Static and Dynamic Mirrleesian Taxation with Non-separable Preferences: A Unified Approach*

Christian Hellwig
Toulouse School of Economics

June 10, 2021

Abstract

I analyze dynamic Mirrlees taxation with preferences that are non-separable between consumption, leisure and type, which determines both ability and consumption needs. I show how to account for non-separable preferences through a simple change in probability measures. I generalize the existing Inverse Euler Equation and optimal static labor tax formulae and provide a unified intuition based on a set of perturbations around the optimal allocations that preserve expected utility and incentive compatibility. Non-separability in preferences gives rise to a new tradeoff between current and future redistribution that is internalized by the planner's solution but not by private savings decisions. This leads to a novel rationale to subsidize (tax) savings and make labor taxes more (less) persistent, when more productive agents also have higher (lower) consumption needs.

1 Introduction

How should optimal taxes be structured to balance redistribution motives and efficiency distortions? Starting with the seminal work of Mirrlees (1971), the existing literature on optimal income taxes captures the efficiency-redistribution tradeoff through a basic asymmetric information friction: workers have private information about their ability or inclination to work.[1] Dynamic life cycle economies augment the tax design problem by a dynamic insurance motive through stochastic evolution of individual types (i.e. abilities or preferences) and aggregate savings.

*I thank Fabrice Collard, Tuuli Vanhapelto and Nicolas Werquin for useful discussions and seminar participants at Toulouse School of Economics for comments. Funding from the French National Research Agency (ANR) under grant ANR-17-EURE-0010 (Investissements d'Avenir) is gratefully acknowledged.

[1] see Mirrlees (1971), Diamond (1998), Saez (2001), among many others.

Existing results on dynamic optimal tax design leverage an assumption that preferences are separable, and in particular that the marginal utility of consumption is independent of leisure or ability types, into a sharp characterization of optimal labor and savings distortions.[2] Optimal savings distortions are characterized by the Inverse Euler Equation, which captures the idea that wealth accumulation has adverse effects on future incentives to work (see e.g. Diamond and Mirrlees, 1978, and Rogerson, 1985, and Golosov, Kocherlakota and Tsyvinski, 2003). In parallel, Farhi and Werning (2013) and Golosov, Troshkin and Tsyvinski (2016) have characterized optimal labor wedges emphasizing smoothing and backloading of tax distortions.

The separability assumption plays a key role in limiting the interaction between static redistribution and dynamic savings margins, which greatly facilitates the characterization of each of the two margins separately. At the same time it imposes strong restrictions on substitution between home production and market work, across periods or wealth effects on labor supply, which in turn limit the scope of applicability for many questions of applied interest involving life-cycle savings, home production, taxation of couples, or insurance against important life cycle risks. Aguiar and Hurst (2005, 2007) document patterns of substitution between market and non-market work and consumption in both the time series and cross section which directly contradict this hypothesis.[3]

I characterize optimal labor and savings wedges in a dynamic Mirrleesian economy with *arbitrary* non-separability in preferences between consumption, types and leisure.[4] I show three main results:

First, I identify a new rationale for savings taxes or subsidies based on the effect of savings on *current* redistribution. This rationale is conceptually distinct from the wealth effect of savings on future incentives that was captured by the original Inverse Euler Equation, and it disappears when preferences are separable. Furthermore, when it is optimal to subsidize savings, the optimal allocation features mean-reverting social mobility and overturns the immiseration property that often characterizes optimal social mobility in dynamic private information economies.[5]

[2]see e.g. Kocherlakota, Golosov and Tsyvinski (2003), Kocherlakota (2005), Albanesi and Sleet (2006), Golosov and Tsyvinski (2006), Farhi and Werning (2013), Golosov, Troshkin, and Tsyvinski (2016), or Stantcheva (2018) for important contributions and Stantcheva (2020) for a recent review of the literature.

[3]see also Aguiar, Hurst and Karabarbounis (2012) for a review and further discussion of consumption patterns and intra- and inter-temporal substitution between time use and consumption goods at the micro level. Benhabib, Rogerson and Wright (1991) have shown how allowing for substitution between home and market production improves upon some key predictions of standard real business cycle models.

[4]Grochulski and Kocherlakota (2010), Köhne and Kuhn (2015) and Köhne (2018) extend the dynamic Mirrlees model to allow for non-time separable preferences through the introduction of habits or durable consumption goods, but they retain the separability of consumption and effort.

[5]See for example Thomas and Worrall (1990), or Atkeson and Lucas (1992).

Savings affect current information rents and redistribution both directly and indirectly because a change in future redistribution also feeds back into current incentives. These two effects go in opposite directions but the direct effect dominates unless types are highly persistent, and in expectation, the effect of savings on current redistribution always dominates the wealth effect of savings on future incentives. If consumption needs or marginal utilities are increasing with the worker's ability type, savings reduce current information rents and facilitate current redistribution: this makes it optimal to subsidize savings. When instead low-ability types have higher consumption needs, it becomes optimal to tax savings. The magnitude of this savings wedge is increasing in the optimal concurrent labor tax and decreasing in the persistence of types.

Second, I provide a new characterization of optimal labor wedges based on an arbitrage between redistribution of consumption and redistribution of leisure. This arbitrage highlights that the optimal allocation doesn't just trade off between incentives and redistribution, but also between the channels through which redistribution takes place: is it more efficient to tax top income earners by asking them to work harder, or by asking them to consume less?

The optimal allocation equalizes the marginal cost of redistribution to the marginal benefit of redistribution through consumption and to the marginal benefit of redistribution through leisure, for each type, i.e. at the optimum, the planner is indifferent between increasing the work load, or reducing the consumption, of the most productive types. This simple observation leads to a natural complement of existing formulae for optimal income taxes from the static Mirrlees model (e.g. Diamond 1998, Saez, 2001) or the dynamic tax formula in Golosov, Troshkin and Tsyvinski (2016), which applies in both static and dynamic economies, and is easy to link to sufficient statistics from the distribution of earnings and consumption.

The marginal cost of redistribution is given by the ratio of the marginal efficiency loss from the labor distortion to the reduction in marginal information rents for a given type. The marginal benefits of redistribution consist of a static component and a dynamic component. The static component is determined by perturbations that transfer consumption or leisure from higher to lower types while preserving expected utility and incentive compatibility. The additional dynamic benefit of tax distortions arises because a commitment to future labor taxes and redistribution also reduces current information rents, thus facilitating current redistribution. The optimal labor wedge arbitrages distortions over time, while internalizing that future labor distortions generate an extra benefit from raising future redistribution. As in Farhi and Werning (2013), the optimal labor taxes are persistent and backloaded over time or with age.

Third, when preferences are non-separable, labor tax-smoothing interacts with the optimal savings wedge, since the latter determines the planner's discount rate between current and future costs and benefits fo redistribution. Taxing or subsidizing savings to increase current redistribution goes hand in hand with increasing or reducing the persistence of labor tax distortions. In addition, the feedback from future to current incentives and redistribution also affects optimal persistence of labor taxes directly.

All my results derive from the following key technical insight: Non-separability in preferences enters the characterization of optimal allocations through a change in the probability measure of types, or Radon-Nikodym derivative, that can be interpreted as an "incentive adjustment" to Pareto weights that is required to respect incentive compatibility. If a local perturbation of consumption allocations relaxes (tightens) incentive compatibility constraints for higher types, then the perturbation increases (reduces) the scope for further redistribution. The change in probability measures captures this through a first-order shift of probability weights towards lower (higher) types. This effect disappears with separable preferences because local perturbations of consumption impact incentive constraints only locally, so the original and incentive-adjusted Pareto weights are the same. A similar change in probabilities results from perturbations to leisure, but the latter always shifts weights towards lower ability types. The change in probability measure appears in some existing characterizations of the optimal labor wedge, but doesn't appear to have been interpreted as such.[6]

This characterization of optimal allocations in turn allows me to interpret optimal labor and savings wedges through elementary perturbations that capture the marginal costs and benefits of redistribution, or the private and social marginal returns to savings, thus providing a unified intuition and generalization of many existing results to arbitrary preferences: Optimal savings wedges are characterized by a generalized Inverse Euler Equation which incorporates the wealth effect of savings on future incentives through the change in probabilities, and adds a static savings wedge to account for the effect of savings on current period incentives. Optimal labor wedges are captured by a dual representation of the costs and benefits of redistribution through either consumption and leisure, which are each governed by their respective change in probabilities. Their combination highlights a novel two-way interaction that arises from concerns for redistribution that are internalized by the planner's solution but not by private decisions. The magnitude of the optimal savings wedge is linked to the magnitude of current labor wedges since the latter is used to "price" current redistribution, and the persistence of labor wedges is linked to the sign and level of the

[6] See, e.g. Hellwig (2007) or for the static Mirrlees model with non-separable preferences, or Golosov, Troshkin and Tsyvinski (2016) for a similar characterization in the dynamic model with non-separable preferences.

savings wedge since the latter determines the optimal tradeoff between current and future resources and efficiency distortions.

From an applied perspective, these results substantially expand the applicability of optimal taxation results. Whenever households tradeoff between market and non-market work or consumption, or when labor supply decisions coordinate multiple household members across different activities and interact with life-cycle savings, it is important to know whether non-separability of preferences makes it easier or harder to provide insurance without distorting incentives. The present analysis answers this question by showing that the tradeoff shifts towards positive savings taxes, lower but more persistent labor taxes, and more redistribution, when needs-based and ability-based redistribution motives are aligned (low types have low ability and high needs), but towards savings subsidies, higher but less persistent labor taxes, and less redistribution in the opposite case (low ability types also have low consumption needs). Understanding how different redistribution motives interact thus becomes crucially important for any quantitative assessment of dynamic tax and insurance policies.

Section 2 introduces the model. Section 3 develops the core idea of incentive-adjusted probabilities. Section 4 reviews the benchmark results with separable preferences. Section 5 discusses optimal savings distortions, section 6 optimal labor wedges. Section 7 concludes with examples.

2 The Model

There are $T < \infty$ periods, and there is a measure-1 continuum of agents. Each agent draws a type sequence $\theta^T \in \left[\underline{\theta}, \overline{\theta}\right]^T$ iid across agents according to cdf $F(\cdot)$ (and pdf $f(\cdot)$) at date 0 (prior to the first period), and $\theta_t \in \left[\underline{\theta}, \overline{\theta}\right]$ is then revealed to the agent at the beginning of period t; in period t, the agent thus knows θ^t. Agents' preferences over consumption and earnings are defined by

$$\mathbb{E}_0 \left\{ \sum_{t=1}^{T} \beta^{t-1} U\left(C_t, Y_t; \theta_t\right) \right\}$$

i.e. an agent's utility in period t is a function of his current type θ_t, current consumption C_t and current earnings Y_t. The function U is twice continuously differentiable, $U_C > 0$, $U_{CC} < 0$, $U_Y < 0$, $U_{YY} < 0$, $U_\theta > 0$ and U otherwise satisfies the usual Inada conditions as C or Y approach 0 or ∞. Consumption and earnings are assumed to be observable but individual types are the agents' private information. Types are private information, but the social planner keeps a record of the agents' past announcements and can borrow or save with a return $R > 0$.

I assume that $f(\cdot)$ is positive over the interior of its support, that $f\left(\theta_t | \theta^{t-1}\right)$ only depends on

t and θ_{t-1} (i.e. θ^t follows a time-dependent Markov chain), conditional densities $f\left(\theta_t|\theta^{t-1}\right)$ are differentiable w.r.t. θ_{t-1}, and that "types are persistent", i.e.

$$\mathcal{J}\left(\theta_t,\theta^{t-1}\right)\equiv\frac{\frac{\partial f\left(\theta_t|\theta^{t-1}\right)}{\partial\theta_{t-1}}}{f\left(\theta_t|\theta^{t-1}\right)}$$

is non-decreasing in θ_t. This formulation includes types that follow a generic linear AR(1) process in which $\frac{\partial f\left(\theta_t|\theta^{t-1}\right)}{\partial\theta_{t-1}}=-\rho\frac{\partial f\left(\theta_t|\theta^{t-1}\right)}{\partial\theta_t}$, with useful polar cases $\rho=0$ (types are independent across time, no persistence) and $\rho=1$ (types follow a random walk, full persistence).

I make a few additional assumptions about $U\left(C,Y;\theta\right)$. The first is the *Strict Single-Crossing Condition:* $-U_Y\left(C,Y;\theta\right)/U_C\left(C,Y;\theta\right)$ is strictly decreasing in θ for all $(C,Y;\theta)$, or

$$\frac{U_{C\theta}}{U_C}-\frac{U_{Y\theta}}{U_Y}>0.$$

This assumption guarantees monotonicity of any incentive-compatible allocation in the static Mirrlees screening problem: on the margin, higher types are more willing to work.

Next, I assume that both consumption and leisure (or earnings) are normal goods, or that the elasticities

$$\mathcal{E}_C\equiv\frac{\partial ln\left(-U_Y/U_C\right)}{\partial lnC}=-\left(\frac{U_{CC}}{U_C}-\frac{U_{CY}}{U_Y}\right)C$$
$$\mathcal{E}_Y\equiv\frac{\partial ln\left(-U_Y/U_C\right)}{\partial lnY}=\left(\frac{U_{YY}}{U_Y}-\frac{U_{CY}}{U_C}\right)Y.$$

are both non-negative for all $(C,Y;\theta)$. These two elasticities will play a key role in my analysis.

Finally, I make assumptions about the planner's motives for redistribution. The signs of $U_{C\theta}$ and $U_{Y\theta}$ play a key role in determining how non-separabilities affect the optimal tax design problem. For illustrative purposes, I will occasionally use the following class of *"weakly separable"* preferences:

$$U\left(C,Y;\theta\right)=\mathcal{U}\left(\gamma\left(C,\underline{C}\left(\theta\right)\right)-n\left(Y,\theta\right)\right)$$

The utility aggregator $\mathcal{U}\left(\cdot\right)$ is strictly increasing and concave. The function $n\left(Y,\theta\right)$ can be interpreted as (a disutility of) hours worked to generate earnings Y.[7] I assume that $n_Y>0$, $n_\theta<0$, and $n_{\theta Y}<0$, so that θ can be interpreted as the agent's labor productivity or disutility of effort, with higher types being more productive. This formulation captures redistribution of effort from less to more productive agents, or equivalently, redistribution of leisure towards less productive agents, i.e. redistribution *"from each according to his ability"*.

[7] While it is convenient for the analysis to define preferences in terms of the observables C and Y, it is straightforward to map the type-contingent preference over earnings into a preference over leisure or hours worked. I will thus refer interchangeably to redistribution of Y as redistribution of earnings or leisure.

The function $\gamma\left(C,\underline{C}\left(\theta\right)\right)$ can be interpreted as a consumption index that includes a type-dependent subsistence consumption $\underline{C}\left(\theta\right)$. I assume that $\gamma_C>0$, $\gamma_{\underline{C}}\leq 0$, and $\gamma_{\underline{C}C}\geq 0$. An increase in subsistence consumption $\underline{C}\left(\theta\right)$ then lowers the agent's utility and increases their marginal utility. This formulation thus introduces redistribution of consumption towards those with the highest marginal utilities, i.e. redistribution *"to each according to his needs"*.

With these assumptions, the cross-partials $U_{C\theta}$ and $U_{Y\theta}$ take the following form:

$$\begin{aligned}\frac{U_{C\theta}}{U_C} &= \frac{\mathcal{U}''}{\mathcal{U}'}\left(\gamma_{\underline{C}}\underline{C}'\left(\cdot\right)-n_\theta\right)+\frac{\gamma_{\underline{C}C}}{\gamma_C}\underline{C}'\left(\cdot\right)\\ \frac{U_{Y\theta}}{U_Y} &= \frac{\mathcal{U}''}{\mathcal{U}'}\left(\gamma_{\underline{C}}\underline{C}'\left(\cdot\right)-n_\theta\right)+\frac{n_{\theta Y}}{n_Y}.\end{aligned}$$

If preferences are separable between consumption and earnings ($\mathcal{U}''=0$), $U_{Y\theta}/U_Y$ captures the motive for redistribution based on ability, and $U_{C\theta}/U_C$ the needs-based redistribution motive. With the ordering assumption that higher types are more productive, it follows that $U_{Y\theta}/U_Y<0$.

The slope of $\underline{C}\left(\cdot\right)$ in turn determines how consumption needs are aligned with ability. If consumption needs are decreasing in type, then the two redistribution motives reinforce each other towards increasing consumption and lowering effort of the lowest types, but they generate opposite incentives effects: higher ability types have less consumption needs which reduces incentives to work. In this case $U_{C\theta}/U_C\leq 0$. The Single-Crossing Condition imposes that $0\geq U_{C\theta}/U_C>U_{Y\theta}/U_Y$, i.e. that ability has a stronger impact on work incentives than needs.

If consumption needs are instead increasing in type, $U_{C\theta}/U_C>0$. In this case, the Single-Crossing Condition holds automatically since needs- and ability-based incentives are both stronger for higher types. In this case, the planner has a motive of demanding higher effort from, and offering higher consumption to high types.

