Optimal Nonlinear Savings Taxation*

Charles Brendon†

July 21, 2023

Abstract

This paper analyses the design of optimal nonlinear savings taxation, in a multi-period consumption-savings economy where consumers face persistent, uninsurable shocks to their marginal utility of consumption. Its main contributions are: (a) to show that shocks of this kind generically justify positive marginal savings taxes, and (b) to characterise these taxes by reference to a limited number of sufficient statistics. This characterisation highlights the novel insights that can result from connecting 'Mirrleesian' and 'sufficient statistics' approaches to dynamic taxation. Specifically, the restrictions implicit in mechanism design analysis allow sufficient statistics approaches to remain manageable, even in high-dimensional choice settings.

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

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

*First draft: October 2021. I am very grateful to Florian Scheuer for providing detailed comments on an earlier draft. I also thank Vasco Carvalho, Paweł Doligalski, Lukas Freund, Andrew Hannon, Christian Hellwig, David Martimort, Ludvig Sinander, Stefanie Stantcheva, Bruno Strulovici and Nicolas Werquin for useful comments and discussions, together with seminar participants in Bristol, Cambridge, EUI and TSE. The first draft was completed when I was a visitor at TSE: I am very grateful for the hospitality shown. Richard Nickl provided important literary motivation.

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

1 Introduction

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

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

The main focus of the paper is on the optimal determination of the savings tax schedule. I show that it is generically optimal to set a positive marginal tax rate on savings in all time periods and for all shock histories – strictly positive away from the extreme ends of the savings distribution. The revenue from this savings tax is used to fund a positive, uniform lump-sum transfer each period, which permits higher within-period consumption for those with high marginal utility, relative to what they would be able to afford under a laissez-faire system.

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

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

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

The simplification relies, in part, on the Markovian structure that I place on the hidden type process. This ensures 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. This independence implies that there is no gain to distorting allocations in later time periods, in order to improve the structure of incentives more than one period previously. This keeps the relevant behavioural considerations for policy design to a manageable level.

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

[1] 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).

1.1 Preview of main characterisation

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

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

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 transferring income to an individual whose savings are s. The welfare weight – an endogenous object that evolves with individuals' wealth levels – varies across types with a common history in proportion to their marginal utility of consumption alone. Thus the policymaker is conditionally utilitarian. It is decreasing in savings, because higher savers have a relatively low contemporaneous marginal utility of consumption.

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

Cutting taxes at s' in the current period may also change savings in the previous period, by an amount proportional to a compensated cross elasticity $\epsilon^s_{-1}(s')$. 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 particular to consumption choice models with imperfect insurance. It is the behavioural response that arises because of a change to the state-contingent profile of returns provided by the savings instrument. Its relevance is intrinsically linked to type persistence in the underlying structural model. It is zero when types are iid.

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

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

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

1.2 Paper outline

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

To aid understanding, the main characterisation is presented constructively, in steps. In Section 5 I use standard methods to characterise constrained-optimal allocations by reference to the costs and benefits of changing *utility* levels for a cross-section of individuals. The resulting expressions are insightful, and reveal useful features about the dynamics of consumption when types are persistent, but they rely heavily on arguments of the utility function. To overcome this, Section 7 defines the set of objects that will be used to construct an alternative sufficient statistics characterisation. The characterisation that follows is given as Theorem 1 in Section 8. Section 9 explores the qualitative properties of optimal savings taxes, notably the result that optimal marginal tax rates are positive. Section 10 explains the limited role for intertemporal cross-elasticities in characterisation of optimal taxes. Section 11 provides an illustrative quantification of marginal tax rates for top savings levels, based on the formula obtained in Section 8, and compares the outcomes with alternatives in the literature. Section 12 concludes.

All proofs are relegated to the appendix.

2 Relation to literature

The basic insurance problem that I study was first proposed and analysed by Atkeson and Lucas (1992), who focused on the properties of constrained-optimal allocations in the presence of unobservable shocks to marginal utility. Their paper gave particular attention to long-run outcomes, showing that the immiseration result of Thomas and Worrall (1990) – whereby measure 1 of agents see their consumption converge to zero over time – carried over to this setting, as well as emphasising that the optimum could not be decentralised via conventional linear pricing. Technically, Atkeson and Lucas assumed an iid type distribution, with types drawn from a finite set period-by-period – a structure retained by most of the literature that explores the sensitivity of their immiseration result.[4] The current paper instead allows for persistent (Markovian) type draws, which has non-trivial implications for consumption dynamics relative to the iid case. I also assume types are drawn from a continuum, which proves crucial in finding a mathematical link from the mechanism design characterisation to behavioural statistics.

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

Much – though not all – of this literature has focused on characterising the differences between constrained-optimal allocations and laissez-faire outcomes, rather than focusing directly on tax instruments.[5] Emphasis in the early papers was on the well-known 'inverse Euler equation' – an expression that implies a distortion relative to savings behaviour under autarky, but does not directly map to any particular tax instrument.[6] Likewise, more recent papers by Farhi and Werning (2013) and Golosov, Troshkin and Tsyvinski (2016) have examined the properties of the 'wedge' between the consumption-labour marginal rate of substitution and the marginal product of labour. But in dynamic settings the link between this wedge and labour income

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

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

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

tax rate is no longer direct. By contrast, the main characterisation in the present paper relates to the marginal savings tax itself – an object directly controlled by policy.

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

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

Away from the dynamic Mirrlees literature, the results in this paper contribute to the influential movement to link policy prescriptions to observable 'sufficient statistics', insofar as possible. From original contributions by Diamond (1998) and Saez (2001), which re-cast the static Mirrlees (1971) model by reference to instruments rather than allocations, this approach now encompasses broad areas of macro and micro policy design.[7] Yet in contrast with the static literature, for dynamic tax problems 'sufficient statistics' and 'mechanism design' approaches are commonly interpreted as alternatives rather than complements – that is, as distinct methods generating distinct policy prescriptions, rather than equivalent formulations linked by a duality relationship.

The reason for this divorce has been a desire to obtain comparably simple policy lessons for dynamic

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

tax environments as for static, and the seeming difficulty of achieving this in a mechanism design setting. Instead, influential papers by Piketty and Saez (2013a) and Saez and Stantcheva (2018) have discarded information-theoretic foundations, in favour of a long-run focus: tax instruments are assumed to be time-invariant, and the effects of any changes are analysed purely by reference to their mechanical, welfare and behavioural effects *in steady state*.[8] This overcomes the need to consider arbitrary intertemporal cross elasticities – the response of savings in period t to taxation in $t+59$, say – by the simplifying assumption that what matters is what happens in the long run.[9] The present paper instead shows that simple, intuitive sufficient statistics characterisations *can* arise from a mechanism design approach, attributing this to the substantial structure placed on preferences and shocks in these settings. Thus I highlight an alternative route to policy insight, by comparison with the more radical focus on long-run outcomes alone.

By offering a novel justification for savings taxation, this paper also contributes to the large general literature on the desirability of intertemporal distortions. Work on this topic has moved on considerably from the classic Chamley (1986) and Judd (1985) zero tax results, due both to the direct assault of Straub and Werning (2020),[10] and the earlier findings that savings taxes could play a useful role in computational Ramsey environments.[11] Yet a common theme in this literature remains that savings distortions are only introduced because of significant limitations elsewhere in the tax system – particularly credit constraints, limits on age-dependent taxation, or arbitrary tax ceilings. It provides few direct arguments *for* savings taxes. In this regard the present paper differs: if savings reveal consumption need, and the government would like to redistribute according to consumption need, then a savings tax is the most direct, appropriate intervention.

Finally, this paper has links to the growing microeconomic mechanism design literature that gives particular attention to the problems implied by type persistence.[12] Current work by Bloedel, Krishna and Strulovici (2020), Bloedel, Krishna and Leukhina (2021) and Makris and Pavan (2020) deploy various settings to explore the dynamics of wedges in problems without quasilinearity. I complement this by showing that type persistence has a direct mapping to the presence of one additional elasticity in the sufficient statistics formulation of optimal tax policy: the compensated responsiveness of savings to a change in insurance.

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

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

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

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

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

Persistence matters to the policymaker only to the extent that this elasticity is non-zero.

3 Model setup

This section and the next present the information-theoretic dynamic social insurance problem that is the starting point for the analysis.

3.1 Preliminaries

Time is discrete but infinite, indexed by the natural numbers and starting in period zero. The economy consists of a measure-1 continuum of individuals, plus a policymaker 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, $\beta \in (0,1)$ is the discount factor, $u : \mathbb{R}^+ \to \mathbb{R}$ or $\mathbb{R}^{++} \to \mathbb{R}$ is the period utility function, and $\alpha_t \in [\underline{\alpha}, \bar{\alpha}] \subset \mathbb{R}^+$ is the idiosyncratic taste disturbance in t, with $\underline{\alpha} > 0$ and $\bar{\alpha} < \infty$. To keep notation compact, we will refer to the interval $[\underline{\alpha}, \bar{\alpha}]$ as A. $\alpha^t \in A^{t+1}$ will denote a complete idiosyncratic history of taste draws up to period t, and $\alpha_t^s \in A^{s-t+1}$ a partial sequence of draws between periods t and s (inclusive). I make the following standard assumption on the utility function:

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

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

The taste parameter is assumed to follow a Markov process, identical through time in all periods except the initial period 0. Conditional on drawing $\alpha_t \in A$ in period t, the distribution of shocks in $t+1$ is denoted

$\Pi(\alpha_{t+1}|\alpha_t)$, 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. I place the following regularity structure on the distributions:

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

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

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

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

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

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

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

Integrating provides a link between the two preceding objects:

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

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

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

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

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

Finally, I place some structure on the informativeness of current type about lagged type. This takes the form of: (a) a monotone likelihood ratio condition, so that higher values for current type are more likely when past type is higher, and (b) a boundedness condition, which says that variations in past type are associated with bounded variations in the pdf for current type. This structure is not used in the characterisation results, but it plays an important role in Theorem 2, confirming that optimal savings taxes are positive.

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}, and bounded relative to the ratio of lagged types:*

$$\frac{\pi(\alpha_{t+1}|\alpha''_t)}{\pi(\alpha_{t+1}|\alpha'_t)} \leq \left(\frac{\alpha''_t}{\alpha'_t}\right)^{\kappa}$$

for some $\kappa > 0$.

It is helpful to view this assumption by reference to the elasticity of the density with respect to lagged type, π^{Δ}:

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

Assumption 4 implies that $\pi^{\Delta}(\alpha_{t+1}|\alpha_t)$ is monotone increasing in α_{t+1}, and bounded above by κ.

3.3 Planner choice

The planner's aim in period 0 is to maximise the average lifetime utility, denoted W_0:

$$W_0 := \int_{\underline{\alpha}}^{\bar{\alpha}} U_0(\alpha_0)\, d\Pi(\alpha_0) \tag{8}$$

The 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. Thus individuals report their type each period, and receive a consumption allocation conditional on their reports to date, $c_t(\alpha^t)$. A complete set of $c_t(\alpha^t)$ functions for all $t \geq 0$ and $\alpha^t \in A^{t+1}$ is referred to as an **allocation**.[13]

The planner's choice is restricted by the resource and incentive compatibility constraints detailed below, plus a technical **boundedness restriction**. This is that there must exist a sequence of history-contingent scalars $\left\{K_t\left(\alpha^{t-1}\right)\right\}_{t\geq 0, \alpha^{t-1}\in A^t}$ such that for all $\alpha^t \in A^{t+1}$, the bound:

$$\left|\mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} \alpha_s u\left(c_s\left(\alpha^s\right)\right)\right| \leq K_t\left(\alpha^{t-1}\right) \tag{9}$$

holds uniformly for every α^t that succeeds α^{t-1}.

Condition (9) rules out that the planner engineers extreme degrees of cross-sectional inequality at any history node. The value of $K_t\left(\alpha^{t-1}\right)$ can be arbitrarily large, and is not itself subject to any uniform boundedness in t or α^{t-1}. The condition is only used actively in Lemma 1, below, as a device to guarantee that information rents exist everywhere. Rather than being assumed *ex-ante*, it could equivalently be checked *ex-post*.

3.4 Resources

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

$$\sum_{t=0}^{\infty} R^{-t}\left[y_t - \int_{\alpha^t} c_t\left(\alpha^t\right) d\Pi_t\left(\alpha^t\right)\right] \geq 0 \tag{10}$$

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

3.5 Incentive compatibility

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

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

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 $t\geq 0$. [14]

Incentive compatibility then requires:

$$\begin{aligned}&\alpha_t' u\left(c_t\left(\alpha^{t-1},\alpha_t'\right)\right)+\beta\int_{\alpha_{t+1}}V_{t+1}\left(\left(\alpha^{t-1},\alpha_t'\right);\alpha_{t+1}\right)d\Pi\left(\alpha_{t+1}|\alpha_t'\right)\\ \geq&\alpha_t' u\left(c_t\left(\alpha^{t-1},\alpha_t''\right)\right)+\beta\int_{\alpha_{t+1}}V_{t+1}\left(\left(\alpha^{t-1},\alpha_t''\right);\alpha_{t+1}\right)d\Pi\left(\alpha_{t+1}|\alpha_t'\right)\end{aligned} \tag{11}$$

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

Note that the true type α_t' affects restriction (11) in two ways. Most directly, it controls the marginal utility of consumption in t. But current type also affects the distribution of future type draws, $\Pi\left(\alpha_{t+1}|\alpha_t'\right)$. This persistence channel complicates the link between preferences and type, relative to a canonical two-good screening problem.

An allocation that satisfies constraints (9), (10) and (11) for all histories and all time periods is described as **incentive-feasible**. The **planner's problem** is to maximise W_0 on the set of incentive-feasible allocations.

4 First-order incentive compatibility

4.1 A relaxed incentive constraint

Condition (11) 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, I thus replace them with a 'first-order' envelope requirement that is necessary for (11) to be true, but not sufficient. The conjecture, to be verified, is that optimal policy for the simplified constraint set will also be optimal for the more complex constraint set.[15]

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

[15]This first-order approach is now used extensively in the dynamic Mirrleesian literature, as discussed in Section 2.

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

$$\omega_t\left(\alpha^{t-1}\right) := \int_{\alpha_t}\left\{\alpha_t u\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right) + \beta\omega_{t+1}\left(\alpha^{t-1},\alpha_t\right)\right\} d\Pi\left(\alpha_t|\alpha_{t-1}\right) \tag{12}$$

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

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

The following result is an application of Theorem 2 in Milgrom and Segal (2002):

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

$$\begin{aligned}\alpha_t' u\left(c_t\left(\alpha_t'\right)\right) + \beta\omega_{t+1}\left(\alpha_t'\right) = {} & \underline{\alpha} u\left(c_t\left(\underline{\alpha}\right)\right) + \beta\omega_{t+1}\left(\underline{\alpha}\right) \\ & + \int_{\underline{\alpha}}^{\alpha_t'}\frac{1}{\alpha_t}\left[\alpha_t u\left(c_t\left(\alpha_t\right)\right) + \beta\omega_{t+1}^{\Delta}\left(\alpha_t\right)\right] d\alpha_t\end{aligned} \tag{14}$$

I refer to equation (14) as the **relaxed incentive constraint**. For an arbitrary type α_t', it decomposes the value of lifetime utility into the value for the lowest type $\underline{\alpha}$, plus the sum of information rents between $\underline{\alpha}$ and α_t'. These information rents are the objects contained within the integral on the last line.

An allocation that satisfies the interiority constraint (9), resource constraint (10) and relaxed incentive constraint (14) 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, then it must be optimal for the main planner's problem. Confirming this inclusion is the central issue in justifying the first-order approach, and the focus of the next subsection.

4.2 Sufficiency

I present two alternative criteria for confirming global incentive compatibility. The first is an 'integral monotonicity condition', of the type introduced in quasilinear settings by Pavan, Segal and Toikka (2014).[16] This has the advantage that it is both necessary and sufficient for (14) to imply global incentive compatibility, but the disadvantage that it depends on properties of the utility function in addition to the allocation. In this regard it makes significant informational demands. A subsequent corollary gives a sufficient – but not necessary – monotonicity condition on the allocation alone. This condition has a particularly clear interpretation by reference to the decentralisation that is introduced later in the paper.

The integral monotonicity condition is:[17]

Proposition 1. *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:*

$$\int_{\alpha'_t}^{\alpha''_t} \frac{1}{\alpha_t^2} \left\{ \mathbb{E}_t \left[\sum_{s=t+1}^{\infty} \beta^{s-t} \left(1 - D_{t,s}\left(\alpha^s\right)\right) \alpha_s \left[u\left(c_s\left(\alpha^{t-1}, \alpha_t, \alpha_{t+1}^s\right)\right) - u\left(c_s\left(\alpha^{t-1}, \alpha''_t, \alpha_{t+1}^s\right)\right)\right] \middle| \alpha_t \right] \right\} d\alpha_t \geq 0 \tag{15}$$

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

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

Thus global incentive compatibility is confirmed so long as future consumption is weakly decreasing in current type, along all subsequent history nodes. In the decentralisation presented in Section 6, this is equivalent to requiring that higher savings in t raise consumption along every subsequent path for type draws – so that consumption is a normal good at every date-state. Thus I refer to an allocation that satisfies the requirements of Corollary 1 a normal allocation:

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

Normality – and the global incentive compatibility that it implies – guarantees that period-t consumption

[16] C.f. their Theorem 3.

[17] To simplify presentation, I adopt the convention for definite integrals that $\int_{\alpha'_t}^{\alpha''_t} \{\cdot\} d\alpha_t$ corresponds to $-\int_{\alpha''_t}^{\alpha'_t} \{\cdot\} d\alpha_t$ when $\alpha'_t > \alpha''_t$.

is non-decreasing in α_t for types with a common history. It has particular relevance because, as I show in Section6, it is also a sufficient condition for a market decentralisation with nonlinear taxes to be possible.

For some purposes it is convenient to strengthen the condition, and neglect the possibility that multiple types bunch at the same consumption value in t. To this end, I define 'strictly normal' allocations as follows:

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

5 Characterising direct allocations

In the present section, I provide a direct characterisation of optimal allocations for the mechanism design problem, trading off the benefits and costs of reducing information rents. The resulting representation depends on the traditional components of a mechanism design setup: utility functions and their arguments, and the latent type process. Yet it is of critical instrumental value to the subsequent sufficient statistics formulation. The main characterisation result in Section 8 is proved via a direct manipulation of the conditions derived here.

