bunch at the same consumption value in t. To this end, I 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 $\delta_t\left(\alpha^{t-1}\right) > 0$ such that $\frac{c_t\left(\alpha^{t-1}, \alpha_t''\right) - c_t\left(\alpha^{t-1}, \alpha_t'\right)}{\alpha_t'' - \alpha_t'} \geq \delta_t\left(\alpha^{t-1}\right)$ for all $\left(\alpha_t', \alpha_t''\right) \in A^2$.

[18] To simplify presentation, I 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 $\alpha_t' > \alpha_t''$.

That is, current consumption is strictly increasing in α_t, by an amount that is bounded below with respect to the change in α_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 $\left(c_t\left(\alpha^{t-1},\underline{\alpha}\right),c_t\left(\alpha^{t-1},\bar{\alpha}\right)\right)$ where a positive measure of types locate.[19]

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, I derive the following:

Proposition 2. *Suppose an interior allocation is optimal in the relaxed problem. For a cross-section of types in t*

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

with a common history α^{t-1}, the following must hold a.e.:

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

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

where η_t is the shadow value of resources for the planner, satisfying:

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

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

$$\lambda_{t+1}=\lambda_t+\mu_t\left(\alpha_t'\right) \tag{20}$$

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

with $\mu_t\left(\alpha_t'\right)$ 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 $\lambda_{t+1}^{\Delta}\left(\alpha_t'\right)-\rho\left(\alpha_t'|\alpha_{t-1}\right)\lambda_t^{\Delta}$ is the shadow cost of raising $\omega_{t+1}^{\Delta}\left(\alpha_t'\right)$, holding constant ω_t^{Δ}.

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.

λ_t and λ_t^{Δ} in these expressions are multipliers deriving from prior incentive restrictions – capturing the shadow costs of changing ω_t and ω_t^{Δ} respectively. As highlighted by Marcet 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 α_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.[20] When shocks are iid, the inverse marginal utility is well-known to follow a quasi-martingale process, with substantial implications for long-run inequality.[21] There has been significant recent debate about the sensitivity of this conclusion to type persistence.[22] In this subsection I 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 α^t, the period-t expected value of the period-s inverse marginal utility of consumption satisfies:*

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

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 α_t, in periods after t. This is the only component that matters in the long run, because $\rho\left(\alpha_r|\alpha_{r-1}\right) \in [0,1)$, and so $\frac{\mathbb{E}_t\left[D_{t,s}\left(\alpha^s\right)\alpha_s\right]}{\mathbb{E}_t\left[\alpha_s\right]} \to 0$ as $s \to \infty$.[23] Since λ_{t+1} follows a martingale,[24] the implication is that shocks to this martingale process control *long-run*

[20] 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 is sufficiently small.

[21] This setting is explored in detail by Farhi and Werning (2007).

[22] See, for instance, Bloedel, Krishna and Strulovici (2020) and Bloedel, Krishna and Leukhina (2021).

[23] Recall that $D_{t,s}\left(\alpha^s\right) := \prod_{r=t+1}^{s}\rho\left(\alpha_r|\alpha_{r-1}\right)$, and α_s is bounded.

[24] C.f. equation (20).

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 α_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.

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.[25] 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 α_t. Type persistence implies that periods after t matter for information rents in t.[26] If it is desirable in t to raise rents at α_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 exent 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 $\frac{\mathbb{E}_t\left[D_{t,s}(\alpha^s)\alpha_s\right]}{\mathbb{E}_t[\alpha_s]}$ is precisely the elasticity of $\mathbb{E}_t\left[\alpha_s\right]$ with respect to α_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:

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

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 $\mu_s\left(\alpha_s\right) = 0$ for all α_s, $s \geq t$. But since μ_s is the multiplier on the relaxed incentive constraint, this implies convergence to a first-best allocation, maximising welfare

[25] See, for example, the discussion in Pavan (2017).

[26] Recall, from equation (15), that the information rent at type α_t is given by:

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

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

on the resource constraint alone. This violates incentive compatibility. Thus for $R \leq \beta^{-1}$, as t becomes large:

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

That is, as time progresses the contemporaneous *expectation of the long-run marginal cost of utility provision* converges almost surely to zero.

6 A consumption-savings decentralisation

6.1 Decentralised choice structure

I focus on a simple market decentralisation of the optimal direct allocation, making use of nonlinear savings taxes. The market structure works as follows. Individuals enter generic period t with a given value of net wealth, M_t, which is normalised to remain positive along the equiulibrium path. This wealth 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 over time, and in the history of past savings decisions, s^{t-1}. This tax is denoted $T_t\left(s_t; s^{t-1}\right)$, or simply $T_t(s_t)$ if context allows, with $T_t'\left(s_t; s^{t-1}\right)$ or $T_t'(s_t)$ to dente the corresponding marginal savings tax where it exists.

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

$$c_t + s_t = M_t \tag{23}$$

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

Given M_0, individuals choose contingent consumption sequences to maximise U_0, subject to (23) and (24), plus a 'no Ponzi' constraint, which must hold along all history paths:

$$\lim_{T\to\infty} R^{-T} M_T \geq 0 \tag{25}$$

As discussed in Appendix A.5, conditions (23) to (25) together imply a forward-looking multi-period budget constraint that must be satisfied by all realised consumption-savings paths from generic period t onwards:

$$M_t = \sum_{r=t}^{\infty} R^{t-r} \left[c_r + T_r \left(s_r; s^{r-1} \right) \right] \tag{26}$$

Different realisations of individuals' type draws over time will give rise to different optimal consumption-savings paths. Where required I denote by $s_t^* \left(\alpha_t; s^{t-1} \right)$ the optimal level of savings in the decentralisation for an individual whose current type is α_t and whose past savings history is given by s^{t-1}.[27]

6.2 Normalising taxes

Consistent with the usual Barro-Ricardian equivalence result, by changing the timing of taxation it is generally possible to construct multiple tax systems that deliver the same market allocation. Hence a degree of normalisation is necessary. I focus on a normalisation with two properties. The first is that the fiscal consequences of marginal savings decisions are fully capitalised upfront. That is, changes in an individual's current savings may give rise to changes in their current tax bill, but they leave unaffected their expected future tax liabilities. The second normalisation is to set to zero the expected future tax liabilities of an agent whose current type takes the maximum value, $\bar{\alpha}$.

Mathematically, these restrictions are easiest to express by defining the expected lifetime tax liability from $t+1$ on for an individual whose period-t type is α_t, and who saves an amount s_t:

$$\mathcal{T}_{t+1} \left(s_t, \alpha_t; s^{t-1} \right) = \mathbb{E}_t \left[\sum_{r=t+1}^{\infty} R^{t+1-r} T_r \left(s_r; s^{t-1}, s_t, s_{t+1}^{r-1} \right) \middle| s^{t-1}, \alpha_t \right] \tag{27}$$

where expectations are taken with respect to future type draws, given the optimal savings choices that they induce.

If $s_t^* \left(\alpha_t; s^{t-1} \right)$ denotes optimal savings for type α_t given history s^{t-1}, the first normalisation is to set:

$$\begin{aligned} \mathcal{T}_{t+1} \left(s_t^* \left(\alpha_t'; s^{t-1} \right), \alpha_t'; s^{t-1} \right) =& \mathcal{T}_{t+1} \left(s_t^* \left(\bar{\alpha}; s^{t-1} \right), \bar{\alpha}; s^{t-1} \right) \\ & - \int_{\alpha_t'}^{\bar{\alpha}} \left. \frac{\partial \mathcal{T}_{t+1} \left(s_t^* \left(\tilde{\alpha}_t; s^{t-1} \right), \alpha_t; s^{t-1} \right)}{\partial \alpha_t} \right|_{\tilde{\alpha}_t = \alpha_t} d\alpha_t \end{aligned} \tag{28}$$

[27] This discussion implicitly assumes a unique such optimum. The formal treatment in the appendix does not.

That is, variations in α_t change expected future taxes only insofar as they affect expectations across future type histories. There is no marginal effect from a change in period-t savings, $s_t^*\left(\alpha_t; s^{t-1}\right)$. This is useful for the results that follow, because it means the impact of an individual's marginal saving decision on public finances can be fully summarised by its impact on contemporaneous tax revenue.

The second normalisation is to set:

$$\mathcal{T}_{t+1}\left(s_t^*\left(\bar{\alpha}; s^{t-1}\right), \bar{\alpha}; s^{t-1}\right) = 0 \tag{29}$$

The agent whose current type is highest – and so will be saving the least in equilibrium – faces an expected lifetime tax bill of zero from period $t+1$ onwards. This normalisation ensures that there is always one period-t type, $\bar{\alpha}$, whose post-tax wealth in t is equal to the period-t expectation of their lifetime consumption from $t+1$ on.

6.3 Tax revenues: an 'impulse response' result

Equation (28) implies an intuitive link between variations in an individual's current type, and changes in their expected lifetime tax payments. Specifically:

Lemma 2. *Suppose condition (28) holds. For any s^{t-1} and $\alpha_t \in A$ with the property that $\frac{d}{d\alpha_t}\left[T\left(s_t^*\left(\alpha_t\right); s^{t-1}\right)\right]$ exists, expected future tax liabilities react to current type as follows:*

$$\alpha_t \frac{d}{d\alpha_t}\left\{\mathbb{E}\left[\sum_{r=t}^{\infty} R^{t-r} T_r\left(s_r\right) \middle| s^{t-1}, \alpha_t\right]\right\} = \mathbb{E}\left[\sum_{r=t}^{\infty} R^{t-r} D_{t,r}\left(\alpha^r\right) \alpha_r \frac{d}{d\alpha_r}\left\{T_r\left(s_r^*\left(\alpha_r\right)\right)\right\} \middle| s^{t-1}, \alpha_t\right] \tag{30}$$

If, in addition, $T_r(\cdot)$ is absolutely continuous in s_r and $s_r^*(\cdot)$ absolutely continuous in α_r for $r \geq t$, this can be written as:

$$\alpha_t \frac{d}{d\alpha_t}\left\{\mathbb{E}\left[\sum_{r=t}^{\infty} R^{t-r} T_r\left(s_r\right) \middle| s^{t-1}, \alpha_t\right]\right\} = \mathbb{E}\left[\sum_{r=t}^{\infty} R^{t-r} D_{t,r}\left(\alpha^r\right) \alpha_r \frac{ds_r^*\left(\alpha_r\right)}{d\alpha_r} T_r'\left(s_r^*\left(\alpha_r\right)\right) \middle| s^{t-1}, \alpha_t\right] \tag{31}$$

Thus the effect on lifetime tax payments of a marginal increase in an individual's *current* type can be expressed a discounted sum of the revenue effects marginal changes in type at all *future* horizons, discounted by the rate of persistence in the shock process (the $D_{t,r}\left(\alpha^r\right)$ term).

In the terminology of Pavan, Segal and Toikka (2014), what matters for lifetime tax payments is the **impulse response** of the shock, captured by $D_{t,r}(\alpha^r)$. When types are iid over time, $D_{t,r}(\alpha^r) = 0$ for $r > t$: a shock to current type has no persistence. The only effect of a change in α_t is to change period-t savings, and the associated level of taxation.[28] Expected future savings are left unaffected, just as they would be if the *same* behavioural change in period-t savings were engineered for a given, invariant type.

When type is persistent, however, a change in true type in t shifts expectations of future type, in a way that necessarily changes the expected level of future taxes. This drives a wedge between the effect on expected lifetime taxes of a change in true period-t type, and the effect of the associated change in period-t savings in isolation – holding type constant.

6.4 Feasibility of decentralisation

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 $\left\{c_t^*\left(\alpha^t\right)\right\}_{t,\alpha^t}$ can be decentralised by a sequence of tax functions $T_t\left(s_t; s^{t-1}\right)$ that satisfy (28) and (29), provided the allocation is normal.*

Like its use in validating the first-order approach, the restriction to normal allocations is sufficient for the decentralisation to be possible, but not necessary. Its main role is to allow a link between the value of individuals' wealth in the decentralisation, and the net-present value of their expected future consumption. If future consumption were increasing in current type α_t along some nodes, it becomes possible that multiple current type reports could imply the same net-present value for future consumption, at least for some future type distributions. This makes it more complicated to assign each type report to a unique level of future wealth, and vice-versa.

It is easy to amend the proof to allow weaker restrictions.[29] The advantage of normality is that it is an ordinal restriction on the allocation, independent of the assumed type process or utility functions. In this way it is consistent with the general focus in this paper on minimising analytical reference to the

[28] Recall that I normalise $D_{t,t}\left(\alpha^t\right)$ to 1.

[29] The proof goes through without amendment so long as the *net-present value* of future consumption is monotone decreasing in α_t for all period-t types. This is a much weaker restriction than decreasingness at almost all future nodes, but at the cost of a reference to the type distribution.

two key unobservables that dominate the mechanism design approach – the utility function and the type process.

7 Sufficient statistics: preliminaries

A central contribution of the present paper is to provide a direct characterisation of the optimal marginal tax rate, $T_t'(s_t)$, by reference to behavioural elasticities and other observable objects, such as the conditional savings distribution, that arise in the decentralisation. This section defines the relevant statistics, and provides intermediate steps in order to map from a utility-based to a sufficient statistics representation of the optimum.

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.

I achieve this by use of a novel, intuitive analytical step, 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.[30] 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

[30] The discussion in Piketty and Saez (2013b) provides detailed treatment.

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 α_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, I 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:*

$$\begin{aligned}&\int_{\underline{c}}^{\bar{c}}\left[1-\frac{\alpha_t(c)\,u'(c)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t(c)\,|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\pi^c\left(c|\alpha^{t-1}\right)dc \\ &+\int_{\underline{c}}^{\bar{c}}\left(\alpha_t(c)\right)^2u''(c)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t(c)\right)}{\eta_t}-\rho\left(\alpha_t(c)\,|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_t}\right\}\left(\frac{d\alpha_t(c)}{dc}\right)^{-1}\pi^c\left(c|\alpha^{t-1}\right)dc \\ =&0\end{aligned}\tag{32}$$

$$\begin{aligned}&-\int_{\underline{c}}^{c'}\left[1-\frac{\alpha_t(c)\,u'(c)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t(c)\,|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\pi^c\left(c|\alpha^{t-1}\right)dc \\ &-\int_{\underline{c}}^{c'}\left(\alpha_t(c)\right)^2u''(c)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t(c)\right)}{\eta_t}-\rho\left(\alpha_t(c)\,|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}\left(\frac{d\alpha_t(c)}{dc}\right)^{-1}\pi^c\left(c|\alpha^{t-1}\right)dc \\ &+\left(\alpha_t(c')\right)^2u'(c')\left(\frac{d\alpha_t(c)}{dc}\right)^{-1}\pi^c\left(c'|\alpha^{t-1}\right)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t(c')\right)}{\eta_t}-\rho\left(\alpha_t(c')\,|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\} \\ =&0\end{aligned}\tag{33}$$

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* α^{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:

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

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 (32) and (33) appear just to be slightly more unweildy alternatives to (17) and (18). But their value is that their components correspond to meaningful behavioural statistics. To this end, the four relevant behavioural statistics are defined in turn here.

The contemporaneous elasticity of savings with respect to the post-tax rate of return: This is denoted ε^s_t. 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:

$$\varepsilon^s_t := \frac{R\left(1 - T'_t(s_t)\right)}{s_t} \frac{ds_t}{dR\left(1 - T'_t(s_t)\right)}$$

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 with a given history of savings, it is the effect on s_t of a marginal increase in M_t, holding constant current and future tax schedules. Like ε^s_t, the value of this statistic will be endogenous to properties of the tax schedule itself. For instance,, individuals who locate in regions where the marginal tax rate is increasing rapidly are less likely to increase their savings in response to higher wealth, relative to their behaviour in a linear setting.