Non-separability in preferences between consumption and earnings ($\mathcal{U}''<0$) does not affect incentives, but strengthens both redistribution motives towards lower types, since $U_\theta=\mathcal{U}'\cdot\left(\gamma_{\underline{C}}\underline{C}'\left(\cdot\right)-n_\theta\right)>0$, i.e. higher types obtain higher utility from any pair (C,Y).

Throughout the paper I assume that $U_{Y\theta}/U_Y<0$, and I distinguish between the case where need-based redistribution is aligned with redistribution based on ability ($0\geq U_{C\theta}/U_C>U_{Y\theta}/U_Y$ for all $(C,Y;\theta)$) and the case where it is not ($U_{C\theta}/U_C\geq 0>U_{Y\theta}/U_Y$ for all $(C,Y;\theta)$).[8] In both cases I assume that $U_{C\theta}$ doesn't change sign, i.e. either $U_{C\theta}\geq 0$ for all $(C,Y;\theta)$, or $U_{C\theta}\leq 0$ for all $(C,Y;\theta)$. The standard model of ability-based redistribution ($U\left(C,Y;\theta\right)=U\left(C,N\right)$ with $N=Y/a\left(\theta\right)$ and $U_{CN}\geq 0$, where $a\left(\theta\right)$ denotes labor productivity) belong to the first case.

[8]It is straight-forward to extend the results to the alternative case in which need-based redistribution is the dominant channel for incentives, adapting the corresponding preference assumptions and their interpretation.

2.1 Social planner's problem

The utilitarian planner designs a sequence $\left\{C\left(\theta^t\right), Y\left(\theta^t\right)\right\}$ to maximize the agents' expected utility, subject to incentive compatibility and break-even conditions. I reformulate this problem as a problem of minimizing the cost of the planner's allocation, subject to promise-keeping and incentive compatibility:

$$K\left(v_0\right)=\min_{\left\{C\left(\theta^t\right), Y\left(\theta^t\right)\right\}} \mathbb{E}_0\left\{\sum_{t=1}^T R^{-t}\left(C\left(\theta^t\right)-Y\left(\theta^t\right)\right)\right\}, \text{ s.t.}$$

$$\mathbb{E}_0\left\{\sum_{t=1}^T \beta^{t-1} U\left(C\left(\theta^t\right), Y\left(\theta^t\right) ; \theta_t\right)\right\} \geq v_0$$

and

$$\mathbb{E}_0\left\{\sum_{t=1}^T \beta^{t-1} U\left(C\left(\theta^t\right), Y\left(\theta^t\right) ; \theta_t\right)\right\} \geq \mathbb{E}_0\left\{\sum_{t=1}^T \beta^{t-1} U\left(C\left(\overrightarrow{\theta}^t\left(\theta^t\right)\right), Y\left(\overrightarrow{\theta}^t\left(\theta^t\right)\right) ; \theta_t\right)\right\}$$

for all types $\theta^T \in[\underline{\theta}, \overline{\theta}]^T$ and announcement strategies $\overrightarrow{\theta}^T:[\underline{\theta}, \overline{\theta}]^T \rightarrow[\underline{\theta}, \overline{\theta}]^T$ that are measurable w.r.t. θ^t in period t.

To define the problem recursively, let

$$v_{t-1}\left(\theta^{t-1}\right)=\mathbb{E}\left\{\sum_{\tau=t}^T \beta^{\tau-t} U\left(C\left(\theta^\tau\right), Y\left(\theta^\tau\right) ; \theta_\tau\right) | \theta^{t-1}\right\}$$

$$w_t\left(\theta^t\right)=U\left(C\left(\theta^t\right), Y\left(\theta^t\right) ; \theta_t\right)+\beta v_t\left(\theta^t\right)$$

I relax incentive compatibility to local IC by which this inequality must only hold for $\theta^{\prime}$ sufficiently close to θ. Following Pavan, Segal and Toikka (2014), Farhi and Werning (2013) or Kapicka (2013), the local IC constraint is

$$\frac{\partial w_t\left(\theta^t\right)}{\partial \theta_t} \equiv \dot{w}_t\left(\theta^t\right)=U_\theta\left(C\left(\theta^t\right), Y\left(\theta^t\right) ; \theta_t\right)+\beta \Delta_t\left(\theta^t\right), \text{ where}$$

$$\Delta_t\left(\theta^t\right)=\int w_{t+1}\left(\theta^{t+1}\right) \frac{\partial f\left(\theta_{t+1} | \theta^t\right)}{\partial \theta_t} d \theta^{t+1}=\mathbb{E}\left\{w_{t+1}\left(\theta^{t+1}\right) \mathcal{J}\left(\theta_{t+1}, \theta^t\right) | \theta^t\right\}.$$

The term $\Delta_t\left(\theta^t\right)$ constitutes a promise or commitment towards future redistribution so as to limit dynamic marginal information rents, i.e. the part of the marginal information rent $\dot{w}_t\left(\theta^t\right)$ that an agent obtains at type θ^t, which is not embedded in the current utility U_θ. With time-independent types, $\mathcal{J}\left(\theta_t, \theta^{t-1}\right)=0$, so the local IC condition reduces to its static counterpart. In that case, incentive provision occurs only within, but not across periods, so the planner incurs no dynamic incentive commitments and Δ can be dropped from the set of state variables.[9]

Using $v = v_{t-1}\left(\theta^{t-1}\right)$ and $\Delta = \Delta_{t-1}\left(\theta^{t-1}\right)$ as state variables, the optimal allocation then minimizes the planner's net present value of transfers $K_t\left(v, \Delta, \theta^{t-1}\right)$ subject to promise keeping and local incentive compatibility constraints and satisfies the following recursive characterization:

$$K_t\left(v, \Delta, \theta^{t-1}\right) = \min \int R^{-1}\left\{C_t\left(\theta^t\right) - Y_t\left(\theta^t\right) + K_{t+1}\left(v_t\left(\theta^t\right), \Delta_t\left(\theta^t\right), \theta^t\right)\right\} f\left(\theta_t|\theta^{t-1}\right) d\theta_t, \text{ s.t.}$$

$$v = \int w_t\left(\theta^t\right) f\left(\theta_t|\theta^{t-1}\right) d\theta_t$$

$$\Delta = \int w_t\left(\theta^t\right) \mathcal{J}\left(\theta_t, \theta^{t-1}\right) f\left(\theta_t|\theta^{t-1}\right) d\theta_t$$

$$w_t\left(\theta^t\right) = U\left(C_t\left(\theta^t\right), Y_t\left(\theta^t\right); \theta_t\right) + \beta v_t\left(\theta^t\right)$$

$$\dot{w}_t\left(\theta^t\right) = U_\theta\left(C_t\left(\theta^t\right), Y_t\left(\theta^t\right); \theta_t\right) + \beta \Delta_t\left(\theta^t\right)$$

The initial state variables v_0 and Δ_0 are then chosen so that $\Delta_0 \in \arg\min_\Delta K_1\left(v_0, \Delta, \theta^{-1}\right)$ and $K\left(v_0\right) = K_1\left(v_0, \Delta_0, \theta^{-1}\right) = 0$, i.e. Δ_0 is set to minimize $K_1\left(v_0, \Delta, \theta^{-1}\right)$, meaning that at date 0 the planner's break-even constraint is satisfied and the planner is free of prior commitments.

This recursive control problem decomposes into a static optimal control problem using $w_t\left(\theta^t\right)$ as the state and $C_t\left(\theta^t\right)$, $Y_t\left(\theta^t\right)$ and $\Delta_t\left(\theta^t\right)$ as control variables and a dynamic programming problem to keep track of of utility promises $v_t\left(\theta^t\right)$ and incentive commitments $\Delta_t\left(\theta^t\right)$. The Hamiltonian is

$$\begin{aligned}
\mathcal{H}_t \quad = \quad & R^{-1}\left\{C_t\left(\theta^t\right) - Y_t\left(\theta^t\right) + K\left(v_t\left(\theta^t\right), \Delta_t\left(\theta^t\right), \theta^t, t+1\right)\right\} f\left(\theta_t|\theta^{t-1}\right) \\
& + \left\{\lambda_t\left(v - w_t\left(\theta^t\right)\right) + \eta_t\left(w_t\left(\theta^t\right)\mathcal{J}\left(\theta_t, \theta^{t-1}\right) - \Delta\right)\right\} f\left(\theta_t|\theta^{t-1}\right) \\
& + \psi_t\left(\theta^t\right)\left\{w_t\left(\theta^t\right) - U\left(C_t\left(\theta^t\right), Y_t\left(\theta^t\right); \theta_t\right) - \beta v_t\left(\theta^t\right)\right\} \\
& + \mu_t\left(\theta^t\right)\left\{U_\theta\left(C_t\left(\theta^t\right), Y_t\left(\theta^t\right); \theta_t\right) + \beta \Delta_t\left(\theta^t\right)\right\}
\end{aligned}$$

The first-order conditions w.r.t. $C_t\left(\cdot\right)$ and $Y_t\left(\cdot\right)$ yield

$$\psi_t\left(\theta^t\right) = \frac{R^{-1}}{U_C\left(\theta^t\right)} f\left(\theta_t|\theta^{t-1}\right) + \mu_t\left(\theta^t\right)\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} = \frac{R^{-1}}{-U_Y\left(\theta^t\right)} f\left(\theta_t|\theta^{t-1}\right) + \mu_t\left(\theta^t\right)\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}$$

where I use the notation $G\left(\theta^t\right) \equiv G\left(C_t\left(\theta^t\right), Y_t\left(\theta^t\right); \theta_t\right)$ for any function G of both the allocation $\left(C_t\left(\theta^t\right), Y_t\left(\theta^t\right)\right)$ and the current type θ_t, to ease notation. These first order conditions define a

[9] As a corollary, it follows that unless types are time-independent, the planner's solution is generically time-inconsistent, since the planner must keep track of prior incentive commitments that are summarized by Δ: at date t agents' incentives are influenced by how the planner promises to reallocate resources at future dates. But once those future dates arrive those prior incentive effects are "sunk".

shadow cost of increasing the utility of agents with type θ^t, $\psi_t\left(\theta^t\right)$, which consists of two components. The first term $f\left(\theta_t|\theta^{t-1}\right)R^{-1}/U_C\left(\theta^t\right)$ or $f\left(\theta_t|\theta^{t-1}\right)R^{-1}/(-U_Y\left(\theta^t\right))$ represents a direct shadow cost of increasing type θ utility through higher consumption or lower earnings. The second term measures how a direct consumption or earnings increase affects $U_\theta\left(\theta^t\right)$ and thereby tightens or relaxes the local incentive compatibility constraint at θ^t. The latter is weighted by the multiplier $\mu_t\left(\theta^t\right)$ and added to the former.

The first-order conditions for $v_t\left(\theta^t\right)$ and $\Delta_t\left(\theta^t\right)$, along with the envelope conditions at $t+1$, yield

$$\beta R\frac{\psi_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)} = \frac{\partial K_{t+1}}{\partial v_t\left(\theta^t\right)} = \lambda_{t+1}$$

$$\beta R\frac{\mu_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)} = -\frac{\partial K_{t+1}}{\partial \Delta_t\left(\theta^t\right)} = \eta_{t+1}$$

These conditions link the marginal costs of promised utility λ_{t+1} and incentive commitments η_{t+1} back in time to marginal costs of promised utility and shadow price of redistribution at θ^t. The initial incentive commitment Δ_0 is freely chosen, which results in $\eta_1 = -\frac{\partial K_1}{\partial \Delta_0} = 0$.

Finally, the multiplier $\mu_t\left(\theta^t\right)$ is given by the solution to the following linear ODE:

$$\begin{aligned}\dot{\mu}_t\left(\theta^t\right) &= -\frac{\partial \mathcal{H}_t}{\partial w_t\left(\theta^t\right)} = \lambda_t f\left(\theta_t|\theta^{t-1}\right) - \eta_t \mathcal{J}\left(\theta_t,\theta^{t-1}\right) f\left(\theta_t|\theta^{t-1}\right) - \psi_t\left(\theta^t\right) \\ &= \lambda_t f\left(\theta_t|\theta^{t-1}\right) - \eta_t \mathcal{J}\left(\theta_t,\theta^{t-1}\right) f\left(\theta_t|\theta^{t-1}\right) - \frac{R^{-1}}{U_C\left(\theta^t\right)} f\left(\theta_t|\theta^{t-1}\right) - \mu_t\left(\theta^t\right)\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}\end{aligned}$$

along with the boundary conditions $\mu\left(\underline{\theta}\right) = \mu\left(\overline{\theta}\right) = 0$, which imply that there are no distortions at the top and the bottom of the type distribution. The FOC for earnings yields an analogous ODE.

Let $\tau_t\left(\theta^t\right) = 1 + U_Y\left(\theta^t\right)/U_C\left(\theta^t\right)$ denote the labor wedge at θ^t, i.e. the distortion between the marginal product and the marginal rate of substitution between consumption and earnings. Combining the first-order conditions for $C_t\left(\cdot\right)$ and $Y_t\left(\cdot\right)$ yields the following static optimality condition:

$$\frac{R^{-1}}{U_C\left(\theta^t\right)}\frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)} = \frac{\mu_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)}\left(\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} - \frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}\right).$$

This static optimality condition succinctly summarizes the static trade-off between efficiency and redistribution at date 1. To substantiate this interpretation, consider a perturbation that leaves a given type θ^t agent's utility unchanged ($\Delta C_1\left(\theta^t\right) = \delta/U_C\left(\theta^t\right)$ and $\Delta Y_1\left(\theta^t\right) = \delta/\left(-U_Y\left(\theta^t\right)\right)$). The planner raises resources

$$\Delta Y_1\left(\theta^t\right) - \Delta C_1\left(\theta^t\right) = \left(\frac{1}{-U_Y\left(\theta^t\right)} - \frac{1}{U_C\left(\theta^t\right)}\right)\delta = \frac{1}{U_C\left(\theta^t\right)}\frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\delta$$

through this perturbation, which have to be weighted by the density $f\left(\theta_t|\theta^{t-1}\right)$. The LHS in the static optimality condition thus represents the planner's resource gain from this perturbation discounted by R^{-1}, or in other words, the marginal efficiency cost of distorting labor supply at θ^t.

The RHS describes how this resource gain can be re-distributed. The perturbation changes the marginal information rent $U_\theta\left(\theta^t\right)$ at θ^t by

$$\Delta U_\theta\left(\theta^t\right) = U_{\theta C}\left(\theta^t\right)\Delta C_1 + U_{\theta Y}\left(\theta^t\right)\Delta Y_1 = \left(\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} - \frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}\right)\delta.$$

From the single-crossing condition, $\Delta U_\theta\left(\theta^t\right) > 0$. Therefore, upon reducing the distortion at θ^t, the planner cannot freely re-distribute the extra resources across all types, but must raise the utility of all types $\theta' > \theta_t$ by an extra $\Delta U_\theta\left(\theta^t\right)$, relative to all types $\theta' < \theta_t$. Hence, at each θ^t, the planner faces a simple tradeoff between an efficiency motive on the LHS and a redistribution motive on the RHS: More redistribution around θ^t must come at the cost of lower efficiency at θ^t, and vice versa. I can thus re-state the static optimality condition as follows:

$$\frac{\mu_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)} = MC_t\left(\theta^t\right) \equiv R^{-1}\frac{\frac{1}{-U_Y\left(\theta^t\right)} - \frac{1}{U_C\left(\theta^t\right)}}{\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} - \frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}}.$$

In other words, the optimal allocation equates the normalized multiplier $\mu_t\left(\theta^t\right)/f\left(\theta_t|\theta^{t-1}\right)$, or *shadow price of redistribution*, to a static *marginal cost of redistribution at* θ that is given by the ratio of the resource loss of marginally increasing the tax distortion, to the reduction in marginal information rents and hence the increase in redistribution that this entails.