5.1 Characterisation result

By conventional optimisation techniques and elementary manipulations, I derive the following:

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

$$\mathbb{E}_{t-1}\left\{\alpha_t\left[1+\lambda_t+\lambda_t^{\Delta}\rho\left(\alpha_t|\alpha_{t-1}\right)\right]-\frac{\eta_t}{u'\left(c_t\left(\alpha_t\right)\right)}\middle|\alpha_t > \alpha_t'\right\} = \frac{\pi\left(\alpha_t'|\alpha_{t-1}\right)}{\left(1-\Pi\left(\alpha_t'|\alpha_{t-1}\right)\right)}\cdot\left(\alpha_t'\right)^2\cdot\left\{\lambda_{t+1}^{\Delta}\left(\alpha_t'\right)-\rho\left(\alpha_t'|\alpha_{t-1}\right)\lambda_t^{\Delta}\right\} \tag{16}$$

$$\mathbb{E}_{t-1}\left\{\alpha_t\left[1+\lambda_t+\lambda_t^{\Delta}\varepsilon^{\alpha}\left(\alpha_{t-1}\right)\right]-\frac{\eta_t}{u'\left(c_t\left(\alpha_t\right)\right)}\right\} = 0 \tag{17}$$

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

$$\eta_t = \left(\beta R\right)^{-t}\frac{\mathbb{E}\left[\alpha_0\right]}{\mathbb{E}\left[\frac{1}{u'\left(c_0\left(\alpha_0\right)\right)}\right]} \tag{18}$$

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

$$\lambda_{t+1} = \lambda_t + \mu_t\left(\alpha_t'\right) \tag{19}$$

$$\lambda_{t+1}^{\Delta} = \rho\left(\alpha_t'|\alpha_{t-1}\right)\lambda_t^{\Delta} - \frac{1-\Pi\left(\alpha_t'|\alpha_{t-1}\right)}{\alpha_t\pi\left(\alpha_t'|\alpha_{t-1}\right)}\mathbb{E}_{t-1}\left[\mu_t\left(\alpha_t'\right)\middle|\alpha_t > \alpha_t'\right] \tag{20}$$

with $\mu_t\left(\alpha_t'\right)$ a mean-zero multiplier defined in the appendix. The conditional distribution and densities are replaced with their unconditional equivalents for period 0*.*

Expressions (16) and (17) are the main objects of interest here. (16) can be interpreted by reference to the costs and benefits of changing information rents at the critical type α_t', for agents with a common history. This yields a direct welfare benefit, mitigated by the direct marginal resource cost of the higher utility – together accounting 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 (14) to remain true. This is captured by the object on the right-hand side.[18]

If instead utility is raised uniformly for all agents with a common history prior to t, then no within-period change to information rents is required. The result is equation (17): the value of raising welfare across types must equal the resource cost of doing so.

λ_t and λ_t^{Δ} in these expressions are cumulated multipliers deriving from prior incentive restrictions – capturing the shadow costs of changing ω_t and ω_t^{Δ} respectively. As highlighted by Marcet and Marimon (2019), the dynamics of these objects correspond to the dynamics of the Pareto weights that the policymaker implicitly attaches to different agents' utility as time progresses.

6 A consumption-savings decentralisation

6.1 Decentralised choice structure

I now consider a simple market decentralisation of the optimal direct allocation, making use of nonlinear savings taxes. The market structure works as follows. Individuals enter period t with a given value of net wealth, M_t, which is normalised to remain positive along the equilibrium path. This wealth can either be allocated to period-t consumption, c_t, or savings, s_t. This choice is observable, and the planner implements a non-linear tax on s_t, which may vary over time, and in the history of past savings decisions, s^{t-1}.

[18] 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}.

This tax is denoted $T_t\left(s_t;s^{t-1}\right)$, or simply $T_t\left(s_t\right)$ if context allows, with $T_t'\left(s_t;s^{t-1}\right)$ or $T_t'\left(s_t\right)$ to denote the corresponding marginal savings tax.

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

$$c_t + s_t = M_t \tag{21}$$

$$M_{t+1} = R\left[s_t - T_t\left(s_t;s^{t-1}\right)\right] \tag{22}$$

Given M_0, individuals choose contingent consumption sequences to maximise U_0, subject to (21) and (22), plus a 'no Ponzi' constraint, which must hold along all history paths:

$$\lim_{T\to\infty} R^{-T} M_T \geq 0 \tag{23}$$

Conditions (21) to (23) are easily shown to imply a forward-looking multi-period budget constraint that must be satisfied by all realised consumption-savings paths from generic period t onwards:

$$M_t = \sum_{r=t}^{\infty} R^{t-r}\left[c_r + T_r\left(s_r;s^{r-1}\right)\right] \tag{24}$$

6.2 Normalising taxes

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

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

$$\mathcal{T}_{t+1}\left(s_t,\alpha_t;s^{t-1}\right) = \mathbb{E}_t\left[\sum_{r=t+1}^{\infty} R^{t+1-r} T_r\left(s_r;s^{t-1},s_t,s_{t+1}^{r-1}\right)\middle| s^{t-1},\alpha_t\right] \tag{25}$$

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

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

$$\mathcal{T}_{t+1}\left(s_t^*\left(\alpha_t';s^{t-1}\right),\alpha_t';s^{t-1}\right)=\mathcal{T}_{t+1}\left(s_t^*\left(\bar{\alpha};s^{t-1}\right),\bar{\alpha};s^{t-1}\right) - \int_{\alpha_t'}^{\bar{\alpha}} \left.\frac{\partial\mathcal{T}_{t+1}\left(s_t^*\left(\tilde{\alpha}_t;s^{t-1}\right),\alpha_t;s^{t-1}\right)}{\partial\alpha_t}\right|_{\tilde{\alpha}_t=\alpha_t} d\alpha_t \tag{26}$$

That is, variations in α_t change expected future taxes only insofar as they affect expectations across future type histories. This ensures the impact of an individual's marginal saving decision on public finances can be fully summarised by its impact on contemporaneous tax revenue.

The second normalisation is to set:

$$\mathcal{T}_{t+1}\left(s_t^*\left(\bar{\alpha};s^{t-1}\right),\bar{\alpha};s^{t-1}\right)=0 \tag{27}$$

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

6.3 Feasibility of decentralisation

Proposition 3 provides conditions under which an incentive-feasible allocation can be decentralised by a tax scheme of this kind.

Proposition 3. *An incentive-feasible allocation $\left\{c_t^*\left(\alpha^t\right)\right\}_{t,\alpha^t}$ can be decentralised by a sequence of tax functions $T_t\left(s_t;s^{t-1}\right)$ that satisfy (26) and (27), provided the allocation is normal.*

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

It is easy to amend the proof to allow weaker restrictions.[19] The advantage of normality is that it is an ordinal restriction on the allocation, independent of the type process and utility function. This is consistent with the general focus in this paper on moving beyond analytical reference to these two key unobservables.

7 Sufficient statistics: definitions

A central contribution of the present paper is to provide a direct characterisation of the optimal marginal tax rate, $T_t'(s_t)$, by reference to the conditional savings distribution that arises in the decentralisation, behavioural statistics, and a set of endogenous welfare weights whose definition links closely to the underlying insurance problem. This section defines the relevant statistics for this characterisation.

7.1 The savings distribution

I denote by $\Pi^s\left(s_t|\alpha^{t-1}\right)$ and $\pi^s\left(s_t|\alpha^{t-1}\right)$, respectively, the realised distribution and density functions for savings in the market decentralisation, for a cross section of individuals with common shock history α^{t-1}. These objects are linked to the underlying type distribution, under the assumption of normality:

$$\Pi^s\left(s_t\left(\alpha_t;\alpha^{t-1}\right)|\alpha^{t-1}\right) = 1-\Pi\left(\alpha_t|\alpha_{t-1}\right) \tag{28}$$

$$\pi^s\left(s_t\left(\alpha_t;\alpha^{t-1}\right)|\alpha^{t-1}\right) = -\pi\left(\alpha_t|\alpha_{t-1}\right)\left(\frac{ds_t\left(\alpha_t;\alpha^{t-1}\right)}{d\alpha_t}\right)^{-1} \tag{29}$$

An important related object in the characterisation is the localised Pareto statistic. I denote this $a_t\left(s_t|\alpha^{t-1}\right)$:

$$a_t\left(s_t|\alpha^{t-1}\right) := \frac{s_t\pi^s\left(s_t|\alpha^{t-1}\right)}{1-\Pi^s\left(s_t|\alpha^{t-1}\right)} \tag{30}$$

For all of these functions, the dependence on α^{t-1} is left implicit whenever the context allows.

7.2 Behavioural statistics

Four relevant behavioural statistics are used in the main characterisation:

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

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

The compensated elasticity of lagged savings, with respect to contemporary returns: This elasticity, denoted $\epsilon_{t-1,t}^s(s_t)$, measures the response of savings in $t-1$ to the reprofiling of period-t insurance that is generated by a cut in the marginal savings tax rate in the interval $(s_t, s_t+\Delta)$, taking the limit as Δ becomes small. This is normalised by a measure of the overall monetary value of the tax cut, $\left(1-\Pi^s\left(s_t|\alpha^{t-1}\right)\right)ds_t$, where $\Pi^s\left(s_t|\alpha^{t-1}\right)$ denotes the relevant conditional distribution of savings in t, and ds_t is the differential measure on which the marginal tax cut is applied. This normalisation ensures that differences in $\epsilon_{t-1,t}^s\left(s_t\right)$ at different s_t values are not driven by mechanical differences in the value of the perturbation. With abuse of differential notation, the idea is to set:

$$\epsilon_{t-1,t}^s\left(s_t\right)\left(1-\Pi^s\left(s_t|\alpha^{t-1}\right)\right)ds_t := \frac{\left(1-T_t'\left(s_t\right)\right)}{s_{t-1}}\left.\frac{ds_{t-1}}{d\left(1-T_t'\left(s_t\right)\right)}\right|_{\text{comp}} \tag{31}$$

More formally, $\epsilon_{t-1,t}^s\left(s_t\right)$ is defined implicitly as the linear component at s_t of the general compensated derivative of s_{t-1}, with respect to an arbitrary perturbation to the tax schedule in t. Suppose that the tax paid in t when choosing to hold wealth M_{t+1} is perturbed to:[20]

$$T_t\left(s\left(M_{t+1}\right)\right) - \Gamma f\left(s\left(M_{t+1}\right)\right) \tag{32}$$

where Γ is a scalar, used only to take derivatives, f an arbitrary bounded, differentiable function, and $s\left(M_{t+1}\right)$ the quantity of savings in t corresponding $t+1$ wealth of M_{t+1}, absent the perturbation. This is equivalent to raising the lump-sum benefit component of the tax system by $f\left(\underline{s}\right)$ ($\underline{s}$ being the lowest saving level for the cohort), and raising the marginal rate of return on savings at each M_{t+1} by a *proportional* amount $f'\left(s\left(M_{t+1}\right)\right)$, per unit movement in Γ.[21] The derivative with respect to this class of tax change is linear in the pointwise marginal tax changes:[22]

$$\left.\frac{ds_{t-1}}{d\Gamma}\right|_{\Gamma=0,\text{ comp}} = f\left(\underline{s}\right)H + \int_{s_t} f'\left(s_t\right)h\left(s_t\right)ds_t \tag{33}$$

with the function $h\left(s_t\right)$ and scalar H independent of the choice of $f\left(\cdot\right)$. Consistent with the heuristic definition in (31), I normalise:

$$h\left(s_t\right) := s_{t-1}\epsilon^s_{t-1,t}\left(s_t\right)\left(1-\Pi^s\left(s_t|s^{t-1}\right)\right) \tag{34}$$

where $\Pi^s\left(s_t|s^{t-1}\right)$ is the induced measure of savings in t, given history of savings s^{t-1}. This is the implicit definition of $\epsilon^s_{t-1,t}\left(s_t\right)$.

$\epsilon^s_{t-1,t}\left(s_t\right)$ is a *compensated* elasticity, viewed from the perspective of $t-1$. It is calculated assuming a uniform compensating adjustment to lifetime utility across period-t states, so that the agent choosing the relevant s_{t-1} on the equilibrium path does not experience any change to their lifetime value from doing so. Thus the behavioural change in $t-1$ that it captures is purely due to the re-profiling of state-by-state utility outcomes in t, and not to a first-order change in the utility value of a given quantity of savings.

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

[20]With nonlinear budget constraints, the direction of a shift in the budget constraint matters for the characterisation of income effects. Defining the perturbations at fixed values for M_{t+1} means that $f\left(\cdot\right)$ describes the change in period-t resources available at each choice for M_{t+1}. This simplifies the analysis considerably.

[21]See the proof of Lemma 6.

[22]This claim is established in the proof of Lemma 3.

consider a marginal reduction in $T_t(\underline{s})$, holding constant the profile of marginal tax rates at higher savings. This tax cut will shift consumption possibilities in t, by an equal amount for all types. The marginal effect on *utility* for type α_t will be $\alpha_t u'(c_t(\alpha_t))$, and in general this will vary in α_t. This implies that the change in the lump-sum component will induce a marginal reprofiling of utility across type draws. Once again, I consider the compensated effect of this reprofiling of insurance, given a uniform adjustment to utility across period-t states that leaves indifferent those agents who choose the relevant s_{t-1} on the equilibrium path.

This compensated income effect is denoted $\left.\frac{ds_{t-1}}{dM_t}\right|_{\text{comp}}$. It is defined as the derivative $\left.\frac{ds_{t-1}}{d\Gamma}\right|_{\Gamma=0,\text{ comp}}$ in the case where the perturbation schedule satisfies $f(s(M_{t+1})) = 1$ for all realised M_{t+1} values. It is thus equated to H in (33).

7.3 Welfare weights

In keeping with the static literature, I will make use of 'social welfare weights' to capture the marginal value to the policymaker of providing an extra unit of income to each type, expressed in units of current resources. For each history α^t, these are defined by:[23]

$$g_t\left(\alpha^t\right) := \alpha_t u'\left(c_t\left(\alpha^t\right)\right)\frac{1+\lambda_t\left(\alpha^{t-1}\right)}{\eta_t} \tag{35}$$

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.

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: $\mathbb{E}\left[\lambda_{t+1}|\alpha^{t-1}\right] = \lambda_t$.[24] In the decentralised allocation, this updating process is exactly capturing 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.[25]

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

[24]See Proposition 2.

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

The link between Pareto weights and wealth in a decentralised market economy has been understood at least since Negishi (1960). The interesting feature of the present context is the fact that the evolution of the wealth distribution over time can be identified with changes in the effective welfare objective of the policymaker. A policymaker in the initial period may seek a radical utilitarian allocation, unconstrained by any initial profile of asset ownership. But as time progresses, respect for the evolving pattern of wealth is a necessary counterpart to respect for past incentive constraints. An optimal plan – from the perspective of the initial period – remains cross-sectionally utilitarian, but only for subsets of individuals who share a common history. Across subgroups, substantial differentiation in treatment will emerge.

The time inconsistency here is evident, and provides a challenge to the plausibility of the commitment assumption. What if a new government would like to redistribute wealth? The present setup sidesteps this question with its commitment assumption, but it is of first-order importance.

8 Sufficient statistics characterisation

8.1 Characterisation

In Appendix B, I prove the following, central characterisation result:

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

$$\mathbb{E}\left[1 - T_0'(s_0)\frac{ds_0}{dM_0} - g_0(s_0)\middle| s_0 \geq s_0'\right] = T_0'(s_0')\,\varepsilon_0^s a_0(s_0') \tag{36}$$

and:

$$\mathbb{E}\left[g_0(s_0)\right] + \mathbb{E}\left[T_0'(s_0)\frac{ds_0}{dM_0}\right] = 1 \tag{37}$$

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

$$\mathbb{E}_{t-1}\left[1 - T_t'(s_t)\frac{ds_t}{dM_t} - g_t(s_t)\middle| s_t \geq s_t'\right] = T_t'(s_t')\,\varepsilon_t^s a_t(s_t') + RT_{t-1}'(s_{t-1})\,s_{t-1}\epsilon_{t-1,t}^s(s_t') \tag{38}$$

and:

$$\mathbb{E}_{t-1}\left[g_t(s_t)\right] + \mathbb{E}_{t-1}\left[T_t'(s_t)\frac{ds_t}{dM_t}\right] + RT_{t-1}'(s_{t-1})\left.\frac{ds_{t-1}}{dM_t}\right|_{comp} = 1 \tag{39}$$

8.2 Discussion

As previewed, equations (36) to (39) can be understood intuitively by reference to marginal changes in the intertemporal budget constraint that links consumption in one period to income in the next.

Condition (36) assesses the effects of a marginal tax cut at s_0', applied in the initial time period. Heuristically, the effects of this can be divided into those *above* s_0', and those *at* s_0'. For those above s_0', the tax cut serves to relax the within-period budget constraint by a uniform amount, and the left-hand side of (36) accounts for this from the policymaker's perspective. There are three components: (1) the direct cost of the transfer, normalised to 1 per agent by construction, minus (2) the additional tax revenue that is received on whatever fraction of the additional income is saved, $T_0'(s_0)\frac{ds_0}{dM_0}$, minus (3) the social welfare value of providing an additional consumption unit, $g_0(s_0)$.

The right-hand side of (36) 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 generates higher tax revenue. This is captured by the object $T_0'(s_0')\, s_0' \varepsilon_0^s$. The Pareto parameter is a weighting term, capturing the measure of individuals affected by substitution effects, relative to the number who experience income effects.

Condition (38) is the equivalent to (36) for $t > 0$, and features an extra term that allows for the impact that changes to tax schedules in t have on savings in $t-1$.[26] Clearly this term depends critically on the sign and magnitude of the (compensated) cross-elasticity $\epsilon_{t-1,t}(s_t')$. This will be discussed in detail in Section 10.

Conditions (37) and (39) describe the consequences of changing the lump-sum component of the social insurance system. They are essentially variants of (36) and (38) respectively, absent contemporaneous substitution effects – and when the relevant intertemporal behavioural effect in (39) is $\left.\frac{ds_{t-1}}{dM_t}\right|_{\text{comp}}$ rather than $\epsilon_{t-1,t}(s_t')$.