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 $\epsilon^s_{t-1,t}(s_t)$, 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 Δ 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, $\left(1 - \Pi^s\left(s_t|\alpha^{t-1}\right)\right)$, where $\Pi^s\left(s_t|\alpha^{t-1}\right)$ 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:[31]

$$\epsilon^s_{t-1,t}\left(s_t\right)\left(1-\Pi^s\left(s_t|\alpha^{t-1}\right)\right) ds_t := \frac{\left(1-T_t'\left(s_t\right)\right)}{s_{t-1}} \left.\frac{ds_{t-1}}{d\left(1-T_t'\left(s_t\right)\right)}\right|_{\text{comp}} \tag{34}$$

$\epsilon^s_{t-1,t}\left(s_t\right)$ 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 α^{t-1} does not experience any change to their continuation value, $\omega_t\left(\alpha^{t-1}\right)$. 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 discussed in more detail later. Ultimately the sign and magnitude of $\epsilon^s_{t-1,t}\left(s_t\right)$ will capture important links between tax cuts, insurance and savings. This 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 $T_t\left(\underline{s}\right)$, holding constant the profile of marginal tax rates at

[31] The formal definition of $\epsilon^s_{t-1,t}(s_t)$ 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.

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 α_t will be $\alpha_t u'(c_t(\alpha_t))$, and in general this will vary in α_t. This implies that the change in the lump-sum component will induce a marginal reprofiling of utility across type draws. Once again, I 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, I denote the compensated effect of higher period-t income on $t-1$ savings by:

$$\left.\frac{ds_{t-1}}{dM_t}\right|_{\text{comp}}$$

7.3 Equivalence results

Making use of these definitions, the following Lemma provides the ingredients to link from expressions (32) and (33) to a sufficient statistics representation:

Lemma 3. *The following relationships hold:*

1. *[Contemporaneous savings elasticity]*

$$T_t'(s_t)\, s_t \varepsilon_t^s = \frac{\lambda_{t+1}^{\Delta}(\alpha_t(c_t))}{\eta_t}(\alpha_t(c_t))^2 u'(c_t)\left(\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1} \tag{35}$$

2. *[Contemporaneous income effect]*

$$T_t'(s_t)\frac{ds_t}{dM_t} = -\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c_t))}{\eta_t}(\alpha_t(c_t))^2 u''(c_t)\left(\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1} \tag{36}$$

3. *[Intertemporal effect of tax cut]*

$$\begin{aligned} RT_{t-1}'(s_{t-1})\, s_{t-1}\epsilon_{t-1,t}^s(s_t') = &-\rho(\alpha_t(c_t')\,|\alpha_{t-1})\,\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}(\alpha_t(c_t'))^2 u'(c_t')\left(\frac{d\alpha_t(c_t')}{dc_t}\right)^{-1}\frac{\pi^c(c_t'|\alpha^{t-1})}{\Pi^c(c_t'|\alpha^{t-1})} \\ &+\frac{1}{\Pi^c(c_t'|\alpha^{t-1})}\int_{\underline{c}}^{c_t'}\left\{\rho(\alpha_t(c_t)\,|\alpha_{t-1})\,\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\alpha_t(c_t)(u'(c_t))\right. \\ &\left.\times\left[1+\frac{c_t u''(c_t)}{u'(c_t)}\left(\frac{c_t}{\alpha_t(c_t)}\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1}\right]\pi^c(c_t|\alpha^{t-1})\right\}dc_t \end{aligned} \tag{37}$$

4. *[Intertemporal income effect]*

$$RT'_{t-1}\left(s_{t-1}\right)\left.\frac{ds_{t-1}}{dM_t}\right|_{comp}=\int_{\underline{c}}^{\bar{c}}\left\{\rho\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\alpha_t\left(c_t\right)\left(u'\left(c_t\right)\right)\right.$$
$$\left.\times\left[1+\frac{c_t u''\left(c_t\right)}{u'\left(c_t\right)}\left(\frac{c_t}{\alpha_t\left(c_t\right)}\frac{d\alpha_t\left(c_t\right)}{dc_t}\right)^{-1}\right]\pi^c\left(c_t|\alpha^{t-1}\right)\right\}dc_t \tag{38}$$

The proof of these relationships, in the Appendix, is algebraically involved but 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 λ_t^{Δ} and λ_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 I 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, I will 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. Ror each history α^t and current consumption level c_t, these are defined by:[32]

$$g_t\left(\alpha^t\right):=\alpha_t u'\left(c_t\left(\alpha^t\right)\right)\frac{1+\lambda_t\left(\alpha^{t-1}\right)}{\eta_t} \tag{39}$$

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

Economically, the most interesting component of the welfare weight is the object $\lambda_t\left(\alpha^{t-1}\right)$. This updates period-by period in response to the shocks that agents receive, with mean-zero innovations: $\mathbb{E}\left[\lambda_{t+1}|\alpha^{t-1}\right]=\lambda_t$.[33] In the decentralised allocation, the updating process will capture the wealth that

[32]Where the mapping between α_t and s_t is bijective, I write $g_t\left(s_t\right)$ in place of $g_t\left(\alpha^t\right)$, leaving history implicit.

[33]See Proposition 2.

agents accumulate along each history branch. Higher values for λ_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.[34]

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 can be identified with changes in the effective welfare objective of the policymaker. A policymaker in the initial period may seek a radical utilitarian allocation, unconstrained 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 will emerge.

The time inconsistency here is evident, and provides a challenge to the plausibility of the commitment assumption.[35] What if a new government would like to redistribute wealth?

8 Sufficient statistics characterisation

8.1 Characterisation

Corollary 2 and Lemma 3 together deliver the main 'sufficient statistics' characterisation result:

Theorem 1. *If a strictly normal allocation is optimal in the relaxed problem, with $c_t\left(\alpha_t\right)$ continuous for any given α^{t-1}, then at $t=0$, for almost all realised savings levels s_0':*

$$\mathbb{E}\left[1-T_0'\left(s_0\right)\frac{ds_0}{dM_0}-g_0\left(s_0\right)\middle| s_0\geq s_0'\right]=T_0'\left(s_0'\right)\varepsilon_0^s\frac{s_0'\pi^s\left(s_0'\right)}{1-\Pi^s\left(s_0'\right)} \tag{40}$$

[34] Formally, the proof of Theorem 2 establishes that λ_t is decreasing in α_{t-1}, for agents with a common history α^{t-2}. Since higher α_{t-1} corresponds to higher consumption in $t-1$, there is a monotone link from savings to Pareto weights.

[35] Brendon and Ellison (2018) propose an alternative solution concept under commitment that delivers stationarity in the Pareto weights.

and:

$$\mathbb{E}\left[g_0\left(s_0\right)\right]=\mathbb{E}\left[1-T_0'\left(s_0\right)\frac{ds_0}{dM_0}\right] \tag{41}$$

Similarly, for $t>0$, any given α^{t-1}, and almost all realised s_t':

$$\mathbb{E}_{t-1}\left[1-T_t'\left(s_t\right)\frac{ds_t}{dM_t}-g_t\left(s_t\right)\middle|s_t\geq s_t'\right]=T_t'\left(s_t'\right)\varepsilon_t^s\frac{s_t'\pi^s\left(s_t'|\alpha^{t-1}\right)}{1-\Pi^s\left(s_t'|\alpha^{t-1}\right)}+RT_{t-1}'\left(s_{t-1}\right)s_{t-1}\epsilon_{t-1,t}^s\left(s_t'\right) \tag{42}$$

and:

$$\mathbb{E}_{t-1}\left[g_t\left(s_t\right)\right]=1-\mathbb{E}_{t-1}\left[T_t'\left(s_t\right)\frac{ds_t}{dM_t}\right]-RT_{t-1}'\left(s_{t-1}\right)\left.\frac{ds_{t-1}}{dM_t}\right|_{comp} \tag{43}$$

.

Proof. Follows from direct substitution of the expressions in Lemma 3 into the conditions in Corollary 2, applying the definition of the social welfare weights. □

8.2 Intuition

As previewed, equations (40) to (43) 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 (40) and (42), 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 (41) and (43), 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 (40) 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 (40) 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, $T_0'\left(s_0\right)\frac{ds_0}{dM_0}$, minus (3) the social welfare value of providing an additional consumption unit, $g_0\left(s_0\right)$. Taken together, these objects give the net fiscal cost of the transfer that high-saving agents receive.

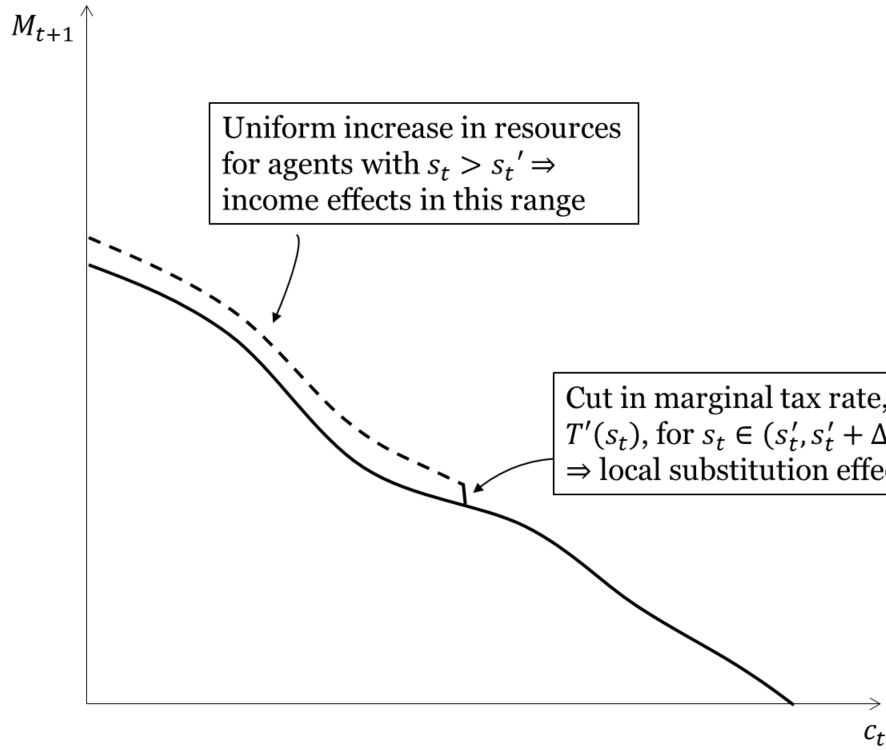

Figure 1: Effects on the budget constraint of a marginal savings tax cut at s_t'

The right-hand side of (40) 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 $T_0'\left(s_0'\right) s_0' \varepsilon_0^s$, interacted with the density of savers affected.

Condition (42) is the equivalent to (40) 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 $RT_{t-1}'\left(s_{t-1}\right) s_{t-1} \epsilon_{t-1,t}\left(s_t'\right)$, 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 (compensated) cross-elasticity $\epsilon_{t-1,t}\left(s_t'\right)$: 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.

The striking feature of Theorem 1 is what is *not* in it. Despite the infinite-horizon, stochastic setting, just two behavioural adjustment margins matter for optimal taxes: the contemporaneous effect on savings, and the effect on savings one period back. This is an extremely helpful simplification. It speaks positively for the practical applicability of dynamic Mirrleesian tax analysis – a feature that is not generally considered its main strength.

Conditions (41) and (43) describe the consequences of shifting the entire budget constraint, rather than just an upper segment. They are essentially variants of (40) and (42) 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 $\epsilon_{t-1,t}\left(s_t'\right)$. 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 $T_t\left(s_t\right)$ 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 low consumption need, and thus attracts a high net fiscal contribution. Optimal policy faces the familiar trade-off between redistributing towards those given preference by the social (and *ex-ante* individual) welfare criterion – 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. As in the static labour income tax literature, 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 'slowly' by comparison with the remaining probability mass in the upper tail. When this is true, 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 endopints if the conditional type density $\pi(\cdot|\cdot)$ is everywhere bounded above zero.[36]

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

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

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

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

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

$$T_t'\left(\bar{s}\right) = \frac{1 - g_t\left(\bar{s}\right) - RT_{t-1}'\left(s_{t-1}\right)\frac{s_{t-1}}{\bar{s}}\epsilon_{t-1,t}^s\left(\bar{s}\right)a_t\left(\bar{s}|\alpha^{t-1}\right)}{\left.\frac{ds_t}{dM_t}\right|_{\bar{s}} + \varepsilon_t^s a_t\left(\bar{s}|\alpha^{t-1}\right)} \tag{44}$$

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.

[36] This follows from equation (75) in the appendix, and the fact that $T_t'\left(s_t\left(\alpha_t\right)\right) = 0$ when $\lambda_{t+1}^{\Delta}\left(\alpha_t\right) = 0$ – a result established in the proof of Theorem 2.

9.3 Optimal transfers

Conditions (41) and (43) 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 $\mathbb{E}\left[g_0\left(s_0\right)\right]$, and the net cost of doing so, $\mathbb{E}\left[1 - T'\left(s_0\right)\frac{ds_0}{dM_0}\right]$. 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 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 α_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:

$$\mathbb{E}_{t-1}\left[g_t\left(s_t\right)\right] < \mathbb{E}_{t-1}\left[1 - T'\left(s_t\right)\frac{ds_t}{dM_t}\right] \tag{45}$$

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

10.1 Insurance incentivises prior saving

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, $\epsilon^s_{t-1,t}\left(s'_t\right)$.

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 α_{t-1}. Type persistence means that those with higher α_{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 $\epsilon^{s}_{t-1,t}\left(s_t'\right)$ in equation (42) 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: $\epsilon^s_{t-1,t}\left(s_t\right) \equiv 0$, 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 (42). Markovian shocks imply that the preferences of two distinct α_{t-1} types across alternative allocations from $t+1$ on are identical, conditional on drawing a particular α_t. Compensated distortions to $t+1$ allocations – which, by definition, hold constant expected lifetime utility for each period-t type draw – can do nothing to induce additional saving in $t-1$ or earlier.

10.2 Intertemporal elasticities: a source of progressivity

Thus it may be desirable to provide additional insurance in t, beyond what is contemporaneously optimal, because this helps to increase savings in $t-1$ – reducing the costs of the tax distortion. Additional insurance is often associated with additional progressivity in the tax system. In this subsection I formalise this link, via the systematic cross-sectional variation that exists in the intertemporal elasticity $\epsilon^s_{t-1,t}\left(s'_t\right)$.

Formally, I show the following:

Proposition 5. *There exists a threshold $\tilde{s}_t$ such that $s_{t-1}\epsilon^s_{t-1,t}\left(s'_t\right)$ is positive for $s'_t < \tilde{s}_t$, and negative for $s'_t > \tilde{s}_t$.*

The implications of this Proposition can be seen by comparing policies that satisfy condition (42) with those that neglect intertemporal cross-elasticities – as would be optimal for a policymaker re-optimising in t. At the optimum, I have:

$$\mathbb{E}_{t-1}\left[1 - T'_t\left(s_t\right)\frac{ds_t}{dM_t} - g_t\left(s_t\right)\middle| s_t \geq s'_t\right] > T'_t\left(s'_t\right)s'_t\varepsilon^s_t a_t\left(s'_t|\alpha^{t-1}\right) \tag{46}$$

for all s'_t below a threshold and

$$\mathbb{E}_{t-1}\left[1 - T'_t\left(s_t\right)\frac{ds_t}{dM_t} - g_t\left(s_t\right)\middle| s_t \geq s'_t\right] < T'_t\left(s'_t\right)s'_t\varepsilon^s_t a_t\left(s'_t|\alpha^{t-1}\right) \tag{47}$$

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 threshold value s_t' – the savings level where taxes are being cut – is changed.