3 Incentive-adjusted Probability Measures

In this section, I introduce the key technical idea that the solution to the multipliers $\mu_t\left(\theta^t\right)$ and λ_t can be defined by means of a change in the probability measure that governs the Markov chain of types. This change captures the notion that incentive compatibility imposes restrictions on the planner's ability to redistribute utility from higher to lower types. Specifically, by combining the linear ODE for $\mu_t\left(\theta^t\right)$ with the first order conditions for consumption and earnings and the boundary conditions $\mu\left(\underline{\theta}\right) = \mu\left(\overline{\theta}\right) = 0$, I obtain two separate, but equivalent characterizations of $\mu_t\left(\theta^t\right)$ and λ_t, once based on $\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}$ and marginal utilities of consumption, and once based on $\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}$ and marginal disutilities of earnings. These characterizations are summarized below in Proposition 1:

Proposition 1 *: The multipliers* $\mu_t\left(\theta^t\right)$ *and* λ_t *can be represented in one of the following two (equivalent) forms:*

(i) In terms of marginal benefits of redistribution through consumption:

$$\frac{\mu_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)} = \widehat{MB}_t\left(\theta^t\right) + \hat{\varrho}_t\left(\theta^t\right)\eta_t, \text{ where}$$

$$\begin{aligned}
\widehat{MB}_t\left(\theta^t\right) &\equiv R^{-1}\frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)}\left\{\hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta',\theta^{t-1}\right)}|\theta' \geq \theta_t,\theta^{t-1}\right) - \hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta^t\right)}|\theta^{t-1}\right)\right\} \\
\hat{f}\left(\theta_t|\theta^{t-1}\right) &\equiv \frac{f\left(\theta_t|\theta^{t-1}\right)\hat{m}\left(\theta^t\right)}{\int_{\underline{\theta}}^{\overline{\theta}} f\left(\theta_t|\theta^{t-1}\right)\hat{m}\left(\theta^t\right)d\theta_t} \text{ with } \hat{m}\left(\theta^t\right) = e^{-\int_{\underline{\theta}}^{\overline{\theta}}\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}d\theta_t}
\end{aligned}$$

$$\hat{\varrho}_t\left(\theta^t\right) = \varrho_t\left(\theta^t\right) + \frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)}\left\{\hat{\mathbb{E}}\left(\varrho_t\left(\theta,\theta^{t-1}\right)\frac{U_{\theta C}\left(\theta,\theta^{t-1}\right)}{U_C\left(\theta,\theta^{t-1}\right)}|\theta \geq \theta_t,\theta^{t-1}\right) - \hat{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}|\theta^{t-1}\right)\right\}$$

(ii) In terms of marginal benefits of redistribution through earnings:

$$\frac{\mu_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)} = \widetilde{MB}_t\left(\theta^t\right) + \widetilde{\varrho}_t\left(\theta^t\right)\eta_t, \text{ where}$$

$$\begin{aligned}
\widetilde{MB}_t\left(\theta^t\right) &\equiv R^{-1}\frac{1-\widetilde{F}\left(\theta_t|\theta^{t-1}\right)}{\widetilde{f}\left(\theta_t|\theta^{t-1}\right)}\left\{\widetilde{\mathbb{E}}\left(\frac{1}{-U_Y\left(\theta',\theta^{t-1}\right)}|\theta' \geq \theta_t,\theta^{t-1}\right) - \widetilde{\mathbb{E}}\left(\frac{1}{-U_Y\left(\theta^t\right)}|\theta^{t-1}\right)\right\} \\
\widetilde{f}\left(\theta_t|\theta^{t-1}\right) &\equiv \frac{f\left(\theta_t|\theta^{t-1}\right)\widetilde{m}\left(\theta^t\right)}{\int_{\underline{\theta}}^{\overline{\theta}} f\left(\theta_t|\theta^{t-1}\right)\widetilde{m}\left(\theta^t\right)d\theta_t} \text{ with } \widetilde{m}\left(\theta^t\right) = e^{-\int_{\underline{\theta}}^{\overline{\theta}}\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}d\theta_t}
\end{aligned}$$

$$\widetilde{\varrho}_t\left(\theta^t\right) = \varrho_t\left(\theta^t\right) + \frac{1-\widetilde{F}\left(\theta_t|\theta^{t-1}\right)}{\widetilde{f}\left(\theta_t|\theta^{t-1}\right)}\left\{\widetilde{\mathbb{E}}\left(\varrho_t\left(\theta,\theta^{t-1}\right)\frac{U_{\theta Y}\left(\theta,\theta^{t-1}\right)}{U_Y\left(\theta,\theta^{t-1}\right)}|\theta \geq \theta_t,\theta^{t-1}\right) - \widetilde{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}|\theta^{t-1}\right)\right\}$$

In addition,

$$\begin{aligned}
\lambda_t &= R^{-1}\hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta^t\right)}|\theta^{t-1}\right) + \eta_t\hat{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}|\theta^{t-1}\right) \\
&= R^{-1}\widetilde{\mathbb{E}}\left(\frac{1}{-U_Y\left(\theta^t\right)}|\theta^{t-1}\right) + \eta_t\widetilde{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}|\theta^{t-1}\right)
\end{aligned}$$

In this proposition $\hat{\mathbb{E}}\left(\cdot\right)$ denotes the expectations operator associated with cdf $\hat{F}\left(\cdot\right)$ and $\widetilde{\mathbb{E}}\left(\cdot\right)$ the expectations operator associated with cdf $\widetilde{F}\left(\cdot\right)$. The main message from Proposition 1 is that non-separabity of preferences enters the characterization of multipliers by adjusting the probability measure over types from $F\left(\cdot\right)$ to $\hat{F}\left(\cdot\right)$ when $U_{\theta C} \neq 0$ or from $F\left(\cdot\right)$ to $\widetilde{F}\left(\cdot\right)$ when $U_{\theta Y} \neq 0$.

The direction of these incentive adjustments depends on the signs of $U_{\theta C}$ and $U_{\theta Y}$. Since $U_{\theta Y}/U_Y < \min\left\{0, U_{\theta C}/U_C\right\}$, $\widetilde{F}\left(\cdot|\theta^{t-1}\right)$ is first-order stochastically dominated by $F\left(\cdot|\theta^{t-1}\right)$ and

$\hat{F}\left(\cdot|\theta^{t-1}\right)$. In addition, $\hat{F}\left(\cdot|\theta^{t-1}\right)$ first-order stochastically dominates $F\left(\cdot|\theta^{t-1}\right)$ when $U_{\theta C} > 0$, and $F\left(\cdot|\theta^{t-1}\right)$ first-order stochastically dominates $\hat{F}\left(\cdot|\theta^{t-1}\right)$ when $U_{\theta C} < 0$. The characterization subsumes the benchmark case with separable preferences ($U_{\theta C} = 0$) as a special case in which $\hat{F}\left(\cdot|\theta^{t-1}\right) = F\left(\cdot|\theta^{t-1}\right)$, i.e. incentive-adjusted and original probability measures coincide.

The ranking of these distributions in terms of first-order dominance is naturally linked to the interpretation of $U_{\theta C}/U_C$ and $U_{\theta Y}/U_Y$ as needs- and ability-based redistribution motives. Since $U_{\theta Y}/U_Y < 0$, increasing utilities by reducing earnings lowers information rents, and thus allows for more redistribution of utility towards lower types. This explains why $\widetilde{F}\left(\cdot|\theta^{t-1}\right)$ overweighs the lower types relative to $F\left(\cdot|\theta^{t-1}\right)$.

Likewise, $\hat{F}\left(\cdot|\theta^{t-1}\right)$ is shifted towards higher (lower) types, whenever higher types have higher (lower) consumption needs. When higher types have lower consumption needs ($U_{\theta C} < 0$), increasing consumption reduces information rents and thus allows for more redistribution towards lower types. In contrast, when consumption needs are increasing in type, extra consumption increases information rents and reduces redistribution towards lower types, in which case $\hat{F}\left(\cdot|\theta^{t-1}\right)$ first-order stochastically dominates $F\left(\cdot|\theta^{t-1}\right)$.

Redistributive Perturbations: These changes in probability measures are naturally linked to a perturbation that identifies the *marginal benefit of redistribution* $\widehat{MB}_t\left(\theta^t\right)$ as the resource gain of a local utility transfer from $\theta' > \theta_t$ to $\theta' < \theta_t$ that preserves incentive compatibility and expected utility.[10] When $\hat{\varrho}_t\left(\theta^t\right)\eta_t = 0$, the optimal allocation reduces to a simple comparison of static costs and benefits of redistribution. When $\hat{\varrho}_t\left(\theta^t\right)\eta_t > 0$, the static marginal benefit is augmented by a backward-looking dynamic component.

Consider the following class of perturbations which transfer consumption from all agents with types $\theta' > \theta$ to all agents with types $\theta' < \theta$

$$\Delta C_1\left(\theta', \theta^{t-1}\right) = \begin{cases} \frac{1}{U_C(\theta', \theta^{t-1})} \cdot \Delta V\left(\theta', \theta^{t-1}\right) < 0 \text{ if } \theta' > \theta_t \\ \frac{1}{U_C(\theta', \theta^{t-1})} \cdot \Delta V\left(\theta', \theta^{t-1}\right) > 0 \text{ if } \theta' < \theta_t \end{cases}$$

This perturbation decreases realized utility by $-\Delta V\left(\theta', \theta^{t-1}\right)$ for all agents with types $\theta' > \theta$ and increases realized utility by $\Delta V\left(\theta', \theta^{t-1}\right)$ for all agents with types $\theta' < \theta$. Set $\Delta V\left(\theta', \theta^{t-1}\right)$ so that $\lim_{\theta' \uparrow \theta} \Delta V\left(\theta', \theta^{t-1}\right) = \hat{\zeta}_- > 0$ and $\lim_{\theta' \downarrow \theta} \Delta V\left(\theta', \theta^{t-1}\right) = \hat{\zeta}_+ < 0$, with $\hat{\zeta}_- - \hat{\zeta}_+ = \delta > 0$, i.e. the change in utility around θ is of size δ.

[10] See Brendon (2013) and Farhi and Werning (2012) for related perturbation arguments in static and dynamic Mirrlees models with separable preferences.

The perturbation changes marginal information rents $U_\theta\left(\theta',\theta^{t-1}\right)$ by

$$\Delta U_\theta\left(\theta',\theta^{t-1}\right)=\frac{U_{\theta C}\left(\theta',\theta^{t-1}\right)}{U_C\left(\theta',\theta^{t-1}\right)}\cdot\Delta V\left(\theta',\theta^{t-1}\right),$$

and it therefore preserves local incentive compatibility if and only if

$$\Delta V'\left(\theta',\theta^{t-1}\right)=\Delta U_\theta\left(\theta',\theta^{t-1}\right)=\frac{U_{\theta C}\left(\theta',\theta^{t-1}\right)}{U_C\left(\theta',\theta^{t-1}\right)}\Delta V\left(\theta',\theta^{t-1}\right).$$

The unique perturbation that preserves local IC for all $\theta'\neq\theta$ is then given by

$$\Delta V\left(\theta',\theta^{t-1}\right)=\begin{cases} e^{\int_\theta^{\theta'}\frac{U_{\theta C}\left(\theta'',\theta^{t-1}\right)}{U_C\left(\theta'',\theta^{t-1}\right)}d\theta''}\cdot\hat{\zeta}_+=\frac{\hat{m}\left(\theta',\theta^{t-1}\right)}{\hat{m}\left(\theta_t,\theta^{t-1}\right)}\hat{\zeta}_+<0 \text{ if } \theta'>\theta_t \\ e^{\int_\theta^{\theta'}\frac{U_{\theta C}\left(\theta'',\theta^{t-1}\right)}{U_C\left(\theta'',\theta^{t-1}\right)}d\theta''}\cdot\hat{\zeta}_-=\frac{\hat{m}\left(\theta',\theta^{t-1}\right)}{\hat{m}\left(\theta_t,\theta^{t-1}\right)}\hat{\zeta}_->0 \text{ if } \theta'<\theta_t\end{cases}$$

For θ' just above θ, the perturbation changes $U_\theta\left(\theta',\theta^{t-1}\right)$ by $\left(U_{\theta C}\left(\theta',\theta^{t-1}\right)/U_C\left(\theta',\theta^{t-1}\right)\right)\cdot\hat{\zeta}_+$, and this adjustment to the slope must be offered to all $\theta''>\theta'$ to restore local incentive compatibility at θ'. But these modifications for any $\theta''>\theta'$ then generate further adjustments for all $\theta'''>\theta''$, and so on. By re-weighting the utility perturbations according to $\hat{m}(\cdot)$, the perturbation preserves local incentive compatibility for all types: the ODE is solved by "integrating up" the cumulative utility changes for higher types that are required as a result of preserving local IC at all lower types.

I complete the perturbation by requiring that the net change to expected utility is 0:

$$0=\int_{\underline{\theta}}^{\overline{\theta}}\Delta V\left(\theta',\theta^{t-1}\right)f\left(\theta'|\theta^{t-1}\right)d\theta'=\int_\theta^{\overline{\theta}}\frac{\hat{m}\left(\theta'\right)}{\hat{m}\left(\theta\right)}\hat{\zeta}_+f\left(\theta'|\theta^{t-1}\right)d\theta'+\int_{\underline{\theta}}^{\theta}\frac{\hat{m}\left(\theta'\right)}{\hat{m}\left(\theta\right)}\hat{\zeta}_-f\left(\theta'|\theta^{t-1}\right)d\theta',$$

which yields $\hat{\zeta}_-=\delta\left(1-\hat{F}\left(\theta_t|\theta^{t-1}\right)\right)$ and $\hat{\zeta}_+=-\delta\hat{F}\left(\theta_t|\theta^{t-1}\right)$. The discounted resource gain of this perturbation is then given by

$$\begin{aligned} & -R^{-1}\int_{\underline{\theta}}^{\overline{\theta}}\Delta C_1\left(\theta',\theta^{t-1}\right)f\left(\theta'|\theta^{t-1}\right)d\theta' \\ = & \ \delta R^{-1}\frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{m}\left(\theta_t,\theta^{t-1}\right)}\left\{\hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta',\theta^{t-1}\right)}|\theta'\geq\theta_t,\theta^{t-1}\right)-\hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta^t\right)}|\theta^{t-1}\right)\right\} \\ = & \ \delta f\left(\theta_t|\theta^{t-1}\right)\widehat{MB}_t\left(\theta^t\right),\end{aligned}$$

which justifies the interpretation of $\widehat{MB}_t\left(\theta^t\right)$ as the marginal benefit of redistribution around θ^t via consumption, i.e. the marginal benefit of transfering consumption around θ^t while maintaining incentive compatibility and keeping expected utility unchanged. This transfer saves resources because higher types have lower marginal utilities.

In the benchmark with separable preferences ($U_{\theta C}(\cdot)=0$), this perturbation transfers utility uniformly from all types $\theta'>\theta_t$ to all type $\theta'<\theta_t$. With separable preferences, this uniform transfer of utilities does not affect local incentive compatibility constraints other than at θ_t, which explains why the original and incentive-adjusted probability measures coincide.

An analogous perturbation argument explains why a local IC and expected-utility preserving perturbation that redistributes leisure or earnings around θ must reweight states according to $\widetilde{m}(\cdot)$, resulting in the above expression $\widetilde{MB}_t\left(\theta^t\right)$ as the marginal benefit of redistribution via earnings.

Tying incentive adjustments to observables: The next proposition decomposes the changes in probability measures $\hat{m}(\cdot)$ and $\widetilde{m}(\cdot)$ into a risk component $U_C\left(\theta^t\right)$ or $-U_Y\left(\theta^t\right)$ that captures the redistribution motive, and an incentive component $\hat{M}\left(\theta^t\right)$ or $\widetilde{M}\left(\theta^t\right)$ that represents the adjustment to the marginal redistribution of resources: preserving incentive compatibility requires that on the margin consumption must be redistributed in proportion to $\hat{M}(\cdot)$ and earnings in proportion to $\widetilde{M}(\cdot)$. Furthermore, the latter only depends on the elasticities $\mathcal{E}_C$ and $\mathcal{E}_Y$ and the distribution of consumption and earnings.

Proposition 2 *The incentive adjustments $\hat{m}(\cdot)$ and $\widetilde{m}(\cdot)$ admit the following representations:*

(i) Consumption-based incentive adjustment: $\hat{m}\left(\theta^t\right)=U_C\left(\theta^t\right)\hat{M}\left(\theta^t\right)$, where

$$\hat{M}\left(\theta^t\right) \quad = \quad e^{\int_{\underline{\theta}}^{\theta_t}\mathcal{E}_C\left(\theta',\theta^{t-1}\right)dlnC\left(\theta',\theta^{t-1}\right)}$$

(ii) Earnings-based incentive adjustment: $\widetilde{m}\left(\theta^t\right)=-U_Y\left(\theta^t\right)\widetilde{M}\left(\theta^t\right)$, where

$$\widetilde{M}\left(\theta^t\right) \quad = \quad e^{-\int_{\underline{\theta}}^{\theta_t}\mathcal{E}_Y\left(\theta',\theta^{t-1}\right)dlnC\left(\theta',\theta^{t-1}\right)}$$

The incentive component $\hat{M}(\cdot)$ is increasing in θ_t so consumption must be redistributed regressively on the margin except when $\mathcal{E}_C=0$. Moreover, $\hat{M}(\cdot)$ only depends on the distribution of consumption and the elasticity $\mathcal{E}_C(\cdot)$, and when the latter is constant, $\hat{M}\left(\theta^t\right)=C\left(\theta^t\right)^{\mathcal{E}_C}$. On the other hand, $\widetilde{M}\left(\theta^t\right)$ is decreasing in θ_t, unless $\mathcal{E}_Y=0$, so the marginal incentive-compatible redistribution of utility via earnings is progressive. It only depends on the distribution of earnings and the earnings elasticity $\mathcal{E}_Y(\cdot)$, and when the latter is constant, $\widetilde{M}\left(\theta^t\right)=Y\left(\theta^t\right)^{-\mathcal{E}_Y}$.