The regularity conditions that Theorem 1 imposes on the consumption-type link merit some comment. The restriction to normal allocations is sufficient both to ensure the validity of the first-order approach (Corollary 1), and the validity of the consumption-savings decentralisation (Proposition 3). Restricting to *strictly* normal allocations additionally rules out the possibility of multiple types bunching at a certain point. This would imply atoms in the savings distribution at the corresponding values. These are possible

[26] In the language of Saez, Slemrod and Giertz (2012), taxes in t exert a 'fiscal externality' on the $t-1$ tax base.

to handle theoretically, but it is cumbersome to do so. Similarly, continuity in $c_t(\alpha_t)$ can be relaxed. When it is, optimal policy must additionally consider the effects of discontinuous jumps in consumption as the tax schedule shifts. Again, the additional insight is not first order, and the possibility is left as an extension.

8.3 Intuition behind the proof

The proof of Theorem 1 is algebraically involved, but it contains at its core a simple economic idea, founded on well-known envelope logic.

Proposition 2 characterised optimal allocations for the mechanism design problem, analysing the costs and benefits of engineering, at the margin, a step change to the cross-sectional profile of *utilities* within a given cohort, by changing information rents at a certain point. Sufficient statistics characterisations instead rely on perturbations to the cross-sectional profile of *wealth*. The proof works by conjecturing and confirming a link from wealth changes to utility changes.

Tax cuts provide a uniform increase in real resources for all individuals whose α_t value is below that of the critical individual, α_t', whose marginal tax rate is cut. Invoking a classic envelope argument, I conjecture that this should imply an increase in *utility* for these individuals that is proportional to their within-period marginal utility, $\alpha_t u'(c_t(\alpha_t))$.

Treating the expressions in Proposition 2 as basis functions, I already have the machinery to express the costs and benefits of raising utility by $\alpha_t u'(c_t(\alpha_t))$ units, across a subset of types. This just requires mapping from step changes in utility – described in the Proposition – to a more complex profile of changes.

This done, the key second step in the proof is to link the resulting expressions to behavioural statistics from the market decentralisation. This is a non-trivial exercise in reverse engineering, and makes heavy use of links, which I prove, between the derivative of allocations with respect to type, and the behavioural effects of changes in prices and insurance. For this it is crucial that types lie on a continuum, and that preferences are separable over time.

9 Properties of optimal taxes

9.1 Positive marginal rates at interior points

The characterisation in Theorem 1 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.*

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

9.1.1 Sketch of proof

Like Theorem 1, Theorem 2 relies on links between the mechanism design representation of the problem, and the market decentralisation. I first show that the marginal tax rate in t for type α_t takes the same sign as $\lambda_{t+1}^{\Delta}(\alpha_t)$ – the multiplier on the NPV information rent constraint, (13), for $t+1$. Using earlier results, I show this satisfies the equation:

$$\lambda_{t+1}^{\Delta}(\alpha_t)\,\alpha_t\pi(\alpha_t|\alpha_{t-1}) = \int_{\alpha_t}^{\bar{\alpha}} \left[\left(1+\lambda_t+\lambda_t^{\Delta}\pi^{\Delta}(\tilde{\alpha}_t|\alpha_{t-1})\right) - (1+\lambda_{t+1}(\tilde{\alpha}_t))\right]\pi(\tilde{\alpha}_t|\alpha_{t-1})\,d\tilde{\alpha}_t \qquad (40)$$

where the integral on the right-hand side is zero over the full range – i.e. when $\alpha_t = \underline{\alpha}$. Given Assumption 4, it follows that a sufficient condition for $\lambda_{t+1}^{\Delta}(\alpha_t)$ to be positive is that $(1+\lambda_{t+1}(\alpha_t))$ – the multiplier on the promise value constraint (12) – is monotone decreasing in α_t.

The central part of the proof is devoted to establishing this. In a stand-alone result, Proposition 6 in the Appendix, I show that $(1+\lambda_{t+1}(\alpha_t))$ is a measure of the marginal cost of providing utility to an agent in the long run, satisfying the relationship:

$$\frac{1+\lambda_{t+1}(\alpha_t)}{\eta_t} = \lim_{s\to\infty}\left\{\frac{1}{\mathbb{E}_t[\alpha_s]}\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}\,u'(c_s(\alpha^s))}\right]\right\} \qquad (41)$$

I then show that the object on the right-hand side must be decreasing in α_t under the assumptions of the problem.

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 reach zero. As in the static labour income tax literature, a critical role is played by the limiting properties of the distribution and density of savings. 'Zero distortion' results generally follow if the density remains positive at endpoints, or converges to zero slowly by comparison with the remaining probability mass in the upper tail. When this is true, the local efficiency costs of taxation become large relative to any insurance benefit. It is possible to place restrictions on the model primitives that guarantee zero top marginal tax rates, as the following result illustrates:

Proposition 4. *Suppose that the type density function satisfies $\pi\left(\underline{\alpha}|\alpha_{t-1}\right) > 0$ (respectively $\pi\left(\bar{\alpha}|\alpha_{t-1}\right) > 0$) for some α_{t-1}. Then the optimal marginal tax rate on savings is zero at the top (bottom) of the conditional savings distribution for individuals whose lagged type was α_{t-1}.*

But when $\pi\left(\underline{\alpha}|\alpha_{t-1}\right) = 0$, the result is no longer guaranteed. As a direct corollary of Theorem 1, I can write an expression for the optimal marginal tax rate at the top of the savings distribution in this case:

Corollary 2. *Given α^{t-1}, the optimal marginal tax rate at the upper limit of the savings distribution, $\bar{s}$, satisfies:*

$$T_t'\left(\bar{s}\right) = \frac{1 - g_t\left(\bar{s}\right) - RT_{t-1}'\left(s_{t-1}\right) s_{t-1} \epsilon_{t-1,t}^{s}\left(\bar{s}\right)}{\left.\frac{ds_t}{dM_t}\right|_{\bar{s}} + \varepsilon_t^s a_t\left(\bar{s}\right)} \tag{42}$$

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

9.3 Optimal transfers

Conditions (37) and (39) 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 fiscal 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 beyond the level where the average welfare weight is unity – the usual benchmark with quasilinear preferences – because the net cost of the transfer is mitigated by tax revenue on the additional savings it induces.

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

$$\mathbb{E}_{t-1}\left[g_t\left(s_t\right)\right] < \mathbb{E}_{t-1}\left[1 - T'\left(s_t\right)\frac{ds_t}{dM_t}\right] \tag{43}$$

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

10 Intertemporal elasticities and type persistence

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. The only statistic preventing the within-period characterisation from being isomorphic to a textbook Saez formula is the elasticity of lagged savings, $\epsilon^{s}_{t-1,t}\left(s'_t\right)$. The purpose of this Section is to discuss its role, and relate it to more familiar intuition.

10.1 Insurance incentivises prior saving

So long as marginal savings taxes are positive in period $t-1$, there are social benefits to inducing more savings in $t-1$, and – by a standard second-best argument – the policymaker should be willing to incur some marginal costs along other dimensions to achieve this. In particular, the tax system should admit distortions that have indirect incentive effects on saving in $t-1$.

When types are persistent and Markovian, this can be effected through changes to the degree of cross-sectional insurance that is provided in period t. Given the link between type and behaviour, those with lower

s_{t-1} have higher values for α_{t-1} – and thus place more subjective weight on high type draws in t. Consider a change to period-t allocations that improves relative outcomes for high α_t draws, for each level of $t-1$ savings. Assume that this is 'compensated', so that it has no effect on expected utility in $t-1$ for those who keep their savings unchanged. But because of type persistence, those who start out saving slightly less would *not* view this same reprofiling with indifference. From their perspective, it improves the attractiveness of the savings package consumed by nearby types, and raises their willingness to save.

The elasticity $\epsilon^s_{t-1,t}\left(s'_t\right)$ in equation (38) captures this effect. It motivates an additional distortion to outcomes from t onwards, relative to an optimal plan from the perspective of t alone. Note that this distortion implies a second, more prosaic source of time inconsistency in the setting, distinct from the more fundamental challenge of a widening wealth distribution. A policymaker re-optimising in t would have no incentive to consider the effect of their choices on savings in $t-1$, which would by now already be determined.

Atkinson and Stiglitz (1976) showed that when consumption goods were independent of labour supply, there were no gains to differential consumption taxation. The counterpart to this result in our setting is provided by the case of iid types. In this case, any change to the profile of insurance at t will be viewed identically by all types in $t-1$, since they all share a common distribution over future outcomes. Thus a compensated change to insurance in t will not affect saving in $t-1$: $\epsilon^s_{t-1,t}\left(s_t\right) \equiv 0$, and only contemporaneous elasticities matter for optimal taxes.

More generally, this line of reasoning suggests that the Markov property of shocks is crucial to ensuring just *two* elasticities feature in (38). Markovian shocks imply that the preferences of two distinct α_{t-1} types across alternative allocations from $t+1$ on are identical, conditional on drawing a particular α_t. Compensated distortions to $t+1$ allocations – which, by definition, hold constant expected lifetime utility for each period-t type draw – cannot influence insurance, or savings, in $t-1$ or earlier.

10.2 Intertemporal elasticities: a source of progressivity

The upshot is that it may be desirable to provide 'excessive' insurance in t, beyond what is contemporaneously optimal, because this helps to mitigate the costs of the tax distortion in $t-1$. Additional insurance is often associated with additional progressivity in the tax system. This can be formalised as follows:

Proposition 5. *When types are persistent and $T'_{t-1}\left(s_{t-1}\right) > 0$, there exists a threshold $\tilde{s}_t$ such that $T'_{t-1}\left(s_{t-1}\right) s_{t-1}\epsilon^s_{t-1,t}\left(s'_t\right)$ is positive for $s'_t < \tilde{s}_t$, and negative for $s'_t > \tilde{s}_t$.*

The implications of this Proposition can be seen by comparing policies that satisfy condition (38) with those that neglect intertemporal cross-elasticities – as would be optimal for a policymaker re-optimising in t. At the optimum, I have:

$$\mathbb{E}_{t-1}\left[1 - T_t'(s_t)\frac{ds_t}{dM_t} - g_t(s_t)\middle| s_t \geq s_t'\right] > T_t'(s_t')\, s_t' \varepsilon_t^s a_t(s_t') \tag{44}$$

for all s_t' below a threshold and

$$\mathbb{E}_{t-1}\left[1 - T_t'(s_t)\frac{ds_t}{dM_t} - g_t(s_t)\middle| s_t \geq s_t'\right] < T_t'(s_t')\, s_t' \varepsilon_t^s a_t(s_t') \tag{45}$$

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'. At least locally, therefore, the re-optimising policymaker would prefer to raise marginal tax rates at low s_t', and cut them at high s_t'. The optimum has 'too much' progressivity, viewed *ex-post*.

Why does the elasticity change sign in the manner described in Proposition 5? A tax cut at s_t' has two conflicting consequences for insurance. The first is to redistribute resources collectively to a group of relatively high savers. Since high savers have a low marginal consumption utility, this worsens the insurance value of savings, reducing $\epsilon_{t-1,t}^s(s_t')$. But the second effect is to improve the nature of insurance *within* the group that benefits. Uniform income provision, which a tax cut implies, generates non-uniform effects on *utility* within the sub-group. Lower savers have higher marginal utility, and so benefit more. This force improves the overall insurance profile of savings, raising the value of $\epsilon_{t-1,t}^s(s_t')$. By similar logic, the compensated income effect, $\left.\frac{ds_{t-1}}{dM_t}\right|_{\text{comp}}$, is positive.

When taxes are cut at a low level of savings, the within-group insurance gains are large relative to cross-group effects, as the within-group dispersion of marginal utilities is large. As the threshold s_t' increases, cross-group effects come to be relatively more significant, worsening the insurance properties of the change and ensuring a lower value for $\epsilon_{t-1,t}^s(s_t')$. This is the reason for the single crossing property established in Proposition 5.

11 An indicative quantification

To give the results a more concrete character, in this penultimate section I provide an indicative quantification of the optimal top marginal savings tax, based on the formula (42).

11.1 Decomposition with no intertemporal effects

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

$$T'_{t}\left(\bar{s}\right)=\frac{1-g_{t}\left(\bar{s}\right)}{\left.\frac{ds_{t}}{dM_{t}}\right|_{\bar{s}}+\varepsilon^{s}_{t}a_{t}\left(\bar{s}\right)} \tag{46}$$

If consumption approaches zero at the upper limit for savings, an empirical value of $\left.\frac{ds_{t}}{dM_{t}}\right|_{\bar{s}}$ equal to 1 is an obvious approximation, corresponding to homotheticity in demand. Since, in the model, savings and consumption covary perfectly conditional on within-period wealth, the value of $a_{t}\left(\bar{s}\right)$ can be informed either by data on the upper tail of the savings distribution or on the lower tail of the consumption distribution. Saez and Stantcheva (2018) fit a Pareto parameter of 1.38 to administrative data on capital *income* for the US, though the well-known correlation between wealth and returns may bias this downwards relative to the parameter for savings alone. For consumption, Toda and Walsh (2015) suggest a value of approximately 4 for the lower Pareto parameter, based on cross-sectional US data. The main issue in applying either of these statistics directly to our setup is that $a_{t}\left(\bar{s}\right)$ is specified as a *conditional* parameter, taken across individuals with common initial wealth. The tail parameter for the population as a whole will be influenced by heterogeneity in initial wealth conditions.

The compensated intertemporal elasticity of consumption, which equals ε^{s}_{t}, is strictly less than the Frisch – for which a value of around 0.5 is standard.[27] Since the difference between the two becomes small as the share of current consumption in marginal expenditure becomes small, I assume this value for the conditional upper tail of savers.

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

Taken together, and leaving $a_t(\bar{s})$ subject to uncertainty, these assumptions would imply a value for the top marginal tax rate on total savings equal to $\frac{1-g_t(\bar{s})}{1+0.5a_t(\bar{s})}$, where $g_t(\bar{s})$ is the value of the social welfare weight at the highest savings level.

Some care is needed quantifying the social welfare weight. Unlike in static Mirrleesian environments, it does not make sense to assume that it approaches zero for the least-favoured types. To see why, recall that $g_t(s_t)$ is directly proportional to the marginal utility of consumption $\alpha_t u'(c_t)$, which is optimally set equal to the marginal value of savings. Even individuals whose draw for α_t is arbitrarily close to zero will have returns to saving that are bounded above zero, because future marginal utility is not likely to remain so low. Put differently: individuals whose present consumption need is exceptionally low may have much greater need in the future. The marginal social welfare weight remains high because of this.

To make progress, recall that equation (39) shows that the cross-sectional average value for g_t, $\bar{g}_t$, is 1 minus the value of tax recovered through income effects at the margin, both in t and $t-1$. Let the latter value be equal to some number $\chi_t \in (0,1)$:

$$\chi_t := \mathbb{E}_{t-1}\left[T_t'(s_t)\frac{ds_t}{dM_t}\right] + RT_{t-1}'(s_{t-1})\frac{ds_{t-1}}{dM_t}\bigg|_{\text{comp}} \tag{47}$$

Simple algebra gives:

$$1-g_t(\bar{s}) = \chi_t + (1-\chi_t)\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) \tag{48}$$

$$= \chi_t + (1-\chi_t)\left(\frac{\mathbb{E}_{t-1}\left[\alpha_t u'(c_t)\right] - \underline{\alpha}u'\left(c_t\left(\underline{\alpha}\right)\right)}{\mathbb{E}_{t-1}\left[\alpha_t u'(c_t)\right]}\right) \tag{49}$$

This varies in the size of the average income effect on tax revenue, χ_t, and the proportional deviation of the lowest type's marginal utility from the average. An individual with type draw $\underline{\alpha}$ in t is assumed to have minimal consumption needs in period t, and can devote all resources to saving. The formula requires an estimate for the impact that such a high level of saving has on their marginal utility of wealth, relative to the average. With iid shocks, time-separable utility and an EIS of 0.5, the proportional difference in marginal utility is approximately twice the proportional difference in lifetime consumption. Thus if average current (annual) consumption is valued generously at ten per cent of average wealth, the term $\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right)$ will be approximately 0.2.[28]

[28]Persistence in α provides a direct, offsetting reason why lower period-t types to value a unit of savings less than the average

		$a_t = 2$	$a_t = 4$	$a_t = 6$	$a_t = 10$
$\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) = 0.05$	$\chi_t = 0.01$	0.030	0.020	0.015	0.010
	$\chi_t = 0.05$	0.049	0.033	0.024	0.016
	$\chi_t = 0.1$	0.073	0.048	0.036	0.024
$\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) = 0.1$	$\chi_t = 0.01$	0.055	0.036	0.027	0.018
	$\chi_t = 0.05$	0.073	0.048	0.036	0.024
	$\chi_t = 0.1$	0.095	0.063	0.048	0.032
$\left(\frac{\bar{g}_t - g_t(\bar{s})}{\bar{g}_t}\right) = 0.2$	$\chi_t = 0.01$	0.104	0.069	0.052	0.035
	$\chi_t = 0.05$	0.12	0.08	0.06	0.04
	$\chi_t = 0.1$	0.14	0.093	0.07	0.047

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

The value of χ_t, the average income effect on tax revenue, is difficult to fix in a partial quantification of this kind, because it depends on the chosen marginal tax rate across the entire range of savings. The general lesson of this section is that a top marginal tax rate as high as 10 per cent of gross savings is unlikely. Since the qualitative factors that matter elsewhere in the distribution tend to push towards a lower marginal rate on lower savings levels,[29] and since the proportion of additional income that is saved will generally be below 100 per cent, 0.1 is a generous upper bound for χ_t.

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

The main quantitative lesson is that the top marginal tax rate on savings could perhaps exceed 5 per cent, but only with some favourable assumptions, particularly on the value of the conditional Pareto parameter and the (endogenous) impact of high savings on lifetime marginal utility. More likely is a number of the order of 1 to 5 per cent. Yet this is still very substantial as a marginal levy on gross savings. It could easily wipe out any net income on marginal savings.

Clearly the numbers depend critically on the qualitative tail properties of the cross-sectional savings dis-

type, but since insurance is incomplete this effect will be limited: the marginal value of saving, even to a low current type, will be dominated by the risk of high type draws in the future.

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

tribution. If – as conjectured by Toda and Walsh (2015) – Pareto tails for the overall savings and consumption distributions only emerge as a consequence of demographic evolution, and are absent *within* cohorts, then $a_t(\bar{s}) = \infty$ would hold, and we would return to zero distortion at the top in the iid case.

