Optimal nonlinear savings taxation*

Charles Brendon†

February 16, 2022

Abstract

This paper analyses the design of optimal nonlinear savings taxation, in a multi-period consumption-savings economy where consumers face persistent, uninsurable shocks to the marginal value that they place on consuming. Its main contribution is to show that shocks of this kind generically justify positive marginal savings taxes, and to characterise these taxes by reference to a limited number of 'sufficient statistics'. The method for obtaining this characterisation is novel, and – at least in this model – shows how 'Mirrleesian' and 'sufficient statistics' approaches to dynamic taxation may be reconnected.

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

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

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

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

1 Introduction

This paper revisits the problem of optimal taxation in dynamic economies with private information. Adapting the dynamic asymmetric information problem due to Atkeson and Lucas (1992), it considers a government that would like to insure individuals against persistent, unobservable shocks over time to their marginal utility of consumption. Though this environment has been studied extensively in the social insurance literature, we derive a number of novel results that clarify the character of optimal policy within it. In doing so, we provide a more general contribution to the links between dynamic Mirrleesian and sufficient statistic analyses of optimal taxation. Specifically, we show that the assumptions embedded in the dynamic asymmetric information problem are sufficient simplifications for optimal nonlinear taxation to abide by tractable elasticity formulae in dynamic settings.

To this end, we first show that constrained-optimal allocations in the Atkeson-Lucas environment can be decentralised in a standard consumption-savings economy, provided savings are subject to a nonlinear tax period-by-period. Given this, we provide a simple, intuitive characterisation of the corresponding tax instruments at any optimum. The marginal tax rate on savings is generically positive in all time periods, limiting to zero at extreme ends of the savings distribution. The revenue from it is used to fund a positive lump-sum transfer each period, which permits higher within-period consumption for those with high 'need', or high marginal utility. Net of the lump-sum component, the expected total tax bill for individuals each period is zero *ex-ante*, and its realised value *ex-post* increases monotonically in their choice of savings.

For any given time period and shock history, we show that the optimal marginal savings tax rate can be written as a function of a small number of behavioural statistics. 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 our 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 our 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] Our results provide a counter-point to this. The characterisation of optimal policy that we offer is not significantly more complex than a textbook Saez formula with income effects. It is written in terms of behavioural objects that are defined independently of the utility function and hidden type process, with the sole exception of social welfare weights – in which a reference to marginal utility is standard.

Indeed, the most remarkable feature of our 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

[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).

marginal tax rate, the marginal effect of higher income on savings, and a compensated elasticity that measures effect of a change to insurance at t on savings at $t-1$.

This simplification can be interpreted through the lens of the Atkinson-Stiglitz theorem. Consistent with the wider literature, we impose a Markovian structure on shocks. This means that *conditional on the next period's type draw*, preferences across alternative insurance schemes more than one period ahead are independent of an individual's current type. It follows that there is no justification for distorting type-specific allocations in $t+1$, so as to imprve allocations in time periods prior to t. This keeps the relevant behavioural considerations for policy design to a manageable level.

The important general lesson, we argue, is that the mechanism design approach imposes structural restrictions on consumer preferences that allow optimal dynamic tax problems to become tractable. Far from complexifying, the Mirrlees approach can provide a powerful, theoretically-grounded basis for *simple* policy advice in dynamic settings – drastically limiting the relevant set of behavioural statistics. Through novel, but elementary, manipulations, we provide a roadmap for achieving this.

1.1 Preview of main characterisation

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

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

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, $T'(s)\frac{ds}{dM}$, less a social welfare weight, $g(s)$, 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, discussed in detail below – is decreasing in savings, because higher savers have a relatively low contemporaneous marginal utility of consumption.

The right-hand side of the equation gives the fiscal benefits of substitution effects that are induced by the tax change. When taxes are cut, savings in the current period increase in proportion to the contemporaneous savings elasticity ε^{s}. This raises revenue in proportion to the marginal tax rate $T'(s')$. Cutting taxes at s' in the current period may also change savings in the previous period, by an amount proportional to a cross elasticity $\epsilon^{s}_{-1}(s')$. This raises additional income in the previous period, in proportion to that period's marginal tax rate $T'_{-1}(s_{-1})$, whose relative value depends on the gross real interest rate R.

The cross elasticity $\epsilon^{s}_{-1}(s')$ is the least conventional of the objects in the characterisation, and is intrinsically linked to type persistence: it is zero when types are iid. We discuss it in detail in the body of the paper.[4] As

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

[4] See Section 9.3.

usual for Mirrleesian problems, the importance of substitution effects relative to income effects depends on the equilibrium distribution of the taxed quantity – here: how many agents locate at the savings level s', relative to those above it? This is captured by the hazard rate $\frac{\pi^s(s')}{1-\Pi^s(s')}$, the ratio of the density of savers at s' to the probability that savings exceed s'.

Equation (1) is also helpful for understanding the result in Theorem 2, 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 raising 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, and this comes at a net cost. Reflecting the individual's own *ex-ante* insurance preferences, this is an undesirable diversion of resources. It is optimal to retain positive marginal savings taxes, as this allows more resources to be directed towards lower savers.

1.2 Paper outline

The rest of the paper is organised as follows. Section 2 provides an overview of related literature. Section 3 introduces the detailed setup of the dynamic information-theoretic problem that we study. Section 4 outlines how nonlinear savings taxes can be used to decentralise incentive-compatible allocations for this environment, and derives sufficient conditions on the allocation for this decentralisation to work. Like much of the optimal taxation literature dating back to Mirrlees (1971), we keep analysis tractable via a 'first-order' approach to incentive compatibility: Section 5 reminds readers of this approach, and provides a novel, intuitive increasingness condition on the allocation that guarantees its validity.

To aid understanding, our main characterisation is presented constructively, in steps. In Section 6 we use standard methods to characterise constrained-optimal allocations by reference to the costs and benefits of changing *utility* levels for a cross-section of individuals. The resulting expressions are insightful, and reveal novel features about the dynamics of consumption when types are persistent, but they rely heavily on arguments of the utility function. Section 7 explains our novel approach for mapping from these utility-based expressions to a more 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 properties of optimal savings taxes, and explains the link between our results and established principles in the optimal taxation literature – notably the Atkinson-Stiglitz theorem. Section 10 concludes.

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

2 Relation to literature

The basic insurance problem that we study was first popularised by Atkeson and Lucas (1992), who focused on the properties of constrained-optimal allocations in the presence of unobservable shocks to marginal utility. Their paper gave particular attention to long-run outcomes, showing that the immiseration result of Thomas and

Worrall (1990) carried over to their setting, as well as emphasising that the optimum could not be decentralised via conventional linear pricing. Technically, Atkeson and Lucas assumed an iid type distribution, with types drawn from a finite set period-by-period – a structure retained by more recent literature that explores the sensitivity of their immiseration result.[5] Our paper instead allows for persistent (Markovian) type draws, which has non-trivial implications for consumption dynamics relative to the iid case. We 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, our paper is situated in the dynamic Mirrleesian public finance tradition, analysing optimal tax systems subject to the deep information frictions that necessitate departures from the Second Welfare Theorem. Unlike our paper, 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.[6] 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.[7] 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 this 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. Like us, 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 ours.

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

[5] See, for instance, Sleet and Yeltekin (2006) and Farhi and Werning (2007).

[6] 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.

[7] The inverse Euler condition had been derived in less general settings by Diamond and Mirrlees (1978) and Rogerson (1985).

microeconomic literature with quasilinear preferences, like Hellwig (2021) we adapt their methodology to our setting of interest.

Away from the dynamic Mirrlees literature, our paper contributes to the growing 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. Yet in contrast with the static literature, for dynamic tax problems 'sufficient statistics' and 'mechanism design' approaches are increasingly treated as rivals rather than complements – methods that generate distinct policy prescriptions, rather than distinct methods for understanding the same prescriptions.

The reason for this separation has been a desire to obtain policy lessons for dynamic tax environments that are as simple as for static, and the seeming difficulty of achieving this in a mechanism design setting. Influential papers by Piketty and Saez (2013a) and Saez and Stantcheva (2018) have thus quite 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 t to taxation in s, say – by asserting that all that matters is what happens in the long run.[9] Our 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 we highlight an alternative route to policy insight from the more radical focus on long-run outcomes alone.

By offering a novel justification for savings taxation, our 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 desirable *faute de mieux*, given limitations elsewhere in the tax system – particularly credit constraints, limits on age-dependent taxation or arbitrary tax ceilings. It provides few arguments *for* savings taxes *per se*. In this regard our 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, our paper has links to the growing microeconomic mechanism design literature that gives particular attention to the problems implied by type persistence.[12] Current papers by Bloedel, Krishna and Strulovici (2020), Bloedel, Krishna and Leukhina (2020) and Makris and Pavan (2020) deploy various settings to explore the dynamics of wedges in problems without quasilinearity. Our paper provides a novel decomposition of consumption dynamics into two distinct multiplier processes – one stationary, one nonstationary – that helps shed light on

[8]Stantcheva (2020) provides 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.

[10]The significance of the Straub-Werning critique remains disputed – see, in particular, Chari, Nicolini and Teles (2020).

[11]Influential references include Aiyagari (1995), İmrohoroğlu (1998), Erosa and Gervais (2002) and Conesa, Kitao and Krueger (2009).

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

the distinct roles played by type persistence and risk aversion in these settings.

3 Model setup

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 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:

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

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

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

Type draws are assumed to be independent across individuals, so there is no aggregate risk. Since there is no other intrinsic source of uncertainty, and no policy reason to introduce one artificially, an agent's consumption in period t will be measurable with respect to their history 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 $\Pi(\alpha_{t+1}|\alpha_t)$, with conditional density $\pi(\alpha_{t+1}|\alpha_t)$. The equivalent (unconditional) objects for period 0 are denoted $\Pi(\alpha_0)$ and $\pi(\alpha_0)$ respectively. We place the following reguarity structure on the distributions:

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

Notice that the density functions may be 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, which 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 elasticity of the distribution of types at $t+1$ with respect to type at t. Specifically, let $\rho(\alpha_{t+1}|\alpha_t)$ be defined by:[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 these two 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 $\varepsilon^{\alpha}(\alpha_t) < 1$. Thus higher current α may raise expectations about future type, but not by so much as the increase in current type.

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 $D_{t,s}(\alpha^s)$:

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

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

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

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

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

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

Finally, we 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.

[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 $\rho(\alpha_{t+1}|\alpha_t) \equiv \rho$.

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

Note that this condition implies that $\pi^{\Delta}(\alpha_{t+1}|\alpha_t)$ 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:

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

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. Individuals report their type each period, and receive a consumption allocation conditional on their reports to date, $c_t(\alpha^t)$. A complete set of $c_t(\alpha^t)$ functions for all $t \geq 0$ and $\alpha^t \in A^{t+1}$ is referred to as an **allocation**.[14]

The planner's choice is restricted by the resource and incentive compatibility constraints detailed below, plus a technical **interiority restriction**. This is defined by a set of scalars $\{K_t\}_{t\geq 0}$ and the bound:

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

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.

3.4 Resources

The resource constraint is the simpler of the two main restrictions. We assume that there is an exogenous, time-invariant world real interest rate, whose gross value is $R \leq \beta^{-1}$. The constraint requires net-present value of consumption to equal the net-present value of endowments:

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

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

[14] The dependence of c_t on α^t will be left implicit where the context allows.

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 $t \geq 0$. The Markov property of shocks means the value of V_{t+1} is unaffected by the truthfulness, or otherwise, of past reports.

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' features in two places in restriction (12). Most directly it affects the marginal utility of consumption in t. But current type is also allowed to influence the distribution of future type draws, $\Pi\left(\alpha_{t+1}|\alpha_t'\right)$. This second channel complicates the variation of preferences as α_t' varies.

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 A consumption-savings decentralisation

The analysis focuses heavily on a decentralisation of the chosen allocation via period-by-period consumption-savings choice. This works as follows. Individuals enter each period t with a given value of net wealth, M_t, normalised to include the value of future endowments. Wealth in t can either be allocated to period-t consumption, c_t, or savings, s_t. This choice is observable, and the planner implements a non-linear tax on s_t, which may vary in the history of past savings decisions s^{t-1}. This tax is denoted $T_t\left(s_t;s^{t-1}\right)$, or simply $T_t\left(s_t\right)$ if context allows. The tax is normalised to equal zero on average, given s^{t-1} and the equilibrium distribution of future choices.