The elasticities $\mathcal{E}_Y\left(\theta^t\right)$ and $\mathcal{E}_C\left(\theta^t\right)$ were defined in section 2. They determine how much the marginal rate of substitution responds to increases in consumption or earnings. These two parameters have natural counter-parts in terms of income and substitution effects of labor supply, which are fully discussed in a static application of this framework in Hellwig and Werquin (2021). Intuitively speaking, $1/\mathcal{E}_Y\left(\theta^t\right)$ represents an elasticity of labor supply to wages, holding consumption constant, and $\mathcal{E}_C\left(\theta^t\right)$ governs the relative strength of income and substitution effects.

When preferences are weakly separable ($U(C,Y;\theta)=\mathcal{U}\left(\gamma\left(C,\underline{C}(\theta)\right)-n(Y,\theta)\right)$), the risk and incentive components both include a factor $\gamma_C(\cdot)$ or $n_Y(\cdot)$ which then cancels out from the product:

$\hat{m}\left(\theta^t\right) = \mathcal{U}'\left(\theta^t\right)\chi\left(\theta^t\right)$ where $\chi\left(\theta^t\right) = e^{\int_{\underline{\theta}}^{\theta_t} \frac{\gamma_{C\underline{C}}\left(\theta',\theta^{t-1}\right)}{\gamma_C\left(\theta',\theta^{t-1}\right)} \underline{C}'\left(\theta',\theta^{t-1}\right) d\theta'}$ and $\widetilde{m}\left(\theta^t\right) = \mathcal{U}'\left(\theta^t\right) p\left(\theta^t\right)$ where $p\left(\theta^t\right) = e^{\int_{\underline{\theta}}^{\theta_t} \frac{n_{Y\theta}\left(\theta',\theta^{t-1}\right)}{n_Y\left(\theta',\theta^{t-1}\right)} d\theta'}$. Hence $\hat{m}\left(\theta^t\right)$ and $\widetilde{m}\left(\theta^t\right)$ decompose into a risk component from the curvature in $\mathcal{U}$ and an incentive component from the non-separabilities between the consumption utility $\gamma\left(C, \underline{C}\left(\cdot\right)\right)$ or effort $n\left(Y,\cdot\right)$ and type θ. When preferences are additively separable, then $\mathcal{U}'\left(\cdot\right) = \chi\left(\cdot\right) = 1$, and the risk and incentive components in the adjusted probability measure just offset each other, $\hat{M}\left(\theta^t\right) = 1/U_C\left(\theta^t\right)$. Also, with weakly separable preferences, the elasticity $\mathcal{E}_C\left(\cdot\right)$ only depends on the curvature of $\gamma_C\left(\cdot\right)$, while the elasticity $\mathcal{E}_Y\left(\cdot\right)$ only depends on the curvature in $n\left(Y,\theta\right)$, in which case $1/\mathcal{E}_Y\left(\cdot\right)$ represents the Frisch elasticity of labor supply.

4 Benchmark: separable preferences

Before developing the general characterization of optimal labor and savings wedges, it will be useful to pause and review the well-known benchmark with separable preferences that has been extensively analyzed in Golosov, Kocherlakota and Tsyvinski (2003), Farhi and Werning (2013), Golosov, Troshkin and Tsyvinski (2016) and summarized by Stantcheva (2020). This can be done simply by setting $U_{\theta C} = U_{CY} = 0$ in the characterization of multipliers given by Proposition 1.

Inverse Euler Equation: With separable preferences, the optimal savings wedge is given by the Inverse Euler Equation

$$U_C\left(\theta^t\right) = \beta R\left\{\mathbb{E}\left(\frac{1}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)\right\}^{-1} < \beta R\mathbb{E}\left(U_C\left(\theta^{t+1}\right)|\theta^t\right).$$

The incentive-compatibility and expected utility preserving perturbation in the previous section generalizes a well-known perturbation argument which explains why, with separable preferences, the returns to savings must be distributed in proportion to $1/U_C\left(\theta^{t+1}\right)$ and therefore the discounted shadow cost of utility at date $t+1$ equals $\lambda_{t+1} = R^{-1}\cdot\mathbb{E}\left(1/U_C\left(\theta^{t+1}\right)\right)$. This expression arises because the planner must redistribute resources at date $t+1$ in such a manner that utility changes are uniform across all types, so as to preserve incentive compatibility. As a result the resources from savings are distributed regressively: higher θ_{t+1} types enjoy higher returns from savings than lower θ_{t+1} types. This observation then leads to the Inverse Euler Equation (IEE) for optimal savings distortions, which augments the standard inter-temporal substitution channel by the wealth effect of savings on future incentives and redistribution. This effect reduces the social returns to savings relative to the private ones and is captured by the savings wedge implied by the IEE.

Smoothing and back-loading of labor taxes: The optimal labor wedge satisfies

$$\frac{R^{-1}}{U_C(\theta^t)}\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)} = \frac{\mu_t(\theta^t)}{f(\theta_t|\theta^{t-1})} = MC_t(\theta^t) = MB_t(\theta^t) + \beta R \varrho_t(\theta^t)\cdot MC_{t-1}(\theta^{t-1})$$

$$\begin{aligned}\text{where } MB_t(\theta^t) &\equiv R^{-1}\frac{1-F(\theta_t|\theta^{t-1})}{f(\theta_t|\theta^{t-1})}\left\{\mathbb{E}\left(\frac{1}{U_C(\theta',\theta^{t-1})}|\theta' \geq \theta_t,\theta^{t-1}\right) - \mathbb{E}\left(\frac{1}{U_C(\theta^t)}|\theta^{t-1}\right)\right\} \\ \varrho_t(\theta^t) &= \frac{1-F(\theta_t|\theta^{t-1})}{f(\theta_t|\theta^{t-1})}\mathbb{E}\left(\mathcal{J}\left(\theta',\theta^{t-1}\right)|\theta' \geq \theta_t,\theta^{t-1}\right) = \frac{\frac{\partial\left(1-F\left(\theta_t|\theta^{t-1}\right)\right)}{\partial\theta_{t-1}}}{f(\theta_t|\theta^{t-1})} > 0\end{aligned}$$

i.e. the planner's solution equates the static marginal cost of redistribution for each θ^t, to the static marginal benefit of redistribution around θ^t, $MB_t(\theta^t)$, and a backwards-looking dynamic marginal benefit $\beta R \varrho_t(\theta^t)\cdot MC_{t-1}(\theta^{t-1})$, which captures the marginal impact of prior incentive commitments on the shadow price of redistribution in the current period. These prior incentive commitments constitute a promise to limit future information rents. They appear in the recursive optimality condition through the term $\beta R \varrho_t(\theta^t)\cdot MC_{t-1}(\theta^{t-1})$, where $\varrho_t(\theta^t) > 0$ denotes the rate of decay in information rents. One can rewrite the incentive commitment constraint as

$$\Delta = \int \left(U_\theta(\theta^t) + \beta\Delta_t(\theta^t)\right)\varrho_t(\theta^t) f(\theta_t|\theta^{t-1})\, d\theta_t.$$

Therefore, $\varrho_t(\theta^t)$ describes the rate at which marginal information rents at θ^t feed back into incentive commitments at θ^{t-1}. $\varrho_t(\theta^t)$ naturally maps into the persistence of the type process at θ^{t-1}: If $\varrho_t(\theta^t) = 0$ for all $\theta_t|\theta^{t-1}$, $\Delta = 0$ and the recursive optimality condition reduces to its static counterpart, since there is no scope for dynamic incentive commitments. When types are perfectly persistent, $\varrho_t(\theta^t) = 1$ and information rents at θ^t are fully passed through to incentive commitments at θ^{t-1}.

The optimal labor distortion trades off the marginal cost of efficiency distortions at θ^t against the marginal redistributional gain. Consider a perturbation $\left(\Delta C_t(\theta^t), \Delta Y_t(\theta^t)\right)$ that marginally increases distortions at θ^t, but leaves a type θ^t agent's utility unchanged. The marginal efficiency loss associated with this perturbation is $MC_t(\theta^t)\cdot\delta$. The marginal benefit of redistribution at θ^t is $MB_t(\theta^t)\cdot\delta$. In addition, the same reduction in marginal information rents at θ^t also feeds back into marginal information rent at θ^{t-1} at a rate $\varrho_t(\theta^t)$. This reduction in marginal information rents can be transformed into a marginal efficiency gain at a rate $\beta MC_{t-1}(\theta^{t-1})$, which must in turn be adjusted to period t by the return R. Hence the dynamic benefit from additional redistribution at t is given by $\delta\cdot\varrho_t(\theta^t)\cdot\beta R\cdot MC_{t-1}(\theta^{t-1})$. Combining these terms yields the recursive characterization of $MC_t(\theta^t)$ given above.

The recursive optimality condition thus smoothes tax distortions optimally between θ^{t-1} and θ^t, where backloading distortions to θ^t generates additional marginal benefits from redistribution around θ^t. Solving the optimality condition backwards and expressing $MC_t\left(\theta^t\right)$ in terms of the labor tax $\tau_t\left(\theta^t\right)$ yields the following expression:

$$\frac{\tau_t\left(\theta^t\right)}{U_{\theta Y}\left(\theta^t\right)} = \sum_{\tau=0}^{t-1} (\beta R)^\tau \prod_{s=0}^{\tau-1} \varrho_{t-s}\left(\theta^{t-s}\right) MB_{t-\tau}\left(\theta^{t-\tau}\right)$$

i.e. at the optimum the current tax distortion is equal to a weighted sum of current and past marginal benefits of redistribution, with weights that take into account the impact of efficiency distortions at θ^t on prior information rents. Hence, the optimal cost-benefit trade-off in the dynamic model sets the cost of current tax distortions against the benefits of higher current and past redistribution: more redistribution in period t reduces prior information rents and thus allows for more redistribution in earlier periods without adversely affecting incentives. The incentive commitment thus constitutes a commitment to future taxes and redistribution to limit current information rents.

Alternative representations: The recursive optimality condition can be related in a straightforward manner to the characterizations provided by Golosov, Troshkin and Tsyvinski (2016) and Farhi and Werning (2013). Multiplying both sides of the optimality condition by $\frac{U_{Y\theta}(\theta^t)}{U_Y(\theta^t)}U_C\left(\theta^t\right)$ gives the characterization provided by Golosov, Troshkin and Tsyvinski (2016):

$$\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)} = \frac{U_{Y\theta}\left(\theta^t\right)}{U_Y\left(\theta^t\right)} RU_C\left(\theta^t\right) MB_t\left(\theta^t\right) + \beta R \varrho_t\left(\theta^t\right) \frac{U_{Y\theta}\left(\theta^t\right)/U_Y\left(\theta^t\right)}{U_{Y\theta}\left(\theta^{t-1}\right)/U_Y\left(\theta^{t-1}\right)} \frac{U_C\left(\theta^t\right)}{U_C\left(\theta^{t-1}\right)} \frac{\tau_{t-1}(\theta^{t-1})}{1-\tau_{t-1}(\theta^{t-1})}$$

Based on this expression, they show that the impact of past labor wedges on the current one increases with $U_C\left(\theta^t\right)$, and therefore decreases with the current type realization.

One obtains the characterization provided by Farhi and Werning (2013) by multiplying both sides of the recursive optimality condition by $\frac{U_{Y\theta}(\theta^t)}{-U_Y(\theta^t)}$ and then taking expectations:[11]

$$\begin{aligned} \mathbb{E}\left(\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)} \frac{1/U_C\left(\theta^t\right)}{\mathbb{E}(1/U_C\left(\theta^t\right)|\theta^{t-1})}|\theta^{t-1}\right) &= Cov\left(\int_{\underline{\theta}}^{\theta_t} \frac{U_{Y\theta}\left(\theta',\theta^{t-1}\right)}{-U_Y\left(\theta',\theta^{t-1}\right)}d\theta', \frac{1/U_C\left(\theta^t\right)}{\mathbb{E}(1/U_C\left(\theta^t\right)|\theta^{t-1})}|\theta^{t-1}\right) \\ &\quad + \mathcal{R}\left(\theta^{t-1}\right) \frac{\tau_{t-1}(\theta^{t-1})}{1-\tau_{t-1}(\theta^{t-1})} \\ \text{where } \mathcal{R}\left(\theta^{t-1}\right) &= \frac{\mathbb{E}(\varrho_t(\theta^t)U_{Y\theta}\left(\theta^t\right)/U_Y\left(\theta^t\right)|\theta^{t-1})}{U_{Y\theta}\left(\theta^{t-1}\right)/U_Y\left(\theta^{t-1}\right)}. \end{aligned}$$

Hence the recursive optimality condition takes expectations of future tax distortions and future marginal benefits of redistribution under an incentive-adjusted probability measure that re-weights

[11] Their representation is more general in that marginal costs and benefits can be multiplied with arbitrary weighting functions before taking expectations. Hence their result also implies the present one by applying a weighting function $\pi(\cdot)$ that places weight only on a singleton of types. The two characterizations are thus equivalent.

returns by $1/U_C\left(\theta^t\right)$, and labor taxes mean-revert with a rate $\mathcal{R}\left(\theta^{t-1}\right)$. The term $\frac{U_{Y\theta}\left(\theta^{t-1}\right)}{U_Y\left(\theta^{t-1}\right)}$ measures the impact of current labor taxes on current information rents, and $\mathbb{E}(\varrho_t(\theta^t)\frac{U_{Y\theta}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}|\theta^{t-1})$ measures the impact of expected future labor taxes on current information rents. Therefore $\mathcal{R}\left(\theta^{t-1}\right)$ trades off between the two in a such a way that a marginal transfer of tax distortions over time does not change current information rents and redistribution.

With isoelastic preferences ($U\left(C,Y;\theta\right)=\gamma\left(C\right)-\kappa\left(Y/A\left(\theta\right)\right)^{1+\mathcal{E}_Y}$) and $A\left(\theta\right)=e^{\phi\theta}$ with $\phi>0$, the expected marginal benefit of future redistribution is $\left(1+\mathcal{E}_Y\right)Cov\left(log\left(A\left(\theta_t\right)\right),\frac{\hat{M}(\theta^t)}{\mathbb{E}(\hat{M}(\theta^t)|\theta^{t-1})}|\theta^{t-1}\right)$, while labor taxes mean-revert with a rate $\mathcal{R}\left(\theta^{t-1}\right)=\mathbb{E}(\varrho_t(\theta^t)|\theta^{t-1})$ that only depends on the stochastic process of types.

Trading off efficiency and redistribution in this manner relies on the fact that an intervention at or around θ only affects incentives locally when preferences are separable. This assumption completely separates the incentive-compatible utility transfers across time from efficient redistribution within each period. Moreover, the characterization of optimal allocations remains incomplete since it provides a single optimality condition to characterize two allocation variables $\left(C_t\left(\theta^t\right),Y_t\left(\theta^t\right)\right)$ for each type. In the remainder of this paper I show how these two results generalize to non-separable preferences, and I complete the characterization of optimal allocations by also considering the benefits of redistribution via earnings or leisure.

5 Optimal Savings Wedges

I now discuss the implications of Proposition 1 for optimal savings wedges. Two changes appear in the expression for λ_t, relative to the separable benchmark: the use of incentive-adjusted probabilities, and the additional term $\eta_t\hat{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}|\theta^{t-1}\right)$ or $\eta_t\widetilde{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}|\theta^{t-1}\right)$.

The use of incentive-adjusted probabilities follows from the fact that marginal returns to savings must be redistributed in proportion to $\hat{M}\left(\cdot\right)$ for consumption and $\widetilde{M}\left(\cdot\right)$ for earnings, in order to preserve local incentive compatibility.

The additional term in the expression for λ_t identifies a novel feedback effect from savings to current marginal information rents. A marginal increase in savings at θ^{t-1}, which is redistributed at $\theta^t|\theta^{t-1}$ in proportion to $1/U_C\left(\theta^t\right)\cdot\hat{m}\left(\theta^t\right)\cdot\delta=\hat{M}\left(\theta^t\right)$ to preserve incentive compatibility, changes marginal information rents by an agent of type θ^t by $U_{\theta C}\left(\theta^t\right)/U_C\left(\theta^t\right)\cdot\hat{m}\left(\theta^t\right)\cdot\delta$, and as discussed above, this change in marginal information rents is passed through to θ^{t-1} at a rate $\varrho_t\left(\theta^t\right)$. Hence the term $\hat{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}|\theta^{t-1}\right)$ measures the feedback from future returns to savings into current information rents, and these information rents are "priced" by η_t, which represents the shadow cost of incentive commitments at θ^{t-1}. Along the same lines, $\widetilde{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}|\theta^{t-1}\right)$ represents the feedback from a labor deferral to current information rents.