A tax cut at s_t' has two consequences for insurance. The first is to redistribute resources as a whole to a group of relatively high savers. Since high savers have a low marginal consumption utility, this worsens the insurance value of savings, and acts as a negative force on $\epsilon^s_{t-1,t}\left(s_t'\right)$. But the second effect is to improve the nature of insurance *within* the group that benefits. Uniform income provision, which a tax cut implies, generates non-uniform effects on *utility* within the sub-group. Lower savers have higher marginal utility, and so benefit more. This force improves the overall insurance profile of savings, raising the value of $\epsilon^s_{t-1,t}\left(s_t'\right)$.

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 $\epsilon^s_{t-1,t}\left(s_t'\right)$. This is the reason for the single crossing property established in Proposition 5.

11 An indicative quantification

To give the results a more concrete character, in this penultimate section I provide an indicative quantification of the optimal top marginal savings tax, based on the formula (44).

The exercise is more straightforward if the intertemporal response of savings in $t-1$ to a reprofiling of insurance in t, captured by the elasticity $\epsilon^s_{t-1,t}\left(s_t'\right)$, is initially assumed to be small, or of low relevance because the lagged marginal tax rate is small. This corresponds, for instance, to a case when type persistence is low. It is also the relevant exercise for the initial period, $t=0$. Given Proposition 5, this approach will tend to bias the resulting figure downwards. The optimality formula simplifies to:

$$T_t'(\bar{s}) = \frac{1 - g_t(\bar{s})}{\left.\frac{ds_t}{dM_t}\right|_{\bar{s}} + \varepsilon_t^s a_t\left(\bar{s}|\alpha^{t-1}\right)} \tag{48}$$

which is biased below the true optimal value by an amount:

$$-\frac{RT_{t-1}'(s_{t-1})\frac{s_{t-1}}{\bar{s}}\epsilon_{t-1,t}^s(\bar{s})\, a_t\left(\bar{s}|\alpha^{t-1}\right)}{\left.\frac{ds_t}{dM_t}\right|_{\bar{s}} + \varepsilon_t^s a_t\left(\bar{s}|\alpha^{t-1}\right)} \geq 0 \tag{49}$$

This bias will be ignored in the initial discussion.

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, corresponding to homotheticity in demand. Since savings and consumption covary perfectly, conditional on within-period wealth, the value of $\varepsilon_t^s a_t\left(\bar{s}|\alpha^{t-1}\right)$ must be 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. But unlike a_t in our formulation, this is based on a sample of the entire population, not a conditional cross section with common wealth. The relevant cross-sectional object is likely to be substantially lower.

The compensated intertemporal elasticity of consumption is strictly less than the Frisch, for which a value of around 0.5 is standard.[37] 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 $\frac{1-g_t(\bar{s})}{3}$ at most, 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 α_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 α_t.

[37] See, for instance, Attanasio and Weber (2010).

To make progress, recall that equation (43) shows that the cross-sectional average value for g_t, $\bar{g}_t$, is 1 minus the value of tax recovered through income effects at the margin, both in t and $t-1$. Let the latter value be equal to some number $\chi_t \in (0,1)$. Simple algebra gives:

$$1 - g_t(\bar{s}) = \chi_t + (1-\chi_t)\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) \tag{50}$$

$$= \chi_t + (1-\chi_t)\left(\frac{\mathbb{E}_{t-1}\left[\alpha_t u'(c_t)\right] - \underline{\alpha} u'\left(c_t\left(\underline{\alpha}\right)\right)}{\mathbb{E}_{t-1}\left[\alpha_t u'(c_t)\right]}\right) \tag{51}$$

This varies in the size of the average income effect on tax revenue, χ_t, and the proportional deviation of the lowest type's marginal utility from the average. 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. The formula requires an estimate for the impact that this has on their marginal utility of wealth, relative to the average. With time-separable utility and an EIS of 0.5, the proportional difference in marginal utility is roughly two times the proportional difference in lifetime consumption. Thus if average current consumption is valued generously at ten per cent of average wealth, the term $\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right)$ will be approximately 0.2.[38]

The value of χ_t, the average income effect on tax revenue, is obviously difficult to fix in a partial quantification of this kind, because it depends on the chosen marginal tax rate across the entire range of savings. The general lesson of this section is that a top marginal tax rate as high as 10 per cent is hard to justify. Since the qualitative factors that matter elsewhere in the distribution tend to push towards a lower marginal rate on lower savings levels,[39] and since the proportion of additional income that is saved will generally be below 100 per cent, 0.1 is a generous upper bound for χ_t.

Given these considerations, Table 1 summarises the implied top marginal tax rates for alternative values of the three relevant statistics: a_t, χ_t and $\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right)$, given $\varepsilon_t^s = 0.5$ and for negligible intertemporal revenue effects:

The main lesson is that the top marginal tax rate on savings could perhaps come close to 10 per cent, but only with some very favourable assumptions, particularly on the value of the cross-sectional Pareto

[38] Persistence in α provides a direct channel for lower period-t types to value a unit of savings less than the average type, but since insurance is incomplete this effect will be limited: the marginal value of saving, even to a low current type, will be dominated by the risk of high type draws in the future.

[39] Specifically, the expected value of the welfare weight for higher savers is increasing as the threshold savings value falls, and the local Pareto parameter will be higher at modal points in the distribution.

		$a_t = 4$	$a_t = 6$	$a_t = 10$
$\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) = 0.05$	$\chi_t = 0.01$	0.020	0.015	0.010
	$\chi_t = 0.05$	0.033	0.024	0.016
	$\chi_t = 0.1$	0.048	0.036	0.024
$\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) = 0.1$	$\chi_t = 0.01$	0.036	0.027	0.018
	$\chi_t = 0.05$	0.048	0.036	0.024
	$\chi_t = 0.1$	0.063	0.048	0.032
$\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) = 0.2$	$\chi_t = 0.01$	0.069	0.052	0.035
	$\chi_t = 0.05$	0.08	0.06	0.04
	$\chi_t = 0.1$	0.093	0.07	0.047

Table 1: Top marginal savings tax rates implied by alternative statistics

parameter. More likely is a number of the order of 1 to 3 per cent. This is still significant relative to a typical risk-free rate of return, but leaves the possibility of some post-tax interest at the margin for even the highest savers. Moreover, even the lower numbers depend critically on the tail properties of the cross-sectional savings distribution. 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. In this case we would return to zero distortion at the top.

Finally, I provide some considerations of the bias due to ignoring intertemporal revenue effects. Again assuming that all additional income is saved by $\underline{\alpha}$ types, this bias is given by:

$$-T'_{t-1}(s_{t-1}) \frac{s_{t-1}\epsilon^s_{t-1,t}(\bar{s})}{\bar{s}\varepsilon^s_t} \frac{R}{1 + \frac{1}{\varepsilon^s_t a_t(\bar{s}|\alpha^{t-1})}} \tag{52}$$

The final fraction here is likely to be in the neighbourhood of 1. Clearly it is less than R, more significantly so the lower the Pareto parameter. The middle ratio is more significant: the ratio of $s_{t-1}\epsilon^s_{t-1,t}(\bar{s})$ to $\bar{s}\varepsilon^s_t$. Since $\bar{s}$ is the highest level of savings in t for those who saved s_{t-1} in the previous period, $\frac{s_{t-1}}{\bar{s}}$ seems likely to be small. The compensated intertemporal savings response to a cut in the upper marginal tax rate is not a commonly estimated object, and there is little to go on except priors. It seems very likely that $\left|\epsilon^s_{t-1,t}(\bar{s})\right| < \varepsilon^s_t$. Over all, an upper bound for the term in (52) equal to $0.5 \cdot T'_{t-1}(s_{t-1})$ seems quite generous. If lagged marginal tax rates are of the same order of magnitude as current, intertemporal considerations could conceivably raise top rates by at most a half of their value in Table 1.

Perhaps more interesting is the qualitative observation that (52) need not go to zero when $a_t \to \infty$. That is, intertemporal savings responses imply there may be a case for distorting period-t savings, even though within-period considerations alone would yield no distortion at the top. This will happen only if $\epsilon^s_{t-1,t}(\bar{s})$ remains strictly negative in such a case: the compensated insurance effects of a tax cut must still be enough to induce a reduction in prior savings. But the conditions for this to hold are mathematically distinct from the conditions for $a_t \to \infty$.[40]

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

[40] If Assumption 4, the monotone likelihood ratio property, holds then $s_{t-1}\epsilon^s_{t-1,t}(\bar{s})$ will be bounded away from zero: c.f. expression (193) below. This assumption does not place any restrictions on a_t.

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 (2021), '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] Greulich, K., S. Laczo and A. Marcet (2022), 'Pareto-Improving Optimal Capital and Labor Taxes', *Journal of Political Economy*, forthcoming.

[24] Hellwig, C. (2021), ‘Static and Dynamic Mirrleesian Taxation with Non-separable Preferences: A Unified Approach’, TSE Working Paper 1224

[25] Hellwig, C., and N. Werquin (2022), ‘A Fair Day’s Pay for a Fair Day’s Work: Optimal Tax Design as Redistributional Arbitrage’, CEPR Working Paper

[26] İmrohoroğlu, S. (1998), ‘A Quantitative Analysis of Capital Income Taxation’, *International Economic Review*, 39(2): 307–28.

[27] Judd K.L. (1985), ‘Redistributive Taxation in a Simple Perfect Foresight Model’, *Journal of Public Economics*, 28(1): 59–83.

[28] Kapička, M. (2013), ‘Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: a First-order Approach’, *Review of Economic Studies*, 80: 1027–54.

[29] Kleven, H. (2021), ‘Sufficient Statistics Revisited’, *Annual Review of Economics*, 13: 515–538.

[30] Kockerlakota, N.R. (2005), ‘Zero Expected Wealth Taxes: a Mirrlees Approach to Dynamic Optimal Taxation’, *Econometrica*, 73: 1587–621

[31] Kocherlakota, N.R. (2010), *The New Dynamic Public Finance*, Princeton University Press.

[32] Makris, M., and A. Pavan (2020), ‘Wedge Dynamics with Evolving Private Information’, Working Paper.

[33] Marcet, A., and R. Marimon (2019), ‘Recursive Contracts’, *Econometrica*, 87(5): 1589–1631.

[34] Milgrom, P., and I. Segal (2002), ‘Envelope Theorems for Arbitrary Choice Sets’, *Econometrica*, 70(2): 583–601.

[35] Mirrlees, J.A. (1971), ‘An Exploration in the Theory of Optimum Income Taxation’, *Review of Economic Studie*s, 38: 175–208.

[36] Negishi, T. (1960), ‘Welfare Economics and Existence of an Equilibrium for a Competitive Economy’, *Metroeconomica*, 12: 92–97.

[37] 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*.

[38] Pavan, A., I. Segal and J. Toikka (2014), 'Dynamic Mechanism Design: A Myersonian Approach', *Econometrica*, 82(2): 601–653.

[39] Piketty, T., and E. Saez (2013a), 'A Theory of Optimal Inheritance Taxation', *Econometrica*, 81(5): 1851–1886

[40] 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.

[41] Rogerson, W.P. (1985), 'The First-order Approach to Principal-Agent Problems', *Econometrica*, 53(6), 1357–67.

[42] Saez, E. (2001), 'Using Elasticities to Derive Optimal Income Tax Rates', *Review of Economic Studies*, 68: 205–229.

[43] Saez, E., and S. Stantcheva (2018), 'A Simpler Theory of Optimal Capital Taxation', *Journal of Public Economics*, 162: 120–142.

[44] Scheuer, F., and I. Werning (2017), 'The Taxation of Superstars', *Quarterly Journal of Economics*, 132: 211–270.

[45] Sleet, C., and S. Yeltekin (2006), 'Credibility and Endogenous Societal Discounting', *Review of Economic Dynamics*, 9: 410–437.

[46] Stantcheva, S. (2017), 'Optimal Taxation and Human Capital Policies over the Life Cycle', *Journal of Political Economy*, 125: 1931– 90.

[47] Stantcheva, S. (2020), 'Dynamic Taxation', *Annual Review of Economics*, 12: 801–31

[48] Straub, L., and I. Werning (2020), 'Positive Long-Run Capital Taxation: Chamley-Judd Revisited', American Economic Review, 110(1): 86–119.

[49] 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.

[50] 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 elemenatary manipulations.

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

$$\alpha_t' u\left(c_t\left(\alpha^{t-1}, \alpha_t''\right)\right) + \beta \int_{\alpha_{t+1}} V\left(\alpha^{t-1}, \alpha_t'', \alpha_{t+1}\right) d\Pi\left(\alpha_{t+1} | \alpha_t'\right) \tag{53}$$

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

$$u\left(c_t\left(\alpha^{t-1}, \alpha_t''\right)\right) + \beta \int_{\alpha_{t+1}} V\left(\alpha^{t-1}, \alpha_t'', \alpha_{t+1}\right) \frac{d\pi\left(\alpha_{t+1}|\alpha_t'\right)}{d\alpha_t'} d\alpha_{t+1} \tag{54}$$

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 α_t'' and α_{t+1}. $\pi\left(\alpha_{t+1}|\alpha_t'\right)$ is continuously differentiable in α_t' by assumption, and α_t' inhabits a compact interval, so $\frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t}$ is also bounded by construction. Taken together this implies that the object in (54) is bounded in absolute value, uniformly across type reports α_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 $V_t\left(\alpha^{t-1}; \alpha_t\right)$ is absolutely continuous in α_t for all t and α^{t-1}, with:

$$\begin{aligned}
&\alpha_t' u\left(c_t\left(\alpha^{t-1}, \alpha_t'\right)\right) + \beta \int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1}, \alpha_t', \alpha_{t+1}\right) d\Pi\left(\alpha_{t+1}|\alpha_t'\right) \\
=&\underline{\alpha} u\left(c_t\left(\alpha^{t-1}, \underline{\alpha}\right)\right) + \beta \int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1}, \underline{\alpha}, \alpha_{t+1}\right) d\Pi\left(\alpha_{t+1}|\underline{\alpha}\right) \\
&+ \int_{\underline{\alpha}}^{\alpha_t'} \frac{1}{\alpha_t}\left\{\alpha_t u\left(c_t\left(\alpha^{t-1}, \alpha_t\right)\right) + \beta\alpha_t \int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1}, \alpha_t, \alpha_{t+1}\right) \frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t} d\alpha_{t+1}\right\} d\alpha_t
\end{aligned} \tag{55}$$

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

Lemma 4. *For all* $(\alpha_t, \alpha'_t) \in A^2$ *and* $\alpha^{t-1} \in A^t$:

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

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:

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

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

$\square$

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

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

Using this and the definition of $\omega_{t+1}\left(\alpha^{t-1}, \alpha_t\right)$, (55) 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 α^{t-1} there must exist α'_t, α''_t such that:

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

or equivalently:

$$u_t\left(\alpha_t''\right)+\frac{\beta}{\alpha_t'}\int_{\alpha_{t+1}} V_{t+1}\left(\alpha_t'',\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t'\right)d\alpha_{t+1} > u_t\left(\alpha_t'\right)+\frac{\beta}{\alpha_t'}\int_{\alpha_{t+1}} V_{t+1}\left(\alpha_t',\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t'\right)d\alpha_{t+1} \tag{56}$$

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

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

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

Applying this result in (56) yields:

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

Or:

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

Applying Lemma 4 and the definition of V_{t+1}, this is equivalent to:

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

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

A.2.2 'Only if'

Suppose integral monotonicity fails for some (α_t', α_t''), i.e.:

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

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

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

Thus global incentive compatibility must be violated for type α_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:

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

subject to the resource constraint:

$$\sum_{t=0}^{\infty} R^{-t} \left[y_t - \int_{\alpha^t} c_t\left(\alpha^t\right) d\Pi_t\left(\alpha^t\right) \right] \geq 0 \tag{57}$$

and the relaxed incentive constraint:

$$\begin{aligned} \alpha_t' u\left(c_t\left(\alpha^{t-1}, \alpha_t'\right)\right) + \beta\omega_{t+1}\left(\alpha^{t-1}, \alpha_t'\right) = {} & \underline{\alpha} u\left(c_t\left(\alpha^{t-1}, \underline{\alpha}\right)\right) + \beta\omega_{t+1}\left(\alpha^{t-1}, \underline{\alpha}\right) \\ & + \int_{\underline{\alpha}}^{\alpha_t'} \frac{1}{\alpha_t} \left[\alpha_t u\left(c_t\left(\alpha^{t-1}, \alpha_t\right)\right) + \beta\omega_{t+1}^{\Delta}\left(\alpha^{t-1}, \alpha_t\right) \right] d\alpha_t \end{aligned} \tag{58}$$

with, for all $t \geq 0$:

$$\omega_{t+1}\left(\alpha^{t}\right) := \int_{\alpha_{t+1}} \left\{\alpha_{t+1} u\left(c_{t+1}\left(\alpha^{t}, \alpha_{t+1}\right)\right) + \beta \omega_{t+2}\left(\alpha^{t}, \alpha_{t+1}\right)\right\} d\Pi\left(\alpha_{t+1}|\alpha_{t}\right) \tag{59}$$

$$\omega^{\Delta}_{t+1}\left(\alpha^{t}\right) := \int_{\alpha_{t+1}} \rho\left(\alpha_{t+1}|\alpha_{t}\right) \cdot \left\{\alpha_{t} u\left(c_{t+1}\left(\alpha^{t}, \alpha_{t+1}\right)\right) + \beta \omega^{\Delta}_{t+2}\left(\alpha^{t}, \alpha_{t+1}\right)\right\} d\Pi\left(\alpha_{t+1}|\alpha_{t}\right) \tag{60}$$

plus the interiority restriction, which is assumed not to bind for the Proposition. We place multiplier η on (57), $\beta^{t} \mu_{t}\left(\alpha^{t-1}, \alpha_{t}\right) d\Pi_{t}\left(\alpha^{t-1}, \alpha_{t}\right)$ on (58), $\beta^{t+1} \lambda_{t+1}\left(\alpha^{t}\right) d\Pi_{t}\left(\alpha^{t}\right)$ on (59) and $\beta^{t+1} \lambda^{\Delta}_{t+1}\left(\alpha^{t}\right) d\Pi_{t}\left(\alpha^{t}\right)$ on (60). Necessary first-order optimality conditions are:

- With respect to $c_{t}\left(\alpha^{t}\right)$, a.e.:

$$0 = \alpha_{t} u'\left(c_{t}\left(\alpha^{t}\right)\right) \cdot \left[\begin{array}{l} 1 + \lambda_{t}\left(\alpha^{t-1}\right) + \lambda^{\Delta}_{t}\left(\alpha^{t-1}\right) \rho\left(\alpha_{t}|\alpha_{t-1}\right) + \mu_{t}\left(\alpha^{t}\right) \\ -\frac{1}{\alpha_{t} \pi\left(\alpha_{t}|\alpha_{t-1}\right)} \int_{\alpha_{t}}^{\bar{\alpha}} \mu_{t}\left(\alpha^{t-1}, \tilde{\alpha}_{t}\right) \pi\left(\tilde{\alpha}_{t}|\alpha_{t-1}\right) d\tilde{\alpha}_{t} \end{array}\right] - (\beta R)^{-t} \eta \tag{61}$$

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

$$0 = -\lambda_{t+1}\left(\alpha^{t}\right) + \lambda_{t}\left(\alpha^{t-1}\right) + \mu_{t}\left(\alpha^{t-1}, \alpha_{t}\right) \tag{62}$$

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

$$\begin{aligned} 0 = & -\lambda^{\Delta}_{t+1}\left(\alpha^{t}\right) + \lambda^{\Delta}_{t}\left(\alpha^{t-1}\right) \rho\left(\alpha_{t}|\alpha_{t-1}\right) \\ & - \frac{1}{\alpha_{t} \pi\left(\alpha_{t}|\alpha_{t-1}\right)} \int_{\alpha_{t}}^{\bar{\alpha}} \mu_{t}\left(\alpha^{t-1}, \tilde{\alpha}_{t}\right) \pi\left(\tilde{\alpha}_{t}|\alpha_{t-1}\right) d\tilde{\alpha} \end{aligned} \tag{63}$$

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

$$0 = \int_{\underline{\alpha}}^{\bar{\alpha}} \mu_{t}\left(\alpha^{t-1}, \alpha_{t}\right) \pi\left(\alpha_{t}|\alpha_{t-1}\right) d\alpha_{t} \tag{64}$$

Throughout here, we normalise $\lambda_{0} = \lambda^{\Delta}_{0} \equiv 0$, and let $\pi\left(\alpha_{t}|\alpha_{t-1}\right)$ be replaced with $\pi\left(\alpha_{0}\right)$ when $t = 0$.

Using (62) and (63) in (61) gives:

$$\frac{(\beta R)^{-t} \eta}{\alpha_{t} u'\left(c_{t}\left(\alpha^{t}\right)\right)} = 1 + \lambda_{t+1}\left(\alpha^{t}\right) + \lambda^{\Delta}_{t+1}\left(\alpha^{t}\right) \tag{65}$$

Condition (61) can be rearranged to:

$$\begin{aligned}
&\frac{(\beta R)^{-t}\,\eta}{u'\left(c_t\left(\alpha^t\right)\right)} - \alpha_t\left[1+\lambda_t\left(\alpha^{t-1}\right)+\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\rho\left(\alpha_t|\alpha_{t-1}\right)\right] \\
=&\alpha_t\mu_t\left(\alpha^{t-1},\alpha_t\right)-\frac{1}{\pi\left(\alpha_t|\alpha_{t-1}\right)}\int_{\alpha_t}^{\bar{\alpha}}\mu_t\left(\alpha^{t-1},\tilde{\alpha}_t\right)\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha}_t
\end{aligned} \tag{66}$$

This can be integrated across all α_t:

$$\begin{aligned}
&\int_{\underline{\alpha}}^{\bar{\alpha}}\left\{\frac{(\beta R)^{-t}\,\eta}{u'\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right)}-\alpha_t\left[1+\lambda_t\left(\alpha^{t-1}\right)+\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\rho\left(\alpha_t|\alpha_{t-1}\right)\right]\right\}\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t \\
=&\int_{\underline{\alpha}}^{\bar{\alpha}}\left\{\alpha_t\mu_t\left(\alpha^{t-1},\alpha_t\right)\pi\left(\alpha_t|\alpha_{t-1}\right)-\int_{\alpha_t}^{\bar{\alpha}}\mu_t\left(\alpha^{t-1},\tilde{\alpha}_t\right)\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha}_t\right\}d\alpha_t
\end{aligned} \tag{67}$$

Integrating by parts, making use of (64):

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

and so:

$$\int_{\underline{\alpha}}^{\bar{\alpha}}\left\{\frac{(\beta R)^{-t}\,\eta}{u'\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right)}-\alpha_t\left[1+\lambda_t\left(\alpha^{t-1}\right)+\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\rho\left(\alpha_t|\alpha_{t-1}\right)\right]\right\}\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t=0 \tag{68}$$

or, using (5):

$$\frac{1}{\mathbb{E}\left[\alpha_t|\alpha_{t-1}\right]}\mathbb{E}\left[\left.\frac{(\beta R)^{-t}\,\eta}{u'\left(c_t\left(\alpha^t\right)\right)}\right|\alpha^{t-1}\right]=\left[1+\lambda_t\left(\alpha^{t-1}\right)+\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\varepsilon^{\alpha}\left(\alpha_{t-1}\right)\right] \tag{69}$$

Rearranging (69) for period 0 gives an expression for η:

$$\eta=\frac{\mathbb{E}\left[\alpha_0\right]}{\mathbb{E}\left[\frac{1}{u'\left(c_0\left(\alpha_0\right)\right)}\right]} \tag{70}$$

Combining (65) and (69) gives expressions for the objects $1+\lambda_{t+1}\left(\alpha^t\right)$ and $\lambda_{t+1}^{\Delta}\left(\alpha^t\right)$:

$$1+\lambda_{t+1}\left(\alpha^t\right)=(\beta R)^{-t}\,\eta\frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\mathbb{E}\left[\alpha_{t+1}|\alpha_t\right]}\mathbb{E}\left[\left.\frac{1}{\beta R u'\left(c_{t+1}\left(\alpha^{t+1}\right)\right)}\right|\alpha^t\right]-\frac{\varepsilon^{\alpha}\left(\alpha_t\right)}{\alpha_t u'\left(c_t\left(\alpha^t\right)\right)}\right\} \tag{71}$$

$$\lambda_{t+1}^{\Delta}\left(\alpha^t\right)=(\beta R)^{-t}\,\eta\frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\alpha_t u'\left(c_t\left(\alpha^t\right)\right)}-\frac{1}{\mathbb{E}\left[\alpha_{t+1}|\alpha_t\right]}\mathbb{E}\left[\left.\frac{1}{\beta R u'\left(c_{t+1}\left(\alpha^{t+1}\right)\right)}\right|\alpha^t\right]\right\} \tag{72}$$

Integrating (66) above any given α_t' gives:

$$\begin{aligned}
&\int_{\alpha_t'}^{\bar{\alpha}} \left\{ \frac{(\beta R)^{-t}\eta}{u'\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right)} - \alpha_t \left[1+\lambda_t\left(\alpha^{t-1}\right)+\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\rho\left(\alpha_t|\alpha_{t-1}\right)\right]\right\}\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t \\
=&\int_{\alpha_t'}^{\bar{\alpha}} \left\{ \alpha_t\mu_t\left(\alpha^{t-1},\alpha_t\right)\pi\left(\alpha_t|\alpha_{t-1}\right) - \int_{\alpha_t}^{\bar{\alpha}} \mu_t\left(\alpha^{t-1},\tilde{\alpha}_t\right)\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha}_t\right\}d\alpha_t
\end{aligned} \tag{73}$$

and integrating by parts, we have:

$$\begin{aligned}
&\int_{\alpha_t'}^{\bar{\alpha}}\int_{\alpha_t}^{\bar{\alpha}} \mu_t\left(\alpha^{t-1},\tilde{\alpha}_t\right)\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha}_t d\alpha_t \\
=&-\alpha_t'\int_{\alpha_t'}^{\bar{\alpha}} \mu_t\left(\alpha^{t-1},\alpha_t\right)\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t + \int_{\alpha_t'}^{\bar{\alpha}} \alpha_t\mu_t\left(\alpha^{t-1},\alpha_t\right)\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t
\end{aligned} \tag{74}$$

so:

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

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 (71) and (72) we immediately have:

$$\frac{1}{\alpha_t u'\left(c_t\right)} = \frac{1+\lambda_{t+1}+\lambda_{t+1}^{\Delta}}{\eta_t} \tag{76}$$

$$\frac{1}{\mathbb{E}_t\left[\alpha_{t+1}\right]}\mathbb{E}_t\left[\frac{1}{\beta R u'\left(c_{t+1}\right)}\right] = \frac{1+\lambda_{t+1}+\lambda_{t+1}^{\Delta}\varepsilon^{\alpha}\left(\alpha_t\right)}{\eta_t} \tag{77}$$

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

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

$$\mathbb{E}_r\left[\frac{\eta_r}{(\beta R)^{s-r}u'\left(c_s\right)}\right] = \left[1+\lambda_{r+1}\right]\mathbb{E}_r\left[\alpha_s\right] + \lambda_{r+1}^{\Delta}\mathbb{E}_r\left[D_{r,s}\left(\alpha^s\right)\alpha_s\right] \tag{78}$$

Then:

$$
\begin{aligned}
&\mathbb{E}_{r-1}\left[\frac{\eta_{r-1}}{(\beta R)^{s-r+1}u'(c_s)}\right] &(79)\\
=&\mathbb{E}_{r-1}\left\{[1+\lambda_{r+1}]\,\mathbb{E}_r[\alpha_s]+\lambda_{r+1}^{\Delta}\mathbb{E}_r\left[D_{r,s}(\alpha^s)\,\alpha_s\right]\right\}\\
=&\int_{\alpha_r}\left\{\left[\rho(\alpha_r|\alpha_{r-1})\,\lambda_r^{\Delta}-\frac{1}{\alpha_r\pi(\alpha_r|\alpha_{r-1})}\int_{\alpha_r}^{\bar{\alpha}}\mu_r(\tilde{\alpha}_r)\,\pi(\tilde{\alpha}_r|\alpha_{r-1})\,d\tilde{\alpha}_r\right]\cdot\mathbb{E}_r\left[D_{r,s}(\alpha^s)\,\alpha_s\right]\right.\\
&\left.+[1+\lambda_r+\mu_r(\alpha_r)]\cdot\mathbb{E}_r[\alpha_s]\right\}\pi(\alpha_r|\alpha_{r-1})\,d\alpha_r &(80)
\end{aligned}
$$

By an identical argument to Lemma 4, we have:

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

Integrating by parts, we therefore have:

$$\int_{\alpha_r}\frac{1}{\alpha_r}\mathbb{E}_r\left[D_{r,s}(\alpha^s)\,\alpha_s\right]\left[\int_{\alpha_r}^{\bar{\alpha}}\mu_r(\tilde{\alpha}_r)\,\pi(\tilde{\alpha}_r|\alpha_{r-1})\,d\tilde{\alpha}_r\right]d\alpha_r=\int_{\alpha_r}\mathbb{E}_r[\alpha_s]\,\mu_r(\tilde{\alpha}_r)\,\pi(\alpha_r|\alpha_{r-1})\,d\alpha_r \tag{81}$$

where we have used condition (64). Using this in the preceeding expression, the terms in μ_r cancel:

$$
\begin{aligned}
\mathbb{E}_{r-1}\left[\frac{\eta_{r-1}}{(\beta R)^{s-r+1}u'(c_s)}\right]&=\int_{\alpha_r}\left\{[1+\lambda_r]\cdot\mathbb{E}_r[\alpha_s]+\rho(\alpha_r|\alpha_{r-1})\,\lambda_r^{\Delta}\cdot\mathbb{E}_r\left[D_{r,s}(\alpha^s)\,\alpha_s\right]\right\}\pi(\alpha_r|\alpha_{r-1})\,d\alpha_r &(82)\\
&=[1+\lambda_r]\cdot\mathbb{E}_{r-1}[\alpha_s]+\lambda_r^{\Delta}\cdot\mathbb{E}_{r-1}\left[D_{r-1,s}(\alpha^s)\,\alpha_s\right]
\end{aligned}
$$

which makes use of the definition of $D_{r-1,s}(\alpha^s)$. Thus we have iterated expectations backwards a period from condition (78). Now, for any $s>0$, condition (77) implies:

$$\mathbb{E}_{s-1}\left[\frac{\eta_{s-1}}{\beta Ru'(c_s)}\right]=[1+\lambda_s]\,\mathbb{E}_{s-1}[\alpha_s]+\lambda_s^{\Delta}\mathbb{E}_{s-1}\left[D_{s-1,s}(\alpha^s)\,\alpha_s\right] \tag{83}$$

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

A.5 The infinite-horizon budget constraint

Conditions (23) and (24) imply that the following condition must be satisfied between all periods $t \geq 0$, and $T > t$, for any realised consumption-savings path:

$$M_t = \sum_{r=t}^{T} R^{t-r} \left[c_r + T_r \left(s_r ; s^{r-1} \right) \right] + R^{t-T-1} M_{T+1} \tag{84}$$

Since consumption is always positive, its net-present value must converge along Π-almost all history paths that are realised in equilibrium – or else feasibility would be violated. In such cases we can take limits to give:

$$M_t = \sum_{r=t}^{\infty} R^{t-r} c_r + \lim_{T \to \infty} \left[\sum_{r=t}^{T} R^{t-r} T_r \left(s_r ; s^{r-1} \right) + R^{t-T-1} M_{T+1} \right] \tag{85}$$

The no-Ponzi constraint implies that the second term inside the limit is weakly positive, though technically it need not converge to zero: it is possible that the agent may only obtain their desired consumption stream by holding wealth into the infinite future whose present value is strictly positive.[41] But since this serves as a *de facto* tax, without loss we can write:

$$M_t = \sum_{r=t}^{\infty} R^{t-r} \left[c_r + T_r \left(s_r ; s^{r-1} \right) \right] \tag{86}$$

where:

$$\sum_{r=t}^{\infty} R^{t-r} T_r \left(s_r ; s^{r-1} \right) := \lim_{T \to \infty} \left[\sum_{r=t}^{T} R^{t-r} T_r \left(s_r ; s^{r-1} \right) + R^{t-T-1} M_{T+1} \right] \tag{87}$$

A.6 Proof of Lemma 2

Equation (31) follows immediately from (30) given the additional differentiability assumed, so I focus exclusively on establishing (30). Suppressing history dependence for simplicity, by definition of $\mathcal{T}$ we have:

$$\mathbb{E} \left[\sum_{r=t}^{\infty} R^{t-r} T_r \left(s_r \right) \middle| s^{t-1}, \alpha_t \right] = T_t \left(s_t^* \left(\alpha_t \right) \right) + R^{-1} \mathcal{T}_{t+1} \left(s_t^* \left(\alpha_t \right), \alpha_t \right) \tag{88}$$

[41] Transversality rules this out in a conventional infinite-horizon consumption-savings problem. Here, period-by-period feasibility is allowed to depend in complex ways on individuals' asset choices, so that it may not be possible to improve on a strategy that retains strictly positive asset holdings at the limit. None of the analysis that follows makes use of this possibility – it is here for completeness only.