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

$$c_t + s_t = M_t \tag{13}$$

$$M_{t+1} = R\left[s_t - T_t\left(s_t; s^{t-1}\right)\right] \tag{14}$$

Given M_0, individuals choose contingent consumption sequences to maximise U_0, subject to (13) and (14), plus a ‘no Ponzi’ constraint:

$$\lim_{t\to\infty} R^{-t} M_t \geq 0 \tag{15}$$

Notice that conditions (13) to (15) together imply a forward-looking infinite-horizon budget constraint that must be satisfied in all periods $r \geq 0$, for any realised consumption-savings path:

$$M_r = \sum_{t=0}^{\infty} R^{-t}\left[c_{t+r} + T_{t+r}\left(s_{t+r}; s^{t+r-1}\right)\right] \tag{16}$$

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

Proposition 1 provides conditions under which an incentive-feasible allocation can be decentralised by a tax scheme of this kind.

Proposition 1. *An incentive-feasible allocation* $\left\{c_t^*\left(\alpha^t\right)\right\}_{t,\alpha^t}$ *can be decentralised by a sequence of tax functions* $T_t\left(s_t; s^{t-1}\right)$ *provided:*

1. $c_t^*\left(\alpha^{t-1}, \alpha_t\right)$ *is increasing in* α_t *for all* t *and* α^{t-1}*, and*

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

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

One of the main contributions of this paper is to show that it is possible to characterise the constrained-optimal allocation in a simple and intuitive way, by reference to economic statistics that arise in this decentralisation. These statistics are, in particular, the marginal savings tax rate given past outcomes, the elasticity of savings with respect to current and future marginal tax rates, the effect on savings of higher wealth, M_t, and the endogenous conditional distribution of savings at each history node.

5 First-order incentive compatibility

5.1 A relaxed incentive constraint

Condition (12) implies a continuum of constraints for every element of A, after every history α^{t-1}. 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, we thus replace them with a 'first-order' envelope requirement that is necessary for (12) to be true, but not suffiicient.[15] This is easiest stated 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:

$$\omega_t\left(\alpha^{t-1}\right) := \int_{\alpha_t} \left\{\alpha_t u\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right) + \beta\omega_{t+1}\left(\alpha^{t-1},\alpha_t\right)\right\} d\Pi\left(\alpha_t|\alpha_{t-1}\right) \tag{17}$$

$$\omega_t^{\Delta}\left(\alpha^{t-1}\right) := \int_{\alpha_t} \rho\left(\alpha_t|\alpha_{t-1}\right)\cdot\left\{\alpha_t u\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right) + \beta\omega_{t+1}^{\Delta}\left(\alpha^{t-1},\alpha_t\right)\right\} d\Pi\left(\alpha_t|\alpha_{t-1}\right) \tag{18}$$

Formally, we then have the following result:

Lemma 1. *Any incentive-feasible allocation satisfies the following envelope condition for all $t \geq 0$, all $\alpha^{t-1} \in A^t$ and all $\alpha_t' \in A$:*

$$\begin{aligned}\alpha_t' u\left(c_t\left(\alpha_t'\right)\right) + \beta\omega_{t+1}\left(\alpha_t'\right) = {} & \underline{\alpha} u\left(c_t\left(\underline{\alpha}\right)\right) + \beta\omega_{t+1}\left(\underline{\alpha}\right) \\ & + \int_{\underline{\alpha}}^{\alpha_t'} \frac{1}{\alpha_t}\left[\alpha_t u\left(c_t\left(\alpha_t\right)\right) + \beta\omega_{t+1}^{\Delta}\left(\alpha_t\right)\right] d\alpha_t\end{aligned} \tag{19}$$

We refer to (19) as the *relaxed incentive compatibility condition.* For an arbitrary type α_t', this expression decomposes the value of U_t into the value for the lowest type $\underline{\alpha}$, plus the sum of 'information rents' between $\underline{\alpha}$ and α_t'. The 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 (19) 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. As is well known, confirming this inclusion is the central issue in justifying the first-order approach.

[15] 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.

5.2 Sufficiency

When will (19) 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, $\Pi(\alpha_{t+1}|\alpha_t)$. 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 distribution of future outcomes that provides unusually high continuation utility in states with a high α_{t+1} draw.

Global incentive compatibility can be confirmed by reference to an 'integral monotonicity condition', of the type introduced in quasilinear settings by Pavan, Segal and Toikka (2014).[16] This has the advantage that it is necessary, as well as sufficient, for (19) to imply global incentive compatibility, but the disadvantage that it depends on properties of the utility function rather than the allocation alone. Thus two subsequent corollaries give sufficient conditions based on ordinal properties of the allocation.

To simplify presentation, we adopt the convention for definite integrals that $\int_{\alpha'_t}^{\alpha''_t}\{\cdot\}\,d\alpha_t$ corresponds to $-\int_{\alpha''_t}^{\alpha'_t}\{\cdot\}\,d\alpha_t$ when $\alpha'_t > \alpha''_t$. We can then show the following:

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

$$\int_{\alpha'_t}^{\alpha''_t}\frac{1}{\alpha_t^2}\left\{\mathbb{E}_t\left[\sum_{s=t+1}^{\infty}\beta^{s-t}\left(1-D_{t,s}\left(\alpha^s\right)\right)\alpha_s\left[u\left(c_s\left(\alpha^{t-1},\alpha_t,\alpha_{t+1}^s\right)\right)-u\left(c_s\left(\alpha^{t-1},\alpha''_t,\alpha_{t+1}^s\right)\right)\right]\middle|\,\alpha_t\right]\right\}d\alpha_t\geq 0 \quad (20)$$

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.*

Non-increasingness here is a more restrictive requirement than condition (20) – hence the absence of an 'only if' statement in this Corollary. But the result is nonetheless useful because it provides a sufficient condition for the first-order approach to work that is independent of the utility function, relating only to ordinal properties of the consumption distribution. In this regard it is a vector generalisation of the usual monotonicity condition in bivariate screening problems.

Corollary 1 also provides simple overlap with the requirements for decentralisation. Notice that if the condition of the Corollary is satisfied then, by incentive compatibility, $c_t\left(\alpha^{t-1},\alpha_t\right)$ must be non-decreasing in α_t. Moreover, if the condition is satisfied and there are distinct α'_t and α''_t with $c_t\left(\alpha^{t-1},\alpha'_t\right)=c_t\left(\alpha^{t-1},\alpha''_t\right)$, then

[16] C.f. their Theorem 3.

$c_s\left(\alpha^{t-1},\alpha_t',\alpha_{t+1}^s\right)=c_s\left(\alpha^{t-1},\alpha_t'',\alpha_{t+1}^s\right)$ must be true along all subsequent paths – again by incentive compatibility. It follows that any allocation satisfying the condition in Corollary 1 will also satisfy the requirements for decentralisation in Proposition 1, and so an allocation of this kind has a simple consumption-savings decentralisation.

This link itself allows for an intuitive interpretation of the non-increasingness condition in Corollary 1. Viewed through the lens of the decentralisation, it is equivalent to requiring that consumption should behave as a *normal good* at all subsequent date-states, as savings in t are increased. That is, notwithstanding the changes to subsequent tax schedules that may be induced at the margin, higher savings in t imply uniformly higher subsequent consumption. Given this, we refer to an allocation that satisfies the requirements of Corollary 1 a **normal allocation**.

A second corollary relates to the iid case, where attention can be restricted to contemporaneous outcomes:

Corollary 2. *Let types by iid over time. Then a relaxed incentive-feasible allocation is incentive-feasible if and only if $c_t\left(\alpha^{t-1},\alpha_t\right)$ is non-decreasing in α_t for all t, α^{t-1}.*

Heuristically, this follows because $D_{t,s}\left(\alpha^s\right)\equiv 0$ with iid types, and the expectation term in (20) is independent of period-t type, so is equivalent to requiring that future value should be non-increasing in α_t. This can only be true if current consumption is non-decreasing in type. Full details are in the appendix.

6 Utility-based characterisation

This section provides a first characterisation of optimal allocations, based on perturbations to the profile of utility across agents. This takes as its main inputs arguments of the within-period utility function and the type distribution.

6.1 Characterisation result

By conventional techniques, we derive the following:

Proposition 3. *Suppose an interior allocation is optimal in the relaxed problem. For a cross-section of types in t with a common history α^{t-1}, the following must hold a.e.:*

$$\begin{aligned}&\mathbb{E}_{t-1}\left\{\alpha_t\left[1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t|\alpha_{t-1}\right)\right]-\frac{\eta_t}{u'\left(c_t\left(\alpha_t\right)\right)}\middle|\alpha_t>\alpha_t'\right\} \\ =&\frac{\pi\left(\alpha_t'|\alpha_{t-1}\right)}{\left(1-\Pi\left(\alpha_t'|\alpha_{t-1}\right)\right)}\cdot\left(\alpha_t'\right)^2\cdot\left\{\lambda_{t+1}^{\Delta}\left(\alpha_t'\right)-\rho\left(\alpha_t'|\alpha_{t-1}\right)\lambda_t^{\Delta}\right\}\end{aligned} \tag{21}$$

$$\mathbb{E}_{t-1}\left\{\alpha_t\left[1+\lambda_t+\lambda_t^{\Delta}\varepsilon^{\alpha}\left(\alpha_{t-1}\right)\right]-\frac{\eta_t}{u'\left(c_t\left(\alpha_t\right)\right)}\right\}=0 \tag{22}$$

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

$$\eta_t=(\beta R)^{-t}\frac{\mathbb{E}\left[\alpha_0\right]}{\mathbb{E}\left[\frac{1}{u'\left(c_0\left(\alpha_0\right)\right)}\right]} \tag{23}$$

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:

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

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 (21) and (22) are the main objects of interest here. (21) 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 – accunting 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 (19) 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.

λ_t and λ_t^Δ are multipliers that derive from prior incentive restrictions – capturing the shadow costs of changing ω_t and ω_t^Δ respectively. Consistent with the well-known work of Marcet and Marimon (2019), the Pareto weights that the policymaker attaches to different agents' utility in t are updated according to the shadow benefits of changes to incentives in $t-1$ and earlier.

6.2 The dynamics of consumption

A key element of the characterisation in Proposition 3 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.[17] When shocks are iid, the inverse marginal utility is well-known to follow a quasi-martingale process, with substantial implications for long-run inequality.[18] There has been significant recent debate about the sensitivity of this conclusion to type persistence.[19] In this subsection we divert briefly from the main argument to show that meaningful immiseration results still go through in our setting with type persistence.

With shock persistence the dynamics are a degree more complex than an iid case, as the following Proposition demonstrates.

Proposition 4. *For all t and s, $s \geq t$, and any history α^t, the period-t expected value of the period-s inverse marginal*

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

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

[19] See, for instance, Bloedel, Krishna and Strulovici (2020).

utility of consumption satisfies:

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

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 $\rho(\alpha_r|\alpha_{r-1}) \in [0,1)$, and so $\frac{\mathbb{E}_t\left[D_{t,s}(\alpha^s)\alpha_s\right]}{\mathbb{E}_t[\alpha_s]} \rightarrow 0$ as $s \rightarrow \infty$.[20] Since λ_{t+1} follows a martingale, the implication is that shocks to this martingale process control *long-run* consumption outcomes. Intuitively, given diminishing marginal utility it is cost-efficient to spread incentives over time. If there is justification for raising the utility of type α_t from $t+1$ on, in return for this type consuming relatively little in t, then this increase should be delivered evenly across all future periods.

These arguments are well understood from the existing social insurance literature. But when λ_{t+1}^{Δ} is non-zero there is an additional component to the short-run expected marginal cost, more in keeping with the dynamic contracting literature in quasilinear settings.[21] This comes from the desire to spread over time the cost of providing information rents at a given α_t. In particular, type persistence implies that periods after t matter for information rents in t. 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 do so 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 $\mathbb{E}_t[\alpha_s]$ 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:

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

Since $\eta_t > 0$, from (23), 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 that convergence in t to a constant, positive value for $(1+\lambda_{t+1})$ is only possible if the allocation converges to a first-best outcome, with $\mu_s(\alpha_s) = 0$ for all α_s, $s \geq t$. But this outcome straightforwardly 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, the long-run expectation of the marginal cost of utility provision converges almost surely to zero.