Combining the dynamic optimality conditions with the characterization for λ_{t+1} and rearranging terms then yields the following expression for the planner's consumption-savings tradeoff:

$$\begin{aligned}\frac{1}{U_C\left(\theta^t\right)}+\frac{R\mu_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)}\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} &= \frac{R\psi_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)}\\ &= (\beta R)^{-1}\hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)+\frac{R\mu_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)}\hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)\frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)\end{aligned}$$

This expression generalizes the IEE to non-separable preferences. The terms $1/U_C\left(\theta^t\right)$ on the LHS and $(\beta R)^{-1}\hat{\mathbb{E}}\left(1/U_C\left(\theta^{t+1}\right)|\theta^t\right)$ on the RHS describe the inter-temporal tradeoff between saving resources at θ^t and redistributing the returns to savings in an incentive-compatible manner at date $t+1$. As before savings reduce future redistribution. The incentive-adjusted expectation amplifies the impact of savings on future incentives when $U_{\theta C}>0$ and dampens it when $U_{\theta C}<0$.

The two new terms that appear in this expression are related to how savings affect information rents and redistribution at θ^t. Savings affect current redistribution through (i) the direct effect of savings on redistribution at θ^t, and (ii) the feedback from future information rents at $\theta^{t+1}|\theta^t$ to redistribution at θ^t. These two effects determine together how savings affect current redistribution, which in turn increases or reduces the resources generated by savings at θ^t. They are both weighted by the normalized shadow price of redistribution at θ^t, $R\mu_t\left(\theta^t\right)/f\left(\theta_t|\theta^{t-1}\right)$.

The direct effect captures the fact that increasing savings and lowering consumption of type θ^t by $\delta/U_C\left(\theta^t\right)$ changes $U_\theta\left(\theta^t\right)$ and therefore redistribution around θ^t by $U_{\theta C}\left(\theta^t\right)/U_C\left(\theta^t\right)$. The resource gain (if $U_{\theta C}\left(\theta^t\right)>0$) or cost (if $U_{\theta C}\left(\theta^t\right)<0$) associated with this change in redistribution appears in the $U_{\theta C}\left(\theta^t\right)/U_C\left(\theta^t\right)$ term on the LHS.

However, this direct effect must be offset against the feedback of savings from future to current information rents, which is given by $\hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)\frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)$ on the RHS. This term inherits the sign of $U_{\theta C}(\cdot)$, i.e. the effect nets out when preferences are separable, it reduces current redistribution when consumption needs are increasing in type ($U_{\theta C}(\cdot)>0$) and increases current redistribution when consumption needs are decreasing in type ($U_{\theta C}(\cdot)<0$). This force works against the direct effect, and its strength depends on the decay of future information rents, $\varrho_{t+1}\left(\theta^{t+1}\right)$.

Theorem 1 shows how these effects combine to determine the overall savings wedge.

Theorem 1 *: An interior solution to the multi-period dynamic optimal taxation problem satisfies* *the following* ***Generalized Inverse Euler Equation*** *in consumption:*

$$U_C\left(\theta^t\right) = \beta R\left(1+\hat{s}_C\left(\theta^t\right)\right)\mathbb{E}\left(U_C\left(\theta^{t+1}\right)\frac{\hat{M}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{M}\left(\theta^{t+1}\right)|\theta^t\right)}|\theta^t\right)$$

where

$$\hat{s}_C\left(\theta^t\right) = \frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\frac{U_{\theta C}\left(\theta^t\right)/U_C\left(\theta^t\right) - \hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)\frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)}{U_{\theta C}\left(\theta^t\right)/U_C\left(\theta^t\right) - U_{\theta Y}\left(\theta^t\right)/U_Y\left(\theta^t\right)}.$$

The optimal type-contingent savings wedge $\hat{S}_C\left(\theta^{t+1}\right) \equiv \left(1+\hat{s}_C\left(\theta^t\right)\right)\frac{\hat{M}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{M}\left(\theta^{t+1}\right)|\theta^t\right)} - 1$ *has expectation* $\mathbb{E}\left(\hat{S}_C\left(\theta^{t+1}\right)|\theta^t\right) = \hat{s}_C\left(\theta^t\right)$. *Finally,* $\hat{s}_C\left(\theta^t\right) \gtreqless 0$ *if and only if*

$$\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} \gtreqless \hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)\frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)$$

Theorem 1 highlights the existence of a new intertemporal tradeoff between redistribution at date t and redistribution at date $t+1$. The dynamic Mirrlees problem augments the standard consumption-savings tradeoff in resources by an additional inter-temporal tradeoff between present and future redistribution that is internalized by the planner, but not by private savings decisions. This tradeoff in turn can lead to optimal savings taxes or subsidies, depending on whether the planner is more concerned about present or future redistribution, as well as by the direction in which savings taxes or subsidies facilitate redistribution.

The optimal type-contingent savings subsidy $\hat{S}_C\left(\theta^{t+1}\right)$ decomposes into a factor $1+\hat{s}_C\left(\theta^t\right)$ that internalizes the combined effect of savings on current redistribution, and a factor $\frac{\hat{M}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{M}\left(\theta^{t+1}\right)|\theta^t\right)}$ that redistributes future marginal returns regressively to preserve incentive compatibility at θ^{t+1}. This second factor is derived from the IC-preserving change in probability measure and averages to 1. It therefore doesn't affect the average savings subsidy, generalizing the result of zero expected wealth taxes from Kocherlakota (2005) to arbitrary non-separable preferences. The first factor internalizes the effect of savings on current redistribution. This factor determines whether savings should, on average, be taxed or subsidized.

Effect of savings on current redistribution: If types are not too persistent (the direct effect dominates), savings should be subsidized if consumption needs are increasing in type ($U_{\theta C} > 0$) and taxed if consumption needs are decreasing in type ($U_{\theta C} < 0$). These prescriptions are reversed if the indirect effect dominates. The magnitude of the wedge is decreasing in the degree of persistence of information rents, concretely the condition for $\hat{s}_C\left(\theta^t\right) \gtreqless 0$ say that savings should be subsidized (taxed) if the impact of savings on current information rents, $U_{\theta C}\left(\theta^t\right)/U_C\left(\theta^t\right)$, is larger than the expected feedback from future information rents, $\hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)\frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)$. This condition can be equivalently re-stated in terms of the marginal impact of the returns on information rents as

$$U_{\theta C}\left(\theta^t\right) \gtreqless \beta R\mathbb{E}\left(\varrho_{t+1}\left(\theta^{t+1}\right)U_{\theta C}\left(\theta^{t+1}\right)\frac{\hat{M}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{M}\left(\theta^{t+1}\right)|\theta^t\right)}|\theta^t\right).$$

The magnitude of the savings wedge thus relates to the current labor tax $\tau_t\left(\theta^t\right)/\left(1-\tau_t\left(\theta^t\right)\right)$, which is used to "price" redistribution at θ^t, and the expected rate of decay $\varrho_{t+1}\left(\theta^{t+1}\right)$ which governs the relative strength of the indirect effect. If types are highly persistent, one reverts to a "benchmark" in which it is optimal to leave savings decisions on average undistorted. In this case the direct and indirect effect on current redistribution just offset so savings have no impact on current information rents and redistribution. But this knife-edge result arises for reasons very different from Kocherlakota (2005), since it is based on two competing forces for savings taxes or subsidies just canceling each other.

Effect of savings on future incentives: The change of probability measure $\frac{\hat{M}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{M}\left(\theta^{t+1}\right)|\theta^t\right)}$ summarizes how returns to savings must be distributed to preserve incentive compatibility. Since $\hat{M}\left(\cdot\right)$ is increasing in θ_{t+1} unless $\mathcal{E}_C = 0$, the returns on savings are regressive, i.e. higher types must earn higher returns to preserve incentive compatibility. This in turn implies that $\mathbb{E}\left(U_C\left(\theta^{t+1}\right)\frac{\hat{M}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{M}\left(\theta^{t+1}\right)|\theta^t\right)}|\theta^t\right) < \mathbb{E}\left(U_C\left(\theta^{t+1}\right)|\theta^t\right)$, i.e. the state-contingent savings subsidy reduces incentives to save by increasing returns for high realizations of θ^{t+1}.

When preferences are separable, $\hat{M}\left(\theta^{t+1}\right) = 1/U_C\left(\theta^{t+1}\right)$, and the representation collapses to the standard IEE. In general, the adjustment to preserve incentive compatibility can be either more or less regressive, depending on whether $\hat{M}\left(\theta^{t+1}\right)$ is steeper or less steep than $1/U_C\left(\theta^{t+1}\right)$, or equivalently whether $U_{\theta C} > 0$ or $U_{\theta C} < 0$.

Finally, if $\mathcal{E}_C = 0$, the generalized IEE coincides with the normal Euler Equation and it is optimal not to distort savings. This case corresponds to a generalized form of GHH preferences.

With weakly separable preferences ($U\left(C, Y; \theta\right) = \mathcal{U}\left(\gamma\left(C, \underline{C}\left(\theta\right)\right) - n\left(Y, \theta\right)\right)$), the inverse Euler equation can be rewritten as

$$U_C\left(\theta^t\right) = \beta R\left(1 + \hat{s}_C\left(\theta^t\right)\right)\mathbb{E}\left(\mathcal{U}'\left(\theta^{t+1}\right)\cdot\chi\left(\theta^{t+1}\right)|\theta^t\right)\left\{\mathbb{E}\left(\frac{\chi\left(\theta^{t+1}\right)}{\gamma_C\left(\theta^{t+1}\right)}|\theta^t\right)\right\}^{-1}$$

This representation illustrates that the marginal returns to savings, after adjusting for incentive compatibility, depend on three elements: (i) the arithmetic expectation of the "outer" marginal utility $\mathcal{U}'\left(\cdot\right)$ which does not affect incentives, (ii) the harmonic expectation of the "inner" marginal utility $\gamma_C(\cdot)$ which does, and (iii) an incentive adjustment to the expectations given by $\chi(\cdot)$ which factors in the non-separability in γ. It is thus important to distinguish between non-separability coming from curvature in the aggregator function $\mathcal{U}(\cdot)$ which is immaterial for incentives and thus treated according to the standard Euler Equation, and curvature and non-separability in the inner function γ, which matters for incentives and is thus treated according to the Inverse Euler Equation, with an incentive adjustment for the non-separability in γ. The result is a "mixed" Euler Equation that combines elements of both the standard and the Inverse Euler Equation in one.

Optimal Social Mobility: Theorem 1 has direct implications for the question of optimal social mobility, originally posed in the work of Thomas and Worrall (1990) and Atkeson and Lucas (1992) who presented dynamic private information economies in which marginal utilities tended to infinity and consumption to 0 with probability 1 for almost all agents. To this end, I set $\beta R = 1$, so that long-run dynamics are not determined by discounting and returns. Then the optimal savings wedge implies the following about the process of Marginal utilities:

Corollary 1 *: Marginal utilities satisfy the following dynamics:*

(i) If $\hat{s}_C(\theta^t) \leq 0$ *for all* θ^t*, then* $\frac{1}{U_C(\theta^t)} \geq \hat{\mathbb{E}}\left(\frac{1}{U_C(\theta^{t+1})}|\theta^t\right)$ *and* $U_C(\theta^t) \leq \mathbb{E}\left(U_C(\theta^{t+1})|\theta^t\right)$ *for all* θ^t*.*

(ii) If $\hat{s}_C(\theta^t) > 0$ *for all* θ^t*, then* $\frac{1}{U_C(\theta^t)} < \hat{\mathbb{E}}\left(\frac{1}{U_C(\theta^{t+1})}|\theta^t\right)$ *for all* θ^t*.*

The first part of the corollary states that if it is optimal to always restrict savings, the inverse marginal utility process (and by extension, consumption) follow a super-martingale under the incentive-adjusted probability measure, and marginal utility follows a sub-martingal under the original measure. This in turn generalizes the force towards immiseration that is present when preferences are separable and inverse Marginal utilities follow a Martingale.

When it is instead optimal to subsidize savings, then the inequalities are reversed and inverse marginal utilities and consumption drift up under the incentive-adjusted probability measure. While this cannot be directly mapped into corresponding dynamics under the original measure, using the representation of $\hat{M}(\cdot)$ from proposition 2 when $\mathcal{E}_C$ is constant and preferences weakly separable yields

$$\mathbb{E}\left(\left(\frac{C(\theta^{t+1})}{C(\theta^t)}\right)^{\mathcal{E}_C}|\theta^t\right) \geq \mathbb{E}\left(\frac{\mathcal{U}'(\theta^{t+1})\cdot\chi(\theta^{t+1})}{\mathcal{U}'(\theta^t)\cdot\chi(\theta^t)}|\theta^t\right)$$

and when $\chi(\cdot)$ is increasing in θ^{t+1} and types are sufficiently mean-reverting, then it is optimal to have a savings subsidy and a consumption process that inherits the mean-reversion from $\mathcal{U}'(\theta^{t+1})\cdot\chi(\theta^{t+1})$, thus breaking the forces towards immiseration. As discussed for example by Farhi and

Werning (2007), the existing immiseration results rely on the fact that at the optimal allocation, temporary shocks to labor productivity have permanent effects on future marginal utilities as a consequence of smoothing incentives over time. Here the temporary co-movement of consumption needs or marginal utilities with types makes it optimal to build some "foregiveness" into optimal consumption allocations: more productive types also have higher consumption needs, which makes it optimal to front-load consumption increases for incentive reasons. This restores mean reversion in the long-run without differential discounting or altruism towards future generations as in Farhi and Werning (2007) or Phelan (2006).

Inter-temporal Earnings Wedge: Along the same lines, one can re-cast optimal inter-temporal allocations of effort or labor by a generalized IEE which yields the following expression:

$$-U_Y(\theta^t) = \beta R \left(1 + \widetilde{s}_Y(\theta^t)\right) \mathbb{E}\left(-U_Y\left(\theta^{t+1}\right) \frac{\widetilde{M}(\theta^{t+1})}{\mathbb{E}\left(\widetilde{M}(\theta^{t+1})|\theta^t\right)}|\theta^t\right)$$

where

$$\widetilde{s}_Y(\theta^t) = \tau_t(\theta^t)\frac{U_{\theta Y}(\theta^t)/U_Y(\theta^t) - \widetilde{\mathbb{E}}\left(\varrho_{t+1}(\theta^{t+1})\frac{U_{\theta Y}(\theta^{t+1})}{U_Y(\theta^{t+1})}|\theta^t\right)}{U_{\theta C}(\theta^t)/U_C(\theta^t) - U_{\theta Y}(\theta^t)/U_Y(\theta^t)}.$$

The generalized IEE for earnings augments the tradeoff in intertemporal allocation of earnings by a tradeoff between current and future redistribution that is internalized by the planner but not by private decisions. This second tradeoff decomposes into the effect on future incentives through the incentive-adjusted probability or returns $\widetilde{M}(\theta^{t+1})$, a direct effect on current redistribution, and the feedback from future to current redistribution. The latter two are captured by $\widetilde{s}_Y(\theta^t)$.

Since redistribution based on ability lowers current information rents ($U_{\theta Y}/U_Y < 0$), $\widetilde{M}(\cdot)$ is decreasing in θ_{t+1}. Therefore, if $-U_Y(\cdot)$ is decreasing in θ_{t+1},[12] $\mathbb{E}\left(-U_Y\left(\theta^{t+1}\right)\frac{\widetilde{M}(\theta^{t+1})}{\mathbb{E}\left(\widetilde{M}(\theta^{t+1})|\theta^t\right)}|\theta^t\right) > \mathbb{E}\left(-U_Y\left(\theta^{t+1}\right)|\theta^t\right)$, so the incentive adjustment favors backloading leisure, unless $n(Y,\theta)$ is linear in Y, or $\mathcal{E}_Y = 0$. On the other hand, if the direct effect dominates the indirect, then $\widetilde{s}_Y(\theta^t) < 0$ which implies that it is optimal to backload earnings, and subsidize current leisure. Since the latter dominates on average, earnings become mean-reverting at the optimal allocation.

[12]Under the weak separability assumption, $-U_Y(\cdot)$ is decreasing in θ_2 whenever $n_Y(\cdot)$ is decreasing in θ_2. This holds automatically whenever $Y(\cdot)$ is not too strongly increasing.

6 Optimal Labor Wedges

I now turn to the caracterization of optimal labor wedges. Combining Proposition 1 with the dynamic optimality conditions yields the following recursive characterization of static labor wedges:

$$MC_t\left(\theta^t\right) = \widehat{MB}_t\left(\theta^t\right) + \hat{\varrho}_t\left(\theta^t\right)\beta R \cdot MC_{t-1}\left(\theta^{t-1}\right) = \widetilde{MB}_t\left(\theta^t\right) + \tilde{\varrho}_t\left(\theta^t\right)\beta R \cdot MC_{t-1}\left(\theta^{t-1}\right)$$

The equation combines the recursive characterization of labor wedges from the case with separable preferences with the observation that the optimal allocation equates the marginal cost of efficiency distortions to the marginal benefit of redistribution through either consumption or earnings - and the two marginal benefits of redistribution must therefore also be equal to each other. In other words, the planner equates marginal costs and benefits of redistribution, and is indifferent between redistribution via consumption or via earnings at each realization of θ. The marginal benefits of redistribution include both a static and a dynamic component, which are based on the same arbitrage between smoothing tax distortions and backloading to internalize the future marginal benefit of redistribution as in the separable model.