11.2 Adding intertemporal effects

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

$$-RT'_{t-1}(s_{t-1})\frac{s_{t-1}\epsilon^s_{t-1,t}(\bar{s})}{1+\varepsilon^s_t a_t(\bar{s})} \tag{50}$$

The final fraction here is likely to be in the neighbourhood of 1. Clearly it is less than R, more significantly so the lower the Pareto parameter. The middle ratio is more significant: the ratio of $s_{t-1}\epsilon^s_{t-1,t}(\bar{s})$ to $\bar{s}\varepsilon^s_t$. Since $\bar{s}$ is the highest level of savings in t for those who saved s_{t-1} in the previous period, $\frac{s_{t-1}}{\bar{s}}$ is likely to be small. The compensated intertemporal savings response to a cut in the upper marginal tax rate is not a commonly estimated object, and there is little to go on except priors. It seems very likely that $\left|\epsilon^s_{t-1,t}(\bar{s})\right| < \varepsilon^s_t$. Over all, an upper bound for the term in (50) equal to $0.5 \cdot T'_{t-1}(s_{t-1})$ seems quite generous. If lagged marginal tax rates are of the same order of magnitude as current, intertemporal considerations could conceivably raise top rates by at most a half of their value in Table 1.

11.3 Literature comparison

It is instructive to compare equation (46) with similar 'sufficient statistics' conditions in the literature. To date, these have been used to analyse optimal savings taxation under two alternative simplifying assumptions: either that outcomes are studied in steady state, or that horizons are limited to two periods.

11.3.1 Saez and Stantcheva (2018)

A leading example of the first approach is Saez and Stantcheva (2018), whose benchmark optimal savings tax formula can be rewritten as:

$$\frac{\tau_K^{top}}{1-\tau_K^{top}} = \frac{1-g_K^{top}}{e_K^{top}a_K^{top}} \tag{51}$$

where τ_K^{top} is the top marginal tax rate on capital income, g_K^{top} the welfare weight placed on the highest savers, e_K^{top} the elasticity of capital income with respect to the top tax rate, and a_K^{top} the Pareto parameter for capital holdings.

Notice that the right-hand side here is identical in form to the right-hand side of (46), except that there are no income effects – which follows from a direct assumption of linear utility in Saez and Stantcheva. On the left-hand side, the $\left(1 - \tau_K^{top}\right)$ in the denominator is a conversion factor, mapping from a substitution response in the taxed good to revenue denominated in a numeraire. Its absence from (46) just reflects the fact that savings have a fixed relative price of 1 in terms of the numeraire good, contemporaneous consumption. Otherwise the formulae seem near-identical.

But the cosmetic similarities mask important differences in policy focus and horizon. Saez and Stantcheva consider a single, universal tax schedule applied to capital income in perpetuity. The purpose of policy in their paper is to engineer the most desirable feasible stationary distribution of savings. The role of taxes in the present paper is, instead, to provide insurance against short-term shocks as they arrive. Wealth differentials open up – and grow – as a part of this.

Thus contrary to the benchmark case in Saez and Stantcheva, the welfare weights for high savers, $g_t(\bar{s})$, will not be zero in (46). Indeed, the weights will not even be monotone in wealth across the population as a whole. Whereas Saez and Stantcheva assign a lower welfare weight to individuals *because* their wealth is high, the planner in the present paper gives more wealth to individuals *so as* to raise their implicit welfare weight. With incomplete markets, wealth divergences are an efficient – though time-inconsistent – device for managing shocks.

The difference in time perspective also drives differences in the definitions of the observable statistics. e_K^{top} in Saez and Stantcheva is a long-term elasticity of the aggregate stock of capital, with respect to changes in the real interest rate. It could be affected by aggregate budgetary and general equilibrium considerations, and the precise horizon over which it is assessed. The elasticity ε_t^s in (46) is a direct behavioural object – defined by reference to a price-taking consumer's choice. Nonetheless, Saez and Stantcheva choose a benchmark value for e_K^{top} of 0.5, identical for the value used for ε_t^s in Table 1.

Finally, the relevant Pareto parameter in (46) is the *conditional* parameter, controlling for initial wealth levels. This differs from the unconditional, population parameter for capital income, α_K^{top}, used in Saez and Stantcheva, whose value is taken from administrative data to be just 1.38.

On the surface, thinner Pareto tails and non-zero limiting values for the welfare weights ought to push the present paper in the direction of lower savings taxation, relative to Saez and Stantcheva. What is surprising is that the range of values in in Table 1 is quite consistent with their proposal to tax capital *income* at around 60 per cent at the top. For a 4 per cent untaxed real interest rate, this is equivalent to taxing savings at around 2.3 per cent. This apparent inconsistency is accounted for by the mechanical consequences of fixing a broader tax base: gross savings rather than net capital income. When all of gross savings can be taxed, even a numerically small optimal rate can be economically significant – and values for sufficient statistics that might appear conservative can imply relatively high tax rates on capital income.

11.3.2 Scheuer and Slemrod (2021)

An alternative way to think about savings taxation is by reference to two-period models. Here the main benchmark in the literature remains Atkinson and Stiglitz (1976), and its implication that savings taxation should be zero in a setting where (a) all cross-sectional heterogeneity derives from labour productivity differences in the first time period, and (b) utility is separable between consumption and leisure when production takes place.

Scheuer and Slemrod (2021) consider a variant on this approach, deriving a sufficient statistics formula for an environment in which there is heterogeneity in initial wealth holdings, (perfectly) correlated with labour productivity, and initial wealth taxes are not possible. Outcomes depart from the Atkinson-Stiglitz theorem because a high-productivity agent has more wealth than – and so saves more than – a low-productivity agent, given the same post-tax *labour* income. This implies a classic justification for taxing wealth as a behavioural complement to leisure.

As in the present paper, the main trade-off that emerges is between reducing cross-sectional inequality and minimising distortions to the intertemporal allocation of consumption. Unlike the present paper, the policymaker has a choice of instruments. Both labour and wealth tax instruments are available, and the wealth tax rate will generally be higher when the tail of the wealth distribution is fatter than the labour income distribution, and/or when the intertemporal elasticity of substitution in consumption (σ in their notation) is small relative to the labour supply elasticity (ε). The equation that they calibrate is a statement of this *relative* trade-off:

$$t_k = \frac{t_y}{1-t_y}\alpha\left[\frac{\sigma\rho_k}{\varepsilon\rho_y}+1-\alpha\right]^{-1}$$

where ρ_k and ρ_y are the Pareto parameters for wealth and labour income respectively, α is a measure of the wealth-to-consumption ratio, and t_y is the labour income tax.

Taking conventional calibrations for the elasticities and Pareto parameters, an upper bound for α, and a labour income tax rate of 50 per cent, the authors obtain an upper bound for the annual wealth tax of 2.7 per cent – only marginally below the assumed pre-tax real interest rate of 3 per cent. Again, it is notable that this fits well within the range of values obtained in Table 1.

12 Conclusion

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

The second main message 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 simplify behavioural responses in a way that keeps them tractable. In a multi-period world, tax design must inevitably make some simplifying assumptions, to avoid being overwhelmed by the multitude of possible cross-period behavioural responses. One option, pursued in the literature already, is to focus exclusively on steady-state outcomes. Though defensible, this is a significant departure from conventional approaches, both positive and normative. This paper suggests that mechanism design offers a theory-guided alternative route.

References

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

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

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

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

[5] Attanasio, O.P., and G. Weber (2010), 'Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy' *Journal of Economic Literature*, 48 (3): 693–751.

[6] Bloedel, A.W., R.V. Krishna and O. Leukhina (2021), 'Insurance and Inequality with Persistent Private Information', Working Paper, Stanford University and Federal Reserve Bank of St. Louis.

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

[8] 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] Ferey, A., B. Lockwood and D. Taubinsky (2021), 'Sufficient Statistics for Nonlinear Tax Systems with Preference Heterogeneity', NBER Working Paper 29582

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

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

[20] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[35] Pavan, A. (2017), 'Dynamic Mechanism Design: Robustness and Endogenous Types', in *Advances in Economics and Econometrics: Proceeding of the 2015 World Congress of the Econometric Society*.

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

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

[38] Piketty, T., and E. Saez (2013b), 'Optimal Labour Income Taxation', Ch 7 in A.J. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.), *Handbook of Public Economics*, Vol 5: 391–474.

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

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

[41] Saez, E., J. Slemrod and S.H. Giertz (2012), 'The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review', *Journal of Economic Literature*, 50(1): 3–50.

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

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

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

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

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

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

[48] 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.

[49] Toda, A.A., and K. Walsh (2015), 'The Double Power Law in Consumption and Implications for Testing Euler Equations', *Journal of Political Economy*, 123(5): 1177–1200.

Online Appendix

A Proofs from Sections 4 to 6

A.1 Proof of Lemma 1

This result is an application of Theorem 2 in Milgrom and Segal (2002), plus elemenatary manipulations. First, note that the utility of type α_t' from type report α_t'' in period t can be written in the form:

$$\alpha_t' u\left(c_t\left(\alpha^{t-1},\alpha_t''\right)\right)+\beta\int_{\alpha_{t+1}} V\left(\alpha^{t-1},\alpha_t'',\alpha_{t+1}\right)d\Pi\left(\alpha_{t+1}|\alpha_t'\right) \tag{52}$$

The boundedness of lifetime utility (constraint (9)) and the differentiability of the conditional density $\pi\left(\alpha_{t+1}|\alpha_t'\right)$ in α_t' (Assumption 2) together imply 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:

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

Constraint (9) implies $u\left(c_t\left(\alpha^{t-1},\alpha_t''\right)\right)$ and $V\left(\alpha^{t-1},\alpha_t'',\alpha_{t+1}\right)$ are bounded for all α_t'' and α_{t+1}. $\frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t}$ is bounded by Assumption 4. Together this that the object in (53) is bounded, uniformly across type reports α_t''. When the allocation satisfies (11), 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:

$$\begin{aligned}
&\alpha_t' u\left(c_t\left(\alpha^{t-1},\alpha_t'\right)\right)+\beta\int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1},\alpha_t',\alpha_{t+1}\right)d\Pi\left(\alpha_{t+1}|\alpha_t'\right) \\
=&\underline{\alpha} u\left(c_t\left(\alpha^{t-1},\underline{\alpha}\right)\right)+\beta\int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1},\underline{\alpha},\alpha_{t+1}\right)d\Pi\left(\alpha_{t+1}|\underline{\alpha}\right) \\
&+\int_{\underline{\alpha}}^{\alpha_t'}\frac{1}{\alpha_t}\left\{\alpha_t u\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right)+\beta\alpha_t\int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1},\alpha_t,\alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\right\}d\alpha_t
\end{aligned} \tag{54}$$

To obtain the representation in the main text, I then use the following:

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

$$\beta\alpha_t\int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1},\alpha_t',\alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}=\mathbb{E}_t\left[\sum_{s=t+1}^{\infty}\beta^{s-t}D_{t,s}\left(\alpha^s\right)\alpha_s u\left(c_s\left(\alpha^{t-1},\alpha_t',...,\alpha_s\right)\right)\middle|\alpha_t\right] \tag{55}$$

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:

$$\beta\alpha_t \int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1}, \alpha'_t, \alpha_{t+1}\right) \frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t} d\alpha_{t+1} \tag{56}$$
$$= \int_{\alpha_{t+1}} \beta\rho\left(\alpha_{t+1}|\alpha_t\right)\left[\alpha_{t+1}u_{t+1}\left(\alpha^{t-1}, \alpha'_t, \alpha_{t+1}\right) + \beta\alpha_{t+1}\int_{\alpha_{t+2}} V_{t+2}\left(\alpha^{t-1}, \alpha'_t, \alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+2}|\alpha_{t+1}\right)}{d\alpha_{t+1}} d\alpha_{t+2}\right] d\Pi\left(\alpha_{t+1}|\alpha_t\right)$$

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, noting that $\rho\left(\alpha_{t+1}|\alpha_t\right) \in (0,1)$ (Assumption 3), and using the boundedness of value in t, the result follows.

□

Using the definition of $\omega^{\Delta}_{t+1}\left(\alpha^{t-1}, \alpha_t\right)$, setting $\alpha'_t = \alpha_t$ in (55) gives:

$$\alpha_t \int_{\alpha_{t+1}} V_{t+1}\left(\alpha^{t-1}, \alpha_t, \alpha_{t+1}\right) \frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t} d\alpha_{t+1} = \omega^{\Delta}_{t+1}\left(\alpha^{t-1}, \alpha_t\right) \tag{57}$$

Using this and the definition of $\omega_{t+1}\left(\alpha^{t-1}, \alpha_t\right)$, (54) collapses to (14).

A.2 Proof of Proposition 1

'If': Suppose that global incentive compatibility fails. Then for some t, α^{t-1} there must exist α'_t, α''_t such that:

$$u_t\left(\alpha''_t\right) + \frac{\beta}{\alpha'_t}\int_{\alpha_{t+1}} V_{t+1}\left(\alpha''_t, \alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha'_t\right) d\alpha_{t+1} > u_t\left(\alpha'_t\right) + \frac{\beta}{\alpha'_t}\int_{\alpha_{t+1}} V_{t+1}\left(\alpha'_t, \alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha'_t\right) d\alpha_{t+1} \tag{58}$$

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

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

$$=\int_{\alpha_t'}^{\alpha_t''}\frac{d}{d\alpha_t}\left\{\frac{1}{\alpha_t}\left[\alpha_t u_t\left(\alpha_t\right)+\beta\int_{\alpha_{t+1}}V_{t+1}\left(\alpha_t,\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\right]\right\}d\alpha_t \tag{59}$$

$$=\int_{\alpha_t'}^{\alpha_t''}\left\{-\frac{1}{\alpha_t^2}\left[\alpha_t u_t\left(\alpha_t\right)+\beta\omega_{t+1}\left(\alpha^{t-1},\alpha_t\right)\right]+\frac{1}{\alpha_t^2}\left[\alpha_t u_t\left(\alpha_t\right)+\beta\omega_{t+1}^{\Delta}\left(\alpha^{t-1},\alpha_t\right)\right]\right\}d\alpha_t \tag{60}$$

$$=-\beta\int_{\alpha_t'}^{\alpha_t''}\frac{1}{\alpha_t^2}\left(\omega_{t+1}\left(\alpha^{t-1},\alpha_t\right)-\omega_{t+1}^{\Delta}\left(\alpha^{t-1},\alpha_t\right)\right)d\alpha_t \tag{61}$$

where the penultimate line has made use of (14).

Applying this result in (58) yields:

$$u_t\left(\alpha_t''\right)+\frac{\beta}{\alpha_t'}\int_{\alpha_{t+1}}V_{t+1}\left(\alpha_t'',\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t'\right)d\alpha_{t+1}>u_t\left(\alpha_t''\right)+\frac{\beta}{\alpha_t''}\int_{\alpha_{t+1}}V_{t+1}\left(\alpha_t'',\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t''\right)d\alpha_{t+1}$$
$$+\beta\int_{\alpha_t'}^{\alpha_t''}\frac{1}{\alpha_t^2}\left[\omega_{t+1}\left(\alpha^{t-1},\alpha_t\right)+\omega_{t+1}^{\Delta}\left(\alpha^{t-1},\alpha_t\right)\right]d\alpha_t \tag{62}$$

Or:

$$\beta\int_{\alpha_t'}^{\alpha_t''}\frac{1}{\alpha_t^2}\int_{\alpha_{t+1}}\left(V_{t+1}\left(\alpha_t'',\alpha_{t+1}\right)-V_{t+1}\left(\alpha_t,\alpha_{t+1}\right)\right)\left(\pi\left(\alpha_{t+1}|\alpha_t\right)-\alpha_t\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}\right)d\alpha_t>0 \tag{63}$$

Applying Lemma 2 and the definition of V_{t+1}, this is equivalent to:

$$\int_{\alpha_t'}^{\alpha_t''}\frac{1}{\alpha_t^2}\left\{\mathbb{E}_t\left[\sum_{s=t+1}^{\infty}\beta^{s-t}\left(1-D_{t,s}\left(\alpha^s\right)\right)\alpha_s\left[u_s\left(\alpha^{t-1},\alpha_t'',...,\alpha_s\right)-u_s\left(\alpha^s\right)\right]\middle|\alpha_t\right]\right\}d\alpha_t>0 \tag{64}$$

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

'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 'if' part in reverse, this is equivalent to the inequality:

$$\alpha_t' u_t\left(\alpha_t''\right)+\beta\int_{\alpha_{t+1}}V_{t+1}\left(\alpha_t'',\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t'\right)d\alpha_{t+1}>\alpha_t' u_t\left(\alpha_t'\right)+\beta\int_{\alpha_{t+1}}V_{t+1}\left(\alpha_t',\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t'\right)d\alpha_{t+1} \tag{65}$$

Thus global incentive compatibility must be violated for type α_t'.