Differentiating this, and making use of (28):

$$\alpha_t \frac{d}{d\alpha_t}\left\{\mathbb{E}\left[\sum_{r=t}^{\infty} R^{t-r} T_r\left(s_r\right)\middle| s^{t-1}, \alpha_t\right]\right\} = \alpha_t \frac{d}{d\alpha_t}\left[T_t\left(s_t^*\left(\alpha_t\right)\right)\right] + R^{-1}\alpha_t \frac{d}{d\alpha_t}\left\{\mathcal{T}_{t+1}\left(s_t^*\left(\alpha_t\right), \alpha_t\right)\right\} \tag{89}$$

$$= \alpha_t \frac{d}{d\alpha_t}\left\{T_t\left(s_t^*\left(\alpha_t\right)\right)\right\} + R^{-1}\alpha_t \left.\frac{\partial \mathcal{T}_{t+1}\left(s_t, \alpha_t\right)}{\partial \alpha_t}\right|_{s_t = s_t^*\left(\alpha_t\right)} \tag{90}$$

From the definition of $\mathcal{T}$, again, we have:

$$\mathcal{T}_{t+1}\left(s_t, \alpha_t\right) = \int_{\alpha_{t+1}} \left[T_{t+1}\left(s_{t+1}^*\left(\alpha_{t+1}\right)\right) + R^{-1}\mathcal{T}_{t+2}\left(s_{t+1}^*\left(\alpha_{t+1}\right), \alpha_{t+1}\right)\right] \pi\left(\alpha_{t+1}|\alpha_t\right) d\alpha_{t+1} \tag{91}$$

So:

$$\alpha_t \left.\frac{\partial \mathcal{T}_{t+1}\left(s_t, \alpha_t\right)}{\partial \alpha_t}\right|_{s_t = s_t^*\left(\alpha_t\right)}$$

$$= \alpha_t \int_{\alpha_{t+1}} \left[T_{t+1}\left(s_{t+1}^*\left(\alpha_{t+1}\right)\right) + R^{-1}\mathcal{T}_{t+2}\left(s_{t+1}^*\left(\alpha_{t+1}\right), \alpha_{t+1}\right)\right] \frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t} d\alpha_{t+1} \tag{92}$$

$$= \int_{\alpha_{t+1}} \alpha_{t+1} \frac{d}{d\alpha_{t+1}}\left[T_{t+1}\left(s_{t+1}^*\left(\alpha_{t+1}\right)\right) + R^{-1}\mathcal{T}_{t+2}\left(s_{t+1}^*\left(\alpha_{t+1}\right), \alpha_{t+1}\right)\right] \rho\left(\alpha_{t+1}|\alpha_t\right) \pi\left(\alpha_{t+1}|\alpha_t\right) d\alpha_{t+1} \tag{93}$$

$$= \int_{\alpha_{t+1}} \left\{\alpha_{t+1} \frac{d}{d\alpha_{t+1}}\left[T_{t+1}\left(s_{t+1}^*\left(\alpha_{t+1}\right)\right)\right] + \alpha_{t+1} \left.\frac{\partial \mathcal{T}_{t+2}\left(s_{t+2}, \alpha_{t+1}\right)}{\partial \alpha_{t+1}}\right|_{s_{t+2} = s_{t+2}^*\left(\alpha_{t+1}\right)}\right\} \rho\left(\alpha_{t+1}|\alpha_t\right) \pi\left(\alpha_{t+1}|\alpha_t\right) d\alpha_{t+1} \tag{94}$$

$$= \mathbb{E}\left[\sum_{r=t+1}^{\infty} R^{t+1-r} D_{t,r}\left(\alpha^r\right) \alpha_r \frac{d}{d\alpha_r}\left\{T_r\left(s_r^*\left(\alpha_r\right)\right)\right\}\middle| s^{t-1}, \alpha_t\right] \tag{95}$$

where the last line uses the definition of $D_{t,r}\left(\alpha^r\right)$. Combining this with (90), and using the normalisation $D_{t,t}\left(\alpha^t\right) = 1$, gives the result.

A.7 Proof of Proposition 4

It is immediate that if (by normality) $c_{t+s}^*\left(\alpha^{t-1}, \alpha_t, \alpha_{t+1}^{t+s}\right)$ is decreasing in α_t for all $t \geq 0$, $s > 0$, $\alpha^{t-1} \in A^t$ and Π-almost all $\alpha_{t+1}^{t+s} \in A^s$, and $\left\{c_t^*\left(\alpha^t\right)\right\}_{t,\alpha^t}$ is incentive-compatible, then $c_t^*\left(\alpha^{t-1}, \alpha_t\right)$ must be increasing in α_t, strictly except on ranges for α_t where $c_{t+s}^*\left(\alpha^{t-1}, \alpha_t, \alpha_{t+1}^{t+s}\right)$ is invariant in α_t along Π-almost all future paths α_{t+1}^{t+s}.

Given this, we proceed constructively, showing how to map from the allocation $\left\{c_t^*\left(\alpha^t\right)\right\}_{t,\alpha^t}$ to tax functions $T_t\left(s^{t-1}, s_t\right)$, 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 $\left\{c_t^*\left(\alpha^t\right)\right\}_{t,\alpha^t}$, 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:

$$M_0 := \sum_{t=0}^{\infty} R^{-t} y_t \tag{96}$$

For all α_0, let the savings level $s_0(\alpha_0)$ then be defined by:

$$s_0(\alpha_0) := M_0 - c_0^*(\alpha_0) \tag{97}$$

and denote the range of s_0 values across α_0 by S_0:

$$S_0 := \{s_0(\alpha_0)\}_{\alpha_0 \in A} \tag{98}$$

Since consumption is increasing, the minimum value for savings is $s_0(\bar{\alpha})$ and its maximum is $s_0(\underline{\alpha})$, and so $S_0 \subseteq \left[s_0(\bar{\alpha}), s_0(\underline{\alpha})\right]$. 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 $\{s_1, ..., s_t\}$, let $T_t(\tilde{s}_0, s_1, ...s_t) > \varepsilon$ for some $\varepsilon > 0$. Combining the budget constraints (23) and (24), we have, along all future consumption paths:

$$0 = M_1(\tilde{s}_0) = \sum_{t=0}^{T} R^{-t}\left[c_{t+1} + T_{t+1}(\tilde{s}_0, ..., s_{t+1})\right] + R^{-T-2} M_{T+2} \tag{99}$$

and so:

$$\sum_{t=0}^{T} R^{-t} c_{t+1} = -\sum_{t=0}^{T} R^{-t} T_{t+1}(\tilde{s}_0, ..., s_{t+1}) - R^{-T-2} M_{T+2} \tag{100}$$

for all $T \geq 0$. By the 'no Ponzi' condition, the final term on the right-hand side satisfies $\lim_{T\to\infty} R^{-T-2} M_{T+2} \geq 0$, 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 the rest of the argument it helps to have a simple expression for the net-present value of future consumption, under the direct mechanism. Denote by $C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)$ the following object:

$$C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right) = \mathbb{E}\left[\sum_{s=t+1}^{\infty} R^{t+1-s} c_s^*\left(\alpha^{t-1}, \tilde{\alpha}_t, \alpha_{t+1}^s\right) \middle| \alpha_t\right] \tag{101}$$

That is, the expected NPV of consumption from $t+1$ onwards, given a period-t report of $\tilde{\alpha}_t$, a true period-t type α_t, and past history α^{t-1}.

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

$$M_1(s_0) := C_1(\alpha_0(s_0), \alpha_0(s_0)) + \int_{\alpha_0(s_0)}^{\bar{\alpha}} \left. \frac{\partial C_1(\tilde{\alpha}_0, \alpha_0)}{\partial \alpha_0} \right|_{\tilde{\alpha}_0 = \alpha_0} d\alpha_0 \tag{102}$$

where $\alpha_0(s_0) : S_0 \to \mathbb{R}$ is the inverse of $s_0(\alpha_0)$. Where $c_0^*(\alpha_0)$ is strictly increasing, $\alpha_0(s_0)$ is singleton-valued. For ranges of α_0 where $c_0^*(\alpha_0)$ is only weakly increasing, it follows by incentive compatibility that $c_s^*\left(\alpha_0, \alpha_1^s\right)$ is stationary in α_0 for Π-almost all α_1^s paths. Hence, at these values:

$$\frac{dC_1(\alpha_0, \alpha_0)}{d\alpha_0} = \left. \frac{\partial C_1(\tilde{\alpha}_0, \alpha_0)}{\partial \alpha_0} \right|_{\tilde{\alpha}_0 = \alpha_0} \tag{103}$$

It follows that when multiple types choose the same savings value, they receive the same $M_1(s_0)$ under formula (102)

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

$$T_0(s_0) := s_0 - R^{-1} M_1(s_0) \tag{104}$$

An inductive argument can then be applied 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}:

$$S^{t-1} := \left\{ s^{t-1}\left(\alpha^{t-1}\right) \right\}_{\alpha^{t-1} \in A^t} \tag{105}$$

and that this mapping has an inverse correspondence$\alpha^{t-1}(s^{t-1})$, with $\alpha^{t-1} : S^{t-1} \to A^t$. 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:

$$s_t\left(\alpha^{t-1}, \alpha_t\right) := M_t\left(s^{t-1}\left(\alpha^{t-1}\right)\right) - c_t^*\left(\alpha^{t-1}, \alpha_t\right) \tag{106}$$

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

$$S_t\left(\alpha^{t-1}\right) := \left\{ s_t\left(\alpha^{t-1}, \alpha_t\right) \right\}_{\alpha_t \in A} \tag{107}$$

Given α^{t-1}, the inverse $\alpha_t(s_t; \alpha^{t-1})$ gives the set of period-t types corresponding to savings choice s_t, for any $s_t \in S_t(\alpha^{t-1})$. It is singleton-valued iff $c_t^*(\alpha^{t-1}, \alpha_t)$ is strictly increasing in α_t, which again follows except on α_t ranges where $c_s^*\left(\alpha^{t-1}, \alpha_t, \alpha_{t+1}^s\right)$ is constant in α_t for Π-almost all α_{t+1}^s, $s > t$. The mapping $s^t(\alpha^t)$ is then given by extending $s^{t-1}\left(\alpha^{t-1}\right)$:

$$s^{t}\left(\alpha^{t}\right):=\left\{s^{t-1}\left(\alpha^{t-1}\right),s_{t}\left(\alpha^{t}\right)\right\} \tag{108}$$

and $\alpha^{t}\left(s^{t}\right)$ by:

$$\alpha^{t}\left(s^{t}\right):=\left\{\alpha^{t-1}\left(s^{t-1}\right),\alpha_{t}\left(s_{t};\alpha^{t-1}\left(s^{t-1}\right)\right)\right\} \tag{109}$$

For all $s_{t}\in S_{t}\left(\alpha^{t-1}\right)$, let $M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right),s_{t}\right)$ then be given by:

$$\begin{aligned}M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right),s_{t}\right):=&C_{t+1}\left(\alpha_{t}\left(s_{t};\alpha^{t-1}\left(s^{t-1}\right)\right),\alpha_{t}\left(s_{t};\alpha^{t-1}\left(s^{t-1}\right)\right);\alpha^{t-1}\left(s^{t-1}\right)\right)\\&+\int_{\alpha_{t}\left(s_{t};\alpha^{t-1}\left(s^{t-1}\right)\right)}^{\bar{\alpha}}\left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_{t},\alpha_{t};\alpha^{t-1}\left(s^{t-1}\right)\right)}{\partial\alpha_{t}}\right|_{\tilde{\alpha}_{t}=\alpha_{t}}d\alpha_{t}\end{aligned} \tag{110}$$

where we again use the fact that if c_{t}^{*} is constant on a range of α_{t} values, then:

$$\frac{dC_{t+1}\left(\alpha_{t},\alpha_{t};\alpha^{t-1}\right)}{d\alpha_{t}}=\left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_{t},\alpha_{t};\alpha^{t-1}\right)}{\partial\alpha_{t}}\right|_{\tilde{\alpha}_{t}=\alpha_{t}} \tag{111}$$

on this range. This leaves the tax function $T_{t}\left(s^{t-1},s_{t}\right)$ to be given by:

$$T_{t}\left(s^{t-1},s_{t}\right):=s_{t}-R^{-1}M_{t+1}\left(s^{t-1},s_{t}\right) \tag{112}$$

for all $s_{t}\in S_{t}\left(\alpha^{t-1}\left(s^{t-1}\right)\right)$.