Non-separability in preferences modifies the static marginal benefits $\widehat{MB}_t\left(\theta^t\right)$ and $\widetilde{MB}_t\left(\theta^t\right)$, and changes the rate of decay in information rents, $\hat{\varrho}_t\left(\theta^t\right)$ and $\tilde{\varrho}_t\left(\theta^t\right)$. The rate of decay of information rents need not be the same for redistribution via consumption or via leisure. Hence the optimal solution also trades off between current and past marginal benefits: if for example $\hat{\varrho}_t\left(\theta^t\right) > \tilde{\varrho}_t\left(\theta^t\right)$, then the optimal allocation will generate more redistribution via consumption than via leisure implying $\widehat{MB}_t\left(\theta^t\right) < \widetilde{MB}_t\left(\theta^t\right)$, and $\widehat{MB}_t\left(\theta^t\right)$ has more persistent effects on current taxes than $\widetilde{MB}_t\left(\theta^t\right)$. Solving the optimality condition backwards and expressing $MC_t\left(\theta^t\right)$ in terms of the labor tax $\tau_t\left(\theta^t\right)$ yields the following expressions:

$$\begin{aligned}
\frac{\tau_t\left(\theta^t\right)}{U_{\theta Y}\left(\theta^t\right) + \left(1-\tau_t\left(\theta^t\right)\right)U_{\theta C}\left(\theta^t\right)} &= \sum_{\tau=0}^{t-1}(\beta R)^{\tau}\prod_{s=0}^{\tau-1}\hat{\varrho}_{t-s}\left(\theta^{t-s}\right)\widehat{MB}_{t-\tau}\left(\theta^{t-\tau}\right) \\
&= \sum_{\tau=0}^{t-1}(\beta R)^{\tau}\prod_{s=0}^{\tau-1}\tilde{\varrho}_{t-s}\left(\theta^{t-s}\right)\widetilde{MB}_{t-\tau}\left(\theta^{t-\tau}\right).
\end{aligned}$$

Hence, how the non-separability affects the rate of decay of information rents is key to understanding the persistence of labor wedges and the respective impact of marginal benefits of redistribution via consumption and earnings.

Decay of information rents: $\hat{\varrho}_t\left(\theta^t\right)$ and $\tilde{\varrho}_t\left(\theta^t\right)$ can be decomposed into (i) the feedback $\varrho_t\left(\theta^t\right)$ from future distortions $MC_t\left(\theta^t\right)$ to information rents at θ^{t-1} and (ii) the feedback from $\widehat{MB}_t\left(\theta^t\right)$ or $\widetilde{MB}_t\left(\theta^t\right)$ to information rents at θ^{t-1}.

Consider the same perturbation as in section 3, which marginally increases distortions and redistribution at θ^t while preserving expected utility and incentive compatibility. As before this perturbation generates a static cost $MC_t\left(\theta^t\right)$ and benefit $\widehat{MB}_t\left(\theta^t\right)$, and $MC_t\left(\theta^t\right)$ feeds into marginal information rents at θ^{t-1} at a rate $\varrho_t\left(\theta^t\right)$. But in addition, $\widehat{MB}_t\left(\theta^t\right)$ also feeds back into information rents at θ^{t-1}. If consumption is redistributed around θ_t in proportion to $\hat{m}\left(\theta',\theta^{t-1}\right)/U_C\left(\theta',\theta^{t-1}\right)$ to preserve incentive compatibility and expected utility, then marginal information rents for $\theta' \neq \theta_t$ change in proportion to $\hat{m}\left(\theta',\theta^{t-1}\right)U_{\theta C}\left(\theta',\theta^{t-1}\right)/U_C\left(\theta',\theta^{t-1}\right)$, and these changes in turn feed back into information rents at θ^{t-1} at a rate $\varrho_t\left(\theta',\theta^{t-1}\right)$. Hence the second term in the expression of $\hat{\varrho}_t\left(\theta^t\right)$ represents the feedback from $\widehat{MB}_t\left(\theta^t\right)$ to information rents at θ^{t-1}. Along the same lines, the second term in $\widetilde{\varrho}_t\left(\theta^t\right)$ represents the feedback from $\widetilde{MB}_t\left(\theta^t\right)$ to information rents at θ^{t-1}.

The representation of $\hat{\varrho}_t\left(\theta^t\right)$ and $\widetilde{\varrho}_t\left(\theta^t\right)$ in Proposition 1 leads to the following comparative statics result:

Proposition 3 *: Comparative statics for $\hat{\varrho}_t\left(\theta^t\right)$ and $\widetilde{\varrho}_t\left(\theta^t\right)$:*

(i) $\hat{\varrho}_t\left(\theta^t\right) \gtreqless \varrho_t\left(\theta^t\right)$ if $\varrho_t\left(\theta,\theta^{t-1}\right)\frac{U_{\theta C}\left(\theta,\theta^{t-1}\right)}{U_C\left(\theta,\theta^{t-1}\right)}$ is increasing/constant/decreasing in θ.

(ii) $\widetilde{\varrho}_t\left(\theta^t\right) \gtreqless \varrho_t\left(\theta^t\right)$ if $\varrho_t\left(\theta,\theta^{t-1}\right)\frac{U_{\theta Y}\left(\theta,\theta^{t-1}\right)}{U_Y\left(\theta,\theta^{t-1}\right)}$ is increasing/constant/decreasing in θ.

(iii) $\hat{\varrho}_t\left(\theta^t\right)$ and $\widetilde{\varrho}_t\left(\theta^t\right)$ converge to $\varrho_t\left(\theta^t\right)$ as θ_t converges to $\overline{\theta}$ or $\underline{\theta}$.

(iv) $\hat{\varrho}_t\left(\theta^t\right) = \widetilde{\varrho}_t\left(\theta^t\right) = \varrho_t\left(\theta^t\right)$ in the special case where $\frac{\partial}{\partial\theta}\varrho_t\left(\theta,\theta^{t-1}\right) = 0$ for all θ and $U_{\theta C}/U_C$ and $U_{\theta Y}/U_Y$ are constant.

Proposition 3 shows how the persistence of labor taxes $\hat{\varrho}_t\left(\theta^t\right)$ or $\widetilde{\varrho}_t\left(\theta^t\right)$ differs from the decay of information rents $\varrho_t\left(\theta^t\right)$. The terms $\varrho_t\left(\theta^t\right)\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}$ and $\varrho_t\left(\theta^t\right)\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}$ measure the feedback from consumption or earnings changes at θ^t to marginal information rents at θ^{t-1}. If this feedback is increasing in θ_t, then the feedback from $\widehat{MB}_t\left(\theta^t\right)$ to information rents at θ^{t-1} is positive and results in an optimal persistence of labor taxes that is higher than the decay of information rents. If instead the feedback from consumption or earnings changes to prior information rents is decreasing in the type realization, then the feedback from $\widehat{MB}_t\left(\theta^t\right)$ to information rents at θ^{t-1} is negative and the optimal persistence of labor taxes that is less than the decay of information rents.

Two special cases deserve to be mentioned: first, at the top and the bottom of the distribution the optimal persistence of labor taxes converges to the decay of information rents since marginal benefits of redistribution vanish at the top and bottom of the distribution. Second, if the decay rate is uniform in θ_t ($\frac{\partial}{\partial\theta}\varrho_t\left(\theta,\theta^{t-1}\right) = 0$ for all θ), and $U_{\theta C}/U_C$ and $U_{\theta Y}/U_Y$ are constant, then the feedback is constant for all types, and therefore the optimal persistence of labor taxes is also uniform and equal to the decay of information rents. In these cases, future redistribution has no effect on current information rents, the optimal allocation satisfies $\widehat{MB}_t\left(\theta^t\right)=\widetilde{MB}_t\left(\theta^t\right)$ and $MC_t\left(\theta^t\right)=\widehat{MB}_t\left(\theta^t\right)+\varrho_t\left(\theta^t\right)\beta R\cdot MC_{t-1}\left(\theta^{t-1}\right)$, i.e. the planner' solution equalizes static marginal benefits of redistribution and the persistence of labor tax distortions only depends on the exogenous decay of information rents.

Marginal Benefits of Redistribution: The marginal benefits of redistribution can be re-written as $\widehat{MB}_t\left(\theta^t\right)=\frac{1-F\left(\theta_t|\theta^{t-1}\right)}{f\left(\theta_t|\theta^{t-1}\right)U_C\left(\theta^t\right)}\cdot\widehat{B}\left(\theta^t\right)$ and $\widetilde{MB}_t\left(\theta^t\right)=\frac{1-F\left(\theta_t|\theta^{t-1}\right)}{f\left(\theta_t|\theta^{t-1}\right)\left(-U_Y\left(\theta^t\right)\right)}\cdot\widetilde{B}\left(\theta^t\right)$, where

$$\begin{aligned}\widehat{B}\left(\theta^t\right) &= \mathbb{E}\left(\frac{\hat{M}\left(\theta'\right)}{\hat{M}\left(\theta_t\right)}|\theta'\geq\theta_t,\theta^{t-1}\right)-\mathbb{E}\left(\frac{\hat{M}\left(\theta'\right)}{\hat{M}\left(\theta_t\right)}|\theta^{t-1}\right)\frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{1-F\left(\theta_t|\theta^{t-1}\right)}\\ \widetilde{B}\left(\theta^t\right) &= \mathbb{E}\left(\frac{\widetilde{M}\left(\theta'\right)}{\widetilde{M}\left(\theta_t\right)}|\theta'\geq\theta_t,\theta^{t-1}\right)-\mathbb{E}\left(\frac{\widetilde{M}\left(\theta'\right)}{\widetilde{M}\left(\theta_t\right)}|\theta^{t-1}\right)\frac{1-\widetilde{F}\left(\theta_t|\theta^{t-1}\right)}{1-F\left(\theta_t|\theta^{t-1}\right)}\end{aligned}$$

When the two marginal benefits have equal persistence, the optimal labor wedge satisfies the following simple characterization:

Theorem 2 *: If $\hat{\varrho}_t\left(\theta^t\right)=\widetilde{\varrho}_t\left(\theta^t\right)=\varrho_t\left(\theta^t\right)$, then the optimal labor wedge satisfies*

$$\frac{-U_Y\left(\theta^t\right)}{U_C\left(\theta^t\right)}=1-\tau_t\left(\theta^t\right)=\frac{\widetilde{B}\left(\theta^t\right)}{\widehat{B}\left(\theta^t\right)}$$

If instead $\hat{\varrho}_t\left(\theta^t\right)>\widetilde{\varrho}_t\left(\theta^t\right)$, then $1-\tau_t\left(\theta^t\right)<\widetilde{B}\left(\theta^t\right)/\widehat{B}\left(\theta^t\right)$, while if $\hat{\varrho}_t\left(\theta^t\right)<\widetilde{\varrho}_t\left(\theta^t\right)$, then $1-\tau_t\left(\theta^t\right)>\widetilde{B}\left(\theta^t\right)/\widehat{B}\left(\theta^t\right)$.

Theorem 2 shows that with equal persistence, the optimal labor wedge corresponds to the ratio of the marginal benefit of redistribution through earnings to the marginal benefit of redistribution through consumption. This condition captures the basic intuition that the planner arbitrages between asking the high income types to work more vs. asking them to consume less, and on the margin the planner's solution must be indifferent between the two. When persistence is not equal, then the ratio of the two marginal benefits serves as an upper or lower bound, given that the planner's solution trades off more static redistribution against higher persistence.

A special case of the theorem arises when types are iid ($\varrho_t\left(\theta^t\right)=0$). In this case the optimal labor wedge only depends on static costs and benefits of redistribution, and the two static optimality conditions can be re-stated as follows:

$$\frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}=A\left(\theta^t\right)\cdot\widehat{B}\left(\theta^t\right)\ \text{ and }\ \tau_t\left(\theta^t\right)=A\left(\theta^t\right)\cdot\widetilde{B}\left(\theta^t\right)$$

$$\text{where } A\left(\theta^t\right) = \left(\frac{U_{C\theta}\left(\theta^t\right)}{U_C\left(\theta^t\right)}-\frac{U_{Y\theta}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}\right)\cdot\frac{1-F\left(\theta_t|\theta^{t-1}\right)}{\theta_t f\left(\theta_t|\theta^{t-1}\right)}$$

The first optimality condition restates and generalizes the well-known ABC formula of Diamond (1998) and Saez (2001) to the present model which allows for redistribution based on both needs and abilities. The second optimality condition complements the first by obtaining the equivalent representation from marginal benefits of redistribution via earnings. The two optimality conditions offer a new interpretation to this representation as equating the marginal efficiency cost of redistribution to the marginal benefit of redistribution via consumption, since the terms $\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)}/A\left(\theta^t\right)$ and $\tau_t\left(\theta^t\right)/A\left(\theta^t\right)$ map directly to the marginal cost of redistribution, rescaled by $U_C\left(\theta^t\right)$ or $-U_Y\left(\theta^t\right)$, and $\widehat{B}\left(\theta^t\right)$ and $\widetilde{B}\left(\theta^t\right)$ correspond to the re-scaled marginal benefits of redistribution via consumption or earnings.

By taking the ratio between the two marginal benefit terms I obtain the alternative representation given by Theorem 2, which allows me to unify the representation of optimal labor wedges in the static and dynamic Mirrlees model.

What's more, recall that in the iso-elastic case $\hat{M}(\cdot)$ and $\widetilde{M}(\cdot)$ directly derive from the distribution of consumption and income. Hence $\widehat{B}\left(\theta^t\right)$ and $\widetilde{B}\left(\theta^t\right)$ are robustly tied to observable statistics and estimates of the elasticities $\mathcal{E}_C(\cdot)$ and $\mathcal{E}_Y(\cdot)$ which relate to income and substitution effects of labor supply. In Hellwig and Werquin (2021), we tie this alternative representation of optimal income taxes to the sufficient statistics approach of Saez (2001).

The above representation also determines how optimal taxes and marginal costs and benefits of redistribution vary with primitive parameters. If $U(C,Y;\theta)=\mathcal{U}(\gamma(C,\underline{C}(\theta))-n(Y,\theta))$, then $\frac{U_{C\theta}}{U_C}=\frac{\mathcal{U}''}{(\mathcal{U}')^2}U_\theta+\frac{\chi'}{\chi}$ and $\frac{U_{Y\theta}}{U_Y}=\frac{\mathcal{U}''}{(\mathcal{U}')^2}U_\theta-\frac{-p'}{p}$, where $\frac{\chi'(\cdot)}{\chi(\cdot)}$ and $\frac{p'}{p}$ were defined in the text following proposition 2. Therefore fixing the allocation, an increase in $\frac{\chi'(\cdot)}{\chi(\cdot)}$ increases $\frac{U_{C\theta}}{U_C}$, making consumption needs more regressive or less progressive. An increase in $\frac{-p'(\cdot)}{p(\cdot)}$ makes productivities or disutilities of effort more sensitive to type, which makes $\frac{U_{Y\theta}}{U_Y}$ more negative. An increase in $\frac{-\mathcal{U}''}{\mathcal{U}'}$ increases the outer curvature or overall risk aversion and reduces both $\frac{U_{C\theta}}{U_C}$ and $\frac{U_{Y\theta}}{U_Y}$ by the same magnitude. I also consider comparative statics w.r.t. $\mathcal{E}_C(\cdot)$ and $\mathcal{E}_Y(\cdot)$ which only enter through their respective marginal benefits.

Proposition 4 *For a given allocation of consumption $C(\cdot)$ or earnings $Y(\cdot)$:*

(i) The (rescaled) marginal costs $A(\theta)^{-1}$ is decreasing in $\frac{\chi'(\cdot)}{\chi(\cdot)}$ and $\frac{-p'(\cdot)}{p(\cdot)}$, and independent of $\mathcal{E}_C(\cdot)$, $\mathcal{E}_Y(\cdot)$, and $\frac{-\mathcal{U}''}{\mathcal{U}'}$.

(ii) The (rescaled) marginal benefit of redistribution via consumption $\widehat{B}(\cdot)$ is increasing in $\mathcal{E}_C(\cdot)$ and $\frac{-\mathcal{U}''}{\mathcal{U}'}$, decreasing in $\frac{\chi'(\cdot)}{\chi(\cdot)}$, and independent of $\mathcal{E}_Y(\cdot)$ and $\frac{-p'(\cdot)}{p(\cdot)}$.

(iii) The (rescaled) marginal benefit of redistribution via earnings $\widetilde{B}(\cdot)$ is increasing in $\frac{-\mathcal{U}''}{\mathcal{U}'}$ and $\frac{-p'(\cdot)}{p(\cdot)}$, decreasing in $\mathcal{E}_Y(\cdot)$, and independent of $\mathcal{E}_C(\cdot)$ and $\frac{\chi'(\cdot)}{\chi(\cdot)}$.