A.3 Proof of Proposition 2

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 constraints:

$$\sum_{t=0}^{\infty} R^{-t}\left[y_t - \int_{\alpha^t} c_t\left(\alpha^t\right) d\Pi_t\left(\alpha^t\right)\right] \geq 0 \tag{66}$$

$$\begin{aligned}\alpha_t' u\left(c_t\left(\alpha^{t-1},\alpha_t'\right)\right)+\beta\omega_{t+1}\left(\alpha^{t-1},\alpha_t'\right) = {} & \underline{\alpha} u\left(c_t\left(\alpha^{t-1},\underline{\alpha}\right)\right)+\beta\omega_{t+1}\left(\alpha^{t-1},\underline{\alpha}\right) \\ & + \int_{\underline{\alpha}}^{\alpha_t'} \frac{1}{\alpha_t}\left[\alpha_t u\left(c_t\left(\alpha^{t-1},\alpha_t\right)\right)+\beta\omega_{t+1}^{\Delta}\left(\alpha^{t-1},\alpha_t\right)\right] d\alpha_t\end{aligned} \tag{67}$$

$$\omega_{t+1}\left(\alpha^t\right) := \int_{\alpha_{t+1}} \left\{\alpha_{t+1} u\left(c_{t+1}\left(\alpha^t,\alpha_{t+1}\right)\right)+\beta\omega_{t+2}\left(\alpha^t,\alpha_{t+1}\right)\right\} d\Pi\left(\alpha_{t+1}|\alpha_t\right) \tag{68}$$

$$\omega_{t+1}^{\Delta}\left(\alpha^t\right) := \int_{\alpha_{t+1}} \rho\left(\alpha_{t+1}|\alpha_t\right)\cdot\left\{\alpha_{t+1} u\left(c_{t+1}\left(\alpha^t,\alpha_{t+1}\right)\right)+\beta\omega_{t+2}^{\Delta}\left(\alpha^t,\alpha_{t+1}\right)\right\} d\Pi\left(\alpha_{t+1}|\alpha_t\right) \tag{69}$$

I place multipliers η on (66), $\beta^t\mu_t\left(\alpha^{t-1},\alpha_t\right) d\Pi_t\left(\alpha^{t-1},\alpha_t\right)$ on (67), $\beta^{t+1}\lambda_{t+1}\left(\alpha^t\right) d\Pi_t\left(\alpha^t\right)$ on (68) and $\beta^{t+1}\lambda_{t+1}^{\Delta}\left(\alpha^t\right) d\Pi_t\left(\alpha^t\right)$ on (69). Taking necessary optimality conditions with respect to $c_t\left(\alpha^t\right)$, $\omega_{t+1}\left(\alpha^t\right)$, $\omega_{t+1}^{\Delta}\left(\alpha^t\right)$ and $c_t\left(\alpha^{t-1},\underline{\alpha}\right)$, in turn:

$$0 = \alpha_t u'\left(c_t\left(\alpha^t\right)\right)\cdot\left[\begin{array}{c} 1+\lambda_t\left(\alpha^{t-1}\right)+\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\rho\left(\alpha_t|\alpha_{t-1}\right)+\mu_t\left(\alpha^t\right) \\ -\frac{1}{\alpha_t\pi\left(\alpha_t|\alpha_{t-1}\right)}\int_{\alpha_t}^{\bar{\alpha}}\mu_t\left(\alpha^{t-1},\tilde{\alpha}_t\right)\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha}_t \end{array}\right] - (\beta R)^{-t}\eta \tag{70}$$

$$0 = -\lambda_{t+1}\left(\alpha^t\right)+\lambda_t\left(\alpha^{t-1}\right)+\mu_t\left(\alpha^{t-1},\alpha_t\right) \tag{71}$$

$$0 = -\lambda_{t+1}^{\Delta}\left(\alpha^t\right)+\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\rho\left(\alpha_t|\alpha_{t-1}\right)-\frac{1}{\alpha_t\pi\left(\alpha_t|\alpha_{t-1}\right)}\int_{\alpha_t}^{\bar{\alpha}}\mu_t\left(\alpha^{t-1},\tilde{\alpha}_t\right)\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha} \tag{72}$$

$$0 = \int_{\underline{\alpha}}^{\bar{\alpha}}\mu_t\left(\alpha^{t-1},\alpha_t\right)\pi\left(\alpha_t|\alpha_{t-1}\right)d\alpha_t \tag{73}$$

where $\lambda_0 = \lambda_0^{\Delta} \equiv 0$, and $\pi(\alpha_t|\alpha_{t-1})$ is replaced with $\pi(\alpha_0)$ when $t = 0$. Using (71) and (72) in (70) gives:

$$\frac{(\beta R)^{-t}\eta}{\alpha_t u'(c_t(\alpha^t))} = 1 + \lambda_{t+1}(\alpha^t) + \lambda_{t+1}^{\Delta}(\alpha^t) \tag{74}$$

Condition (70) can be rearranged to:

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

This can be integrated across all α_t to give:

$$\int_{\underline{\alpha}}^{\bar{\alpha}}\left\{\frac{(\beta R)^{-t}\eta}{u'(c_t(\alpha^{t-1},\alpha_t))} - \alpha_t\left[1 + \lambda_t\left(\alpha^{t-1}\right) + \lambda_t^{\Delta}\left(\alpha^{t-1}\right)\rho(\alpha_t|\alpha_{t-1})\right]\right\}\pi(\alpha_t|\alpha_{t-1})\,d\alpha_t = 0 \tag{76}$$

where I have used (73) and integrated by parts as follows:

$$\int_{\underline{\alpha}}^{\bar{\alpha}}\int_{\alpha_t}^{\bar{\alpha}}\mu_t\left(\alpha^{t-1},\tilde{\alpha}_t\right)\pi(\tilde{\alpha}_t|\alpha_{t-1})\,d\tilde{\alpha}_t d\alpha_t = \int_{\underline{\alpha}}^{\bar{\alpha}}\alpha_t\mu_t\left(\alpha^{t-1},\alpha_t\right)\pi(\alpha_t|\alpha_{t-1})\,d\alpha_t \tag{77}$$

or, using (5):

$$\frac{1}{\mathbb{E}[\alpha_t|\alpha_{t-1}]}\mathbb{E}\left[\left.\frac{(\beta R)^{-t}\eta}{u'(c_t(\alpha^t))}\right|\alpha^{t-1}\right] = \left[1 + \lambda_t\left(\alpha^{t-1}\right) + \lambda_t^{\Delta}\left(\alpha^{t-1}\right)\varepsilon^{\alpha}(\alpha_{t-1})\right] \tag{78}$$

Rearranging (78) for period 0 gives an expression for η:

$$\eta = \frac{\mathbb{E}[\alpha_0]}{\mathbb{E}\left[\frac{1}{u'(c_0(\alpha_0))}\right]} \tag{79}$$

Combining (74) and (78) gives expressions for the objects $1 + \lambda_{t+1}(\alpha^t)$ and $\lambda_{t+1}^{\Delta}(\alpha^t)$:

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

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

Integrating (75) above any given α_t' gives:

$$\int_{\alpha_t'}^{\bar{\alpha}} \left\{ \frac{(\beta R)^{-t} \eta}{u'\left(c_t\left(\alpha^{t-1}, \alpha_t\right)\right)} - \alpha_t \left[1 + \lambda_t\left(\alpha^{t-1}\right) + \lambda_t^{\Delta}\left(\alpha^{t-1}\right) \rho\left(\alpha_t | \alpha_{t-1}\right)\right]\right\} \pi\left(\alpha_t | \alpha_{t-1}\right) d\alpha_t \tag{82}$$
$$= \int_{\alpha_t'}^{\bar{\alpha}} \left\{ \alpha_t \mu_t\left(\alpha^{t-1}, \alpha_t\right) \pi\left(\alpha_t | \alpha_{t-1}\right) - \int_{\alpha_t}^{\bar{\alpha}} \mu_t\left(\alpha^{t-1}, \tilde{\alpha}_t\right) \pi\left(\tilde{\alpha}_t | \alpha_{t-1}\right) d\tilde{\alpha}_t \right\} d\alpha_t$$

and integrating by parts, I have:

$$\int_{\alpha_t'}^{\bar{\alpha}} \int_{\alpha_t}^{\bar{\alpha}} \mu_t\left(\alpha^{t-1}, \tilde{\alpha}_t\right) \pi\left(\tilde{\alpha}_t | \alpha_{t-1}\right) d\tilde{\alpha}_t d\alpha_t = -\alpha_t' \int_{\alpha_t'}^{\bar{\alpha}} \mu_t\left(\alpha^{t-1}, \alpha_t\right) \pi\left(\alpha_t | \alpha_{t-1}\right) d\alpha_t + \int_{\alpha_t'}^{\bar{\alpha}} \alpha_t \mu_t\left(\alpha^{t-1}, \alpha_t\right) \pi\left(\alpha_t | \alpha_{t-1}\right) d\alpha_t \tag{83}$$

so, using (72):

$$\int_{\alpha_t'}^{\bar{\alpha}} \left\{ \frac{(\beta R)^{-t} \eta}{u'\left(c_t\left(\alpha^{t-1}, \alpha_t\right)\right)} - \alpha_t \left[1 + \lambda_t\left(\alpha^{t-1}\right) + \lambda_t^{\Delta}\left(\alpha^{t-1}\right) \rho\left(\alpha_t | \alpha_{t-1}\right)\right]\right\} \pi\left(\alpha_t | \alpha_{t-1}\right) d\alpha_t \tag{84}$$
$$= -\pi\left(\alpha_t' | \alpha_{t-1}\right) \left(\alpha_t'\right)^2 \left\{ \lambda_{t+1}^{\Delta}\left(\alpha^t\right) - \rho\left(\alpha_t' | \alpha_{t-1}\right) \lambda_t^{\Delta}\left(\alpha^{t-1}\right) \right\}$$

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

A.4 Proof of Proposition 3

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

Given this, I show how to map from target allocation $\left\{c_t^*\left(\alpha^t\right)\right\}_{t,\alpha^t}$ to tax functions $T_t\left(s^{t-1}, s_t\right)$, such that every budget-feasible sequence of savings choices implies a consumption sequence that is consistent with the target allocation for some type history, and every consumption sequence that is realised for some type history under the target allocation can be chosen in the decentralisation. This implies the set of choices under the decentralisation is the same as under the direct mechanism at every history node, and so the decentralisation must implement the target allocation.

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

$$M_0 := \sum_{t=0}^{\infty} R^{-t} y_t \tag{85}$$

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

$$s_0(\alpha_0) := M_0 - c_0^*(\alpha_0) \tag{86}$$

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

$$S_0 := \{s_0(\alpha_0)\}_{\alpha_0 \in A} \tag{87}$$

Since consumption is increasing, $S_0 \subseteq \left[s_0(\bar{\alpha}), s_0(\underline{\alpha})\right]$. Denote by S_0^c the complement of S_0 in $\mathbb{R}$. For all $\tilde{s}_0 \in S_0^c$, let $T_0(\tilde{s}_0) = \tilde{s}_0$, so that $M_1(\tilde{s}_0) = 0$, and for all $t > 0$ and subsequent savings choices $\{s_1, ..., s_t\}$, let $T_t(\tilde{s}_0, s_1, ...s_t) > \varepsilon$ for some $\varepsilon > 0$. Combining (21) and (22), I have, along all future consumption paths:

$$0 = M_1(\tilde{s}_0) = \sum_{t=0}^{T} R^{-t} \left[c_{t+1} + T_{t+1}(\tilde{s}_0, ..., s_{t+1})\right] + R^{-T-2} M_{T+2} \tag{88}$$

and so:

$$\sum_{t=0}^{T} R^{-t} c_{t+1} = -\sum_{t=0}^{T} R^{-t} T_{t+1}(\tilde{s}_0, ..., s_{t+1}) - R^{-T-2} M_{T+2} \tag{89}$$

for all $T \geq 0$. By the 'no Ponzi' condition, $\lim_{T\to\infty} R^{-T-2} M_{T+2} \geq 0$, and so the positive bound on taxes implies the right-hand side must be negative for large enough T. But this would require negative consumption in at least one period. It follows that $\tilde{s}_0$ will not be chosen.

I now denote by $C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)$ the following object:

$$C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right) = \mathbb{E}\left[\sum_{s=t+1}^{\infty} R^{t+1-s} c_s^*\left(\alpha^{t-1}, \tilde{\alpha}_t, \alpha_{t+1}^s\right) \middle| \alpha_t\right] \tag{90}$$

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

$$M_1(s_0) := C_1(\alpha_0(s_0), \alpha_0(s_0)) + \int_{\alpha_0(s_0)}^{\bar{\alpha}} \left.\frac{\partial C_1(\tilde{\alpha}_0, \alpha_0)}{\partial \alpha_0}\right|_{\tilde{\alpha}_0 = \alpha_0} d\alpha_0 \tag{91}$$

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

$$\frac{dC_1(\alpha_0, \alpha_0)}{d\alpha_0} = \left.\frac{\partial C_1(\tilde{\alpha}_0, \alpha_0)}{\partial \alpha_0}\right|_{\tilde{\alpha}_0 = \alpha_0} \tag{92}$$

Thus when multiple types choose the same savings value, they receive the same $M_1(s_0)$ under formula (91). Given $M_1(s_0)$, define $T_0(s_0)$ by:

$$T_0(s_0) := s_0 - R^{-1} M_1(s_0) \tag{93}$$

An inductive argument can then be applied for 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}:

$$S^{t-1} := \left\{s^{t-1}\left(\alpha^{t-1}\right)\right\}_{\alpha^{t-1} \in A^t} \tag{94}$$

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:

$$s_t\left(\alpha^{t-1}, \alpha_t\right) := M_t\left(s^{t-1}\left(\alpha^{t-1}\right)\right) - c_t^*\left(\alpha^{t-1}, \alpha_t\right) \tag{95}$$

By the 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 $S_t\left(\alpha^{t-1}\right)$:

$$S_t\left(\alpha^{t-1}\right) := \left\{s_t\left(\alpha^{t-1}, \alpha_t\right)\right\}_{\alpha_t \in A} \tag{96}$$

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

$$s^t\left(\alpha^t\right) := \left\{s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^t\right)\right\} \tag{97}$$

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

$$\alpha^t\left(s^t\right) := \left\{\alpha^{t-1}\left(s^{t-1}\right), \alpha_t\left(s_t; \alpha^{t-1}\left(s^{t-1}\right)\right)\right\} \tag{98}$$

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

$$\begin{aligned} M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\right) :=& C_{t+1}\left(\alpha_t\left(s_t; \alpha^{t-1}\left(s^{t-1}\right)\right), \alpha_t\left(s_t; \alpha^{t-1}\left(s^{t-1}\right)\right); \alpha^{t-1}\left(s^{t-1}\right)\right) \\ &+ \int_{\alpha_t\left(s_t; \alpha^{t-1}\left(s^{t-1}\right)\right)}^{\bar{\alpha}} \left. \frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\left(s^{t-1}\right)\right)}{\partial \alpha_t} \right|_{\tilde{\alpha}_t = \alpha_t} d\alpha_t \end{aligned} \tag{99}$$

where I again use the fact that if c_t^* is constant on a range of α_t values, then:

$$\frac{dC_{t+1}\left(\alpha_t, \alpha_t; \alpha^{t-1}\right)}{d\alpha_t} = \left. \frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t} \right|_{\tilde{\alpha}_t = \alpha_t} \tag{100}$$

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

$$T_t\left(s^{t-1}, s_t\right) := s_t - R^{-1} M_{t+1}\left(s^{t-1}, s_t\right) \tag{101}$$

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

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

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

$$M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \bar{\alpha}\right)\right) = C_{t+1}\left(\bar{\alpha}, \bar{\alpha}; \alpha^{t-1}\right) \geq 0 \tag{102}$$

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

$$M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t''\right)\right) - M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t'\right)\right) + \int_{\alpha_t'}^{\alpha_t''} \left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t = \alpha_t} d\alpha_t$$

$$= \left[C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t''; \alpha^{t-1}\right)\right] + \left[C_{t+1}\left(\alpha_t', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right)\right] \tag{103}$$

$$= \left[C_{t+1}\left(\alpha_t'', \alpha_t'; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right)\right] + \left[C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t'', \alpha_t'; \alpha^{t-1}\right)\right] \tag{104}$$

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

$$\lim_{\alpha_t'' \to \alpha_t'} \left[\frac{C_{t+1}\left(\alpha_t', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right)}{\alpha_t'' - \alpha_t'}\right] = \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{1}{\alpha_t'' - \alpha_t'} \int_{\alpha_t'}^{\alpha_t''} \left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t = \alpha_t} d\alpha_t\right] \tag{105}$$

or:

$$\lim_{\alpha_t'' \to \alpha_t'} \left[\frac{C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t'', \alpha_t'; \alpha^{t-1}\right)}{\alpha_t'' - \alpha_t'}\right] = \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{1}{\alpha_t'' - \alpha_t'} \int_{\alpha_t'}^{\alpha_t''} \left.\frac{\partial C_{t+1}\left(\tilde{\alpha}_t, \alpha_t; \alpha^{t-1}\right)}{\partial \alpha_t}\right|_{\tilde{\alpha}_t = \alpha_t} d\alpha_t\right] \tag{106}$$

or both. Thus either:

$$\lim_{\alpha_t'' \to \alpha_t'} \left[\frac{M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t''\right)\right) - M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t'\right)\right)}{\alpha_t'' - \alpha_t'}\right] \tag{107}$$

$$= \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{C_{t+1}\left(\alpha_t'', \alpha_t''; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t''; \alpha^{t-1}\right)}{\alpha_t'' - \alpha_t'}\right] \leq 0$$

or:

$$\lim_{\alpha_t'' \to \alpha_t'} \left[\frac{M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t''\right)\right) - M_{t+1}\left(s^{t-1}\left(\alpha^{t-1}\right), s_t\left(\alpha^{t-1}, \alpha_t'\right)\right)}{\alpha_t'' - \alpha_t'}\right] \tag{108}$$

$$= \lim_{\alpha_t'' \to \alpha_t'} \left[\frac{C_{t+1}\left(\alpha_t'', \alpha_t'; \alpha^{t-1}\right) - C_{t+1}\left(\alpha_t', \alpha_t'; \alpha^{t-1}\right)}{\alpha_t'' - \alpha_t'}\right] \leq 0$$

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

It follows that the decentralisation provides the same choice set as the direct mechanism at all nodes, and so will implement it. Using (24), it is straightforward additionally to confirm that the normalisations (26) and (27) are satisfied by the proposed construction.

B Proof of Theorem 1

Overview

The proof works in the following two steps:

1. I start by 'reversing' the envelope condition – using the results in Proposition 2 to state the effects of utility perturbations that raise the welfare of types above a given threshold *in proportion to marginal utilities*. This is Corollary 3.

2. I then prove, in Lemma 3, that the algebraic objects featuring in Corollary 3 link directly to the behavioural objects defined in Section 7 of the main paper.

Both steps, though algebraically involved, are based on elementary techniques.