[20] Recall that $D_{t,s}(\alpha^s) := \prod_{r=t+1}^{s}\rho(\alpha_r|\alpha_{r-1})$.

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

7 Sufficient statistics: preliminaries

This section provides a number of crucial concepts and intermediate steps in order to map from the utility-based representation of optimal policy in Proposition 3 to a sufficient statistics representation.

7.1 Strict normality

For the optimal allocation to be characterised by sufficient statistics that relate to the consumption-savings decentralisation, this decentralisation must itself be valid. Proposition 1 established that this is true when current consumption is non-decreasing in current type, conditional on history. Similarly, Corollary 1 provided a sufficient condition for the validity of the first-order approach, requiring that future consumption should be non-increasing in current type. There is a clear link between the two: by incentive compatibility, non-increasingness of future consumption in current type must imply non-decreasingness of current consumption in current type.

Both of these conditions allowed for the possibility that a positive measure of types (with identical history) could bunch at a common consumption level. Bunching implies that non-decreasingness restrictions on consumption are either binding, or on the cusp of doing so. It is possible to handle these restrictions analytically, but the analysis is clearer without them. For this reason, we focus attention on 'strictly normal' allocations, as defined below.

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 $\left(\alpha_t', \alpha_t''\right) \in A^2$.

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. The link to normality, as defined above, can be seen from the the following:

Proposition 5. *If an allocation is strictly normal and relaxed incentive-compatible, then for all t, α^{t-1} and $\alpha_t'' > \alpha_t$, consumption from $t+1$ onwards satisfies the following inequality:*

$$\mathbb{E}_t\left[\sum_{s=t+1}^{\infty} \frac{\beta^{s-t}\alpha_s u'\left(c_s\left(\alpha^{t-1}, \alpha_t'', \alpha_{t+1}^s\right)\right)}{\alpha_t' u'\left(c_t\left(\alpha^{t-1}, \alpha_t''\right)\right)} \cdot \left(c_s\left(\alpha^{t-1}, \alpha_t', \alpha_{t+1}^s\right) - c_s\left(\alpha^{t-1}, \alpha_t'', \alpha_{t+1}^s\right)\right) \middle| \alpha_t'\right] \geq \delta_t\left(\alpha^{t-1}\right)\left(\alpha_t'' - \alpha_t\right)$$

where $\delta_t\left(\alpha^{t-1}\right)$ is the same constant used to define strict normality.

Thus a strictly normal allocation must see consumption increase strictly in savings, for some positive-measure subset of date-states.

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.[22] This rules out multiple types locating at a single point.

[22] Since 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 $c_t\left(\alpha^{t-1}, \cdot\right)$.

7.2 Towards sufficient statistics: intuition and integration

Equation (21) 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 dependent on a particular cardinalisation of the utility function. It would be preferable to characterise, as far as possible, by reference to measurable objects: behavioural elasticities and observable distributions.

We achieve this by use of a novel, intuitive analytical step, which links the utility-based mechanism design characterisation – set out in Section 6 – to the effects of changing the decentralised tax schedules outlined in Section 4. The logic is as follows. Sufficient statistics characterisations 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. 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.[23] 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 3 also describes the costs and benefits of a simple step change, but in the cross-sectional profile of *utilities* rather than incomes. Its key components are costs and benefits 'per unit change in utility'. 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. The main conceptual insight in this paper is that an understanding of the former can be achieved by starting from the latter.

More specifically, a unit 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. Now, suppose that we were to analyse the effects of a marginal increase in utility for all types below some α_t' by the amount $\alpha_t u'(c_t(\alpha_t))$. This is no longer a simple step change in utility, since $\alpha_t u'(c_t(\alpha_t))$ varies in α_t, but a statement of its effects can be constructed from the statements in Proposition 3, using elementary manipulations. It will implicitly describe the costs and benefits of cutting the marginal savings tax, at the savings level of type α_t'. Relating the component arguments to intuitive behavioural statistics will remain a challenge, but we will at least be starting from the right place.

Thus we obtain the following corollary to Proposition 3:

Corollary 3. *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:*

[23] The discussion in Piketty and Saez (2013b) provides detailed treatment. Heuristic analysis of optimal top tax rates usually considers a marginal tax cut on *all* incomes above a certain threshold, but this can be constructed from local cuts at all points in the upper range.

$$\int_{\underline{c}}^{\bar{c}}\left[1-\frac{\alpha_t(c)\,u'(c)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t(c)\,|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\pi^c\left(c|\alpha^{t-1}\right)dc \tag{24}$$

$$+\int_{\underline{c}}^{\bar{c}}\left(\alpha_t(c)\right)^2u''(c)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t(c)\right)}{\eta_t}-\rho\left(\alpha_t(c)\,|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_t}\right\}\left(\frac{d\alpha_t(c)}{dc}\right)^{-1}\pi^c\left(c|\alpha^{t-1}\right)dc$$

$$=0$$

$$-\int_{\underline{c}}^{c'}\left[1-\frac{\alpha_t(c)\,u'(c)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t(c)\,|\alpha_{t-1}\right)\right\}}{\eta_t}\right]\pi^c\left(c|\alpha^{t-1}\right)dc \tag{25}$$

$$-\int_{\underline{c}}^{c'}\left(\alpha_t(c)\right)^2u''(c)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t(c)\right)}{\eta_t}-\rho\left(\alpha_t(c)\,|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}\left(\frac{d\alpha_t(c)}{dc}\right)^{-1}\pi^c\left(c|\alpha^{t-1}\right)dc$$

$$+\left(\alpha_t(c')\right)^2u'(c')\left(\frac{d\alpha_t(c)}{dc}\right)^{-1}\pi^c\left(c'|\alpha^{t-1}\right)\left\{\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t(c')\right)}{\eta_t}-\rho\left(\alpha_t(c')\,|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}$$

$$=0$$

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

The restriction to continuous $c_t\left(\alpha_t\right)$ functions is made for convenience only: the appendix shows how to incorporate jumps in consumption as type increases. Note also that the realised consumption density will equal the realised savings density, 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 $s_t':=M_t\left(\alpha^{t-1}\right)-c_t'$, so $\frac{ds_t}{dc_t}=-1$.

7.3 Relevant behavioural statistics

By themselves, conditions (24) and (25) are not particularly intuitive statements. Their value becomes evident only when their components can be linked to behavioural statistics. The characterisation in Section 8 will do precisely this. It makes use of four distinct behavioural statistics, defined by reference to the decentralisation of Section 4. These are defined in turn here.

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

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

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, $T_t\left(s_t\right)$, remains unchanged to first order when the marginal tax rate, $T_t'\left(s_t\right)$, changes.

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

The compensated elasticity of lagged savings, with respect to contemporary returns: This is is a more unconventional object, capturing the complementarities that may exist between insurance and prior saving. Denoted $\epsilon_{t-1,t}^s(s_t)$, it is the response of savings in $t-1$ to the change in the profile of insurance at t that is generated by a tax cut at s_t. It is calculated assuming compensation that leaves ω_t constant (in a way that does not affect insurance), so that the relevant behavioural change in $t-1$ is purely due to the re-profiling of state-by-state utility outcomes in t.

The formal definition of $\epsilon_{t-1,t}^s(s_t)$ is technically involved, because of a 'zero measure' issue. Heuristically, but with abuse, it satisfies:

$$\epsilon_{t-1,t}^s\left(s_t\right)\pi^s\left(s_t|\alpha^{t-1}\right)ds_t := \frac{\left(1-T_t'\left(s_t\right)\right)}{s_{t-1}}\left.\frac{ds_{t-1}}{d\left(1-T_t'\left(s_t\right)\right)}\right|_{\text{comp}} \tag{26}$$

More formally, $\epsilon_{t-1,t}^s\left(s_t\right)$ is defined implicitly from the Fréchet derivative of s_{t-1} with respect to an arbitrary profile of changes to the tax schedule in t, given compensation. Let $T_t\left(s\left(M_{t+1}\right)\right)$ denote the tax paid in t when the consumer carries M_{t+1} units of wealth into period $t+1$, with $s\left(M_{t+1}\right)$ the corresponding period-t savings, and suppose that for each M_{t+1} the tax schedule is perturbed to $T_t\left(s\left(M_{t+1}\right),\Gamma\right)$:[24]

$$T_t\left(s\left(M_{t+1}\right),\Gamma\right) := \left[T_t\left(s\left(M_{t+1}\right)\right)-\Gamma f\left(s\left(M_{t+1}\right)\right)\right] \tag{27}$$

where $\Gamma\in\mathbb{R}$ and $f\left(\cdot\right)$ is an arbitrary bounded differentiable function on the interval of realised savings, with $f\left(s\left(\underline{M}\right)\right)=0$ for the lowest realised value of M_{t+1}. The compensated derivative of s_{t-1} with respect to Γ, evaluated at $\Gamma=0$, will be linear in the derivative of $f\left(\cdot\right)$, $f'\left(s_t\right)$.[25] That is:

$$\left.\frac{ds_{t-1}}{d\Gamma}\right|_{\Gamma=0,\text{ comp}} = \int_{s_t} f'\left(s_t\right)g\left(s_t\right)ds_t \tag{28}$$

[24] Formally defining the perturbations as functions of M_{t+1} rather than s_t ensures that $f\left(\cdot\right)$ defines the 'rightward' shift in the budget constraint linking c_t to M_{t+1}, and so $f\left(\cdot\right)$ corresponds to the magnitude of the income effect, in units of period-t income. This makes the characterisation of $\epsilon_{t-1,t}^s\left(s_t\right)$ possible by reference to the different income and substitution effects at t, and the impact these have on information rents.

[25] This is established formally in the proof of Lemma 2.

with the function $g(s_t)$ independent of the choice of $f(\cdot)$. Since $f'(s_t)$ is precisely the proportional increase in the rate of return on s_t per unit change in Γ, we define:

$$g(s_t) := s_{t-1}\epsilon^s_{t-1,t}(s_t)\,\pi^s\left(s_t|\alpha^{t-1}\right) \tag{29}$$

thereby implicitly defining $\epsilon^s_{t-1,t}(s_t)$. Thus, the aggregate change in s_{t-1} induced by an arbitrary change to marginal tax rates at t is the integral of the changes induced piecewise by tax cuts at each point, and $\epsilon^s_{t-1,t}(s_t)$ is constructed to capture the response at each point.

How and why savings at $t-1$ should respond to a re-profiling of insurance in t is an issue to which we return later. Ultimately the sign and magnitude of $\epsilon^s_{t-1,t}(s_t)$ will capture important links between tax cuts, insurance and savings. Through arguments that connect closely to the well-known Atkinson-Stiglitz theorem, 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 $T_t(\underline{s})$, holding constant the profile of marginal rates at higher savings. This tax cut will shift consumption possibilities in t, by an amount that is uniform across all savings levels. The marginal effect on utility will equal $\alpha_t u'(c_t(\alpha_t))$ for all agents. In general this will vary in α_t, implying a change in relative utilities across types. Even if compensation is applied so that ω_t is held constant, the changed insurance profile may affect local incentives to save at $t-1$.

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

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

7.4 Equivalence results

Making use of these definitions, the following Lemma provides the ingredients to link from expressions (24) and (25) to a sufficient statistics representation:

Lemma 2. *The following relationships hold:*

1. *Revenue raised from within-period substitution effects:*

$$T'_t(s_t)\,s_t\varepsilon^s_t = \frac{\lambda^\Delta_{t+1}(\alpha_t(c_t))}{\eta_t}(\alpha_t(c_t))^2 u'(c_t)\left(\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1} \tag{30}$$

2. *Revenue raised from within-period income effects:*

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

3. *Revenue raised from cross-period effects of tax cuts:*

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

4. *Revenue raised from cross-period effects of income transfers:*

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

The proof of these relationships, in the Appendix, is based on elementary manipulations. It exploits two important features of the problem. The first is the duality between welfare maximisation and cost minimisation when designing policy. This enables the multipliers 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 we exploit is the separability of consumption utility over time. This limits the dependence of contemporaneous choice on decisions in other periods, and guarantees the existence of relatively simple cross-relationships between different behavioural statistics.