Proposition 4 shows how non-separabilities in preferences alter the tradeoff between efficiency and redistribution. It translates the stochastic dominance ordering of $\hat{F}(\cdot)$, $F(\cdot)$ and $\widetilde{F}(\cdot)$, as well as the elasticities $\mathcal{E}_C(\cdot)$ and $\mathcal{E}_Y(\cdot)$ into comparative statics of re-scaled marginal costs and benefits of redistribution for a given type-contingent consumption or earnings profile. These comparative statics then identify welfare-improving changes to the optimal allocation that can be implemented through a combination of the perturbations that were presented above:

If $\mathcal{E}_C(\cdot)$ increases, the marginal benefits of redistributing consumption increase, and the optimal allocation therefore shifts towards higher labor taxes, more redistribution via consumption and less redistribution via earnings. If $\mathcal{E}_Y(\cdot)$ increases, the marginal benefits of redistributing earnings are reduced, so the optimal allocation shifts towards less redistribution of earnings and more redistribution of consumption, while also opting for labor taxes.[13]

If $\chi'(\cdot)/\chi(\cdot)$ increases, it is optimal to increase redistribution via earnings and compensate with a combination of higher labor taxes and/or less redistribution via consumption, in line with the intuition that an increase in $\chi'(\cdot)/\chi(\cdot)$ results in a less progressive or more regressive motive of redistribution based on consumption needs.

If $-p'(\cdot)/p(\cdot)$ increases, productivities or disutilities of effort are more dispersed so the gains from redistribution via earnings increase. It's then optimal to increase labor taxes and redistribution via earnings. Redistribution via consumption may increase or decrease depending on whether the combined effect of the tax increase and the increase in $a'(\cdot)/a(\cdot)$ results in higher or lower marginal costs of efficiency distortions.

If $-\mathcal{U}''/\mathcal{U}'$ increases, additional curvature in utility strengthens both redistribution motives. The planner gains from increasing labor taxes and redistribution via both consumption and earnings.

To recap, increases in $\mathcal{E}_C(\cdot)$, $-\mathcal{U}''/\mathcal{U}'$ and $-p'(\cdot)/p(\cdot)$ or a reduction in $\mathcal{E}_Y(\cdot)$ unambiguously shift the tradeoff between efficiency and redistribution towards higher taxes and more redistribution, but the effect of $\chi'(\cdot)/\chi(\cdot)$, which captures changes in the motive for redistribution based on consumption needs, is more subtle: An increase in $\chi'(\cdot)/\chi(\cdot)$ reduces both the marginal benefits of redistribution based on consumption and the marginal cost of efficiency distortions. This leads to more redistribution via earnings, but whether this is compensated by a reduction of redistribution via consumption or an increase in taxes, or a combination of both, is ambiguous and depends on the relative effect of $\chi'(\cdot)/\chi(\cdot)$ on the marginal cost of efficiency distortions and the marginal benefit of consumption-based redistribution.

[13]This comparative static of the labor supply elasticity $\mathcal{E}_Y(\cdot)$ holds $\frac{-p'(\cdot)}{p(\cdot)}$ constant. In practice, an increase in the Frisch elasticity would also change productivities $\frac{-p'(\cdot)}{p(\cdot)}$.

Optimal Persistence of Labor wedges: My final result links optimal tax-smoothing to the sign and magnitude of the savings wedge, which completes the connection between the two wedges. Theorem 3 generalizes the tax-smoothing representation of Farhi and Werning (2013) to non-separable preferences and provides an analogous representation based on redistribution through earnings.[14]

Theorem 3 *: The optimal labor wedge satisfies the following two recursive characterizations:*

(i) Based on marginal benefits of redistribution through consumption:

$$\hat{\mathbb{E}}\left(\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)}\frac{1/U_C\left(\theta^t\right)}{\hat{\mathbb{E}}(1/U_C\left(\theta^t\right)|\theta^{t-1})}|\theta^{t-1}\right) = \widehat{Cov}\left(log\left(\frac{\hat{m}(\theta^t)}{\widetilde{m}(\theta^t)}\right), \frac{1/U_C\left(\theta^t\right)}{\hat{\mathbb{E}}(1/U_C\left(\theta^t\right)|\theta^{t-1})}|\theta^{t-1}\right) + \frac{1}{1+\hat{s}_C(\theta^{t-1})}\hat{\mathcal{R}}\left(\theta^{t-1}\right)\frac{\tau_{t-1}(\theta^{t-1})}{1-\tau_{t-1}(\theta^{t-1})}$$

$$where\ \hat{\mathcal{R}}\left(\theta^{t-1}\right) = \frac{\hat{\mathbb{E}}\left(\hat{\varrho}_t\left(\theta^t\right)\cdot\left(\frac{U_{\theta C}\left(\theta^t\right)}{U_C(\theta^t)} - \frac{U_{\theta Y}\left(\theta^t\right)}{U_Y(\theta^t)}\right)|\theta^{t-1}\right)}{\frac{U_{\theta C}(\theta^{t-1})}{U_C(\theta^{t-1})} - \frac{U_{\theta Y}(\theta^{t-1})}{U_Y(\theta^{t-1})}}$$

(ii) Based on marginal benefits of redistribution through earnings:

$$\widetilde{\mathbb{E}}\left(\tau_t\left(\theta^t\right)\frac{1/\left(-U_Y\left(\theta^t\right)\right)}{\widetilde{\mathbb{E}}\left(1/\left(-U_Y\left(\theta^t\right)\right)|\theta^{t-1}\right)}|\theta^{t-1}\right) = \widetilde{Cov}\left(log\left(\frac{\hat{m}(\theta^t)}{\widetilde{m}(\theta^t)}\right), \frac{1/\left(-U_Y\left(\theta^t\right)\right)}{\widetilde{\mathbb{E}}\left(1/\left(-U_Y\left(\theta^t\right)\right)|\theta^{t-1}\right)}|\theta^{t-1}\right) + \frac{1}{1+\widetilde{s}_Y\left(\theta^{t-1}\right)}\widetilde{\mathcal{R}}\left(\theta^{t-1}\right)\tau_t(\theta^{t-1})$$

$$where\ \widetilde{\mathcal{R}}\left(\theta^{t-1}\right) = \frac{\widetilde{\mathbb{E}}\left(\widetilde{\varrho}_t(\theta^t)\cdot\left(\frac{U_{\theta C}\left(\theta^t\right)}{U_C(\theta^t)} - \frac{U_{\theta Y}\left(\theta^t\right)}{U_Y(\theta^t)}\right)|\theta^{t-1}\right)}{\frac{U_{\theta C}(\theta^{t-1})}{U_C(\theta^{t-1})} - \frac{U_{\theta Y}(\theta^{t-1})}{U_Y(\theta^{t-1})}}$$

This representation incorporates two changes relative to the benchmark with separable preferences. First, expected future tax distortions and marginal benefits of redistribution are computed using incentive-adjusted probability measures. Expected future tax distortions can be re-stated as $\hat{\mathbb{E}}\left(\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)}\frac{1/U_C\left(\theta^t\right)}{\hat{\mathbb{E}}(1/U_C(\theta^t)|\theta^{t-1})}|\theta^{t-1}\right) = \mathbb{E}\left(\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)}\frac{\hat{M}(\theta^t)}{\mathbb{E}(\hat{M}(\theta^t)|\theta^{t-1})}|\theta^{t-1}\right)$ and $\widetilde{\mathbb{E}}\left(\tau_t\left(\theta^t\right)\frac{1/\left(-U_Y\left(\theta^t\right)\right)}{\widetilde{\mathbb{E}}(1/(-U_Y(\theta^t))|\theta^{t-1})}|\theta^{t-1}\right) = \mathbb{E}\left(\tau_t\left(\theta^t\right)\frac{\widetilde{M}(\theta^t)}{\mathbb{E}\left(\widetilde{M}(\theta^t)|\theta^{t-1}\right)}|\theta^{t-1}\right)$, which re-weights the distortions according to the incentive-adjusted returns $\hat{M}(\theta^t)$ or $\widetilde{M}(\theta^t)$. The expected marginal benefits of redistribution can be re-stated along similar lines.

Second the persistence of the labor wedge is determined not just by the tradeoff between future and current information rents that is captured by $\hat{\mathcal{R}}\left(\theta^{t-1}\right)$ and $\widetilde{\mathcal{R}}\left(\theta^{t-1}\right)$, but also by the

[14]The dynamic tax formula of Golosov, Troshkin and Tsyvinski (2016) follows from the recursive optimality condition along the same lines as in the separable case.

optimal savings wedge $1+\hat{s}_C(\theta^{t-1})$ or $1+\widetilde{s}_Y\left(\theta^{t-1}\right)$. The savings wedge enters optimal tax-smoothing because the planner's solution discounts consumption between θ^t and θ^{t-1} by a factor $R\left(1+\hat{s}_C\left(\theta^{t-1}\right)\right)\frac{\hat{M}(\theta^t)}{\mathbb{E}\left(\hat{M}(\theta^t)|\theta^{t-1}\right)}$ and earnings by a factor $R\left(1+\widetilde{s}_Y\left(\theta^{t-1}\right)\right)\frac{\widetilde{M}(\theta^t)}{\mathbb{E}\left(\widetilde{M}(\theta^t)|\theta^{t-1}\right)}$, while the private return on savings is given by R. The planner's solution internalizes dynamic tradeoffs between redistribution at θ^t and θ^{t-1} that are not internalized by private savings decisions. The same wedge enters the resource tradeoff between current and future tax distortions. The savings wedge can thus either reinforce or dampen labor-tax smoothing: a savings tax increases the persistence of labor taxes, while a savings subsidy reduces it. Which one arises is determined as before by the net impact of savings on current information rents.

As before, the terms $\hat{\mathcal{R}}\left(\theta^{t-1}\right)$ and $\widetilde{\mathcal{R}}\left(\theta^{t-1}\right)$ smooth labor wedges inter-temporally in such a way that a marginal transfer of distortions leaves current information rents unchanged. As shown in proposition 3 labor wedges become more (less) persistent when feedback from future to current information rents is increasing (decreasing) in θ_t.

Through $1+\hat{s}_C\left(\theta^{t-1}\right)$ and $\hat{\mathcal{R}}\left(\theta^{t-1}\right)$, as well as $1+\widetilde{s}_Y\left(\theta^{t-1}\right)$ and $\widetilde{\mathcal{R}}\left(\theta^{t-1}\right)$, the persistence of labor taxes thus depends on the sign, level and slope of $U_{\theta C}/U_C$ and $U_{\theta Y}/U_Y$, which govern the strength and direction of the need-based and ability based redistribution motives.

7 Examples

In this section I present four examples to illustrate the different possibilities that the main results allow for: (i) with GHH preferences, the wealth effect on incentives disappears, and the characterization of savings and labor distortions are uniquely determined from static concerns about redistribution. (ii) With preferences that incorporate type dependent marginal utilities or consumption needs (subsistence consumption), the optimal savings tax or subsidy is linked to how consumption needs vary with type. (iii) In an extension with type-dependent volatility, the persistence of labor wedges becomes state-dependent.

1. GHH Preferences: Suppose that $U(C,Y;\theta)=\mathcal{U}\left(C-\kappa\left(Y/A\left(\theta\right)\right)^{1+\mathcal{E}_Y}\right)$, $A(\theta)=e^{\phi\theta}$ with $\phi>0$, and $\frac{-\mathcal{U}''(\cdot)}{\mathcal{U}'(\cdot)}\left(C-n\right)=\psi>0$. Then, the incentive-adjusted probability takes the form $\hat{m}\left(\theta^t\right)=\mathcal{U}'\left(\theta^t\right)$, and $\hat{M}(\theta^t)=1$, i.e. the wealth effect on incentives disappears and the adjustment only incorporates risk aversion due to $\mathcal{U}(\cdot)$. It is straight-forward to check that $\frac{U_{\theta C}}{U_C}-\frac{U_{\theta Y}}{U_Y}=\left(1+\mathcal{E}_Y\right)\phi>0$ and $\frac{U_{\theta C}}{U_C}=-Z\left(\theta^t\right)\psi\left(1+\mathcal{E}_Y\right)\phi$, where $Z\left(\theta^t\right)=n\left(\theta^t\right)/\left(C\left(\theta^t\right)-n\left(\theta^t\right)\right)$. Then the Generalized IEE reduces to a standard Euler equation with a savings wedge:

$$\mathcal{U}'\left(\theta^t\right)=\beta R\left(1+\hat{s}_C\left(\theta^t\right)\right)\mathbb{E}\left(\mathcal{U}'\left(\theta^{t+1}\right)|\theta^t\right),$$

where

$$\hat{s}_C(\theta^{t-1}) = -\frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\psi\left(Z\left(\theta^t\right) - \hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)Z\left(\theta^{t+1}\right)|\theta^t\right)\right).$$

If $Z\left(\theta^t\right) > \hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)Z\left(\theta^{t+1}\right)|\theta^t\right)$, it is optimal to tax savings. The recursive tax-smoothing equation yields

$$\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)} = (1+\mathcal{E}_Y)\,\phi\frac{\hat{F}\left(\theta_t|\theta^{t-1}\right)-F\left(\theta_t|\theta^{t-1}\right)}{f\left(\theta_t|\theta^{t-1}\right)} + \frac{\mathcal{U}'\left(\theta^t\right)}{\mathbb{E}\left(\mathcal{U}'\left(\theta^t\right)|\theta^{t-1}\right)}\frac{\hat{\varrho}_t\left(\theta^t\right)}{1+\hat{s}_C(\theta^{t-1})}\frac{\tau_{t-1}(\theta^{t-1})}{1-\tau_{t-1}(\theta^{t-1})},$$

where the optimal decay of information rents is

$$\hat{\varrho}_t\left(\theta^t\right) = \varrho_t\left(\theta^t\right) - \frac{(1+\mathcal{E}_Y)\,\phi\psi}{f\left(\theta_t|\theta^{t-1}\right)}\int_{\theta_t}^{\overline{\theta}}\left(\varrho_t\left(\theta',\theta^{t-1}\right)Z\left(\theta',\theta^{t-1}\right) - \hat{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)Z\left(\theta^t\right)|\theta^{t-1}\right)\right)\hat{f}\left(\theta'|\theta^{t-1}\right)d\theta'.$$

Taking expectations of the recursive tax-smoothing equation yields the Farhi-Werning representation where $\mathbb{E}\left(\frac{\hat{F}\left(\theta_t|\theta^{t-1}\right)-F\left(\theta_t|\theta^{t-1}\right)}{f\left(\theta_t|\theta^{t-1}\right)}|\theta^{t-1}\right) = -Cov\left(\theta_t, \frac{\mathcal{U}'\left(\theta^t\right)}{\mathbb{E}\left(\mathcal{U}'\left(\theta^t\right)|\theta^{t-1}\right)}|\theta^{t-1}\right)$.

As these expressions show, with GHH preferences the adjusted probability measure only includes a risk component $\frac{\mathcal{U}'\left(\theta^t\right)}{\mathbb{E}\left(\mathcal{U}'\left(\theta^t\right)|\theta^{t-1}\right)}$, but the incentive component $\hat{M}(\cdot)$ disappears. If types are not too persistent, risk aversion introduces a motive for taxing savings and making labor taxes more persistent than at the separable benchmark. If in addition $Z\left(\theta^t\right)$ is decreasing and $\varrho_t\left(\theta^t\right)$ independent of θ_t, the GHH preferences result in higher feedback of information rents and higher optimal persistence of the labor wedge.[15]

2. Isoelastic preferences: Suppose preferences take the form

$$U\left(C,Y;\theta\right) = \frac{1}{1-\mathcal{E}_C}\left(\frac{C}{\underline{C}\left(\theta\right)}\right)^{1-\mathcal{E}_C} - \kappa\left(Y/A\left(\theta\right)\right)^{1+\mathcal{E}_Y},$$

with $\mathcal{E}_C \geq 1$ and $\mathcal{E}_Y > 0$.[16] Suppose also that $A\left(\theta\right) = e^{\phi\theta}$, $\phi > 0$, and $\underline{C}\left(\theta\right) = \underline{C}e^{\Gamma\theta}$ where Γ denotes the sensitivity of consumption needs to types, and Γ can be positive or negative. It follows that $\hat{m}\left(\theta^t\right) = \underline{C}\left(\theta_t\right)^{\mathcal{E}_C-1}$, $\widetilde{m}\left(\theta^t\right) = A\left(\theta_t\right)^{-(1+\mathcal{E}_Y)}$, $\hat{M}(\theta^t) = C\left(\theta^t\right)^{\mathcal{E}_C}$, and $\widetilde{M}(\theta^t) = Y\left(\theta^t\right)^{-\mathcal{E}_Y}$. The Generalized Inverse Euler Equation take the form

$$C\left(\theta^t\right)^{-\mathcal{E}_C}\underline{C}\left(\theta_t\right)^{\mathcal{E}_C-1} = \beta R\left(1+\hat{s}_C\left(\theta^t\right)\right)\frac{\mathbb{E}\left(\underline{C}\left(\theta_{t+1}\right)^{\mathcal{E}_C-1}|\theta^t\right)}{\mathbb{E}\left(C\left(\theta^{t+1}\right)^{\mathcal{E}_C}|\theta^t\right)},$$

where the savings wedges satisfy

$$\hat{s}_C\left(\theta^t\right) = \frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\frac{(\mathcal{E}_C-1)\,\Gamma\left(1-\hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)|\theta^t\right)\right)}{(1+\mathcal{E}_Y)\,\phi+(\mathcal{E}_C-1)\,\Gamma}$$

[15] $Z\left(\theta^t\right)$ is decreasing in θ_t whenever $n\left(\theta^t\right)$ is decreasing in θ_t.