Setp 1: Statement

Corollary 3. *If a strictly normal allocation is optimal in the relaxed problem, with $c_t(\alpha_t)$ continuous at a given history node α^{t-1}, then the following two expressions are true:*

$$0 = \int_{\underline{c}}^{\bar{c}} \left[1 - \frac{\alpha_t(c)\, u'(c) \left\{1 + \lambda_t + \lambda_t^{\Delta} \rho(\alpha_t(c)\,|\alpha_{t-1})\right\}}{\eta_t}\right] \pi^c\left(c|\alpha^{t-1}\right) dc$$
$$+ \int_{\underline{c}}^{\bar{c}} (\alpha_t(c))^2 u''(c) \left\{\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c))}{\eta_t} - \rho(\alpha_t(c)\,|\alpha_{t-1})\, \beta R \frac{\lambda_t^{\Delta}}{\eta_t}\right\} \left(\frac{d\alpha_t(c)}{dc}\right)^{-1} \pi^c\left(c|\alpha^{t-1}\right) dc \tag{109}$$

$$0 = -\int_{\underline{c}}^{c'} \left[1 - \frac{\alpha_t(c)\, u'(c) \left\{1 + \lambda_t + \lambda_t^{\Delta} \rho(\alpha_t(c)\,|\alpha_{t-1})\right\}}{\eta_t}\right] \pi^c\left(c|\alpha^{t-1}\right) dc$$
$$- \int_{\underline{c}}^{c'} (\alpha_t(c))^2 u''(c) \left\{\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c))}{\eta_t} - \rho(\alpha_t(c)\,|\alpha_{t-1})\, \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\} \left(\frac{d\alpha_t(c)}{dc}\right)^{-1} \pi^c\left(c|\alpha^{t-1}\right) dc$$
$$+ (\alpha_t(c'))^2 u'(c') \left(\frac{d\alpha_t(c)}{dc}\right)^{-1} \pi^c\left(c'|\alpha^{t-1}\right) \left\{\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c'))}{\eta_t} - \rho(\alpha_t(c')\,|\alpha_{t-1})\, \beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\} \tag{110}$$

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

Step 1: Proof

Whenever $c_t(\alpha)$ is continuous in α on an open interval $(\alpha', \alpha'') \subset [\underline{\alpha}, \bar{\alpha}]$, with $c' = \lim_{\alpha \searrow \alpha'}(c_t(\alpha))$ and $c'' = \lim_{\alpha \nearrow \alpha''}(c_t(\alpha))$, I can integrate (16) across this range:

$$\int_{c'}^{c''} \left\{ \int_{\alpha_t(c_t)}^{\bar{\alpha}} \left[\frac{1}{u'(c_t(\alpha))} - \frac{\alpha\left\{1 + \lambda_t + \lambda_t^{\Delta}\rho(\alpha|\alpha_{t-1})\right\}}{\eta_t} \right] \pi(\alpha|\alpha_{t-1})\, d\alpha \right\} \frac{du'(c_t)}{dc_t} dc_t \tag{111}$$
$$= -\int_{c'}^{c''} (\alpha_t(c_{t.}))^2 \left\{ \frac{\lambda_{t+1}^{\Delta}(\alpha_t(c_t))}{\eta_t} - \rho(\alpha_t(c_t)|\alpha_{t-1})\,\beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}} \right\} \pi(\alpha_t(c_t)|\alpha_{t-1}) \frac{du'(c_t)}{dc_t} dc_t$$

Integrating the left-hand side by parts, I can show:

$$\int_{c'}^{c''} \left\{ \int_{\alpha_t(c_t)}^{\bar{\alpha}} \left[\frac{1}{u'(c_t(\alpha))} - \frac{\alpha\left\{1 + \lambda_t + \lambda_t^{\Delta}\rho(\alpha|\alpha_{t-1})\right\}}{\eta_t} \right] \pi(\alpha|\alpha_{t-1})\, d\alpha \right\} \frac{du'(c_t)}{dc_t} dc_t \tag{112}$$
$$= \int_{c'}^{c''} \left[1 - \frac{\alpha_t(c_t)\, u'(c_t)\left\{1 + \lambda_t + \lambda_t^{\Delta}\rho(\alpha_t(c_t)|\alpha_{t-1})\right\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t)|\alpha_{t-1})\, dc_t \tag{113}$$
$$+ u'(c'') \int_{c''}^{\bar{c}} \left[\frac{1}{u'(c_t)} - \frac{\alpha_t(c_t)\left\{1 + \lambda_t + \lambda_t^{\Delta}\rho(\alpha_t(c_t)|\alpha_{t-1})\right\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t)|\alpha_{t-1})\, dc_t$$
$$- u'(c') \int_{c'}^{\bar{c}} \left[\frac{1}{u'(c_t)} - \frac{\alpha_t(c_t)\left\{1 + \lambda_t + \lambda_t^{\Delta}\rho(\alpha_t(c_t)|\alpha_{t-1})\right\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t)|\alpha_{t-1})\, dc_t$$

where strict normality guarantees $\frac{d\alpha_t(c_t)}{dc_t}$ is defined a.e.. Condition (111) thus becomes:

$$\int_{c'}^{c''} \left[1 - \frac{\alpha_t(c_t)\, u'(c_t)\left\{1 + \lambda_t + \lambda_t^{\Delta}\rho(\alpha_t(c_t)|\alpha_{t-1})\right\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t)|\alpha_{t-1})\, dc_t \tag{114}$$
$$+ u'(c'') \int_{c''}^{\bar{c}} \left[\frac{1}{u'(c_t)} - \frac{\alpha_t(c_t)\left\{1 + \lambda_t + \lambda_t^{\Delta}\rho(\alpha|\alpha_{t-1})\right\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t)|\alpha_{t-1})\, dc_t$$
$$- u'(c') \int_{c'}^{\bar{c}} \left[\frac{1}{u'(c_t)} - \frac{\alpha_t(c_t)\left\{1 + \lambda_t + \lambda_t^{\Delta}\rho(\alpha|\alpha_{t-1})\right\}}{\eta_t} \right] \frac{d\alpha_t(c_t)}{dc_t} \pi(\alpha_t(c_t)|\alpha_{t-1})\, dc_t$$
$$= -\int_{c'}^{c''} (\alpha_t(c_{t.}))^2 u''(c_t) \left\{ \frac{\lambda_{t+1}^{\Delta}(\alpha_t(c_t))}{\eta_t} - \rho(\alpha_t(c_t)|\alpha_{t-1})\,\beta R \frac{\lambda_t^{\Delta}}{\eta_{t-1}} \right\} \pi(\alpha_t(c_t)|\alpha_{t-1})\, dc_t$$

With no discontinuities in $c_t(\alpha_t)$, (76) can be applied over the entire range, to give:

$$0=\int_{\underline{c}}^{\bar{c}}\left[1-\frac{\alpha_t(c_t)u'(c_t)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho(\alpha_t(c_t)|\alpha_{t-1})\right\}}{\eta_t}\right]\pi^c\left(c_t|\alpha^{t-1}\right)dc_t \tag{115}$$
$$+\int_{\underline{c}}^{\bar{c}}(\alpha_t(c_{t.}))^2u''(c_t)\left\{\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c_t))}{\eta_t}-\rho(\alpha_t(c_t)|\alpha_{t-1})\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}\left(\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1}\pi^c\left(c_t|\alpha^{t-1}\right)dc_t$$

where $\pi^c\left(c_t|\alpha^{t-1}\right):=\frac{d\alpha_t(c_t)}{dc_t}\pi(\alpha_t(c_t)|\alpha_{t-1})$ measures the realised density of consumption at c_t, conditional on history. This will equal $\pi^s\left(s_t|\alpha^{t-1}\right)$ at the corresponding savings level, since $\frac{ds_t}{dc_t}=-1$ (given history). Similarly, for each $c'\in\left(\underline{c},\bar{c}\right)$, making use of (84):

$$0=\int_{c'}^{\bar{c}}\left[1-\frac{\alpha_t(c_t)u'(c_t)\left\{1+\lambda_t+\lambda_t^{\Delta}\rho(\alpha_t(c_t)|\alpha_{t-1})\right\}}{\eta_t}\right]\pi^c\left(c_t|\alpha^{t-1}\right)dc_t \tag{116}$$
$$+(\alpha_t(c'))^2u'(c')\pi^c\left(c'|\alpha^{t-1}\right)\left(\frac{d\alpha_t(c')}{dc'}\right)^{-1}\left\{\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c'))}{\eta_t}-\rho(\alpha_t(c')|\alpha_{t-1})\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}$$
$$+\int_{c'}^{\bar{c}}(\alpha_t(c_{t.}))^2u''(c_t)\left\{\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c_t))}{\eta_t}-\rho(\alpha_t(c_t)|\alpha_{t-1})\beta R\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right\}\left(\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1}\pi^c\left(c_t|\alpha^{t-1}\right)dc_t$$

as given above.

Step 2: Statement

Lemma 3. *The following four relationships hold:*

$$T_t'(s_t)s_t\varepsilon_t^s=\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c_t))}{\eta_t}(\alpha_t(c_t))^2u'(c_t)\left(\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1} \tag{117}$$
$$T_t'(s_t)\frac{ds_t}{dM_t}=-\frac{\lambda_{t+1}^{\Delta}(\alpha_t(c_t))}{\eta_t}(\alpha_t(c_t))^2u''(c_t)\left(\frac{d\alpha_t(c_t)}{dc_t}\right)^{-1} \tag{118}$$

$$RT_t'(s_t)\,s_t\epsilon_{t,t+1}^s\left(s_{t+1}'\right)=-\frac{1}{\Pi^c\left(c_{t+1}'|\alpha^t\right)}\left\{\rho\left(\alpha_{t+1}\left(c_{t+1}'\right)|\alpha_t\right)\beta R\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\left(\alpha_{t+1}\left(c_{t+1}'\right)\right)^2u'\left(c_{t+1}'\right)\left(\frac{d\alpha_{t+1}\left(c_{t+1}'\right)}{dc_{t+1}}\right)^{-1}\pi^c\left(c_{t+1}'|\alpha^t\right)-\right.$$
$$\left.\int_{\underline{c}}^{c_{t+1}'}\left\{\rho\left(\alpha_{t+1}\left(c_{t+1}\right)|\alpha_t\right)\beta R\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\alpha_{t+1}\left(c_{t+1}\right)u'\left(c_{t+1}\right)\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\left(\frac{c_{t+1}}{\alpha_{t+1}\left(c_{t+1}\right)}\frac{d\alpha_{t+1}\left(c_{t+1}\right)}{dc_{t+1}}\right)^{-1}\right]\pi^c\left(c_{t+1}|\alpha^t\right)\right\}dc_{t+1}\right\} \tag{119}$$

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

Step 2: Proof

(a) Proof of Conditions (117) and (118)

I start with two Lemmata:

Lemma 4. *The marginal tax rate satisfies:*

$$T_t'\left(s\left(\alpha_t\right)\right)=\frac{\lambda_{t+1}^{\Delta}\left(\alpha_t\right)}{\eta_t}\left\{\alpha_tu'\left(c_t\left(\alpha_t\right)\right)-\beta R\left(1-T_t'\left(s_t\left(\alpha_t\right)\right)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\right\} \tag{121}$$

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

Proof. Under condition (26), $T_t'(s_t)$ is normalised to be 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 $T_t'(s_t)$ can be obtained from the direct allocation by comparing the shadow cost of providing the utility delivered by a unit increase in savings, and the shadow cost of providing the same quantity of utility in t. By construction, savings raise period-$t+1$ lifetime utility ω_{t+1} at the margin by the amount:

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

and raise ω^{Δ}_{t+1} by the amount:

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

Duality with cost minimisation implies the marginal resource cost of per unit increase in ω_{t+1} can be obtained by dividing the policymaker's marginal value of an increase to ω_{t+1} 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^{Δ}_{t+1} by a unit is:

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

The direct marginal resource gain from a unit increase in savings (reduction in consumption) in t is R^{-t} in present value. Combining, and multiplying by R^t, I obtain:

$$\begin{aligned}T_t'\left(s_t\right)=&1-\frac{\left(1+\lambda_{t+1}\left(\alpha^t\right)\right)}{\eta\left(\beta R\right)^{-t-1}}\left(1-T_t'\left(s_t\right)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\\&-\frac{\lambda^{\Delta}_{t+1}\left(\alpha^t\right)}{\eta\left(\beta R\right)^{-t-1}}\left(1-T_t'\left(s_t\right)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\end{aligned}\tag{122}$$

Expressions for $\frac{\left(1+\lambda_{t+1}\left(\alpha^t\right)\right)}{\eta_t}$ and $\frac{\lambda^{\Delta}_{t+1}\left(\alpha^t\right)}{\eta_t}$, with $\eta_t=\eta\left(\beta R\right)^{-t}$, are given in (80) and (81) respectively. Substituting in (80) gives:

$$\begin{aligned}T_t'\left(s_t\right)=&1-\frac{1}{1-\varepsilon^{\alpha}\left(\alpha_t\right)}\left\{\frac{1}{\mathbb{E}\left[\alpha_{t+1}|\alpha_t\right]}\mathbb{E}\left[\frac{1}{\beta Ru'\left(c_{t+1}\left(\alpha^{t+1}\right)\right)}\middle|\alpha^t\right]-\frac{\varepsilon^{\alpha}\left(\alpha_t\right)}{\alpha_t u'\left(c_t\left(\alpha^t\right)\right)}\right\}\cdot\alpha_t u'\left(c_t\left(\alpha^t\right)\right)\\&-\frac{\lambda^{\Delta}_{t+1}\left(\alpha^t\right)}{\eta_t}\beta R\left(1-T_t'\left(s_t\right)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\end{aligned}\tag{123}$$

where I have used the consumer optimality condition:

$$\alpha_t u'\left(c_t\left(\alpha^t\right)\right)=\beta R\left(1-T_t'\left(s_t\right)\right)\int_{\alpha_{t+1}}V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\tag{124}$$

Using (81), the first line on the right-hand side of (123) can be seen to equal $\frac{\lambda^{\Delta}_{t+1}(\alpha^t)}{\eta_t} \cdot \alpha_t u'\left(c_t\left(\alpha^t\right)\right)$, so:

$$T'_t\left(s\left(\alpha_t\right)\right) = \frac{\lambda^{\Delta}_{t+1}\left(\alpha_t\right)}{\eta_t}\left\{\alpha_t u'\left(c_t\right) - \beta R\left(1 - T'_t\left(s_t\right)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t \frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\right\} \quad (125)$$

as stated. □

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

$$\left\{\alpha_t u'\left(c_t\right) - \beta R\left(1 - T'_t\left(s_t\right)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t \frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\right\}^{-1} = -\frac{\frac{ds_t}{dM_t}}{\alpha_t^2 u''\left(c_t\right)\frac{dc_t}{d\alpha_t}} \quad (126)$$

$$\left\{\alpha_t u'\left(c_t\right) - \beta R\left(1 - T'_t\left(s_t\right)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\alpha_t \frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1}\right\}^{-1} = \frac{s_t \varepsilon^s_t}{\alpha_t^2 u'\left(c_t\right)\frac{dc_t}{d\alpha_t}} \quad (127)$$

Proof. Start with (126). 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:

$$\frac{dc_t}{d\alpha_t} + \frac{dc_t}{dM_t}\frac{dM_t}{d\alpha_t} = \frac{dM_t}{d\alpha_t} \quad (128)$$

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:

$$\alpha_t u'\left(c_t\right) = \beta R\left(1 - T'\left(s_t\right)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \quad (129)$$

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

Differentiating (129) with respect to α_t, given constant savings, yields:

$$u'\left(c_t\right) - \beta R\left(1 - T'_t\left(s_t\right)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1} + \alpha_t u''\left(c_t\right)\left[\frac{dc_t}{d\alpha_t} + \frac{dc_t}{dM_t}\frac{dM_t}{d\alpha_t}\right] = 0 \quad (130)$$

$$u'\left(c_t\right) - \beta R\left(1 - T'_t\left(s_t\right)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1};\alpha_{t+1}\right)\frac{d\pi\left(\alpha_{t+1}|\alpha_t\right)}{d\alpha_t}d\alpha_{t+1} + \alpha_t u''\left(c_t\right)\frac{dM_t}{d\alpha_t} = 0 \quad (131)$$

Rearranging (128):

$$\frac{dM_t}{d\alpha_t} = \frac{\frac{dc_t}{d\alpha_t}}{1 - \frac{dc_t}{dM_t}} = \frac{\frac{dc_t}{d\alpha_t}}{\frac{ds}{dM_t}} \tag{132}$$

Plugging this into the previous expression, trivial manipulations give (126).

Reasoning in a similar way for (127), 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:

$$\frac{dc_t}{d\left(1 - T'(s_t)\right)} + \frac{dc_t}{dM_t}\frac{dM_t}{d\left(1 - T'(s_t)\right)} = \frac{dM_t}{d\left(1 - T'(s_t)\right)} \tag{133}$$

Differentiating (129) with respect to $\left(1 - T_t'(s_t)\right)$ under this joint change gives:

$$0 = -\beta R \int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1}; \alpha_{t+1}\right) \pi\left(\alpha_{t+1} | \alpha_t\right) d\alpha_{t+1} + \alpha_t u''(c_t)\left[\frac{dc_t}{d\left(1 - T_t'(s_t)\right)} + \frac{dc_t}{dM_t}\frac{dM_t}{d\left(1 - T_t'(s_t)\right)}\right] \tag{134}$$

$$= -\alpha_t u'(c_t)\frac{1}{1 - T_t'(s_t)} + \alpha_t u''(c_t)\frac{\frac{dc_t}{d\left(1-T_t'(s_t)\right)}}{\frac{ds_t}{dM_t}} \tag{135}$$

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

$$\frac{s_t \varepsilon_t^s}{u'(c_t)} = -\frac{\frac{ds_t}{dM_t}}{u''(c_t)} \tag{136}$$

and so (127) follows, given (126). □

The results in these Lemmata immediately give conditions (117) and (118).