7.5 Welfare weights

In keeping with the static literature, we make use of 'social welfare weights' to capture the marginal value to the policymaker of providing an extra unit of income to each type, expressed in units of current resources. In our setting these weights are defined for history α^t and current consumption level c_t by:[26]

$$g_t\left(\alpha^t\right) := \alpha_t u'\left(c_t\left(\alpha^t\right)\right)\frac{1+\lambda_t\left(\alpha^{t-1}\right)}{\eta_t} \tag{33}$$

[26] If the mapping between α_t and s_t is bijective, we may sometimes write $g_t(s_t)$ in place of $g_t\left(\alpha^t\right)$, leaving history implicit.

That is, the subjective marginal utility of consumption, $\alpha_t u'(c_t)$, 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. Heuristically, the role of $\eta_t$is to convert from marginal utility units into resource units.

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: $\mathbb{E}\left[\lambda_{t+1}|\alpha^{t-1}\right] = \lambda_t$. In the decentralised allocation, the updating process will capture the wealth that 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.[27]

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 non-stationary manner in which the weights evolve – in parallel with the evolution of the wealth distribution. 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 implicitly incorporated into the societal objective. An optimal plan remains cross-sectionally utilitarian, for any subset of individuals who share a common history. Across subgroups, however, substantial differentiation in treatment is likely to emerge. The time inconsistency here is evident, and provides a challenge to the plausibility of the commitment assumption.[28]

Proposition 4 provides an alternative way to interpret the weight:

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

This is the ratio of the expected marginal cost of providing lifetime utility to type α_t, viewed from $t-1$, relative to the realised cost.

8 Sufficient statistics characterisation

8.1 Characterisation

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

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

$$\mathbb{E}\left[1 - T_0'\left(s_0\right)\frac{ds_0}{dM_0} - g_0\left(s_0\right)\middle| s_0 \geq s_0'\right] = T_0'\left(s_0'\right)\varepsilon_0^s\frac{s_0'\pi^s\left(s_0'\right)}{1-\Pi^s\left(s_0'\right)} \tag{34}$$

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

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

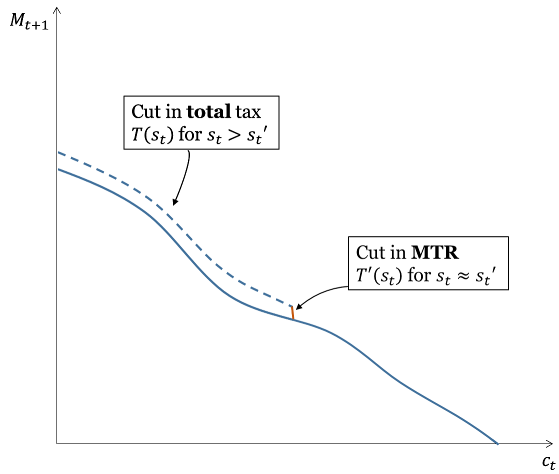

Figure 1: Effects on the budget constraint of a marginal savings tax cut at s_t'

and:

$$\mathbb{E}\left[g_0\left(s_0\right)\right] = \mathbb{E}\left[1 - T_0'\left(s_0\right)\frac{ds_0}{dM_0}\right] \tag{35}$$

Similarly, for $t > 0$, any given α^{t-1}, and for all $s_t' \in \left(\underline{s_t}, \bar{s_t}\right)$:

$$\mathbb{E}_{t-1}\left[1 - T_t'\left(s_t\right)\frac{ds_t}{dM_t} - g_t\left(s_t\right)\middle| s_t \geq s_t'\right] = \left[T_t'\left(s_t'\right)\varepsilon_t^s + RT_{t-1}'\left(s_{t-1}\right)\frac{s_{t-1}}{s_t'}\epsilon_{t-1,t}^s\left(s_t'\right)\right]\frac{s_t'\pi^s\left(s_t'|\alpha^{t-1}\right)}{1 - \Pi^s\left(s_t'|\alpha^{t-1}\right)} \tag{36}$$

and:

$$\mathbb{E}_{t-1}\left[g_t\left(s_t\right)\right] = \mathbb{E}_{t-1}\left[1 - T_t'\left(s_t\right)\frac{ds_t}{dM_t}\right] - RT_{t-1}'\left(s_{t-1}\right)\frac{ds_{t-1}}{dM_t}\bigg|_{comp} \tag{37}$$

.

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

8.2 Intuition

As previewed, equations (34) to (37) 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 (34) and (36), 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 (35) and (37), 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 (34) 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 (34) accounts for this from the policymaker's perspective. There are three components: (1) the direct cost of the transfer, normalised to 1 per agent by construction, minus (2) the additional tax revenue that is received on whatever fraction of the additional income is saved, $T_0'(s_0)\frac{ds_0}{dM_0}$, minus (3) the social welfare value of providing an additional consumption unit to the agent in question, $g_0(s_0)$. Taken together, these objects give the net fiscal cost of the transfer that high-saving agents receive.

The right-hand side of (34) 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 (36) is the equivalent to (34) for $t > 0$. Relative to the period-0 version, it has an extra term that allows for the impact that changes to tax schedules in t have on savings in $t-1$. This is the object $RT_{t-1}'(s_{t-1})s_{t-1}\epsilon_{t-1,t}\left(s_t'\right)$, 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 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 9.3, and we defer further comment for now.

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

Conditions (35) and (37) describe the consequences of shifting the entire budget constraint, rather than just an upper segment. They are essentially variants of (34) and (36) respectively, when the substitution effects due to marginal tax cuts are dropped – 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 average value of the welfare weight in each period, 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 4 that the average value of $T_t(s_t)$ is constructed to be zero, given the history s^{t-1}. Since the marginal rate is positive, an optimal social insurance scheme must therefore provide a positive transfer (negative T) at the lowest savings level, which is then taxed away as savings increase.

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

Returning briefly to the utility-based characterisation, the fact that marginal tax rates are positive at interior points also provides some insight into the expected evolution of the inverse marginal utility of consumption, described in Section 6.2. Recall from section 6.2 that we have:

$$\frac{1}{\mathbb{E}_t\left[\alpha_s\right]}\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}\,u'\left(c_s\right)}\right] = \frac{1+\lambda_{t+1}}{\eta_t} + \frac{\mathbb{E}_t\left[D_{t,s}\left(\alpha^s\right)\alpha_s\right]}{\mathbb{E}_t\left[\alpha_s\right]}\frac{\lambda^{\Delta}_{t+1}}{\eta_t} \tag{38}$$

where $\frac{\mathbb{E}_t\left[D_{t,s}(\alpha^s)\alpha_s\right]}{\mathbb{E}_t[\alpha_s]}$ is the elasticity of $\mathbb{E}_t\left[\alpha_s\right]$ with respect to α_t, which converges to zero as s becomes large. The object on the left-hand side measures the period-t cost of providing a unit of lifetime utility in t, by raising consumption in period $s > t$. Its value, on the right-hand side, combines a 'permanent' component that is constant in s, $\frac{1+\lambda_{t+1}}{\eta_t}$, and a 'transitory' component $\frac{\mathbb{E}_t\left[D_{t,s}(\alpha^s)\alpha_s\right]}{\mathbb{E}_t[\alpha_s]}\frac{\lambda^{\Delta}_{t+1}}{\eta_t}$.

Marginal savings taxes are positive if and only if λ^{Δ}_{t+1} takes a positive value.[29] Thus a corollary to Theorem 2 is that the marginal cost of providing lifetime utility is greater in the short run than the long run, at all interior α_t. Marginal cost is increasing in consumption, so this is equivalent to consumption falling over time, relative to the dynamics at a first best – where marginal cost is constant. Downward drift in consumption comes about precisely because the tax system is reducing the consumer's return on savings below the technological rate R.

9.2 Limiting outcomes

The general finding of strictly positive marginal savings tax rates need not extend to endpoints of the type distribution, where limiting rates may instead reach zero. A critical role is played by the limiting properties of the distribution and density of savings, and it is possible for optimal marginal rates to reach zero at the upper and lower limits. Zero tax results generally follow if the density remains positive at endpoints, or converges to zero relatively slowly, since this implies that the local efficiency costs of taxation become large relative to any redistributive benefit. For optimal top tax rates, the key ratio is well summarised by the Pareto statistic, which we denote $a_t\left(s_t|\alpha^{t-1}\right)$:

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

[29] See the proof of Theorem 2.

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

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

Corollary 4. *Given α^{t-1}, the optimal marginal tax rate at the upper limit of the savings distribution, $\bar{s}$, satisfies:*

$$T_t'\left(\bar{s}\right) = \frac{1 - g_t\left(\bar{s}\right) - RT_{t-1}'\left(s_{t-1}\right)\frac{s_{t-1}}{\bar{s}}\epsilon_{t-1,t}^s\left(\bar{s}\right)a_t\left(\bar{s}|\alpha^{t-1}\right)}{\left.\frac{ds_t}{dM_t}\right|_{\bar{s}} + \varepsilon_t^s a_t\left(\bar{s}|\alpha^{t-1}\right)} \tag{39}$$

In keeping with the static literature on optimal tax design, it is possible to use this equation to obtain approximate estimates for upper marginal savings tax rates, given a shock history. The two problematic objects to

9.3 Intertemporal elasticities: Atkinson-Stiglitz revisited

The most significant theoretical insight from Theorem 1 is the limited extent to which the conventional Saez (2001) condition needs to change when moving from a simple two-good screening problem to an infinite-horizon, persistent-type setting. Given an updated set of set of social welfare weights, optimal taxation in period t can be characterised by reference only to a contemporaneous elasticity of savings, ε_t^s, and the elasticity of lagged savings, $\epsilon_{t-1,t}^s\left(s_t'\right)$. The first of these is a conventional behavioural elasticity – the response of saving to a change in the rate of return. The second is a more unusual object, and the purpose of this subsection is to discuss its role, and to 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, and the associated monotone likelihood property, means that those with higher α_{t-1} place relatively more weight in t on high values for α_t. Marginal utility in t is increasing in α_t, and so this in turn implies that those with lower s_{t-1} place greater 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 insurance at t. A policy that marginally improves the profile of insurance in t that savings provide, holding constant expected utility, will raise the marginal attractiveness of savings in $t-1$.

Thus the presence of $\epsilon_{t-1,t}^s\left(s_t'\right)$ in equation (36) is precisely to capture the effect of the posited 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 couterpart 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: $\epsilon^s_{t-1,t}(s_t) \equiv 0$, and only contemporaneous elasticities matter.

This line of reasoning also indicates that the Markov property of shocks is crucial to ensuring that just two elasticities feature in (36). 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. Thus there are no gains to distorting $t+1$ allocations in order to improve screening in $t-1$, beyond the distortions already implied by different distributions across α_t.

9.3.1 Cross-sectional variation in $\epsilon^s_{t-1,t}\left(s'_t\right)$

Consistent with this discussion, we show that $\epsilon^s_{t-1,t}\left(s'_t\right)$ will exhibit systematic cross-sectional variation, in a manner that contributes to greater progressivity in the marginal tax rate in t, and thus a greater degree of insurance *ex-post*. Formally, we have the following result:

Proposition 6. *The statistic* $\frac{s_{t-1}}{s'_t}\epsilon^s_{t-1,t}\left(s'_t\right)\frac{s'_t\pi^s\left(s'_t|\alpha^{t-1}\right)}{1-\Pi^s\left(s'_t|\alpha^{t-1}\right)}$ *is monotonically decreasing in* $s'_t \in \left(\underline{s_t}, \bar{s_t}\right)$. *It is positive for sufficiently low* s'_t, *and negative for sufficiently high* s'_t.

The implications of this result can be seen by comparing policies that satisfy condition (36) with those that neglect intertemporal cross-elasticities – as would be optimal for a policymaker re-optimising in t. At our optimum, we have:

$$\mathbb{E}_{t-1}\left[1 - T'_t\left(s_t\right)\frac{ds_t}{dM_t} - g_t\left(s_t\right)\middle| s_t \geq s'_t\right] > T'_t\left(s'_t\right)s'_t\varepsilon^s_t\frac{\pi^s\left(s'_t|\alpha^{t-1}\right)}{1-\Pi^s\left(s'_t|\alpha^{t-1}\right)} \tag{40}$$

for all s'_t below a threshold and

$$\mathbb{E}_{t-1}\left[1 - T'_t\left(s_t\right)\frac{ds_t}{dM_t} - g_t\left(s_t\right)\middle| s_t \geq s'_t\right] < T'_t\left(s'_t\right)s'_t\varepsilon^s_t\frac{\pi^s\left(s'_t|\alpha^{t-1}\right)}{1-\Pi^s\left(s'_t|\alpha^{t-1}\right)} \tag{41}$$

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 reason this occurs relates to the ex-ante insurance properties of a tax cut, and how these change with variation in the critical value s_t' where the cut takes place. To see the logic intuitively, recall first that allocatons in t are being perturbed for a particular history path, with given prior savings level s_{t-1}'. Now suppose that a tax cut in t were applied at a very low value for s_t', so that all, or almost all, types are able to benefit from a uniform increase in period-t income. Since marginal utility is highest for those saving the least, the (compensated) effect of this is to redistribute utility towards low savers in period t.