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

Any allocation that is not part of the direct mechanism must impliy a saving choice $\tilde{s}_{t}\in S_{t}^{c}\left(\alpha^{t-1}\left(s^{t-1}\right)\right)$ at some history node that violates the agent's no-Ponzi condition, and so the decentralised allocation does not permit a greater set of options than the direct mechanism. To show that the decentralised allocation provides no fewer options, we must confirm that the no-Ponzi condition is satisfied by the induced value of M_{t}, for all admissible sequences of type reports in the centralised scheme. For this, note that for all t, α^{t-1}:

$$M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right),s_{t}\left(\alpha^{t-1},\bar{\alpha}\right)\right)=C_{t+1}\left(\bar{\alpha},\bar{\alpha};\alpha^{t-1}\right)\geq 0 \tag{113}$$

Thus a sufficient condition for M_{t+1} always to be positive is that $M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right),s_{t}\left(\alpha^{t-1},\alpha_{t}\right)\right)$ should be decreasing in α_t. For any pait of types $\alpha_t'' > \alpha_t'$, we have:

$$
\begin{aligned}
& M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t''\right)\right) - M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t'\right)\right) && (114)\\
=& C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right) - \int_{\alpha_t'}^{\alpha_t''} \left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t=\alpha_t} d\alpha_t \\
=& \left[C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t''; \alpha^{t-1}\right)\right] + \left[C_{t+1}\left(\alpha_t', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right)\right] && (115)\\
& - \int_{\alpha_t'}^{\alpha_t''} \left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t=\alpha_t} d\alpha_t \\
=& \left[C_{t+1}\left(\alpha_t'', \alpha_t'; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right)\right] + \left[C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t'', \alpha_t'; \alpha^{t-1}\right)\right] && (116)\\
& - \int_{\alpha_t'}^{\alpha_t''} \left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t=\alpha_t} d\alpha_t
\end{aligned}
$$

Since $C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)$ is monotone increasing in its first argument, it can be discontinuous at at most countably many values for $\tilde{\alpha}_t$, and these values are independent of α_t by the continuity of π (Assumption 2). It follows that either:

$$
\lim_{\alpha_t'' \to \alpha_t'} \left[\frac{C_{t+1}\left(\alpha_t', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right)}{\alpha_t'' - \alpha_t'}\right] = \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{1}{\alpha_t'' - \alpha_t'} \int_{\alpha_t'}^{\alpha_t''} \left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t=\alpha_t} d\alpha_t\right] \tag{117}
$$

or:

$$
\lim_{\alpha_t'' \to \alpha_t'} \left[\frac{C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t'', \alpha_t'; \alpha^{t-1}\right)}{\alpha_t'' - \alpha_t'}\right] = \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{1}{\alpha_t'' - \alpha_t'} \int_{\alpha_t'}^{\alpha_t''} \left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t=\alpha_t} d\alpha_t\right] \tag{118}
$$

or both. Thus either:

$$
\begin{aligned}
& \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t''\right)\right) - M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t'\right)\right)}{\alpha_t'' - \alpha_t'}\right] && (119)\\
=& \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t''; \alpha^{t-1}\right)}{\alpha_t'' - \alpha_t'}\right] \leq 0
\end{aligned}
$$

or:

$$
\begin{aligned}
& \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t''\right)\right) - M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t'\right)\right)}{\alpha_t'' - \alpha_t'}\right] && (120)\\
=& \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{C_{t+1}\left(\alpha_t'', \alpha_t'; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right)}{\alpha_t'' - \alpha_t'}\right] \leq 0
\end{aligned}
$$

were the inequalities follow from the normality assumption. It follows that $M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t\right)\right)$ is indeed decreasing in α_t.

It follows that the decentralisation provides the same choice set as the direct mechanism at all nodes, and so will implement it. It remains only to confirm that the normalisations (28) and (29) are satisfied by the proposed construction. For this, note from (26) that:

$$M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t\right)\right) = C_{t+1}\left(\alpha_t, \alpha_t; \alpha^{t-1}\right) + \mathcal{T}_{t+1}\left(s_t\left(\alpha^{t-1}, \alpha_t\right), \alpha_t; s^{t-1}\left(\alpha^{t-1}\right)\right) \tag{121}$$

Normalisation (29) is thus an immediate consequence of (113). Taking a partial derivative with respect to true type, we also have:

$$\left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t=\alpha_t} = -\left.\frac{\partial \mathcal{T}_{t+1}\left(s_t, \alpha_t; s^{t-1}\left(\alpha^{t-1}\right)\right)}{\partial \alpha_t}\right|_{s_t=s_t\left(\alpha^{t-1}, \alpha_t\right)} \tag{122}$$

Using this and (121) in (102) and (110) shows that these latter two conditions are equivalent to imposing (28) for the first and subsequent periods.

A.8 Corollary 2: details

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

$$\begin{aligned} &\int_{c'}^{c''}\left\{\int_{\alpha_t(c_t)}^{\bar{\alpha}}\left[\frac{1}{u'\left(c_t\left(\alpha\right)\right)} - \frac{\alpha\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\pi\left(\alpha|\alpha_{t-1}\right)d\alpha\right\}\frac{du'\left(c_t\right)}{dc_t}dc_t \\ &= -\int_{c'}^{c''}\left(\alpha_t\left(c_{t.}\right)\right)^2\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t\left(c_t\right)\right)}{\eta_t} - \rho\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}\pi\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\frac{du'\left(c_t\right)}{dc_t}dc_t \end{aligned} \tag{123}$$

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

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

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:

$$\begin{aligned}
&\int_{c'}^{c''}\left[1-\frac{\alpha_t\left(c_t\right)u'\left(c_t\right)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\frac{d\alpha_t\left(c_t\right)}{dc_t}\pi\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)dc_t \\
&+u'\left(c''\right)\int_{c''}^{\bar{c}}\left[\frac{1}{u'\left(c_t\right)}-\frac{\alpha_t\left(c_t\right)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\frac{d\alpha_t\left(c_t\right)}{dc_t}\pi\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)dc_t \\
&-u'\left(c'\right)\int_{c'}^{\bar{c}}\left[\frac{1}{u'\left(c_t\right)}-\frac{\alpha_t\left(c_t\right)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\frac{d\alpha_t\left(c_t\right)}{dc_t}\pi\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)dc_t \\
=&-\int_{c'}^{c''}\left(\alpha_t\left(c_{t.}\right)\right)^2u''\left(c_t\right)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t\left(c_t\right)\right)}{\eta_t}-\rho\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}\pi\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)dc_t
\end{aligned} \tag{124}$$

Suppose first that there are no discontinuities in $c_t\left(\alpha_t\right)$. Making use of (68), over the entire range (124) gives:

$$\begin{aligned}
&\int_{\underline{c}}^{\bar{c}}\left[1-\frac{\alpha_t\left(c_t\right)u'\left(c_t\right)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\pi^c\left(c_t|\alpha^{t-1}\right)dc_t \\
&+\int_{\underline{c}}^{\bar{c}}\left(\alpha_t\left(c_{t.}\right)\right)^2u''\left(c_t\right)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t\left(c_t\right)\right)}{\eta_t}-\rho\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}\left(\frac{d\alpha_t\left(c_t\right)}{dc_t}\right)^{-1}\pi^c\left(c_t|\alpha^{t-1}\right)dc_t \\
&=0
\end{aligned} \tag{125}$$

And for each $c'\in\left(\underline{c},\bar{c}\right)$, making use of (75):

$$\begin{aligned}
&\int_{c'}^{\bar{c}}\left[1-\frac{\alpha_t\left(c_t\right)u'\left(c_t\right)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\pi^c\left(c_t|\alpha^{t-1}\right)dc_t \\
&+\left(\alpha_t\left(c'\right)\right)^2u'\left(c'\right)\pi^c\left(c'|\alpha^{t-1}\right)\left(\frac{d\alpha_t\left(c'\right)}{dc'}\right)^{-1}\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t\left(c'\right)\right)}{\eta_t}-\rho\left(\alpha_t\left(c'\right)|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\} \\
&+\int_{c'}^{\bar{c}}\left(\alpha_t\left(c_{t.}\right)\right)^2u''\left(c_t\right)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t\left(c_t\right)\right)}{\eta_t}-\rho\left(\alpha_t\left(c_t\right)|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}\left(\frac{d\alpha_t\left(c_t\right)}{dc_t}\right)^{-1}\pi^c\left(c_t|\alpha^{t-1}\right)dc_t \\
&=0
\end{aligned} \tag{126}$$

as given in the text.

A.8.1 Allowing discontinuities

Discontinuities in $c_t\left(\alpha_t\right)$ can be included with minimal additional manipulation. Since $c_t\left(\alpha_t\right)$ is monotone, the set of α_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 $c^u\left(\alpha_t\right)=\lim_{\alpha\searrow\alpha_t}\left(c_t\left(\alpha\right)\right)$ and $c^l\left(\alpha_t\right)=\lim_{\alpha\nearrow\alpha_t}\left(c_t\left(\alpha\right)\right)$ denote the upper and lower limits for consumption respectively. Summing (124) across intervals, over the entire range we have:

$$\int_{\underline{c}}^{\bar{c}} \left[1 - \frac{\alpha_t (c_t) u' (c_t) \left\{1 + \lambda_t + \lambda_t^{\Delta} \rho (\alpha_t (c_t) | \alpha_{t-1})\right\}}{\eta_t}\right] \pi^c \left(c_t | \alpha^{t-1}\right) dc_t \tag{127}$$
$$+ \sum_{\alpha_t \in \mathcal{A}} \left(u' (c^u (\alpha_t)) - u' \left(c^l (\alpha_t)\right)\right) \pi (\alpha_t | \alpha_{t-1}) (\alpha_t)^2 \left\{\frac{\lambda_{t+1}^{\Delta} (\alpha_t)}{\eta_t} - \rho (\alpha_t | \alpha_{t-1}) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}$$
$$+ \int_{c'}^{c''} (\alpha_t (c_{t.}))^2 u'' (c_t) \left\{\frac{\lambda_{t+1}^{\Delta} (\alpha_t (c_t))}{\eta_t} - \rho (\alpha_t (c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\} \pi^c \left(c_t | \alpha^{t-1}\right) dc_t$$
$$= 0$$

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

$$\int_{c'}^{\bar{c}} \left[1 - \frac{\alpha_t (c_t) u' (c_t) \left\{1 + \lambda_t + \lambda_t^{\Delta} \rho (\alpha_t (c_t) | \alpha_{t-1})\right\}}{\eta_t}\right] \pi^c \left(c_t | \alpha^{t-1}\right) dc_t \tag{128}$$
$$+ \sum_{\alpha_t \in \mathcal{A} \cap (\alpha_t (c'), \bar{\alpha})} \left(u' (c^u (\alpha_t)) - u' \left(c^l (\alpha_t)\right)\right) \pi (\alpha_t | \alpha_{t-1}) (\alpha_t)^2 \left\{\frac{\lambda_{t+1}^{\Delta} (\alpha_t)}{\eta_t} - \rho (\alpha_t | \alpha_{t-1}) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}$$
$$+ (\alpha_t (c'))^2 u' (c') \pi^c \left(c' | \alpha^{t-1}\right) \left(\frac{d\alpha_t (c')}{dc'}\right)^{-1} \left\{\frac{\lambda_{t+1}^{\Delta} (\alpha_t (c'))}{\eta_t} - \rho (\alpha_t (c') | \alpha_{t-1}) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}$$
$$+ \int_{c'}^{\bar{c}} (\alpha_t (c_{t.}))^2 u'' (c_t) \left\{\frac{\lambda_{t+1}^{\Delta} (\alpha_t (c_t))}{\eta_t} - \rho (\alpha_t (c_t) | \alpha_{t-1}) \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\} \left(\frac{d\alpha_t (c_t)}{dc_t}\right)^{-1} \pi^c \left(c_t | \alpha^{t-1}\right) dc_t$$
$$= 0$$

A.9 Definitions of $\epsilon_{t-1,t}^s (s_t)$ and $\left.\frac{ds_{t-1}}{dM_t}\right|_{\mathbf{comp}}$

$\epsilon_{t-1,t}^s (s_t)$ 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:

$$M_{t+1} = R [s (M_{t+1}) - T (s (M_{t+1}))] \tag{129}$$

Notice that:

$$s'(M_{t+1}) = \frac{1}{R(1 - T'(s(M_{t+1})))} \tag{130}$$

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:[42]

$$T_t(s(M_{t+1})) - \Gamma f(s(M_{t+1})) \tag{131}$$

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:

$$M_{t+1} = R[s_t - T_t(s(M_{t+1})) + \Gamma f(s(M_{t+1}))] \tag{132}$$

Combining with (129), this implies that realising M_{t+1} requires savings equal to:

$$s_t = s(M_{t+1}) - \Gamma f(s(M_{t+1})) \tag{133}$$

and thus consumption equal to:

$$\begin{aligned} c_t &= M_t - s(M_{t+1}) + \Gamma f(s(M_{t+1})) \\ &= c(M_{t+1}) + \Gamma f(s(M_{t+1})) \end{aligned} \tag{134}$$

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 Γ.

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

$$\frac{ds_t}{dM_{t+1}} = s'(M_{t+1})(1 - \Gamma f'(s(M_{t+1}))) = \frac{1 - \Gamma f'(s(M_{t+1}))}{R(1 - T'(s(M_{t+1})))} \tag{135}$$

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

$$\left[\frac{ds_t}{dM_{t+1}}\right]^{-1} = \frac{R(1 - T'(s(M_{t+1})))}{1 - \Gamma f'(s(M_{t+1}))} \tag{136}$$

[42]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 simplecharacterisation of $\epsilon^s_{t-1,t}(s_t)$ by reference to income and substitution effects at t, and the impact these have on information rents.

The derivative of this with respect to Γ is:

$$\frac{d}{d\Gamma}\left\{\left[\frac{ds_t}{dM_{t+1}}\right]^{-1}\right\} = \frac{f'\left(s\left(M_{t+1}\right)\right) R\left(1 - T'\left(s\left(M_{t+1}\right)\right)\right)}{\left[1 - \Gamma f'\left(s\left(M_{t+1}\right)\right)\right]^2} \tag{137}$$

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

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

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

The proof of Lemma 3, below, establishes that the compensated derivative of s_{t-1} with respect to Γ, evaluated at $\Gamma = 0$, will be linear in $f'\left(s_t\right)$.[43] That is:

$$\left.\frac{ds_{t-1}}{d\Gamma}\right|_{\Gamma=0,\text{ comp}} = f\left(\underline{s}\right) H + \int_{s_t} f'\left(s_t\right) h\left(s_t\right) ds_t \tag{138}$$

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

$$h\left(s_t\right) := s_{t-1}\epsilon^s_{t-1,t}\left(s_t\right)\left(1 - \Pi^s\left(s_t | s^{t-1}\right)\right) \tag{139}$$

where $\Pi^s\left(s_t | s^{t-1}\right)$ is the induced measure of savings in t, given history of savings s^{t-1}. This implicitly defines $\epsilon^s_{t-1,t}\left(s_t\right)$. 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 $\epsilon^s_{t-1,t}\left(s_t\right)$ is constructed to capture the response at each point.

The compensated income effect $\left.\frac{ds_{t-1}}{dM_t}\right|_{\text{comp}}$ is defined as the derivative $\left.\frac{ds_{t-1}}{d\Gamma}\right|_{\Gamma=0,\text{ comp}}$ in the simpler case where the perturbation schedule satisfies $f\left(s\left(M_{t+1}\right)\right) = 1$ for all realised M_{t+1} values. It is thus equated to H in (138).

A.10 Proof of Lemma 3

Conditions 1 and 2

We start with two Lemmata:

[43] This is established formally in the proof of Lemma 3: c.f. equation (171) below.