(b) Proof of conditions (119) and (120)

Condition (119) is obtained by analysing general, offsetting perturbations to the marginal value of saving. Two expressions for this can be stated. First, from the consumer optimality condition:

$$\alpha_t u'(c_t) = \beta R\left(1 - T_t'(s_t)\right)\int_{\alpha_{t+1}} V_{M,t+1}\left(M_{t+1}; \alpha_{t+1}\right)\pi\left(\alpha_{t+1}|\alpha_t\right) d\alpha_{t+1} \tag{137}$$

Second, by differentiating the relaxed incentive constraint and rearranging:

$$\alpha_t u'(c_t) = \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}} \beta \left\{ \frac{1}{\alpha_t} \omega_{t+1}^{\Delta}(\alpha_t) - \frac{d\omega_{t+1}(\alpha_t)}{d\alpha_t} \right\} \tag{138}$$

The right-hand sides of (137) and (138) must be equal at any realised allocation. Denote this object $\gamma_t(\alpha_t)$, i.e.:

$$\gamma_t(\alpha_t) = R\left(1 - T_t'(s_t)\right) \int_{\alpha_{t+1}} V_{M,t+1}(M_{t+1}; \alpha_{t+1}) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \tag{139}$$

$$= \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}} \beta \left\{ \frac{1}{\alpha_t} \omega_{t+1}^{\Delta}(\alpha_t) - \frac{d\omega_{t+1}(\alpha_t)}{d\alpha_t} \right\} \tag{140}$$

I am interested in the response of $c_t(\alpha_t)$ to an arbitrary differential change in the consumer's constraint set, generically denoted by Δ_i. I have:

$$\alpha_t u''(c_t(\alpha_t)) \frac{dc_t(\alpha_t)}{d\Delta_i} = \frac{d\gamma_t(\alpha_t)}{d\Delta_i} \tag{141}$$

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

$$\frac{d\gamma_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{d\gamma_t(\alpha_t)}{d\Delta_j} = 0 \tag{142}$$

Using (142) in (141):

$$\frac{dc_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{dc_t(\alpha_t)}{d\Delta_j} = 0 \tag{143}$$

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

$$\frac{ds_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{ds_t(\alpha_t)}{d\Delta_j} = 0 \tag{144}$$

and so:

$$\frac{d\gamma_t(\alpha_t)}{d\Delta_i} + \Gamma \frac{d\gamma_t(\alpha_t)}{d\Delta_j} = \left. \frac{d\gamma_t(\alpha_t)}{d\Delta_i} \right|_{s_t, c_t} + \Gamma \left. \frac{d\gamma_t(\alpha_t)}{d\Delta_j} \right|_{s_t, c_t} = 0 \tag{145}$$

This means that Γ can be recovered as a ratio of derivatives of $\gamma_t(\alpha_t)$ taken under the assumption of fixed savings and consumption in t, without loss. Thus (143) gives:

$$\frac{dc_t(\alpha_t)}{d\Delta_i} = \frac{\left.\frac{d\gamma_t(\alpha_t)}{d\Delta_i}\right|_{s_t,c_t}}{\left.\frac{d\gamma_t(\alpha_t)}{d\Delta_j}\right|_{s_t,c_t}} \frac{dc_t(\alpha_t)}{d\Delta_j} \tag{146}$$

I take these derivatives of $\gamma_t(\alpha_t)$ for two types of change to the consumer's budget constraint in t. The first, simpler, is a change to post-tax marginal returns, $\left(1 - T_t'(s_t)\right)$. From (139):

$$\left.\frac{d\gamma_t(\alpha_t)}{d\left(1 - T_t'(s_t)\right)}\right|_{s_t,c_t} = \frac{1}{1 - T_t'(s_t)}\gamma_t(\alpha_t) = \frac{1}{1 - T_t'(s_t)}\alpha_t u'(c_t) \tag{147}$$

The second change is a 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:

$$c_{t+1} = M_{t+1} - s(M_{t+2}) \tag{148}$$

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

$$M_{t+2} \equiv R\left[s(M_{t+2}) - T_{t+1}(s(M_{t+2}))\right] \tag{149}$$

Consider perturbations of the form:

$$c_{t+1} = M_{t+1} - s(M_{t+2}) + \Gamma f(s(M_{t+2})) \tag{150}$$

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 (140), since c_t and ω_{t+1} are being held constant, I can write:

$$\left.\frac{d\gamma_t(\alpha_t)}{d\Gamma}\right|_{s_t,c_t} = \frac{1}{\frac{dc_t(\alpha_t)}{d\alpha_t}}\beta\frac{1}{\alpha_t}\frac{d\omega_{t+1}^{\Delta}(\alpha_t)}{d\Gamma} \tag{151}$$

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

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

$$\begin{aligned}\frac{d\omega^{\Delta}_{t+1}(\alpha_t)}{d\Gamma} =& \int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1}))\Bigg\{\alpha^2_{t+1}u'(c_{t+1})\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}\rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t) \\ &-\frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}\int_{\underline{\alpha}}^{\alpha_{t+1}}\left[\tilde{\alpha}_{t+1}(u'(c_{t+1}))+\tilde{\alpha}^2_{t+1}u''(c_{t+1})\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho(\tilde{\alpha}_{t+1}|\alpha_t)\pi(\tilde{\alpha}_{t+1}|\alpha_t)\,d\tilde{\alpha}_{t+1}\Bigg\}d\alpha_{t+1} \\ &+f(s_{t+1}(\bar{\alpha}))\int_{\alpha_{t+1}}\left\{\alpha_{t+1}(u'(c_{t+1}))+\alpha^2_{t+1}u''(c_{t+1})\frac{dc_{t+1}}{d\alpha_{t+1}}\right\}\rho(\alpha_{t+1}|\alpha_t)\pi(\alpha_{t+1}|\alpha_t)\,d\alpha_{t+1}\end{aligned} \tag{152}$$

Proof. I take in turn the income and substitution effects of the pertubation at $t+1$, and then sum. The income effect will be proportional to the increase in c_{t+1} at each M_{t+2} along the budget constraint, given by:

$$\left.\frac{dc_{t+1}}{d\Gamma}\right|_{M_{t+2}} = f(s(M_{t+2})) \tag{153}$$

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

$$\frac{dc_{t+1}}{dM_{t+2}} = -s'(M_{t+2})(1-\Gamma f'(s(M_{t+2}))) = -\frac{1-\Gamma f'(s(M_{t+2}))}{R\left[1-T'_{t+1}(s(M_{t+2}))\right]} \tag{154}$$

The effect on this as Γ changes is:

$$\frac{d}{d\Gamma}\left[\frac{dc_{t+1}}{dM_{t+2}}\right] = \frac{f'(s(M_{t+2}))}{R\left[1-T'_{t+1}(s(M_{t+2}))\right]} \tag{155}$$

$$= f'(s(M_{t+2}))\left(1-T'_{t+1}(s(M_{t+2}))\right)\frac{d}{d\left(1-T'_{t+1}(s(M_{t+2}))\right)}\left.\left[\frac{dc_{t+1}}{dM_{t+2}}\right]\right|_{\Gamma=0} \tag{156}$$

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

$$\left.\frac{d\omega^{\Delta}_{t+1}}{d\Gamma}\right|_{\text{sub}} = \int_{\alpha_{t+1}} \left\{ -\alpha_{t+1} u'(c_{t+1}) + \beta R \left(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))\right) \int_{\alpha_{t+1}} V_M(\alpha_{t+1}) \frac{d\pi(\alpha_{t+2}|\alpha_{t+1})}{d\alpha_{t+1}} d\alpha_{t+2} \right\} \tag{157}$$

$$\times \frac{ds_{t+1}(\alpha_{t+1})}{d\left(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))\right)} \left(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))\right) f'(s_{t+1}(\alpha_{t+1})) \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1}$$

$$= \int_{\alpha_{t+1}} \left\{ -\alpha^2_{t+1} u'(c_{t+1}) \frac{dc_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} \frac{1}{\varepsilon^s_{t+1} s_{t+1}(\alpha_{t+1})} \right\} \tag{158}$$

$$\times \frac{ds_{t+1}(\alpha_{t+1})}{d\left(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))\right)} \left(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))\right) f'(s_{t+1}(\alpha_{t+1})) \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1}$$

$$= - \int_{\alpha_{t+1}} \alpha^2_{t+1} u'(c_{t+1}) \frac{dc_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \tag{159}$$

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

Proceeding similarly, the income effect on ω^{Δ}_{t+1} is:

$$\left.\frac{d\omega^{\Delta}_{t+1}}{d\Gamma}\right|_{\text{inc}} = \int_{\alpha_{t+1}} f(s_{t+1}(\alpha_{t+1})) \rho(\alpha_{t+1}|\alpha_t) \left\{ \alpha_{t+1}(u'(c_{t+1})) + \frac{ds_{t+1}(\alpha_{t+1})}{dM_{t+1}} \left[-\alpha_{t+1} u'(c_{t+1}) \right.\right. \tag{160}$$

$$\left.\left. + \beta R \left(1 - T'_{t+1}(s_{t+1}(\alpha_{t+1}))\right) \int_{\alpha_{t+1}} V_M(\alpha_{t+1}) \frac{d\pi(\alpha_{t+2}|\alpha_{t+1})}{d\alpha_{t+1}} d\alpha_{t+2} \right] \right\} \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1}$$

$$= \int_{\alpha_{t+1}} f(s_{t+1}(\alpha_{t+1})) \left\{ \alpha_{t+1}(u'(c_{t+1})) + \alpha^2_{t+1} u''(c_{t+1}) \frac{dc_{t+1}}{d\alpha_{t+1}} \right\} \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \tag{161}$$

$$= f(s_{t+1}(\bar{\alpha})) \int_{\alpha_{t+1}} \left\{ \alpha_{t+1}(u'(c_{t+1})) + \alpha^2_{t+1} u''(c_{t+1}) \frac{dc_{t+1}}{d\alpha_{t+1}} \right\} \rho(\alpha_{t+1}|\alpha_t) \pi(\alpha_{t+1}|\alpha_t) d\alpha_{t+1} \tag{162}$$

$$- \int_{\alpha_{t+1}} f'(s_{t+1}(\alpha_{t+1})) \frac{ds_{t+1}(\alpha_{t+1})}{d\alpha_{t+1}}$$

$$\times \left\{ \int_{\underline{\alpha}}^{\alpha_{t+1}} \left[\tilde{\alpha}_{t+1}(u'(c_{t+1})) + \tilde{\alpha}^2_{t+1} u''(c_{t+1}) \frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}} \right] \rho(\tilde{\alpha}_{t+1}|\alpha_t) \pi(\tilde{\alpha}_{t+1}|\alpha_t) d\tilde{\alpha}_{t+1} \right\} d\alpha_{t+1}$$

where the second equality makes use of (126). Combining (159) and (162), the result follows. □

Applying (146), (147) and (151), I have:

$$\frac{dc_t}{d\Gamma} = \left[\frac{\alpha_t u'(c_t)}{\left(1 - T'_t(s_t)\right)} \right]^{-1} \frac{dc_t}{d\left(1 - T'_t(s_t)\right)} \frac{1}{\frac{dc_t}{d\alpha_t}} \beta \frac{1}{\alpha_t} \frac{d\omega^{\Delta}_{t+1}}{d\Gamma} \tag{163}$$

So:

$$\frac{1}{s_t}\frac{ds_t}{d\Gamma}=\left[\frac{\alpha_t u'(c_t)}{\left(1-T_t'(s_t)\right)}\right]^{-1}\frac{1}{s_t}\frac{ds_t}{d\left(1-T_t'(s_t)\right)}\frac{1}{\frac{dc_t}{d\alpha_t}}\beta\frac{1}{\alpha_t}\frac{d\omega_{t+1}^{\Delta}}{d\Gamma} \tag{164}$$

$$\begin{aligned}=&\int_{\alpha_{t+1}} f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\varepsilon_t^s\frac{\alpha_{t+1}\frac{ds_{t+1}}{d\alpha_{t+1}}}{\alpha_t\frac{ds_t}{d\alpha_t}}\left\{-\frac{\beta\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)\right. \\ &\left.+\frac{1}{\alpha_{t+1}}\int_{\underline{\alpha}}^{\alpha_{t+1}}\frac{\beta\tilde{\alpha}_{t+1}\left(u'\left(c_{t+1}\right)\right)}{\alpha_t u'\left(c_t\right)}\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho\left(\tilde{\alpha}_{t+1}|\alpha_t\right)\pi\left(\tilde{\alpha}_{t+1}|\alpha_t\right)d\tilde{\alpha}_{t+1}\right\}d\alpha_{t+1} \\ &+f\left(s_{t+1}\left(\bar{\alpha}\right)\right)\varepsilon_t^s\frac{1}{\alpha_t\frac{dc_t}{d\alpha_t}}\int_{\alpha_{t+1}}\frac{\beta\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1}\end{aligned} \tag{165}$$

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

$$\frac{1}{s_t}\frac{ds_t}{d\Gamma}\equiv\int_{\alpha_{t+1}} f'\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\epsilon_{t,t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)\Pi\left(\alpha_{t+1}|\alpha_t\right)\left(-\frac{ds_{t+1}\left(\alpha_{t+1}\right)}{d\alpha_{t+1}}\right)d\alpha_{t+1}+f\left(s_{t+1}\left(\bar{\alpha}\right)\right)\frac{1}{s_t}\left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}} \tag{166}$$

$\epsilon_{t,t+1}$ is thus given by:

$$\begin{aligned}\epsilon_{t,t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right)=&-\varepsilon_t^s\frac{\alpha_{t+1}}{\Pi\left(\alpha_{t+1}|\alpha_t\right)\alpha_t\frac{ds_t}{d\alpha_t}}\left\{-\frac{\beta\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)\right. \\ &\left.+\frac{1}{\alpha_{t+1}}\int_{\underline{\alpha}}^{\alpha_{t+1}}\frac{\beta\tilde{\alpha}_{t+1}\left(u'\left(c_{t+1}\right)\right)}{\alpha_t u'\left(c_t\right)}\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\times\rho\left(\tilde{\alpha}_{t+1}|\alpha_t\right)\pi\left(\tilde{\alpha}_{t+1}|\alpha_t\right)d\tilde{\alpha}_{t+1}\right\}\end{aligned} \tag{167}$$

So:

$$\begin{aligned}&RT_t'\left(s_t\right)s_t\epsilon_{t,t+1}\left(s_{t+1}\left(\alpha_{t+1}\right)\right) \\ =&-RT_t'\left(s_t\right)s_t\varepsilon_t^s\frac{\alpha_{t+1}}{\Pi\left(\alpha_{t+1}|\alpha_t\right)\alpha_t\frac{ds_t}{d\alpha_t}}\left\{-\frac{\beta\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)\right. \\ &\left.+\frac{1}{\alpha_{t+1}}\int_{\underline{\alpha}}^{\alpha_{t+1}}\frac{\beta\tilde{\alpha}_{t+1}\left(u'\left(c_{t+1}\right)\right)}{\alpha_t u'\left(c_t\right)}\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho\left(\tilde{\alpha}_{t+1}|\alpha_t\right)\pi\left(\tilde{\alpha}_{t+1}|\alpha_t\right)d\tilde{\alpha}_{t+1}\right\}\end{aligned} \tag{168}$$

$$\begin{aligned}=&-\beta R\alpha_{t+1}u'\left(c_{t+1}\right)\rho\left(\alpha_{t+1}|\alpha_t\right)\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\frac{\alpha_{t+1}\pi\left(\alpha_{t+1}|\alpha_t\right)}{\Pi\left(\alpha_{t+1}|\alpha_t\right)} \\ &+\frac{1}{\Pi\left(\alpha_{t+1}|\alpha_t\right)}\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\int_{\underline{\alpha}}^{\alpha_{t+1}}\beta R\tilde{\alpha}_{t+1}\left(u'\left(c_{t+1}\right)\right)\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\tilde{\alpha}_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\tilde{\alpha}_{t+1}}\right]\rho\left(\tilde{\alpha}_{t+1}|\alpha_t\right)\pi\left(\tilde{\alpha}_{t+1}|\alpha_t\right)d\tilde{\alpha}_{t+1}\end{aligned} \tag{169}$$

Changing the unit of integration and using $\pi\left(\alpha_{t+1}|\alpha_t\right)\frac{d\alpha_{t+1}(c_{t+1})}{dc_{t+1}} = \pi^c\left(c_{t+1}|\alpha^t\right)$ gives condition (119).

Turning to condition (120), I have already established:

$$\frac{1}{s_t}\left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}} = \varepsilon_t^s \frac{1}{\alpha_t \frac{dc_t}{d\alpha_t}} \int_{\alpha_{t+1}} \frac{\beta\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{170}$$

So:

$$RT_t'\left(s_t\right)\left.\frac{ds_t}{dM_{t+1}}\right|_{\text{comp}} = \frac{\lambda_{t+1}^\Delta}{\eta_t}\alpha_t u'\left(c_t\right)\int_{\alpha_{t+1}} \frac{\beta R\alpha_{t+1}u'\left(c_{t+1}\right)}{\alpha_t u'\left(c_t\right)}\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{171}$$

$$= \frac{\lambda_{t+1}^\Delta}{\eta_t}\int_{\alpha_{t+1}} \beta R\alpha_{t+1}u'\left(c_{t+1}\right)\left[1+\frac{c_{t+1}u''\left(c_{t+1}\right)}{u'\left(c_{t+1}\right)}\frac{\alpha_{t+1}}{c_{t+1}}\frac{dc_{t+1}}{d\alpha_{t+1}}\right]\rho\left(\alpha_{t+1}|\alpha_t\right)\pi\left(\alpha_{t+1}|\alpha_t\right)d\alpha_{t+1} \tag{172}$$

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

C Proofs from Sections 9 and 10

C.1 Proof of Theorem 2

Equation (118), above, gives:

$$T_t'\left(s_t\right)\frac{ds_t}{dM_t} = -\frac{\lambda_{t+1}^\Delta\left(\alpha_t\right)}{\eta_t}\left(\alpha_t\right)^2 u''\left(c_t\right)\left(\frac{dc_t}{d\alpha_t}\right) \tag{173}$$

The utility function is time-separable and concave in consumption at each date-state, which together imply $\frac{ds_t}{dM_t} > 0$. Concavity further gives $u''\left(c_t\right) < 0$, strict normality implies $\frac{dc_t}{d\alpha_t} > 0$, and $\eta_t > 0$ from (18). It follows that $T_t'\left(s_t\right)$ has the same sign as $\lambda_{t+1}^\Delta\left(\alpha_t\left(c_t\right)\right)$.