Because of type persistence, this redistribution is relatively appealing *ex-ante* to those whose savings start out marginally lower than s_{t-1}', compared to those at s_{t-1}. Te perturbation gives these agents an incentive to raise their savings up to s_{t-1}', to benefit from the more appealing profile of returns. Higher savers in $t-1$, by contrast, are not attracted by the change: they place greater weight on states where future savings are high, which have become less profitable. This means that the intertemporal behavioural response to a tax cut on low s_t' is positive. This raises the benefits of lower taxes at s_t', reducing the optimal $T_t'\left(s_t'\right)$.

By contrast, suppose that taxes are only cut at a value s_t' close to the upper threshold $\bar{s}_t$. With compensation, the net effect of this change is to redistribute utility towards relatively high savers. Given persistence, this is relatively appealing in $t-1$ to those with savings marginally higher than s_{t-1}', who are incentivised to reduce their savings. This makes the tax cut relatively costly, since the deterred savings reduce fiscal revenue in $t-1$ – and so the optimal $T_t'\left(s_t'\right)$ tends to be higher.

9.4 Optimal transfers

Conditions (35) and (37) 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 the reasons just explained. 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. That is, at the optimum:

$$\mathbb{E}_{t-1}\left[g_t\left(s_t\right)\right] < \mathbb{E}_{t-1}\left[1-T'\left(s_t\right)\frac{ds_t}{dM_t}\right] \tag{42}$$

Once more, the general message is that type persistence motivates a more generous insurance scheme, because insurance acts as a complement to past savings – and past savings are inefficiently low.

10 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. Our paper suggests that mechanism design offers a theory-guided alternative route.

References

[1] Aiyagari, S.R. (1995), 'Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting', *Journal of Political Economy*, 103(6): 1158–75.

[2] Albanesi, S., and C. Sleet (2006), 'Dynamic Optimal Taxation with Private Information', *Review of Economic Studies*, 73(1): 1–30.

[3] Atkeson, A., and R.E. Lucas (1992), 'On Efficient Distribution With Private Information', *Review of Economic Studies*, 59(3): 427–453.

[4] Atkinson, A.B., and J.E. Stiglitz (1976) , 'The Design of Tax Structure: Direct versus Indirect Taxation', *Journal of Public Economics*, 6: 55–75.

[5] Bloedel, A.W., R.V. Krishna and O. Leukhina (2020), 'Insurance and Inequality with Persistent Private Information', Working Paper, Stanford University and Federal Reserve Bank of St. Louis.

[6] Bloedel. A.W., R.V. Krishna and B. Strulovici (2020), 'Persistent Private Information Revisited', Working Paper.

[7] Brendon, C.F., and M. Ellison (2018), 'Time-Consistently Undominated Policies', Working Paper.

[8] Chamley, C. (1986), 'Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives', *Econometrica*, 54(3): 607–22.

[9] Chari, V.V., J.P. Nicolini and P. Teles (2020), 'Optimal Capital Taxation Revisited', *Journal of Monetary Economics*, (116): 147–165.

[10] Conesa, J.C., S. Kitao and D. Krueger (2009), 'Taxing Capital? Not a Bad Idea After All!', *American Economic Review*, 99(1): 25– 48.

[11] Diamond, P. (1998), 'Optimal Taxation: An Example with a U-shaped Pattern of Optimal Marginal Tax Rates', *American Economic Review*, 88: 83–95.

[12] Diamond, P., and J.A. Mirrlees (1978), 'A Model of Social Insurance with Variable Retirement', *Journal of Public Economics*, 10: 295-336.

[13] 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.

[14] Erosa, A., and M. Gervais (2002), 'Optimal Taxation in Life-Cycle Economies', *Journal of Economic Theory*, 105(2): 338–69.

[15] Farhi, E., and I. Werning (2007), 'Inequality and Social Discounting', *Journal of Political Economy*, 115(3): 365–402.

[16] Farhi, E., and I. Werning (2013), 'Insurance and Taxation over the Life Cycle', *Review of Economic Studies*, 810(2): 596–635.

[17] Golosov, M., N.R. Kocherlakota and A. Tsyvinski (2003), 'Optimal Indirect and Capital Taxation', *Review of Economic Studies*, 70(3): 569–87.

[18] Golosov, M., M. Troshkin and A. Tsyvinski (2016), 'Redistribution and Social Insurance', *American Economic Review*, 106(2): 359–386.

[19] 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.

[20] Golosov, M., A. Tsyvinski and N. Werquin (2014), 'A Variational Approach to the Analysis of Tax Systems', NBER Working Paper 20780.

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

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

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

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

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

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

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

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

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

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

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

[32] 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.*

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

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

[35] 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.

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

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

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

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

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

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

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

[43] 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.

A Appendix

A.1 Proof of Proposition 1

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:

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

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

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

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

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

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

For all $\tilde{s}_0 \in S_0^c$, let $T_0\left(\tilde{s}_0\right) = \tilde{s}_0$, so that $M_1\left(\tilde{s}_0\right) = 0$, and for all $t > 0$ and subsequent savings choices $\{s_1, \ldots, s_t\}$, let $T_t\left(\tilde{s}_0, s_1, \ldots s_t\right) > \varepsilon$ for some $\varepsilon > 0$. Combining the budget constraints (13) and (14), we have, along all future consumption paths:

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

and so:

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

for all $T \geq 0$. By the 'no Ponzi', the final term on the right-hand side satisfies $\lim_{T\to\infty} R^{-T-2}M_{T+2} \geq 0$, 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 all $s_0 \in S_0$, let $M_1(s_0)$ be given by:

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

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, and expectations are with respect to the evolution of types subsequent to period 0 given this α_0. More generally $\alpha_0(s_0)$ may take values from a convex interval in A, and in this case expectations satisfy Bayes's rule in the obvious way. Condition 2 in the Proposition guarantees that the continuation allocation is identical across types within any such set, and so they do not need to be separated in their market treatment.

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

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

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

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

and that this mapping has an inverse correspondence$\alpha^{t-1}\left(s^{t-1}\right)$, 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 $M_t\left(s^{t-1}\left(\alpha^{t-1}\right)\right)$ corresponding to each $\alpha^{t-1} \in A^t$. For all $\alpha_t \in A$, let $s_t\left(\alpha^{t-1}, \alpha_t\right)$ be given by:

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

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

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

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

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

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

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

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

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

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

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

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.

A.2 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:

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

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:

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

Constraint (10) implies that $u\left(c_t\left(\alpha^{t-1}, \alpha_t''\right)\right)$ and $V\left(\alpha^{t-1}, \alpha_t'', \alpha_{t+1}\right)$ are bounded for all 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 (44) 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:

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

To obtain the representation in the main text, we then make use of the following definition and subsequent Lemma:

Definition. For all α^s, $s>t$:

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

and:

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

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

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

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. □

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

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

Using this and the definition of $\omega_{t+1}\left(\alpha^{t-1},\alpha_t\right)$, (45) collapses to (19).

A.3 Proof of Proposition 2

A.3.1 '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:

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

where $u_t\left(\alpha_t\right)$ 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:

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

where the penultimate line has made use of the relaxed incentive constraint (19).

Applying this result in (46) yields:

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

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

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

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

A.3.2 'Only if'

Suppose integral monotonicity fails for some $\left(\alpha_t',\alpha_t''\right)$, 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.3.3 Corollary 2

This corollary can be established directly by a simplified version of the proof of the main proposition when types are iid. Suppose that global incentive compatibility fails. Then for some t and history α^{t-1} there must exist α_t', α_t'' such that:

$$\alpha_t'u_t\left(\alpha_t''\right)+\beta\omega_{t+1}\left(\alpha_t''\right)>\alpha_t'u_t\left(\alpha_t'\right)+\beta\omega_{t+1}\left(\alpha_t'\right)$$

where future values are now independent of period-t type. The relaxed incentive constraint (19) in the iid case gives:

$$\alpha_t''u_t\left(\alpha_t''\right)+\beta\omega_{t+1}\left(\alpha_t''\right)=\alpha_t'u_t\left(\alpha_t'\right)+\beta\omega_{t+1}\left(\alpha_t'\right)+\int_{\alpha_t'}^{\alpha_t''}u_t\left(\alpha_t\right)d\alpha_t$$

Using this in the prior inequality gives:

$$\int_{\alpha_t'}^{\alpha_t''} \left[u\left(\alpha_t''\right) - u_t\left(\alpha_t\right)\right] d\alpha_t\left(\alpha_t'\right) < 0$$

This contradicts non-decreasingness. Likewise, working backwards, a failure of non-decreasingness implies global incentive compatibility will be violated for some pair of types.

A.4 Proof of Proposition 3

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:

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

and the relaxed incentive constraint:

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

with, for all $t \geq 0$:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Condition (51) can be rearranged to:

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

This can be integrated across all α_t:

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

Integrating by parts, making use of (54):

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

and so:

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

or, using (5):

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

Rearranging (59) for period 0 gives an expression for η:

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

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

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

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

Integrating (56) above any given α_t' gives:

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

and integrating by parts, we have:

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

so:

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

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

A.5 Proof of Proposition 4

The proof of Proposition 3 has already established the result for $s = t$ and $s = t + 1$. From (61) and (62) we immediately have:

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

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

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

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

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

Then:

$$\mathbb{E}_{r-1}\left[\frac{\eta_{r-1}}{(\beta R)^{s-r+1} u'(c_s)}\right] \tag{69}$$

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

By an identical argument to Lemma 3, 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:

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

where we have used condition (54). Using this in the preceeding expression, the terms in μ_r cancel:

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

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

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

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

A.6 Proof of Proposition 5

Relaxed incentive compatibility and normality imply full incentive compatibility, by Corollary 1. Incentive compatibility for type α_t'' implies:

$$\alpha_t'\left[u_t(\alpha_t'') - u(\alpha_t')\right] + \mathbb{E}_t\left\{\sum_{s=t+1}^{\infty}\beta^{s-t}\alpha_s\left[u_s(\alpha_t'',\alpha_{t+1}^s) - u_s(\alpha_t';\alpha_{t+1}^s)\right]\middle|\alpha_t'\right\} \leq 0$$

where $u_t(\alpha_t)$ is shorthand for $u\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right)$. Since $c_t(\alpha_t'') > c_t(\alpha_t')$, by the concavity of the within-period utility function:

$$u_t(\alpha_t'') - u(\alpha_t') \geq u'\left(c_t(\alpha_t'')\right)\left(c_t(\alpha_t'') - c_t(\alpha_t')\right) > 0$$

and since $c_s\left(\alpha_t'',\alpha_{t+1}^s\right) \leq c_s\left(\alpha_t',\alpha_{t+1}^s\right)$ for $s > t$:

$$0 \geq u_s\left(\alpha_t'',\alpha_{t+1}^s\right) - u_s\left(\alpha_t';\alpha_{t+1}^s\right) \geq u'\left(c_s\left(\alpha_t'';\alpha_{t+1}^s\right)\right)\left(c_s\left(\alpha_t'';\alpha_{t+1}^s\right) - c_s\left(\alpha_t';\alpha_{t+1}^s\right)\right)$$