Lemma 5. *The marginal tax rate satisfies:*

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

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

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

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

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 ω_{t+1} at the margin by the amount:

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

and raise ω_{t+1}^{Δ} by the amount:

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

The marginal resource cost of increasing ω_{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 ω_{t+1} in the main problem by the resource multiplier:

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

Similarly, the marginal resource cost of increasing ω_{t+1}^{Δ} by a unit is:

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

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:

$$\begin{aligned}
R^{-t}T_t'(s_t) =& R^{-t} - \frac{\beta^{t+1}\left(1+\lambda_{t+1}\left(\alpha^t\right)\right)}{\eta} R\left(1-T_t'(s_t)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \\
&- \frac{\beta^{t+1}\lambda_{t+1}^{\Delta}\left(\alpha^t\right)}{\eta} R\left(1-T_t'(s_t)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}
\end{aligned}$$

Or:

$$\begin{aligned}
T_t'(s_t) =& 1 - \frac{\left(1+\lambda_{t+1}\left(\alpha^t\right)\right)}{\eta\left(\beta R\right)^{-t-1}}\left(1-T_t'(s_t)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \\
&- \frac{\lambda_{t+1}^{\Delta}\left(\alpha^t\right)}{\eta\left(\beta R\right)^{-t-1}}\left(1-T_t'(s_t)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}
\end{aligned}$$

Expressions for $\frac{(1+\lambda_{t+1}(\alpha^t))}{\eta_t}$ and $\frac{\lambda_{t+1}^{\Delta}(\alpha^t)}{\eta_t}$, with $\eta_t = \eta\left(\beta R\right)^{-t}$, follow from (71) and (72) respectively:

$$\frac{1+\lambda_{t+1}\left(\alpha^t\right)}{\eta_t} = \frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\mathbb{E}\left[\alpha_{t+1}|\alpha_t\right]}\mathbb{E}\left[\left.\frac{1}{\beta R u'\left(c_{t+1}\left(\alpha^{t+1}\right)\right)}\right|\alpha^t\right] - \frac{\varepsilon^{\alpha}\left(\alpha_t\right)}{\alpha_t u'\left(c_t\left(\alpha^t\right)\right)}\right\} \tag{140}$$

$$\frac{\lambda_{t+1}^{\Delta}\left(\alpha^t\right)}{\eta_t} = \frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\alpha_t u'\left(c_t\left(\alpha^t\right)\right)} - \frac{1}{\mathbb{E}\left[\alpha_{t+1}|\alpha_t\right]}\mathbb{E}\left[\left.\frac{1}{\beta R u'\left(c_{t+1}\left(\alpha^{t+1}\right)\right)}\right|\alpha^t\right]\right\} \tag{141}$$

Substituting in (140) gives:

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

where we have used the consumer optimality condition:

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

Rearranging the first line:

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

So:

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

as stated. □

Lemma 6. *The contemporaneous income effect and labour supply elasticity satisfy, respectively:*

$$\left\{\alpha_t u'(c_t)-\beta R\left(1-T_t'(s_t)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\right\}^{-1}=-\frac{\frac{ds_t}{dM_t}}{\alpha_t^2 u''(c_t)\frac{dc_t}{d\alpha_t}}\tag{142}$$
$$\left\{\alpha_t u'(c_t)-\beta R\left(1-T_t'(s_t)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\right\}^{-1}=\frac{s_t\varepsilon_t^s}{\alpha_t^2 u'(c_t)\frac{dc_t}{d\alpha_t}}\tag{143}$$

Proof. Start with (142). Given the decentralised scheme, and holding constant actions prior to t, consider a joint marginal change to α_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:

$$\frac{dc_t}{d\alpha_t}+\frac{dc_t}{dM_t}\frac{dM_t}{d\alpha_t}=\frac{dM_t}{d\alpha_t}\tag{144}$$

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:

$$\alpha_t u'(c_t)=\beta R\left(1-T'(s_t)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\tag{145}$$

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 (145) with respect to α_t, given constant savings, yields:

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

Rearranging (144):

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

Plugging this into the previous expression, trivial manipulations give (142).

Reasoning in a similar way for (143), consider the effect of a change to $\left(1-T_t'\left(s_t\right)\right)$ 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 $\frac{dM_t}{d(1-T'(s_t))}$ to solve:

$$\frac{dc_t}{d\left(1-T'\left(s_t\right)\right)}+\frac{dc_t}{dM_t}\frac{dM_t}{d\left(1-T'\left(s_t\right)\right)}=\frac{dM_t}{d\left(1-T'\left(s_t\right)\right)} \tag{146}$$

Differentiating (145) with respect to $\left(1-T_t'\left(s_t\right)\right)$ under this joint change gives:

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

Rearranging, and noting $\frac{dc_t}{d\left(1-T_t'\left(s_t\right)\right)}=-\frac{ds_t}{d\left(1-T_t'\left(s_t\right)\right)}$:

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

and so (143) follows, given (142). □

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:

$$\alpha_t u'\left(c_t\right)=\beta R\left(1-T_t'\left(s_t\right)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{147}$$

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:

$$\alpha_t u'(c_t) = \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}} \beta \left\{ \frac{1}{\alpha_t} \omega_{t+1}^{\Delta}(\alpha_t) - \frac{d\omega_{t+1}(\alpha_t)}{d\alpha_t} \right\} \tag{148}$$

The right-hand sides of (147) and (148) thus give alternative expressions for the shadow value of savings at the chosen allocation. Denote this object $\gamma_t(\alpha_t)$, i.e.:

$$\gamma_t(\alpha_t) := R\left(1 - T_t'(s_t)\right) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \tag{149}$$

$$= \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}} \beta \left\{ \frac{1}{\alpha_t} \omega_{t+1}^{\Delta}(\alpha_t) - \frac{d\omega_{t+1}(\alpha_t)}{d\alpha_t} \right\} \tag{150}$$

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

$$\alpha_t u''(c_t(\alpha_t)) \frac{dc_t(\alpha_t)}{d\Delta_i} = \frac{d\gamma_t(\alpha_t)}{d\Delta_i} \tag{151}$$

Consider two such changes, Δ_i and Δ_j, plus a scalar Γ, with the property:

$$\frac{d\gamma_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{d\gamma_t(\alpha_t)}{d\Delta_j} = 0 \tag{152}$$

Using (152) in (151):

$$\frac{dc_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{dc_t(\alpha_t)}{d\Delta_j} = 0 \tag{153}$$

So long as the perturbations Δ_i and Δ_j are constructed to leave M_t unaffected, this last result in turn implies:

$$\frac{ds_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{ds_t(\alpha_t)}{d\Delta_j} = 0 \tag{154}$$

and so

$$\frac{d\gamma_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{d\gamma_t(\alpha_t)}{d\Delta_j} = \left.\frac{d\gamma_t(\alpha_t)}{d\Delta_i}\right|_{s_t,c_t} + \Gamma \left.\frac{d\gamma_t(\alpha_t)}{d\Delta_j}\right|_{s_t,c_t} = 0 \tag{155}$$

– 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, (151) gives:

$$\begin{aligned}\frac{dc_t(\alpha_t)}{d\Delta_i} &= -\Gamma\frac{dc_t(\alpha_t)}{d\Delta_j} \\ &= \frac{\left.\frac{d\gamma_t(\alpha_t)}{d\Delta_i}\right|_{s_t}}{\left.\frac{d\gamma_t(\alpha_t)}{d\Delta_j}\right|_{s_t}}\frac{dc_t(\alpha_t)}{d\Delta_j}\end{aligned} \tag{156}$$

We now take the derivatives of $\gamma_t(\alpha_t)$ for two changes to the consumer's budget constraint, as viewed in t. The first is a simple change to contemporaneous post-tax returns, $\left(1 - T_t'(s_t)\right)$. From (149):

$$\begin{aligned}\left.\frac{d\gamma_t(\alpha_t)}{d\left(1 - T_t'(s_t)\right)}\right|_{s_t} &= \frac{1}{1 - T_t'(s_t)}\gamma_t(\alpha_t) \\ &= \frac{1}{1 - T_t'(s_t)}\alpha_t u'(c_t)\end{aligned} \tag{157}$$

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

$$c_{t+1} = M_{t+1} - s(M_{t+2}) \tag{158}$$

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

$$M_{t+2} \equiv R\left[s(M_{t+2}) - T_{t+1}(s(M_{t+2}))\right] \tag{159}$$

We will focus on perturbations of the form:

$$c_{t+1} = M_{t+1} - s(M_{t+2}) + \Gamma f(s(M_{t+2})) \tag{160}$$

for an arbitrary bounded, a.e. differentiable function f and scalar Γ. 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 Γ away from zero. Taking the derivative from (150), since c_t and ω_{t+1} are being held constant we can write:

$$\left.\frac{d\gamma_t(\alpha_t)}{d\Gamma}\right|_{s_t,c_t} = \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}}\beta\frac{1}{\alpha_t}\frac{d\omega_{t+1}^{\Delta}(\alpha_t)}{d\Gamma} \tag{161}$$

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

$$\begin{aligned}\frac{d\omega_{t+1}^{\Delta}(\alpha_t)}{d\Gamma} = \int_{\alpha_{t+1}} & f'(s_{t+1}(\alpha_{t+1}))\left\{\alpha_{t+1}^2 u'(c_{t+1})\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}\rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t)\right. \\ & \left.-\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}\int_{\underline{\alpha}}^{\alpha_{t+1}}\left[\tilde{\alpha}_{t+1}(u'(c_{t+1}))+\tilde{\alpha}_{t+1}^2 u''(c_{t+1})\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho(\tilde{\alpha}_{t+1}|\alpha_t)\pi(\tilde{\alpha}_{t+1}|\alpha_t)\,d\tilde{\alpha}_{t+1}\right\}d\alpha_{t+1} \\ & + f(s_{t+1}(\bar{\alpha}))\int_{\alpha_{t+1}}\left\{\alpha_{t+1}(u'(c_{t+1}))+\alpha_{t+1}^2 u''(c_{t+1})\frac{dc_{t+1}}{d\alpha_{t+1}}\right\}\rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t)\,d\alpha_{t+1}\end{aligned} \tag{162}$$

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

$$\left.\frac{dc_{t+1}}{d\Gamma}\right|_{M_{t+2}} = f(s(M_{t+2})) \tag{163}$$

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:

$$\begin{aligned}\frac{dc_{t+1}}{dM_{t+2}} &= -s'(M_{t+2})(1-\Gamma f'(s(M_{t+2}))) && (164)\\ &= -\frac{1-\Gamma f'(s(M_{t+2}))}{R\left[1-T_{t+1}'(s(M_{t+2}))\right]} && (165)\end{aligned}$$

The effect on this as Γ changes is:

$$\frac{d}{d\Gamma}\left[\frac{dc_{t+1}}{dM_{t+2}}\right] = \frac{f'(s(M_{t+2}))}{R\left[1-T_{t+1}'(s(M_{t+2}))\right]} \tag{166}$$

Notice that this is equal to:

$$f'(s(M_{t+2}))\left(1-T_{t+1}'(s(M_{t+2}))\right)\frac{d}{d\left(1-T_{t+1}'(s(M_{t+2}))\right)}\left.\left[\frac{dc_{t+1}}{dM_{t+2}}\right]\right|_{\Gamma=0} \tag{167}$$

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 ω^{Δ}_{t+1} to the extent that consumption is deferred:

$$\begin{aligned}
\left.\frac{d\omega^{\Delta}_{t+1}}{d\Gamma}\right|_{\text{sub}} =& \int_{\alpha_{t+1}} \left\{-\alpha_{t+1}u'\left(c_{t+1}\right) + \beta R\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)\int_{\alpha_{t+1}} V_M\left(\alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+2}|\alpha_{t+1}\right)}{d\alpha_{t+1}}d\alpha_{t+2}\right\} \\
&\times \frac{ds_{t+1}\left(\alpha_{t+1}\right)}{d\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)}\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \\
=& \int_{\alpha_{t+1}} \left\{-\alpha^2_{t+1}u'\left(c_{t+1}\right)\frac{dc_{t+1}\left(\alpha_{t+1}\right)}{d\alpha_{t+1}}\frac{1}{\varepsilon^s_{t+1}s_{t+1}\left(\alpha_{t+1}\right)}\right\} \\
&\times \frac{ds_{t+1}\left(\alpha_{t+1}\right)}{d\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)}\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \\
=& -\int_{\alpha_{t+1}} \alpha^2_{t+1}u'\left(c_{t+1}\right)\frac{dc_{t+1}\left(\alpha_{t+1}\right)}{d\alpha_{t+1}}f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}
\end{aligned} \tag{168}$$

where the intermediate line makes use of Lemma 6, and we have used the fact that the specified perturbation is the equivalent of a change in $\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)$ by $\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)$ units.

Similarly, the income effect on ω^{Δ}_{t+1} will be:

$$\begin{aligned}
\left.\frac{d\omega^{\Delta}_{t+1}}{d\Gamma}\right|_{\text{inc}} =& \int_{\alpha_{t+1}} f\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\left\{\alpha_{t+1}\left(u'\left(c_{t+1}\right)\right) + \frac{ds_{t+1}\left(\alpha_{t+1}\right)}{dM_{t+1}}\left[-\alpha_{t+1}u'\left(c_{t+1}\right)\right.\right. \\
&\left.\left.+\beta R\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)\int_{\alpha_{t+1}} V_M\left(\alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+2}|\alpha_{t+1}\right)}{d\alpha_{t+1}}d\alpha_{t+2}\right]\right\}\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \\
=& \int_{\alpha_{t+1}} f\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\left\{\alpha_{t+1}\left(u'\left(c_{t+1}\right)\right) + \alpha^2_{t+1}u''\left(c_{t+1}\right)\frac{dc_{t+1}}{d\alpha_{t+1}}\right\}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \\
=& f\left(s_{t+1}\left(\bar{\alpha}\right)\right)\int_{\alpha_{t+1}} \left\{\alpha_{t+1}\left(u'\left(c_{t+1}\right)\right) + \alpha^2_{t+1}u''\left(c_{t+1}\right)\frac{dc_{t+1}}{d\alpha_{t+1}}\right\}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \\
&- \int_{\alpha_{t+1}} f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\frac{ds_{t+1}\left(\alpha_{t+1}\right)}{d\alpha_{t+1}} \\
&\times \left\{\int_{\underline{\alpha}}^{\alpha_{t+1}}\left[\tilde{\alpha}_{t+1}\left(u'\left(c_{t+1}\right)\right) + \tilde{\alpha}^2_{t+1}u''\left(c_{t+1}\right)\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho\left(\tilde{\alpha}_{t+1}|\alpha_t\right)\pi\left(\tilde{\alpha}_{t+1}|\alpha_t\right)d\tilde{\alpha}_{t+1}\right\}d\alpha_{t+1}
\end{aligned} \tag{169}$$

where the second equality makes use of (142). Taking substitution and income effects, (168) and (169), together, we obtain the result. □

Applying (156), (157) and (161), we have:

$$\frac{dc_t}{d\Gamma} = \left[\frac{\alpha_t u'\left(c_t\right)}{\left(1-T'_t\left(s_t\right)\right)}\right]^{-1}\frac{dc_t}{d\left(1-T'_t\left(s_t\right)\right)}\frac{1}{\frac{dc_t}{d\alpha_t}}\beta\frac{1}{\alpha_t}\frac{d\omega^{\Delta}_{t+1}}{d\Gamma} \tag{170}$$

So:

$$\begin{aligned}\frac{1}{s_t}\frac{ds_t}{d\Gamma} &= \left[\frac{\alpha_t u'(c_t)}{\left(1-T_t'(s_t)\right)}\right]^{-1}\frac{1}{s_t}\frac{ds_t}{d\left(1-T_t'(s_t)\right)}\frac{1}{\frac{dc_t}{d\alpha_t}}\beta\frac{1}{\alpha_t}\frac{d\omega_{t+1}^{\Delta}}{d\Gamma}\\ &= \int_{\alpha_{t+1}} f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\varepsilon_t^s\frac{\alpha_{t+1}\frac{ds_{t+1}}{d\alpha_{t+1}}}{\alpha_t\frac{ds_t}{d\alpha_t}}\left\{-\frac{\beta\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)\right. \qquad (171)\\ &\quad\left.+\frac{1}{\alpha_{t+1}}\int_{\underline{\alpha}}^{\alpha_{t+1}}\frac{\beta\tilde{\alpha}_{t+1}\left(u'(c_{t+1})\right)}{\alpha_t u'(c_t)}\left[1+\frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho\left(\tilde{\alpha}_{t+1}|\alpha_t\right)\pi\left(\tilde{\alpha}_{t+1}|\alpha_t\right)d\tilde{\alpha}_{t+1}\right\}d\alpha_{t+1}\\ &\quad+f\left(s_{t+1}\left(\bar{\alpha}\right)\right)\varepsilon_t^s\frac{1}{\alpha_t\frac{dc_t}{d\alpha_t}}\int_{\alpha_{t+1}}\frac{\beta\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)}\left[1+\frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\end{aligned}$$

A unit change in Γ changes the slope of the $t+1$ budget constraint at s_{t+1} by $\left(1-T_{t+1}'\left(s_{t+1}\right)\right)f'\left(s_{t+1}\right)$ units, and shifts it uniformly for all higher savings levels by the same amount. Consistent with the definition in Appendix A.9, we have:

$$\begin{aligned}\frac{1}{s_t}\frac{ds_t}{d\Gamma} \equiv& \int_{\alpha_{t+1}} f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\epsilon_{t,t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\Pi\left(\alpha_{t+1}|\alpha_t\right)\left(-\frac{ds_{t+1}\left(\alpha_{t+1}\right)}{d\alpha_{t+1}}\right)d\alpha_{t+1} \qquad (172)\\ &+f\left(s_{t+1}\left(\bar{\alpha}\right)\right)\frac{1}{s_t}\left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}}\end{aligned}$$