A combination of (71) and (72) gives:

$$\lambda_{t+1}^\Delta\left(\alpha^t\right) - \lambda_t^\Delta\left(\alpha^{t-1}\right)\rho\left(\alpha_t|\alpha_{t-1}\right) = -\frac{1}{\alpha_t\pi\left(\alpha_t|\alpha_{t-1}\right)}\int_{\alpha_t}^{\bar{\alpha}}\left[\left(1+\lambda_{t+1}\left(\tilde{\alpha}_t\right)\right)-\left(1+\lambda_t\right)\right]\pi\left(\tilde{\alpha}_t|\alpha_{t-1}\right)d\tilde{\alpha}_t \tag{174}$$

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

$$\lambda_{t+1}^{\Delta}(\alpha_t)\alpha_t\pi(\alpha_t|\alpha_{t-1}) = \int_{\alpha_t}^{\bar{\alpha}}\left[\left(1+\lambda_t+\lambda_t^{\Delta}\pi^{\Delta}(\tilde{\alpha}_t|\alpha_{t-1})\right)-(1+\lambda_{t+1}(\tilde{\alpha}_t))\right]\pi(\tilde{\alpha}_t|\alpha_{t-1})\,d\tilde{\alpha}_t \tag{175}$$

with:

$$\int_{\underline{\alpha}}^{\bar{\alpha}}\left[\left(1+\lambda_t+\lambda_t^{\Delta}\pi^{\Delta}(\alpha_t|\alpha_{t-1})\right)-(1+\lambda_{t+1}(\alpha_t))\right]\pi(\alpha_t|\alpha_{t-1})\,d\alpha_t = 0 \tag{176}$$

Since $\pi^\Delta(\alpha_t|\alpha_{t-1})$ is monotone increasing in α_t, and $\lambda_0^\Delta = 0$, a sufficient condition for the right-hand side of (175) to be weakly positive for all t and all histories is that $(1+\lambda_{t+1}(\alpha_t))$ should be non-increasing in α_t. I have:

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

where convergence is uniform across α_t: $\frac{\mathbb{E}_t\left[D_{t,s}(\alpha^s)\alpha_s\right]}{\mathbb{E}_t[\alpha_s]} \leq \bar{\rho}^{s-t}$ where $\bar{\rho} = \sup_{\alpha,\alpha'}[\rho(\alpha'|\alpha)] < 1$. Thus I wish to show non-increasingness in:

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

$\mathbb{E}_t[\alpha_s]$ is weakly increasing in α_t, and invariant in α_t at the limit as s becomes large, 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. So consider the difference:

$$\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t'',\ldots,\alpha_s\right)\right)}\middle|\alpha_t''\right]-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',\ldots,\alpha_s\right)\right)}\middle|\alpha_t'\right] \tag{177}$$

for $\alpha_t'' > \alpha_t'$. I wish to show that this is nonpositive for s sufficiently large. I have:

$$\begin{aligned}&\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t'',\ldots,\alpha_s\right)\right)}\middle|\alpha_t''\right]-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',\ldots,\alpha_s\right)\right)}\middle|\alpha_t'\right]\\ &=\left\{\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t'',\ldots,\alpha_s\right)\right)}\middle|\alpha_t''\right]-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',\ldots,\alpha_s\right)\right)}\middle|\alpha_t''\right]\right\}\\ &+\left\{\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',\ldots,\alpha_s\right)\right)}\middle|\alpha_t''\right]-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',\ldots,\alpha_s\right)\right)}\middle|\alpha_t'\right]\right\}\end{aligned} \tag{178}$$

The first difference term is weakly negative, by normality. For the second, succesive integrations by parts yields:

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

$$=\int_{\alpha_t'}^{\alpha_t''}\frac{1}{\alpha_t}\left\{\sum_{r=t+1}^{s-1}\mathbb{E}_t\left[D_{t,r}\left(\alpha^r\right)\alpha_r\frac{d}{d\alpha_r}\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',...,\alpha_r,...,\alpha_s\right)\right)}\right]\middle|\alpha_t\right]\right\}d\alpha_t \quad (179)$$

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

By normality, the terms in the second line here are non-positive. For the last, I can write:

$$\mathbb{E}_t\left[D_{t,s-1}\left(\alpha^{s-1}\right)\int_{\alpha_s}\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',...,\alpha_r,...,\alpha_s\right)\right)}\right]\pi^{\Delta}\left(\alpha_s|\alpha_{s-1}\right)\pi\left(\alpha_s|\alpha_{s-1}\right)d\alpha_s\middle|\alpha_t\right]$$

$$\leq\bar{\rho}^{s-1-t}\kappa\mathbb{E}_t\left[\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',...,\alpha_r,...,\alpha_s\right)\right)}\right]\middle|\alpha_t\right] \quad (180)$$

$$=\bar{\rho}^{s-1-t}\kappa\left\{\mathbb{E}_t\left[\alpha_s\right]\frac{1+\lambda_{t+1}}{\eta_t}+\mathbb{E}_t\left[D_{t,s}\left(\alpha^s\right)\alpha_s\right]\frac{\lambda_{t+1}^{\Delta}}{\eta_t}\right\} \quad (181)$$

where the last line uses Proposition 6 in Appendix D, and κ is the upper bound on $\pi^{\Delta}\left(\alpha_s|\alpha_{s-1}\right)$ implied by Assumption 4. Since $\bar{\rho}<1$, this term converges to zero uniformly in s. Hence:

$$\lim_{s\to\infty}\left\{\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t'',...,\alpha_s\right)\right)}\middle|\alpha_t''\right]-\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',...,\alpha_s\right)\right)}\middle|\alpha_t'\right]\right\}\leq 0 \quad (182)$$

From this, I have:

$$\lim_{s\to\infty}\left\{\frac{1}{\mathbb{E}_t\left[\alpha_s|\alpha_t''\right]}\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t'',...,\alpha_s\right)\right)}\middle|\alpha_t''\right]\right\}-\frac{1}{\mathbb{E}_t\left[\alpha_s|\alpha_t'\right]}\lim_{s\to\infty}\left\{\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t}u'\left(c_s\left(\alpha_t',...,\alpha_s\right)\right)}\middle|\alpha_t'\right]\right\}\leq 0 \quad (183)$$

and so:

$$\frac{1+\lambda_{t+1}\left(\alpha_t''\right)}{\eta_t}-\frac{1+\lambda_{t+1}\left(\alpha_t'\right)}{\eta_t}\leq 0 \quad (184)$$

Thus $1+\lambda_{t+1}\left(\alpha_t\right)$ is weakly decreasing in α_t, for all t and α^{t-1}, implying that $\lambda_{t+1}^{\Delta}\left(\alpha_t\right)\geq 0$. This leaves two possibilities:

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

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

Case 1. implies $\lambda_{t+1}^{\Delta}(\alpha_t) > 0$ everywhere except endpoints, from (175) and the fact that the integral in (175) is zero over the full range, and so strictly positive taxes except at endpoints.

It remains to rule out case 2. If this were true but $\lambda_t^{\Delta} > 0$, persistence can still imply $\lambda_{t+1}^{\Delta} > 0$. The conclusion will only be affected if it implies $\lambda_{t+1}^{\Delta} = 0$ everywhere – and so zero taxes for interior values of α_t. Suppose this were true. From the solution for λ_{t+1}^{Δ}, (81), the implication is:

$$\frac{1}{\alpha_t u'(c_t)} = \frac{1}{\mathbb{E}_t[\alpha_{t+1}]}\mathbb{E}_t\left[\frac{1}{\beta R u'(c_{t+1})}\right] \tag{185}$$

whilst constancy of $(1+\lambda_{t+1}(\alpha_t))$ implies that both sides of this equation are constant in α_t, from equation (80). Suppose for now that types are persistent ($\rho(\alpha|\alpha') > 0$ for all interior α, all α'). For the right-hand side of (185) to be constant in α_t, and given nomality (i.e. that c_{t+1} is falling in α_t), a necessary requirement is that the partial derivatives due to persistence are weakly positive for all α_t:

$$\int_{\alpha_{t+1}}\left\{\frac{1}{u'(c_{t+1})} - \frac{\alpha_{t+1}}{\mathbb{E}_t[\alpha_{t+1}]}\mathbb{E}_t\left[\frac{1}{u'(c_{t+1})}\right]\right\}\frac{\frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t}}{\pi(\alpha_{t+1}|\alpha_t)}\pi(\alpha_{t+1}|\alpha_t)\,d\alpha_{t+1} \geq 0 \tag{186}$$

But since $\lambda_{t+1}^{\Delta} = 0$, and we have already established $\lambda_s^{\Delta} \geq 0$ for all s, condition (84) implies:

$$\int_{\alpha'_{t+1}}^{\bar{\alpha}}\left\{\frac{1}{u'(c_{t+1})} - \alpha_{t+1}\frac{1+\lambda_{t+1}}{\eta_t}\right\}\pi(\alpha_{t+1}|\alpha_t)\,d\alpha_{t+1} \leq 0 \tag{187}$$

for all α'_{t+1}, with:

$$\frac{1+\lambda_{t+1}}{\eta_t} = \frac{1}{\mathbb{E}_t[\alpha_{t+1}]}\mathbb{E}_t\left[\frac{1}{u'(c_{t+1})}\right] \tag{188}$$

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:

$$\int_{\alpha_{t+1}}\left\{\frac{1}{u'(c_{t+1})} - \frac{\alpha_{t+1}}{\mathbb{E}_t[\alpha_{t+1}]}\mathbb{E}_t\left[\frac{1}{u'(c_{t+1})}\right]\right\}\frac{\frac{d\pi(\alpha_{t+1}|\alpha_t)}{d\alpha_t}}{\pi(\alpha_{t+1}|\alpha_t)}\pi(\alpha_{t+1}|\alpha_t)\,d\alpha_{t+1} \leq 0 \tag{189}$$

with the inequality strict (a contradiction) unless the object in curly brackets is zero everywhere. But this is easily seen to imply $\lambda_{t+2}^{\Delta} = 0$ everywhere, and by induction $\lambda_{t+s}^{\Delta} = 0$ at all successor nodes. This implies a first-best allocation, with $\alpha_t u'(c_t)$ constant over time and histories, which is 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:

$$\frac{1}{\alpha_t u'(c_t)} = \frac{1}{\mathbb{E}_t[\alpha_s]}\mathbb{E}_t\left[\frac{1}{(\beta R)^{s-t} u'(c_s)}\right] \tag{190}$$

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. This completes the proof.

C.2 Proof of Proposition 4

The proof of Theorem 2 established that $T_t'(s_t(\alpha_t)) = 0$ when $\lambda_{t+1}^{\Delta}(\alpha_t) = 0$. Equation (84) gives:

$$\frac{1}{\pi(\alpha_t'|\alpha_{t-1})}\int_{\alpha_t'}^{\bar{\alpha}}\left\{\frac{(\beta R)^{-t}\eta}{u'(c_t(\alpha^{t-1},\alpha_t))} - \alpha_t\left[1+\lambda_t\left(\alpha^{t-1}\right)+\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\rho(\alpha_t|\alpha_{t-1})\right]\right\}\pi(\alpha_t|\alpha_{t-1})\,d\alpha_t \tag{191}$$
$$= -\left(\alpha_t'\right)^2\left\{\lambda_{t+1}^{\Delta}\left(\alpha^t\right) - \rho\left(\alpha_t'|\alpha_{t-1}\right)\lambda_t^{\Delta}\left(\alpha^{t-1}\right)\right\}$$

When $\alpha_t' = \underline{\alpha}$, condition (76) implies that the integral is zero, and since $\pi\left(\underline{\alpha}|\alpha_{t-1}\right)$ is positive the right-hand side must equal zero. A positive value for $\pi\left(\underline{\alpha}|\alpha_{t-1}\right)$ also implies that $\rho\left(\underline{\alpha}|\alpha_{t-1}\right) = 0$, and the result follows. An identical proof applies for $\alpha_t' = \bar{\alpha}$.

C.3 Proof of Proposition 5

From Lemma 3, $RT_{t-1}'(s_{t-1})s_{t-1}\epsilon_{t-1,t}\left(s_t'\right)$ is equal to:

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

Switching to express arguments in terms of α_t, this equals:

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

Integration by parts gives the following relationship:

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

So:

$$s_{t-1}\epsilon_{t-1,t}\left(s_t'\right)=\beta\frac{\left(\frac{\lambda_t^{\Delta}}{\eta_{t-1}}\right)}{T_{t-1}'\left(s_{t-1}\right)}\mathbb{E}_{t-1}\left[\alpha_t u'\left(c_t\left(\alpha_t\right)\right)\cdot\frac{\alpha_{t-1}\frac{d\pi\left(\alpha_t|\alpha_{t-1}\right)}{d\alpha_{t-1}}}{\pi\left(\alpha_t|\alpha_{t-1}\right)}\middle|\alpha_t\leq\alpha_t'\right] \tag{193}$$

The objects $\alpha_t u'\left(c_t\left(\alpha_t\right)\right)$ and $\frac{\alpha_{t-1}\frac{d\pi\left(\alpha_t|\alpha_{t-1}\right)}{d\alpha_{t-1}}}{\pi\left(\alpha_t|\alpha_{t-1}\right)}$ are non-decreasing in α_t. Of these, $\frac{\frac{d\pi\left(\alpha_t|\alpha_{t-1}\right)}{d\alpha_{t-1}}}{\pi\left(\alpha_t|\alpha_{t-1}\right)}$ is zero in expectation, and so crosses zero once, whilst $\alpha_t u'\left(c_t\left(\alpha_t\right)\right)$ is strictly positive. The proof of Theorem 2 established that $\frac{\lambda_t^{\Delta}}{\eta_{t-1}}$ and $T_{t-1}'\left(s_{t-1}\right)$ will both be strictly positive for the case of interest. Thus for sufficiently low α_t' both sides of the expression must be negative, whist positive correlation between the components implies it is positive for sufficiently high α_t'. If positive for α_t', the expectation term must remain positive for $\alpha_t''>\alpha_t'$, since $\frac{\frac{d\pi\left(\alpha_t|\alpha_{t-1}\right)}{d\alpha_{t-1}}}{\pi\left(\alpha_t|\alpha_{t-1}\right)}$ must be positive on the higher type range. The result follows.

D The dynamics of consumption

The following auxiliary result is used in the proof of Theorem 2:

Proposition 6. *For all t and s, $s\geq t$, and any history α^t, the period-t expected value of the period-s inverse marginal utility of consumption satisfies:*

$$\frac{1}{\mathbb{E}_t\left[\alpha_s\right]}\mathbb{E}_t\left[\frac{1}{\left(\beta R\right)^{s-t}u'\left(c_s\right)}\right]=\frac{1+\lambda_{t+1}}{\eta_t}+\frac{\mathbb{E}_t\left[D_{t,s}\left(\alpha^s\right)\alpha_s\right]}{\mathbb{E}_t\left[\alpha_s\right]}\frac{\lambda_{t+1}^{\Delta}}{\eta_t} \tag{194}$$

Proof. Conditions (80) and (81) immediately give the result for for $s=t$ and $s=t+1$:

$$\frac{1}{\alpha_t u'\left(c_t\right)}=\frac{1+\lambda_{t+1}+\lambda_{t+1}^{\Delta}}{\eta_t} \tag{195}$$

$$\frac{1}{\mathbb{E}_t\left[\alpha_{t+1}\right]}\mathbb{E}_t\left[\frac{1}{\beta R u'\left(c_{t+1}\right)}\right]=\frac{1+\lambda_{t+1}+\lambda_{t+1}^{\Delta}\varepsilon^{\alpha}\left(\alpha_t\right)}{\eta_t} \tag{196}$$

and note that $\varepsilon^{\alpha}(\alpha_t) = \frac{\mathbb{E}_t\left[D_{t,t+1}(\alpha^{t+1})\alpha_{t+1}\right]}{\mathbb{E}_t[\alpha_{t+1}]}$. The proof then works recursively. Suppose that, for $r < s$:

$$\mathbb{E}_r\left[\frac{\eta_r}{(\beta R)^{s-r} u'(c_s)}\right] = [1+\lambda_{r+1}]\,\mathbb{E}_r[\alpha_s] + \lambda_{r+1}^{\Delta}\mathbb{E}_r\left[D_{r,s}(\alpha^s)\alpha_s\right] \tag{197}$$

Then:

$$\mathbb{E}_{r-1}\left[\frac{\eta_{r-1}}{(\beta R)^{s-r+1} u'(c_s)}\right] = \mathbb{E}_{r-1}\left\{[1+\lambda_{r+1}]\,\mathbb{E}_r[\alpha_s] + \lambda_{r+1}^{\Delta}\mathbb{E}_r\left[D_{r,s}(\alpha^s)\alpha_s\right]\right\} \tag{198}$$

$$= \int_{\alpha_r}\left\{\left[\rho(\alpha_r|\alpha_{r-1})\lambda_r^{\Delta} - \frac{1}{\alpha_r\pi(\alpha_r|\alpha_{r-1})}\int_{\alpha_r}^{\bar{\alpha}}\mu_r(\tilde{\alpha}_r)\pi(\tilde{\alpha}_r|\alpha_{r-1})\,d\tilde{\alpha}_r\right]\cdot\mathbb{E}_r\left[D_{r,s}(\alpha^s)\alpha_s\right] + [1+\lambda_r+\mu_r(\alpha_r)]\cdot\mathbb{E}_r[\alpha_s]\right\}\pi(\alpha_r|\alpha_{r-1})\,d\alpha_r \tag{199}$$

By an identical argument to Lemma 2, I have:

$$\frac{d}{d\alpha_r}\left[\mathbb{E}_r[\alpha_s]\right] = \frac{1}{\alpha_r}\mathbb{E}_r\left[D_{r,s}(\alpha^s)\alpha_s\right] \tag{200}$$

So integrating by parts and using (73), I have:

$$\int_{\alpha_r}\frac{1}{\alpha_r}\mathbb{E}_r\left[D_{r,s}(\alpha^s)\alpha_s\right]\left[\int_{\alpha_r}^{\bar{\alpha}}\mu_r(\tilde{\alpha}_r)\pi(\tilde{\alpha}_r|\alpha_{r-1})\,d\tilde{\alpha}_r\right]d\alpha_r = \int_{\alpha_r}\mathbb{E}_r[\alpha_s]\,\mu_r(\tilde{\alpha}_r)\,\pi(\alpha_r|\alpha_{r-1})\,d\alpha_r \tag{201}$$

Using this in (199), the terms in μ_r cancel, and I have:

$$\mathbb{E}_{r-1}\left[\frac{\eta_{r-1}}{(\beta R)^{s-r+1} u'(c_s)}\right] = [1+\lambda_r]\cdot\mathbb{E}_{r-1}[\alpha_s] + \lambda_r^{\Delta}\cdot\mathbb{E}_{r-1}\left[D_{r-1,s}(\alpha^s)\alpha_s\right] \tag{202}$$

which makes use of the definition of $D_{r-1,s}(\alpha^s)$. Thus I have iterated expectations backwards a period from condition (197). Now, for any $s > 0$, condition (196) implies:

$$\mathbb{E}_{s-1}\left[\frac{\eta_{s-1}}{\beta R u'(c_s)}\right] = [1+\lambda_s]\,\mathbb{E}_{s-1}[\alpha_s] + \lambda_s^{\Delta}\mathbb{E}_{s-1}\left[D_{s-1,s}(\alpha^s)\alpha_s\right] \tag{203}$$

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