Using these:

$$\begin{aligned}
&\alpha_t' u'\left(c_t(\alpha_t'')\right)\left(c_t(\alpha_t'') - c_t(\alpha_t')\right) + \mathbb{E}_t\left\{\sum_{s=t+1}^{\infty}\beta^{s-t}\alpha_s u'\left(c_s\left(\alpha_t'';\alpha_{t+1}^s\right)\right)\left(c_s\left(\alpha_t'';\alpha_{t+1}^s\right) - c_s\left(\alpha_t';\alpha_{t+1}^s\right)\right)\middle|\alpha_t''\right\}\\
\leq&\alpha_t'\left[u_t(\alpha_t'') - u(\alpha_t')\right] + \mathbb{E}_t\left\{\sum_{s=t+1}^{\infty}\beta^{s-t}\alpha_s\left[u_s\left(\alpha_t'',\alpha_{t+1}^s\right) - u_s\left(\alpha_t';\alpha_{t+1}^s\right)\right]\middle|\alpha_t''\right\}\\
\leq&0
\end{aligned}$$

And so:

$$\begin{aligned}
&\alpha_t' u'\left(c_t(\alpha_t')\right)\left(c_t(\alpha_t'') - c_t(\alpha_t')\right)\\
\leq&\mathbb{E}_t\left\{\sum_{s=t+1}^{\infty}\beta^{s-t}\alpha_s u'\left(c_s\left(\alpha_t';\alpha_{t+1}^s\right)\right)\left(c_s\left(\alpha_t';\alpha_{t+1}^s\right) - c_s\left(\alpha_t'';\alpha_{t+1}^s\right)\right)\middle|\alpha_t''\right\}
\end{aligned}$$

By strict normality:

$$\left(c_t(\alpha_t'') - c_t(\alpha_t')\right) \geq \delta_t\left(\alpha^{t-1}\right)\left(\alpha_t'' - \alpha_t'\right)$$

So:

$$\mathbb{E}_t\left\{\sum_{s=t+1}^{\infty}\frac{\beta^{s-t}\alpha_s u'\left(c_s\left(\alpha_t';\alpha_{t+1}^s\right)\right)}{\alpha_t' u'\left(c_t\left(\alpha_t'\right)\right)}\left(c_s\left(\alpha_t';\alpha_{t+1}^s\right)-c_s\left(\alpha_t'';\alpha_{t+1}^s\right)\right)\middle|\alpha_t''\right\}\geq\delta_t\left(\alpha^{t-1}\right)\left(\alpha_t''-\alpha_t'\right)$$

A.7 Corollary 3

Monotonicity of $c_t(\alpha_t)$ implies that $c_t(\alpha_t)$ is continuous a.e.. Suppose it is continuous on some open interval $(\alpha',\alpha'')\subset\left[\underline{\alpha},\bar{\alpha}\right]$, and write $c'=\lim_{\alpha\searrow\alpha'}(c_t(\alpha))$ and $c''=\lim_{\alpha\nearrow\alpha''}(c_t(\alpha))$ (i.e. limits as α approaches from above and below respectively). Integrating (21) across this range gives:

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

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

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

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

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

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:

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

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

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

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

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

as given in the text.

A.7.1 Allowing discontinuities

Discontinuities in $c_t(\alpha_t)$ can be included with minimal additional manipulation. Since $c_t(\alpha_t)$ is monotone, the set of 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 $c^u(\alpha_t) = \lim_{\alpha \searrow \alpha_t}(c_t(\alpha))$ and $c^l(\alpha_t) = \lim_{\alpha \nearrow \alpha_t}(c_t(\alpha))$ denote the upper and lower limits for consumption respectively. Summing (75) across intervals, over the entire range we have:

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

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

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

A.8 Proof of Lemma 2

Conditions 1 and 2

We start with two Lemmata:

Lemma 4. *The marginal tax rate satisfies:*

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

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

Proof. Recall that the decentralisation in Proposition 1 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:

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

Or:

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

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

$$
\frac{1+\lambda_{t+1}\left(\alpha^t\right)}{\eta_t} = \frac{1}{1-\varepsilon^{\alpha}(\alpha_t)}\left\{\frac{1}{\mathbb{E}\left[\alpha_{t+1}|\alpha_t\right]}\mathbb{E}\left[\frac{1}{\beta R u'\left(c_{t+1}\left(\alpha^{t+1}\right)\right)}\middle|\alpha^t\right] - \frac{\varepsilon^{\alpha}(\alpha_t)}{\alpha_t u'\left(c_t\left(\alpha^t\right)\right)}\right\} \tag{80}
$$

$$
\frac{\lambda_{t+1}^{\Delta}\left(\alpha^t\right)}{\eta_t} = \frac{1}{1-\varepsilon^{\alpha}(\alpha_t)}\left\{\frac{1}{\alpha_t u'\left(c_t\left(\alpha^t\right)\right)} - \frac{1}{\mathbb{E}\left[\alpha_{t+1}|\alpha_t\right]}\mathbb{E}\left[\frac{1}{\beta R u'\left(c_{t+1}\left(\alpha^{t+1}\right)\right)}\middle|\alpha^t\right]\right\} \tag{81}
$$

Substituting in (80) gives:

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

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:

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

So:

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

as stated. □

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

$$\left\{\alpha_t u'(c_t) - \beta R\left(1 - T_t'(s_t)\right) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1};\alpha_{t+1})\,\alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1}\right\}^{-1} = -\frac{\frac{ds_t}{dM_t}}{\alpha_t^2 u''(c_t)\frac{dc_t}{d\alpha_t}} \tag{82}$$

$$\left\{\alpha_t u'(c_t) - \beta R\left(1 - T_t'(s_t)\right) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1};\alpha_{t+1})\,\alpha_t \frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t} d\alpha_{t+1}\right\}^{-1} = \frac{s_t \varepsilon_t^s}{\alpha_t^2 u'(c_t)\frac{dc_t}{d\alpha_t}} \tag{83}$$

Proof. Start with (82). 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:

$$\frac{dc_t}{d\alpha_t} + \frac{dc_t}{dM_t}\frac{dM_t}{d\alpha_t} = \frac{dM_t}{d\alpha_t} \tag{84}$$

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:

$$\alpha_t u'(c_t) = \beta R\left(1 - T'(s_t)\right) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1};\alpha_{t+1})\,\pi(\alpha_{t+1}|\alpha_t)\, d\alpha_{t+1} \tag{85}$$

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

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

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

Rearranging (84):

$$\frac{dM_t}{d\alpha_t} = \frac{\frac{dc_t}{d\alpha_t}}{1 - \frac{dc_t}{dM_t}} = \frac{\frac{dc_t}{d\alpha_t}}{\frac{ds}{dM_t}}$$

Plugging this into the previous expression, trivial manipulations give (82).

Reasoning in a similar way for (83), consider the effect of a change to $\left(1 - T_t'(s_t)\right)$ 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:

$$\frac{dc_t}{d\left(1 - T'(s_t)\right)} + \frac{dc_t}{dM_t}\frac{dM_t}{d\left(1 - T'(s_t)\right)} = \frac{dM_t}{d\left(1 - T'(s_t)\right)} \tag{86}$$

Differentiating (85) with respect to $\left(1-T_t'(s_t)\right)$ under this joint change gives:

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

$$-\alpha_t u'\left(c_t\right)\frac{1}{1-T_t'\left(s_t\right)}+\alpha_t u''\left(c_t\right)\frac{\frac{dc_t}{d\left(1-T_t'(s_t)\right)}}{\frac{ds_t}{dM_t}}=0$$

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

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

and so (83) follows, given (82). □

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:

$$\alpha_t u'\left(c_t\right)=\beta R\left(1-T_t'\left(s_t\right)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{87}$$

Second, differentiating the relaxed incentive constraint:

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

or:

$$\alpha_t u'\left(c_t\right)=\frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}}\beta\left\{\frac{1}{\alpha_t}\omega_{t+1}^{\Delta}\left(\alpha_t\right)-\frac{d\omega_{t+1}\left(\alpha_t\right)}{d\alpha_t}\right\} \tag{88}$$

The right-hand sides of (87) and (88) thus give alternative expressions for the shadow value of savings at the chosen allocation. Denote this object $\gamma_t\left(\alpha_t\right)$, i.e.:

$$\gamma_t\left(\alpha_t\right):=R\left(1-T_t'\left(s_t\right)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{89}$$

$$=\frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}}\beta\left\{\frac{1}{\alpha_t}\omega_{t+1}^{\Delta}\left(\alpha_t\right)-\frac{d\omega_{t+1}\left(\alpha_t\right)}{d\alpha_t}\right\} \tag{90}$$

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:

$$\alpha_t u''(c_t(\alpha_t)) \frac{dc_t(\alpha_t)}{d\Delta_i} = \frac{d\gamma_t(\alpha_t)}{d\Delta_i} \tag{91}$$

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

$$\frac{d\gamma_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{d\gamma_t(\alpha_t)}{d\Delta_j} = 0 \tag{92}$$

Using (92) in (91):

$$\frac{dc_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{dc_t(\alpha_t)}{d\Delta_j} = 0 \tag{93}$$

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

$$\frac{ds_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{ds_t(\alpha_t)}{d\Delta_j} = 0 \tag{94}$$

and so

$$\frac{d\gamma_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{d\gamma_t(\alpha_t)}{d\Delta_j} = \left.\frac{d\gamma_t(\alpha_t)}{d\Delta_i}\right|_{s_t, c_t} + \Gamma \left.\frac{d\gamma_t(\alpha_t)}{d\Delta_j}\right|_{s_t, c_t} = 0 \tag{95}$$

– 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, (91) gives:

$$\begin{aligned} \frac{dc_t(\alpha_t)}{d\Delta_i} &= -\Gamma \frac{dc_t(\alpha_t)}{d\Delta_j} \\ &= \frac{\left.\frac{d\gamma_t(\alpha_t)}{d\Delta_i}\right|_{s_t}}{\left.\frac{d\gamma_t(\alpha_t)}{d\Delta_j}\right|_{s_t}} \frac{dc_t(\alpha_t)}{d\Delta_j} \end{aligned} \tag{96}$$

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 (89):

$$\begin{aligned} \left.\frac{d\gamma_t(\alpha_t)}{d\left(1 - T_t'(s_t)\right)}\right|_{s_t} &= \frac{1}{1 - T_t'(s_t)} \gamma_t(\alpha_t) \\ &= \frac{1}{1 - T_t'(s_t)} \alpha_t u'(c_t) \end{aligned} \tag{97}$$

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:

$$c_{t+1} = M_{t+1} - s(M_{t+2}) \tag{98}$$

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

$$M_{t+2} \equiv R\left[s(M_{t+2}) - T_{t+1}(s(M_{t+2}))\right] \tag{99}$$

We will focus on perturbations of the form:

$$c_{t+1} = M_{t+1} - s(M_{t+2}) + \Gamma f(s(M_{t+2})) \tag{100}$$

for an arbitrary bounded, a.e. differentiable function f and scalar Γ. The focus of interest will be differential movements in Γ away from zero. Taking the derivative from (90), since c_t and ω_{t+1} are being held constant we can write:

$$\left.\frac{d\gamma_t(\alpha_t)}{d\Gamma}\right|_{s_t,c_t} = \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}}\beta\frac{1}{\alpha_t}\frac{d\omega^{\Delta}_{t+1}(\alpha_t)}{d\Gamma} \tag{101}$$

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

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

$$\begin{aligned}\frac{d\omega^{\Delta}_{t+1}(\alpha_t)}{d\Gamma} =& \int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1}))\Bigg\{\alpha^2_{t+1}u'(c_{t+1})\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}\rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t) \\ &-\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}\int_{\underline{\alpha}}^{\alpha_{t+1}}\left[\tilde{\alpha}_{t+1}(u'(c_{t+1})) + \tilde{\alpha}^2_{t+1}u''(c_{t+1})\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho(\tilde{\alpha}_{t+1}|\alpha_t)\pi(\tilde{\alpha}_{t+1}|\alpha_t)\,d\tilde{\alpha}_{t+1}\Bigg\}\,d\alpha_{t+1} \\ &+ f(s_{t+1}(\bar{\alpha}))\int_{\alpha_{t+1}}\left\{\alpha_{t+1}(u'(c_{t+1})) + \alpha^2_{t+1}u''(c_{t+1})\frac{dc_{t+1}}{d\alpha_{t+1}}\right\}\rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t)\,d\alpha_{t+1}\end{aligned} \tag{102}$$

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

$$\left.\frac{dc_{t+1}}{d\Gamma}\right|_{M_{t+2}} = f(s(M_{t+2})) \tag{103}$$

To assess the magnitude of the substitution effect, note that the slope of the budget constraint at each point is:

$$\frac{dc_{t+1}}{dM_{t+2}} = -s'(M_{t+2})(1 - \Gamma f'(s(M_{t+2}))) \tag{104}$$

$$= -\frac{1 - \Gamma f'(s(M_{t+2}))}{R\left[1 - T'_{t+1}(s(M_{t+2}))\right]} \tag{105}$$

The effect on this as Γ changes is thus:

$$\frac{d}{d\Gamma}\left[\frac{dc_{t+1}}{dM_{t+2}}\right] = \frac{f'(s(M_{t+2}))}{R\left[1 - T'_{t+1}(s(M_{t+2}))\right]} \tag{106}$$

Notice that this is equal to:

$$f'\left(s\left(M_{t+2}\right)\right)\left(1-T'_{t+1}\left(s\left(M_{t+2}\right)\right)\right)\frac{d}{d\left(1-T'_{t+1}\left(s\left(M_{t+2}\right)\right)\right)}\left[\frac{dc_{t+1}}{dM_{t+2}}\right]\Bigg|_{\Gamma=0} \tag{107}$$

i.e. the perturbation changes the slope of the budget constraint by the equivalent of $f'\left(s\left(M_{t+2}\right)\right)\left(1-T'_{t+1}\left(s\left(M_{t+2}\right)\right)\right)$ times a change in the post-tax rate of return.