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

$$\begin{aligned}\epsilon_{t,t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right) =& -\varepsilon_t^s\frac{\alpha_{t+1}}{\Pi\left(\alpha_{t+1}|\alpha_t\right)\alpha_t\frac{ds_t}{d\alpha_t}}\left\{-\frac{\beta\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)\right. \qquad (173)\\ &+\frac{1}{\alpha_{t+1}}\int_{\underline{\alpha}}^{\alpha_{t+1}}\frac{\beta\tilde{\alpha}_{t+1}\left(u'(c_{t+1})\right)}{\alpha_t u'(c_t)}\left[1+\frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\\ &\left.\times\rho\left(\tilde{\alpha}_{t+1}|\alpha_t\right)\pi\left(\tilde{\alpha}_{t+1}|\alpha_t\right)d\tilde{\alpha}_{t+1}\right\} \qquad (174)\end{aligned}$$

So:

$$RT_t'\left(s_t\right)s_t\epsilon_{t,t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right) \tag{175}$$

$$=-RT_t'\left(s_t\right)s_t\varepsilon_t^s\frac{\alpha_{t+1}}{\Pi\left(\alpha_{t+1}|\alpha_t\right)\alpha_t\frac{ds_t}{d\alpha_t}}\left\{-\frac{\beta\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)\right. \tag{176}$$

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

$$=\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\alpha_t u'\left(c_t\right)\left\{-\frac{\beta R\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\rho\left(\alpha_{t+1}|\alpha_t\right)\frac{\alpha_{t+1}\pi\left(\alpha_{t+1}|\alpha_t\right)}{\Pi\left(\alpha_{t+1}|\alpha_t\right)}\right. \tag{177}$$

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

$$=-\beta R\alpha_{t+1}u'\left(c_{t+1}\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\frac{\alpha_{t+1}\pi\left(\alpha_{t+1}|\alpha_t\right)}{\Pi\left(\alpha_{t+1}|\alpha_t\right)} \tag{178}$$

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

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

For condition 4, we have established:

$$\frac{1}{s_t}\left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}}=\varepsilon_t^s\frac{1}{\alpha_t\frac{dc_t}{d\alpha_t}}\int_{\alpha_{t+1}}\frac{\beta\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{179}$$

So:

$$RT_t'\left(s_t\right)\left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}}=\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\alpha_t u'\left(c_t\right)\int_{\alpha_{t+1}}\frac{\beta R\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{180}$$

$$=\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\int_{\alpha_{t+1}}\beta R\alpha_{t+1}u'\left(c_{t+1}\right)\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{181}$$

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

A.11 Proof of Theorem 2

Equation (36), above, gives:

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

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 $u''\left(c_t\right)<0$, 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 8. *For all t and α^{t-1}, the long-run expectation of variation in the inverse, discounted marginal utility of consumption is bounded, i.e.:*

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

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

$$\begin{aligned}&\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',...,\bar{\alpha}\right)\right)}-\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',...,\underline{\alpha}\right)\right)}\right]\\ =&\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}}\frac{\bar{\alpha}}{\mathbb{E}_s\left[\alpha_{s+1}|\bar{\alpha}\right]}\mathbb{E}_s\left[\frac{1}{\beta R u'\left(c_{s+1}\left(\alpha_t',...,\bar{\alpha},\alpha_{s+1}\right)\right)}\middle|\bar{\alpha}\right]\right]\\ &-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}}\frac{\underline{\alpha}}{\mathbb{E}_s\left[\alpha_{s+1}|\underline{\alpha}\right]}\mathbb{E}_s\left[\frac{1}{\beta R u'\left(c_{s+1}\left(\alpha_t',...,\underline{\alpha},\alpha_{s+1}\right)\right)}\middle|\underline{\alpha}\right]\right]\end{aligned}\tag{182}$$

If the right-hand side of (182) is unbounded in s then so is the expression:

$$\begin{aligned}&\frac{1}{(\beta R)^{s-t}}\mathbb{E}_t\left[\mathbb{E}_s\left[\frac{1}{\beta R u'\left(c_{s+1}\left(\alpha_t',...,\bar{\alpha},\alpha_{s+1}\right)\right)}\middle|\bar{\alpha}\right]-\frac{\underline{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\bar{\alpha}\right]}{\bar{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\underline{\alpha}\right]}\mathbb{E}_s\left[\frac{1}{\beta R u'\left(c_{s+1}\left(\alpha_t',...,\underline{\alpha},\alpha_{s+1}\right)\right)}\middle|\underline{\alpha}\right]\right]\\ =&\frac{1}{(\beta R)^{s+1-t}}\mathbb{E}_t\left[\int_{\alpha_{s+1}}\left[\frac{1}{u'\left(c_{s+1}\left(\alpha_t',...,\bar{\alpha},\alpha_{s+1}\right)\right)}\pi\left(\alpha_{s+1}|\bar{\alpha}\right)-\frac{1}{u'\left(c_{s+1}\left(\alpha_t',...,\underline{\alpha},\alpha_{s+1}\right)\right)}\frac{\underline{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\bar{\alpha}\right]}{\bar{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\underline{\alpha}\right]}\pi\left(\alpha_{s+1}|\underline{\alpha}\right)\right]d\alpha_{s+1}\right]\end{aligned}\tag{183}$$

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

$$\frac{1}{(\beta R)^{s+1-t}}\mathbb{E}_t\left[\int_{\alpha_{s+1}}\frac{1}{u'\left(c_{s+1}\left(\alpha_t',...,\underline{\alpha},\alpha_{s+1}\right)\right)}\left[\frac{\pi\left(\alpha_{s+1}|\bar{\alpha}\right)-\pi\left(\alpha_{s+1}|\underline{\alpha}\right)\frac{\underline{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\bar{\alpha}\right]}{\bar{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\underline{\alpha}\right]}}{\pi\left(\alpha_{s+1}|\underline{\alpha}\right)}\right]\pi\left(\alpha_{s+1}|\underline{\alpha}\right)d\alpha_{s+1}\right]\tag{184}$$

Moreover, continuity of the density implies that the object:

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

is bounded in α_{s+1}. Thus unboundedness of (184) in s implies:

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

This implies:

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

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

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

But since consumption is increasing in α_s, we have:

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

Thus:

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

must imply

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

But:

$$\begin{aligned}\lim_{s\to\infty}\left\{\frac{1}{\mathbb{E}_t\left[\alpha_s\right]}\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\right)}\right]\right\}&=\frac{1+\lambda_{t+1}\left(\alpha_t\right)}{\eta_t}\\&=\frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\mathbb{E}_t\left[\alpha_{t+1}\right]}\mathbb{E}_t\left[\frac{1}{\beta R u'\left(c_{t+1}\right)}\right]-\frac{\varepsilon^{\alpha}\left(\alpha_t\right)}{\alpha_t u'\left(c_t\right)}\right\}\end{aligned}\tag{185}$$

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

□

Turning to the main argument, a combination of (62) and (63) gives:

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

Or, using the definitions of ρ and π^{Δ}:

$$\lambda^{\Delta}_{t+1}\left(\alpha_t\right)\alpha_t\pi\left(\alpha_t|\alpha_{t-1}\right)=\int_{\alpha_t}^{\bar{\alpha}}\left[\left(1+\lambda_t+\lambda^{\Delta}_t\pi^{\Delta}\left(\tilde{\alpha}_t|\alpha_{t-1}\right)\right)-\left(1+\lambda_{t+1}\left(\tilde{\alpha}_t\right)\right)\right]\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha}_t\tag{186}$$

with:

$$\int_{\underline{\alpha}}^{\bar{\alpha}}\left[\left(1+\lambda_t+\lambda^{\Delta}_t\pi^{\Delta}\left(\alpha_t|\alpha_{t-1}\right)\right)-\left(1+\lambda_{t+1}\left(\alpha_t\right)\right)\right]\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t=0\tag{187}$$

Since $\pi^{\Delta}\left(\alpha_t|\alpha_{t-1}\right)$ is monotone increasing in α_t, and $\lambda^{\Delta}_0=0$, a sufficient condition for the right-hand side of (186) to be weakly positive for all t and all histories is that $(1+\lambda_{t+1}(\alpha_t))$ should be non-increasing in α_t. Lemma 8 enables this to be established. We have:

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

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

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

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

Clearly $\mathbb{E}_t[\alpha_s]$ is weakly increasing in α_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:

$$\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha''_t,...,\alpha_s\right)\right)}\middle|\alpha''_t\right]-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha'_t,...,\alpha_s\right)\right)}\middle|\alpha'_t\right]\tag{188}$$

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

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

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

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

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

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

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

Lemma 8 implies that the expectation is finite as $s \to \infty$, and so this object converges to zero as s becomes large. Since all other components of the difference (188) are weakly negative, uniform convergence guarantees that $\frac{1+\lambda_{t+1}\left(\alpha_t''\right)}{\eta_t} - \frac{1+\lambda_{t+1}\left(\alpha_t'\right)}{\eta_t} \leq 0$.

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

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

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

The first case implies $\lambda^{\Delta}_{t+1}(\alpha_t) > 0$ everywhere except endpoints, from (186) and the fact that the integral in (186) is zero over the full range.

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

$$\frac{1}{\alpha_t u'(c_t)} = \frac{1}{\mathbb{E}_t[\alpha_{t+1}]}\mathbb{E}_t\left[\frac{1}{\beta R u'(c_{t+1})}\right] \tag{189}$$

with both objects constant in α_t. Suppose for now that types are persistent ($\rho(\alpha|\alpha') > 0$ for all type pairs). For the right-hand side of (189) to be constant in α_t, and given nomality, a necessary requirement is that the partial derivatives due to persistence are weakly positive for all α_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:

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

But since $\lambda^{\Delta}_{t+1} = 0$, and we have already established $\lambda^{\Delta}_s \geq 0$ for all s, condition (75) implies:

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

for all α'_{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, $\frac{\frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t}}{\pi(\alpha_{t+1}|\alpha_t)}$ is a strictly increasing function, and so:

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

with the inequality strict (a contradiction) unless the object in curly brackets is zero everywhere. This in turn would imply that $\lambda^{\Delta}_{t+2} = 0$ everywhere. Repeating the argument, this would imply $\lambda^{\Delta}_{t+3} = 0$ at all successor nodes, and so on. Thus the only possibility consistent with $\lambda^{\Delta}_{t+1} = 0$ at interior points is that λ^{Δ} 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 $\lambda^{\Delta}_{t+1}(\alpha_t) = 0$ for all α_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 α_t. But if the right-hand side is constant in α_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 α_t to keep the left-hand side constant.

A.12 Proof of Proposition 5

From Lemma 3, $RT'_{t-1}(s_{t-1}) s_{t-1} \epsilon_{t-1,t}(s'_t)$ is equal to:

$$\begin{aligned} &-\rho(\alpha_t(c'_t)|\alpha_{t-1}) \beta R \frac{\lambda^{\Delta}_t}{\eta_{t-1}} (\alpha_t(c'_t))^2 u'(c'_t) \left(\frac{d\alpha_t(c'_t)}{dc_t} \right)^{-1} \frac{\pi^c(c'_t|\alpha^{t-1})}{\Pi^c(s'_t|\alpha^{t-1})} \\ &+ \frac{1}{\Pi^c(s'_t|\alpha^{t-1})} \int_{\underline{c}}^{c'_t} \rho(\alpha_t(c_t)|\alpha_{t-1}) \beta R \frac{\lambda^{\Delta}_t}{\eta_{t-1}} \left[\alpha_t(c_t)(u'(c_t)) + (\alpha_t(c_t))^2 u''(c_t) \left(\frac{d\alpha_t(c_t)}{dc_t} \right)^{-1} \right] \pi^c\left(c_t|\alpha^{t-1}\right) dc_t \end{aligned} \tag{190}$$

Switching to express arguments in terms of α_t:

$$\begin{aligned} &-\rho(\alpha_t|\alpha_{t-1}) \beta R \frac{\lambda^{\Delta}_t}{\eta_{t-1}} (\alpha_t)^2 u'\left(c_t(\alpha'_t)\right) \frac{\pi(\alpha'_t|\alpha_{t-1})}{\Pi(\alpha'_t|\alpha_{t-1})} \\ &+ \frac{1}{\Pi(\alpha'_t|\alpha_{t-1})} \int_{\underline{\alpha}}^{\alpha'_t} \rho(\alpha_t|\alpha_{t-1}) \beta R \frac{\lambda^{\Delta}_t}{\eta_{t-1}} \left[\alpha_t(u'(c_t(\alpha_t))) + \alpha_t^2 u''(c_t(\alpha_t)) \frac{dc_t(\alpha_t)}{d\alpha_t} \right] \pi(\alpha_t|\alpha_{t-1}) d\alpha_t \end{aligned} \tag{191}$$

Integration by parts gives the following relationship:

$$\begin{aligned} &\int_{\underline{\alpha}}^{\alpha'_t} \rho(\alpha_t|\alpha_{t-1}) \beta R \frac{\lambda^{\Delta}_t}{\eta_{t-1}} \left[\alpha_t(u'(c_t(\alpha_t))) + \alpha_t^2 u''(c_t(\alpha_t)) \frac{dc_t(\alpha_t)}{d\alpha_t} \right] \pi(\alpha_t|\alpha_{t-1}) d\alpha_t \\ &- \rho(\alpha_t|\alpha_{t-1}) \beta R \frac{\lambda^{\Delta}_t}{\eta_{t-1}} (\alpha_t)^2 u'\left(c_t(\alpha'_t)\right) \pi(\alpha'_t|\alpha_{t-1}) \\ =&\beta R \frac{\lambda^{\Delta}_t}{\eta_{t-1}} \int_{\underline{\alpha}}^{\alpha'_t} \alpha_t u'(c_t(\alpha_t)) \frac{\alpha_{t-1} \frac{d\pi(\alpha_t|\alpha_{t-1})}{d\alpha_{t-1}}}{\pi(\alpha_t|\alpha_{t-1})} \pi(\alpha_t|\alpha_{t-1}) d\alpha_t \end{aligned} \tag{192}$$

So:

$$s_{t-1}\epsilon_{t-1,t}(s'_t) = \beta \frac{\left(\frac{\lambda^{\Delta}_t}{\eta_{t-1}}\right)}{T'_{t-1}(s_{t-1})} \mathbb{E}_{t-1} \left[\alpha_t u'(c_t(\alpha_t)) \frac{\alpha_{t-1} \frac{d\pi(\alpha_t|\alpha_{t-1})}{d\alpha_{t-1}}}{\pi(\alpha_t|\alpha_{t-1})} \middle| \alpha_t \le \alpha'_t \right] \tag{193}$$

The expectation term contains two objects that are monotone increasing in α_t, under the maintained assumptions (including MLRP). The term $\frac{\frac{d\pi(\alpha_t|\alpha_{t-1})}{d\alpha_{t-1}}}{\pi(\alpha_t|\alpha_{t-1})}$ is zero in expectation, and so crosses zero once, whilst $\alpha_t u'(c_t(\alpha_t))$ is strictly positive. From Lemma 5:

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

where the inequality follows a step used in the proof of Lemma 8. Thus for sufficiently low α'_t (corresponding to high s'_t) both sides of the expression must be negative, whist positive correlation between the components implies it is positive for sufficiently high α'_t (low s'_t). Once positive for a given α'_t, it is clear that the expectation term must remain positive for all $\alpha''_t > \alpha'_t$, since $\frac{\frac{d\pi(\alpha_t|\alpha_{t-1})}{d\alpha_{t-1}}}{\pi(\alpha_t|\alpha_{t-1})}$ must be positive on the higher type range. The result follows.