Substitution effects have an impact on ω^{Δ}_{t+1} to the extent that consumption is deferred:

$$\begin{aligned}
\left.\frac{d\omega^{\Delta}_{t+1}}{d\Gamma}\right|_{\text{sub}} &= \int_{\alpha_{t+1}}\left\{-\alpha_{t+1}u'\left(c_{t+1}\right)+\beta R\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)\int_{\alpha_{t+1}}V_M\left(\alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+2}|\alpha_{t+1}\right)}{d\alpha_{t+1}}d\alpha_{t+2}\right\}\\
&\times\frac{ds_{t+1}\left(\alpha_{t+1}\right)}{d\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)}\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\\
&= \int_{\alpha_{t+1}}\left\{-\alpha^2_{t+1}u'\left(c_{t+1}\right)\frac{dc_{t+1}\left(\alpha_{t+1}\right)}{d\alpha_{t+1}}\frac{1}{\varepsilon^s_{t+1}s_{t+1}\left(\alpha_{t+1}\right)}\right\}\\
&\times\frac{ds_{t+1}\left(\alpha_{t+1}\right)}{d\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)}\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\\
&= -\int_{\alpha_{t+1}}\alpha^2_{t+1}u'\left(c_{t+1}\right)\frac{dc_{t+1}\left(\alpha_{t+1}\right)}{d\alpha_{t+1}}f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}
\end{aligned} \tag{108}$$

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

$$\begin{aligned}
\left.\frac{d\omega^{\Delta}_{t+1}}{d\Gamma}\right|_{\text{inc}} &= \int_{\alpha_{t+1}}f\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\left\{\alpha_{t+1}\left(u'\left(c_{t+1}\right)\right)+\frac{ds_{t+1}\left(\alpha_{t+1}\right)}{dM_{t+1}}\left[-\alpha_{t+1}u'\left(c_{t+1}\right)\right.\right.\\
&\left.\left.+\beta R\left(1-T'_{t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\right)\int_{\alpha_{t+1}}V_M\left(\alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+2}|\alpha_{t+1}\right)}{d\alpha_{t+1}}d\alpha_{t+2}\right]\right\}\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\\
&= \int_{\alpha_{t+1}}f\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\left\{\alpha_{t+1}\left(u'\left(c_{t+1}\right)\right)+\alpha^2_{t+1}u''\left(c_{t+1}\right)\frac{dc_{t+1}}{d\alpha_{t+1}}\right\}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\\
&= f\left(s_{t+1}\left(\bar{\alpha}\right)\right)\int_{\alpha_{t+1}}\left\{\alpha_{t+1}\left(u'\left(c_{t+1}\right)\right)+\alpha^2_{t+1}u''\left(c_{t+1}\right)\frac{dc_{t+1}}{d\alpha_{t+1}}\right\}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\\
&- \int_{\alpha_{t+1}}f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\frac{ds_{t+1}\left(\alpha_{t+1}\right)}{d\alpha_{t+1}}\\
&\times\left\{\int_{\underline{\alpha}}^{\alpha_{t+1}}\left[\tilde{\alpha}_{t+1}\left(u'\left(c_{t+1}\right)\right)+\tilde{\alpha}^2_{t+1}u''\left(c_{t+1}\right)\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho\left(\tilde{\alpha}_{t+1}|\alpha_t\right)\pi\left(\tilde{\alpha}_{t+1}|\alpha_t\right)d\tilde{\alpha}_{t+1}\right\}d\alpha_{t+1}
\end{aligned} \tag{109}$$

where the second equality makes use of (82). Taking substitution and income effects, (108) and (109), together, we obtain the result. □

Applying (96), (97) and (101), we have:

$$\frac{dc_t}{d\Gamma} = \left[\frac{\alpha_t u'(c_t)}{\left(1 - T_t'(s_t)\right)} \right]^{-1} \frac{dc_t}{d\left(1 - T_t'(s_t)\right)} \frac{1}{\frac{dc_t}{d\alpha_t}} \beta \frac{1}{\alpha_t} \frac{d\omega_{t+1}^{\Delta}}{d\Gamma} \tag{110}$$

So:

$$\begin{aligned}
\frac{1}{s_t}\frac{ds_t}{d\Gamma} &= \left[\frac{\alpha_t u'(c_t)}{\left(1 - T_t'(s_t)\right)} \right]^{-1} \frac{1}{s_t} \frac{ds_t}{d\left(1 - T_t'(s_t)\right)} \frac{1}{\frac{dc_t}{d\alpha_t}} \beta \frac{1}{\alpha_t} \frac{d\omega_{t+1}^{\Delta}}{d\Gamma} \\
&= \int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \varepsilon_t^s \frac{\alpha_{t+1}\frac{ds_{t+1}}{d\alpha_{t+1}}}{\alpha_t \frac{ds_t}{d\alpha_t}} \left\{ -\frac{\beta\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)} \rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t) \right. \\
&\quad \left. + \frac{1}{\alpha_{t+1}} \int_{\underline{\alpha}}^{\alpha_{t+1}} \frac{\beta\tilde{\alpha}_{t+1}(u'(c_{t+1}))}{\alpha_t u'(c_t)} \left[1 + \frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})} \frac{\tilde{\alpha}_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \rho(\tilde{\alpha}_{t+1}|\alpha_t)\pi(\tilde{\alpha}_{t+1}|\alpha_t) d\tilde{\alpha}_{t+1} \right\} d\alpha_{t+1} \\
&\quad + f(s_{t+1}(\bar{\alpha}))\varepsilon_t^s \frac{1}{\alpha_t \frac{dc_t}{d\alpha_t}} \int_{\alpha_{t+1}} \frac{\beta\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)} \left[1 + \frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})} \frac{\alpha_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\alpha_{t+1}} \right] \rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1}
\end{aligned} \tag{111}$$

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

$$\begin{aligned}
\frac{1}{s_t}\frac{ds_t}{d\Gamma} \equiv &\int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \epsilon_{t,t+1}(s_{t+1}(\alpha_{t+1})) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \\
&+ f(s_{t+1}(\bar{\alpha})) \frac{1}{s_t} \left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}}
\end{aligned} \tag{112}$$

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:

$$\begin{aligned}
\epsilon_{t,t+1}(s_{t+1}(\alpha_{t+1})) = &\varepsilon_t^s \frac{\alpha_{t+1}\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}}{\alpha_t \frac{ds_t}{d\alpha_t}} \left\{ -\frac{\beta\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)} \rho(\alpha_{t+1}|\alpha_t) \right. \\
&+ \frac{1}{\alpha_{t+1}\pi(\alpha_{t+1}|\alpha_t)} \int_{\underline{\alpha}}^{\alpha_{t+1}} \frac{\beta\tilde{\alpha}_{t+1}(u'(c_{t+1}))}{\alpha_t u'(c_t)} \left[1 + \frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})} \frac{\tilde{\alpha}_{t+1}}{c_{t+1}} \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \qquad (113) \\
&\left. \times \rho(\tilde{\alpha}_{t+1}|\alpha_t)\pi(\tilde{\alpha}_{t+1}|\alpha_t) d\tilde{\alpha}_{t+1} \right\} \qquad (114)
\end{aligned}$$

So:

$$
\begin{aligned}
&RT_t'(s_t)\, s_t \epsilon_{t,t+1}(s_{t+1}(\alpha_{t+1})) && (115)\\
=&RT_t'(s_t)\, s_t \varepsilon_t^s \frac{\alpha_{t+1}\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}}{\alpha_t \frac{ds_t}{d\alpha_t}} \left\{-\frac{\beta\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)}\rho(\alpha_{t+1}|\alpha_t)\right. && (116)\\
&\left.+\frac{1}{\alpha_{t+1}\pi(\alpha_{t+1}|\alpha_t)}\int_{\underline{\alpha}}^{\alpha_{t+1}} \frac{\beta\tilde{\alpha}_{t+1}(u'(c_{t+1}))}{\alpha_t u'(c_t)}\left[1+\frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho(\tilde{\alpha}_{t+1}|\alpha_t)\pi(\tilde{\alpha}_{t+1}|\alpha_t)\,d\tilde{\alpha}_{t+1}\right\}\\
=&-\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\alpha_t u'(c_t)\,\alpha_{t+1}\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}\left\{-\frac{\beta R\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)}\rho(\alpha_{t+1}|\alpha_t)\right. && (117)\\
&\left.+\frac{1}{\alpha_{t+1}\pi(\alpha_{t+1}|\alpha_t)}\int_{\underline{\alpha}}^{\alpha_{t+1}} \frac{\beta R\tilde{\alpha}_{t+1}(u'(c_{t+1}))}{\alpha_t u'(c_t)}\left[1+\frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho(\tilde{\alpha}_{t+1}|\alpha_t)\pi(\tilde{\alpha}_{t+1}|\alpha_t)\,d\tilde{\alpha}_{t+1}\right\}\\
=&\beta R\alpha_{t+1}u'(c_{t+1})\,\rho(\alpha_{t+1}|\alpha_t)\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\alpha_{t+1}\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} && (118)\\
&-\frac{\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}}{\pi(\alpha_{t+1}|\alpha_t)}\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\int_{\underline{\alpha}}^{\alpha_{t+1}}\beta R\tilde{\alpha}_{t+1}(u'(c_{t+1}))\left[1+\frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho(\tilde{\alpha}_{t+1}|\alpha_t)\pi(\tilde{\alpha}_{t+1}|\alpha_t)\,d\tilde{\alpha}_{t+1}
\end{aligned}
$$

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

For condition 4, we have established:

$$
\frac{1}{s_t}\left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}} = \varepsilon_t^s \frac{1}{\alpha_t\frac{dc_t}{d\alpha_t}}\int_{\alpha_{t+1}} \frac{\beta\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)}\left[1+\frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t)\,d\alpha_{t+1} \quad (119)
$$

So:

$$
RT_t'(s_t)\left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}} = \frac{\lambda_{t+1}^{\Delta}}{\eta_t}\alpha_t u'(c_t)\int_{\alpha_{t+1}} \frac{\beta R\alpha_{t+1}u'(c_{t+1})}{\alpha_t u'(c_t)}\left[1+\frac{c_{t+1}u''(c_{t+1})}{u'(c_{t+1})}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t)\,d\alpha_{t+1} \quad (120)
$$

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

A.9 Proof of Theorem 2

Equation (31), above, gives:

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

The utility function is time-separable and concave in consumption at each date-state, which together straightforwardly imply $\frac{ds_t}{dM_t} > 0$. Concavity further gives $u''(c_t) < 0$, the strict increasingness assumption implies $\frac{dc_t}{d\alpha_t} > 0$, and $\eta > 0$ from (23). It follows that $T_t'(s_t)$ has the same sign as $\lambda_{t+1}^{\Delta}(\alpha_t(c_t))$, and so we focus on signing the latter object.

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

Lemma 7. *For all t and 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:

$$\begin{aligned}&\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}\,u'\left(c_s\left(\alpha_t',...,\bar{\alpha}\right)\right)}-\frac{1}{(\beta R)^{s-t}\,u'\left(c_s\left(\alpha_t',...,\underline{\alpha}\right)\right)}\right]\\ =&\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}}\frac{\bar{\alpha}}{\mathbb{E}_s\left[\alpha_{s+1}|\bar{\alpha}\right]}\mathbb{E}_s\left[\frac{1}{\beta Ru'\left(c_{s+1}\left(\alpha_t',...,\bar{\alpha},\alpha_{s+1}\right)\right)}\middle|\bar{\alpha}\right]\right]\\ &-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}}\frac{\underline{\alpha}}{\mathbb{E}_s\left[\alpha_{s+1}|\underline{\alpha}\right]}\mathbb{E}_s\left[\frac{1}{\beta Ru'\left(c_{s+1}\left(\alpha_t',...,\underline{\alpha},\alpha_{s+1}\right)\right)}\middle|\underline{\alpha}\right]\right]\end{aligned}\tag{121}$$

If the right-hand side of (121) is unbounded in s then so is the expression:

$$\begin{aligned}&\frac{1}{(\beta R)^{s-t}}\mathbb{E}_t\left[\mathbb{E}_s\left[\frac{1}{\beta Ru'\left(c_{s+1}\left(\alpha_t',...,\bar{\alpha},\alpha_{s+1}\right)\right)}\middle|\bar{\alpha}\right]-\frac{\underline{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\bar{\alpha}\right]}{\bar{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\underline{\alpha}\right]}\mathbb{E}_s\left[\frac{1}{\beta Ru'\left(c_{s+1}\left(\alpha_t',...,\underline{\alpha},\alpha_{s+1}\right)\right)}\middle|\underline{\alpha}\right]\right]\\ =&\frac{1}{(\beta R)^{s+1-t}}\mathbb{E}_t\left[\int_{\alpha_{s+1}}\left[\frac{1}{u'\left(c_{s+1}\left(\alpha_t',...,\bar{\alpha},\alpha_{s+1}\right)\right)}\pi\left(\alpha_{s+1}|\bar{\alpha}\right)-\frac{1}{u'\left(c_{s+1}\left(\alpha_t',...,\underline{\alpha},\alpha_{s+1}\right)\right)}\frac{\underline{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\bar{\alpha}\right]}{\bar{\alpha}\mathbb{E}_s\left[\alpha_{s+1}|\underline{\alpha}\right]}\pi\left(\alpha_{s+1}|\underline{\alpha}\right)\right]d\alpha_{s+1}\right]\end{aligned}\tag{122}$$

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

$$\frac{1}{(\beta R)^{s+1-t}}\mathbb{E}_t\left[\int_{\alpha_{s+1}}\frac{1}{u'\left(c_{s+1}\left(\alpha_t',...,\underline{\alpha},\alpha_{s+1}\right)\right)}\left[\frac{\pi\left(\alpha_{s+1}|\bar{\alpha}\right)-\pi\left(\alpha_{s+1}|\underline{\alpha}\right)\frac{\underline{\alpha}\mathbb{E}_s[\alpha_{s+1}|\bar{\alpha}]}{\bar{\alpha}\mathbb{E}_s[\alpha_{s+1}|\underline{\alpha}]}}{\pi\left(\alpha_{s+1}|\underline{\alpha}\right)}\right]\pi\left(\alpha_{s+1}|\underline{\alpha}\right)d\alpha_{s+1}\right]\tag{123}$$

Moreover, continuity of the density implies that the object:

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

is bounded in α_{s+1}. Thus unboundedness of (123) 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:

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

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

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

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:

$$\begin{aligned}\lim_{s\to\infty}\left\{\frac{1}{\mathbb{E}_t\left[\alpha_s\right]}\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\right)}\right]\right\}&=\frac{1+\lambda_{t+1}\left(\alpha_t\right)}{\eta_t}\\&=\frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\mathbb{E}_t\left[\alpha_{t+1}\right]}\mathbb{E}_t\left[\frac{1}{\beta Ru'\left(c_{t+1}\right)}\right]-\frac{\varepsilon^{\alpha}\left(\alpha_t\right)}{\alpha_t u'\left(c_t\right)}\right\}\end{aligned}\tag{124}$$

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 (52) and (53) gives:

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

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

$$\lambda_{t+1}^{\Delta}\left(\alpha_t\right)\alpha_t\pi\left(\alpha_t|\alpha_{t-1}\right)=\int_{\alpha_t}^{\bar{\alpha}}\left[\left(1+\lambda_t+\lambda_t^{\Delta}\pi^{\Delta}\left(\tilde{\alpha}_t|\alpha_{t-1}\right)\right)-\left(1+\lambda_{t+1}\left(\tilde{\alpha}_t\right)\right)\right]\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha}_t\tag{125}$$

with:

$$\int_{\underline{\alpha}}^{\bar{\alpha}}\left[\left(1+\lambda_t+\lambda_t^{\Delta}\pi^{\Delta}\left(\alpha_t|\alpha_{t-1}\right)\right)-\left(1+\lambda_{t+1}\left(\alpha_t\right)\right)\right]\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t=0\tag{126}$$

Since $\pi^{\Delta}\left(\alpha_t|\alpha_{t-1}\right)$ is monotone increasing in α_t, and $\lambda_0^{\Delta}=0$, a sufficient condition for the right-hand side of (125) to be weakly positive for all t and all histories is that $\left(1+\lambda_{t+1}\left(\alpha_t\right)\right)$ should be non-increasing in α_t. Lemma 7 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 $\bar{\rho}=\sup_{\alpha,\alpha'}[\rho(\alpha'|\alpha)]<1$.

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:

$$\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha''_t,...,\alpha_s\right)\right)}\middle|\alpha''_t\right]-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha'_t,...,\alpha_s\right)\right)}\middle|\alpha'_t\right] \tag{127}$$

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:

$$\int_{\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \left[\frac{1}{(\beta R)^{s-t} u'\left(c_s\left(\alpha_t', \ldots, \alpha_s\right)\right)} \right] \pi\left(\alpha_s | \alpha_{s-1}\right) d\alpha_s \ldots \pi\left(\alpha_{t+2} | \alpha_{t+1}\right) d\alpha_{t+2} \right\} \left[\pi\left(\alpha_{t+1} | \alpha_t''\right) - \pi\left(\alpha_{t+1} | \alpha_t'\right) \right] d\alpha_{t+1}$$

$$= \int_{\alpha_t'}^{\alpha_t''} \int_{\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \left[\frac{1}{(\beta R)^{s-t} u'\left(c_s\left(\alpha_t', \ldots, \alpha_s\right)\right)} \right] \pi\left(\alpha_s | \alpha_{s-1}\right) d\alpha_s \ldots \pi\left(\alpha_{t+2} | \alpha_{t+1}\right) d\alpha_{t+2} \right\} \frac{d\pi\left(\alpha_{t+1} | \alpha_t\right)}{d\alpha_t} d\alpha_{t+1} d\alpha_t$$

$$= \int_{\alpha_t'}^{\alpha_t''} \frac{1}{\alpha_t} \int_{\alpha_{t+1}} \alpha_{t+1} \frac{d}{d\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \left[\frac{1}{(\beta R)^{s-t} u'\left(c_s\left(\alpha_t', \ldots, \alpha_s\right)\right)} \right] \pi\left(\alpha_s | \alpha_{s-1}\right) d\alpha_s \ldots \pi\left(\alpha_{t+2} | \alpha_{t+1}\right) d\alpha_{t+2} \right\}$$
$$\times \rho\left(\alpha_{t+1} | \alpha_t\right) \pi\left(\alpha_{t+1} | \alpha_t\right) d\alpha_{t+1} d\alpha_t$$

$$= \int_{\alpha_t'}^{\alpha_t''} \frac{1}{\alpha_t} \int_{\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \alpha_{t+1} \frac{d}{d\alpha_{t+1}} \left[\frac{1}{(\beta R)^{s-t} u'\left(c_s\left(\alpha_t', \ldots, \alpha_s\right)\right)} \right] \pi\left(\alpha_s | \alpha_{s-1}\right) d\alpha_s \ldots \pi\left(\alpha_{t+2} | \alpha_{t+1}\right) d\alpha_{t+2} \right\}$$
$$\times \rho\left(\alpha_{t+1} | \alpha_t\right) \pi\left(\alpha_{t+1} | \alpha_t\right) d\alpha_{t+1} d\alpha_t$$
$$+ \int_{\alpha_t'}^{\alpha_t''} \frac{1}{\alpha_t} \int_{\alpha_{t+1}} \left\{ \int_{\alpha_{t+2}} \cdots \int_{\alpha_s} \left[\frac{1}{(\beta R)^{s-t} u'\left(c_s\left(\alpha_t', \ldots, \alpha_s\right)\right)} \right] \pi\left(\alpha_s | \alpha_{s-1}\right) d\alpha_s \ldots \alpha_{t+1} \frac{d\pi\left(\alpha_{t+2} | \alpha_{t+1}\right)}{d\alpha_{t+1}} d\alpha_{t+2} \right\}$$
$$\times \rho\left(\alpha_{t+1} | \alpha_t\right) \pi\left(\alpha_{t+1} | \alpha_t\right) d\alpha_{t+1} d\alpha_t$$

$$= \cdots$$

$$= \int_{\alpha_t'}^{\alpha_t''} \frac{1}{\alpha_t} \left\{ \sum_{r=t+1}^{s} \mathbb{E}_t \left[D_{t,r}\left(\alpha^r\right) \alpha_r \frac{d}{d\alpha_r} \left[\frac{1}{(\beta R)^{s-t} u'\left(c_s\left(\alpha_t', \ldots, \alpha_r, \ldots, \alpha_s\right)\right)} \right] \middle| \alpha_t \right] \right\} d\alpha_t$$

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

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

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:

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

Lemma 7 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 (127) 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. $\lambda_{t+1}\left(\alpha_t\right)$ is constant in α_t

The first case implies $\lambda^{\Delta}_{t+1}(\alpha_t) > 0$ everywhere except endpoints, from (125) and the fact that the integral in (125) 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:

$$\frac{1}{\alpha_t u'(c_t)} = \frac{1}{\mathbb{E}_t[\alpha_{t+1}]}\mathbb{E}_t\left[\frac{1}{\beta R u'(c_{t+1})}\right] \tag{128}$$

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 (128) 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 (65) 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 $\lambda_{t+1}^{\Delta}(\alpha_t) = 0$ 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.10 Proof of Proposition 6

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

$$\begin{aligned}
&-\rho\left(\alpha_t\left(c'_t\right)|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\left(\alpha_t\left(c'_t\right)\right)^2 u'\left(c'_t\right)\left(\frac{d\alpha_t\left(c'_t\right)}{dc_t}\right)^{-1}\frac{\pi^c\left(c'_t|\alpha^{t-1}\right)}{1-\Pi^s\left(s'_t|\alpha^{t-1}\right)} \\
&+\frac{1}{1-\Pi^s\left(s'_t|\alpha^{t-1}\right)}\int_{\underline{c}}^{c'_t}\rho\left(\alpha_t(c_t)|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\left[\alpha_t(c_t)(u'(c_t))+(\alpha_t(c_t))^2 u''(c_t)\left(\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1}\right]\pi^c\left(c_t|\alpha^{t-1}\right)dc_t
\end{aligned} \tag{129}$$

Switching to express arguments in terms of α_t:

$$\begin{aligned}
&-\rho\left(\alpha_t|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}(\alpha_t)^2 u'\left(c_t\left(\alpha'_t\right)\right)\frac{\pi\left(\alpha'_t|\alpha_{t-1}\right)}{\Pi\left(\alpha'_t|\alpha_{t-1}\right)} \\
&+\frac{1}{\Pi\left(\alpha'_t|\alpha_{t-1}\right)}\int_{\underline{\alpha}}^{\alpha'_t}\rho\left(\alpha_t|\alpha_{t-1}\right)\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\left[\alpha_t\left(u'\left(c_t\left(\alpha_t\right)\right)\right)+\alpha_t^2 u''\left(c_t\left(\alpha_t\right)\right)\frac{dc_t\left(\alpha_t\right)}{d\alpha_t}\right]\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t
\end{aligned} \tag{130}$$

Integration by parts gives the following relationship:

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

So:

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

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, whilst $\alpha_t u'(c_t(\alpha_t))$ is strictly positive. 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'). From Lemma 4:

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

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