[16] It is possible to rewrite these preferences as $U\left(C,Y;\theta\right) = \underline{C}\left(\theta\right)^{\mathcal{E}_C-1}\left\{\frac{1}{1-\mathcal{E}_C}C^{1-\mathcal{E}_C} - \kappa\left(Y/\bar{A}\left(\theta\right)\right)^{1+\mathcal{E}_Y}\right\}$,where $\bar{A}\left(\theta\right) = \underline{C}\left(\theta\right)^{\frac{\mathcal{E}_C-1}{1+\mathcal{E}_Y}}A\left(\theta\right)$. Hence the iso-elastic preference model has an equivalent reinterpretation as including a shock to time preference rates.

Hence it is optimal to subsidize savings if $1 > \hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)|\theta^t\right)$, and to tax them otherwise. Rewriting the IEEs yields:

$$\mathbb{E}\left(\left(\frac{C\left(\theta^{t+1}\right)}{C\left(\theta^{t}\right)}\right)^{\mathcal{E}_C}|\theta^t\right) = \beta R\left(1+\hat{s}_C\left(\theta^t\right)\right)\mathbb{E}\left(\left(\frac{\underline{C}\left(\theta_{t+1}\right)}{\underline{C}\left(\theta_{t}\right)}\right)^{\mathcal{E}_C-1}|\theta^t\right),$$

Therefore expected consumption growth is an increasing function of the savings subsidy $\hat{s}_C\left(\theta^t\right)$ and the expected growth in consumption needs $\underline{C}\left(\theta_{t+1}\right)/\underline{C}\left(\theta_t\right)$. Then, if types are fully persistent ($\varrho_{t+1}\left(\theta^{t+1}\right)=1$) or consumption needs independent of type ($\Gamma = 0$), the savings subsidy is 0 and consumption growth is independent of the current type, resulting in divergence of consumption profiles. These dynamics are reinforced when $\Gamma < 0$, which leads to a savings tax that frontloads consumption and reduces future consumption growth. In addition consumption growth is increasing in θ_t resulting in even more divergence and polarization of consumption over time.

If instead $\Gamma > 0$ and $\hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)|\theta^t\right) < 1$, then it is optimal to subsidize savings and consumption needs introduce a force towards mean-reversion. The resulting consumption process is mean-reverting around a positive growth trend. With $\Gamma < 0$ on the other hand, the savings tax frontloads consumption and lowers average consumption growth. Furthermore, consumption growth is increasing in θ_t resulting in more divergence of consumption over time.

3. Subsistence consumption needs: Suppose preferences take the form

$$U\left(C,Y;\theta\right) = \frac{1}{1-\chi}\left(C-\underline{C}\left(\theta\right)\right)^{1-\chi} - \kappa\left(Y/A\left(\theta\right)\right)^{1+\mathcal{E}_Y},$$

with $\chi \geq 0$, $A\left(\theta\right) = e^{\phi\theta}$, $\phi > 0$, and $\underline{C}\left(\theta\right) = \underline{C}e^{\Gamma\theta}$ where Γ denotes as before the sensitivity of consumption needs to types, and Γ can be positive or negative. It then follows that $\frac{U_{\theta C}}{U_C} = \Gamma\left(\mathcal{E}_C\left(\theta^t\right)-\chi\right)$, $\frac{U_{\theta Y}}{-U_Y} = \left(1+\mathcal{E}_Y\right)\phi$, $\hat{m}\left(\theta^t\right) = e^{\Gamma\int_{\underline{\theta}}^{\theta}\left(\mathcal{E}_C\left(\theta',\theta^{t-1}\right)-\chi\right)d\theta'}$ and $\hat{M}(\theta^t) = e^{\int_{\underline{\theta}}^{\theta}\mathcal{E}_C\left(\theta',\theta^{t-1}\right)d\theta'}$, where $\mathcal{E}_C\left(\theta^t\right) = \chi\frac{C\left(\theta^t\right)}{C\left(\theta^t\right)-\underline{C}\left(\theta^t\right)}$. The Generalized IEE takes the form

$$\left(C\left(\theta^t\right)-\underline{C}\left(\theta^t\right)\right)^{-\chi} = \beta R\left(1+\hat{s}_C\left(\theta^t\right)\right)\left\{\hat{\mathbb{E}}\left(\left(C\left(\theta^{t+1}\right)-\underline{C}\left(\theta^{t+1}\right)\right)^{\chi}|\theta^t\right)\right\}^{-1}$$

where the savings wedge takes the form

$$\hat{s}_C\left(\theta^t\right) = \frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\Gamma\frac{\mathcal{E}_C\left(\theta^t\right)-\chi-\hat{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\left(\mathcal{E}_C\left(\theta^{t+1}\right)-\chi\right)|\theta^t\right)}{\left(1+\mathcal{E}_Y\right)\phi+\Gamma\left(\mathcal{E}_C\left(\theta^t\right)-\chi\right)}$$

Hence if $\Gamma\left(\mathcal{E}_C\left(\theta^t\right)-\chi-\hat{\mathbb{E}}\left(\varrho_t\left(\theta^t\right)\left(\mathcal{E}_C\left(\theta^{t+1}\right)-\chi\right)|\theta^t\right)\right) < 0$ it is optimal to tax savings, in the opposite case it is optimal to subsidize them.

Suppose that $C\left(\cdot\right)/\underline{C}\left(\cdot\right)$ is increasing in θ_t and hence $\mathcal{E}_C\left(\theta^t\right)$ is decreasing in θ_t. This holds whenever Γ is negative or at most small and positive, i.e. consumption needs are decreasing or not too strongly increasing in θ_t. Then, if $\Gamma < 0$, the persistence parameters satisfy $\hat{\varrho}_t\left(\theta^t\right) > \widetilde{\varrho}_t\left(\theta^t\right) = \varrho_t\left(\theta^t\right)$, and therefore $\widehat{MB}_t\left(\theta^t\right) < \widetilde{MB}_t\left(\theta^t\right)$, i.e. at the optimal allocation, marginal benefits of redistribution through consumption are smaller, but more persistent than marginal benefits of redistribution through earnings - the decreasing consumption needs strengthen the planner's motive to redistribute consumption towards lower types, along with a positive tax on savings. If instead $\Gamma > 0$, the opposite conclusion holds ($\hat{\varrho}_t\left(\theta^t\right) < \widetilde{\varrho}_t\left(\theta^t\right) = \varrho_t\left(\theta^t\right)$ and $\widehat{MB}_t\left(\theta^t\right) > \widetilde{MB}_t\left(\theta^t\right)$), i.e. the optimal allocation shifts towards more redistribution via earnings, and it becomes optimal to subsidize savings.

As in the case with GHH preferences, the endogenous persistence and the savings wedge are mutually reinforcing, leading to higher $\hat{\varrho}_t\left(\theta^t\right)$ and positive savings taxes when $\Gamma < 0$. But the two effects can also result in savings subsidies and less persistence than at the separable benchmark if $\Gamma > 0$ (consumption needs are increasing in type).

4. Type-dependent volatility: Suppose as in example 2 that preferences are isoelastic,

$$U\left(C, Y; \theta\right) = \frac{1}{1-\mathcal{E}_C}\left(\frac{C}{\underline{C}\left(\theta\right)}\right)^{1-\mathcal{E}_C} - \kappa\left(Y/A\left(\theta\right)\right)^{1+\mathcal{E}_Y},$$

with $\mathcal{E}_C \geq 1$, $A\left(\theta\right) = e^{\phi\theta}$, $\phi > 0$, and $\underline{C}\left(\theta\right) = \underline{C}e^{\Gamma\theta}$ where Γ can be positive or negative. Suppose that $\theta_t = \mu\left(\theta_{t-1}\right) + \sigma\left(\theta_{t-1}\right)v_t$, where v_t is iid over time, and $\mu'\left(\theta_t\right) \in [0,1]$. In this case, $\varrho_t\left(\theta^t\right)$ is no longer uniform for all θ_t, but instead

$$\varrho_t\left(\theta^t\right) = \mu'\left(\theta_{t-1}\right) + \sigma'\left(\theta_{t-1}\right)\frac{\theta_t - \mu\left(\theta_{t-1}\right)}{\sigma\left(\theta_{t-1}\right)}$$

Therefore, $\varrho_t\left(\theta^t\right)$ is increasing in θ_t if $\sigma'\left(\theta_{t-1}\right) > 0$ (higher types face higher uncertainty),, and $\varrho_t\left(\theta^t\right)$ is decreasing in θ_t if $\sigma'\left(\theta_{t-1}\right) < 0$ (higher types face less uncertainty). We then have

$$\hat{\varrho}_t\left(\theta^t\right) = \varrho_t\left(\theta^t\right) + \mathcal{E}_C\Gamma\frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)}\frac{\sigma'\left(\theta_{t-1}\right)}{\sigma\left(\theta_{t-1}\right)}\left\{\hat{\mathbb{E}}\left(\theta_t|\theta \geq \theta_t, \theta^{t-1}\right) - \hat{\mathbb{E}}\left(\theta_t|\theta^{t-1}\right)\right\}$$

$$\widetilde{\varrho}_t\left(\theta^t\right) = \varrho_t\left(\theta^t\right) + \left(1+\mathcal{E}_Y\right)\phi\frac{1-\widetilde{F}\left(\theta_t|\theta^{t-1}\right)}{\widetilde{f}\left(\theta_t|\theta^{t-1}\right)}\frac{\sigma'\left(\theta_{t-1}\right)}{\sigma\left(\theta_{t-1}\right)}\left\{\widetilde{\mathbb{E}}\left(\theta_t|\theta \geq \theta_t, \theta^{t-1}\right) - \widetilde{\mathbb{E}}\left(\theta_t|\theta^{t-1}\right)\right\}$$

Proposition 3 then implies that $\widetilde{\varrho}_t\left(\theta^t\right) < \varrho_t\left(\theta^t\right)$ if $\sigma'\left(\theta_{t-1}\right) < 0$ and $\widetilde{\varrho}_t\left(\theta^t\right) > \varrho_t\left(\theta^t\right)$ if $\sigma'\left(\theta_{t-1}\right) > 0$, while the sign of $\hat{\varrho}_t\left(\theta^t\right) - \varrho_t\left(\theta^t\right)$ depends on $\Gamma\sigma'\left(\theta_{t-1}\right)$. This example shows that type dependent volatilities also contribute the persistence of the labor wedge: the feedback of information rents is increasing with the uncertainty about the current type realization, which increases persistence in labor wedges when $\Gamma > 0$ and $\sigma'\left(\theta_{t-1}\right) > 0$.

8 Conclusion

This paper aims to provide a general analysis of optimal dynamic taxation with non-separable preferences, and explore to what extent the core insights from the existing literature survive or generalize. First, I have shown how to incorporate non-separability in the analysis by means of a simple change in probability measures that captures the need for regressive redistribution of resources, and/or either progressive or regressive redistribution of utilities, as a means to preserve incentive compatibility. This incentive adjustment depends on preference parameters such as risk aversion, labor supply elasticities and consumption needs.

Second, by applying this incentive adjustment to the optimality condition for redistribution through both consumption and earnings, I obtain a double representation of the optimal labor and savings wedges, which completes the existing representations that are solely based on redistribution through consumption. This double representation captures the basic intuition that the planner can transfer utility from high to low types by transfering either consumption or leisure, and at the optimal allocation, the planner is indifferent between the two.

Third, I have shown how non-separability generates feedback from future allocations to current information rents. This feedback generates a new inter-temporal tradeoff between current and future redistribution and tax distortions which in turn modifies the main results from the existing literature on optimal savings and labor tax distortions. Optimal savings taxes may be positive or negative, depending on how savings internalize a tradeoff between current and future redistribution. The dynamics of labor taxes inherit the same savings wedge that determines how the planner's solution discounts resources from one period to the next. Finally the optimal persistence of labor taxes also depends on how tax-smoothing feeds back into current information rents.

One stark benchmark result stands out: when preferences are "sufficiently" isoelastic and types are highly persistent then the rationale for savings taxes or subsidies again disappears, and labor taxes are highly persistent with an upwards drift that is independent of assumptions about agent's preferences. Whether or not the theoretical results presented here also provide a strong quantitative case for taxing or subsidizing savings then depends on whether this benchmark comes close to reality or not. Answering this question, or related ones that bring the present results closer to applications, is an important direction for future work.

Finally, my results do not address the issue of tax implementation. In Hellwig (2021), I propose an implementation using history-dependent labor taxes with no private savings, and show that private savings decisions do not matter for the implementation of the optimal (or more generally, any) incentive compatible allocation, but only serve to fulfill future tax obligations, extending the principle of Ricardian Equivalence to dynamic Mirrlees models. Even when savings are allowed the resulting labor and savings taxes need not match the wedges of the present characterization one-for-one: with dynamic information rents, the static labor supply decision includes a forward-looking element that alters the mapping from static labor wedges to tax implementations. Future work will have to explore the importance of these dynamic information rents for optimal tax design.

References

[1] Aguiar, Mark, and Eric Hurst (2005). "Consumption versus Expenditure." Journal of Political Economy, Vol 113(5), 919-948.

[2] Aguiar, Mark, and Eric Hurst (2007). "Life-Cycle Prices and Production." The American Economic Review, Vol 97(5), 1533-1559.

[3] Aguiar, Mark, Eric Hurst (2007), and Loukas Karabarbounis (2012). "Recent Developments in the Economics of Time Use." Annual Review of Economics, Vol 4(1), 373-397.

[4] Albanesi, Stefania, and Christopher Sleet (2006). "Dynamic Optimal Taxation With Private Information." The Review of Economic Studies, Vol 73(1), 1-30.

[5] Atkeson, Andrew, and Robert Lucas (1992). "On Efficient Distribution with Private Information." The Review of Economic Studies, Vol 59(3), 427-453.

[6] Benhabib, Jess, Richard Rogerson, and Randall Wright (1991). "Homework in Macroeconomics: Household Production and Aggregate Fluctuations." Journal of Political Economy, Vol 99(6), 1166-1187.

[7] Brendon, Charles (2013). "Efficiency, Equity, and Optimal Income Taxation", working paper, Cambridge University.

[8] Diamond, Peter (1998). "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." The American Economic Review, Vol 88(1), 83-95.

[9] Diamond, Peter and James Mirrlees (1978). "A Model of Social Insurance With Variable Retirement." Journal of Public Economics, Vol 10(3), 295-336.

[10] Farhi, Emmanuel and Ivan Werning (2007). "Inequality and Social Discounting." Journal of Political Economy, Vol 117(3), 365-402.

[11] Farhi, Emmanuel and Ivan Werning (2011). "Capital Taxation: Quantitative Explorations of the Inverse Euler Equation." Journal of Political Economy, Vol 120(3), 398-445.

[12] Farhi, Emmanuel and Ivan Werning (2013). "Insurance and Taxation over the Life Cycle." The Review of Economic Studies, Vol 80(2), 596-635.

[13] Golosov, Mikhail, Narayana Kocherlakota and Aleh Tsyvinski (2003). "Optimal Indirect and Capital Taxation." The Review of Economic Studies, Vol 70(3), 569-587.

[14] Golosov, Mikhail, Maxim Troshkin and Aleh Tsyvinski (2016). "Redistribution and Social Insurance." The American Economic Review, Vol 106(2), 359-386.

[15] Golosov, Mikhail, and Aleh Tsyvinski (2006). "Designing Optimal Disability Insurance: A Case for Asset Testing." Journal of Political Economy, Vol 114(2), 257-279.

[16] Grochulski, Borys, and Narayana Kocherlakota (2010). "Nonseparable Preferences and Optimal Social Security Systems." Journal of Economic Theory, Vol 145(6), 2055-2077.

[17] Hellwig, Martin (2007). "A Contribution to the Theory of Optimal Utilitarian Income Taxation." Journal of Public Economics, Vol 91(7-8), 1449-1477.

[18] Hellwig, Christian (2021). "Barro meets Mirrlees: Ricardian Equivalence in Dynamic Private Information Economies", working paper, Toulouse School of Economics.

[19] Hellwig, Christian, and Nicolas Werquin (2021). "Sufficient statistics for optimal income taxes in dynamic economies", work in progress, Toulouse School of Economics

[20] Kapicka, Marek (2013). "Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-order Approach." The Review of Economic Studies, Vol 80(3), 1027-1054.

[21] Kocherlakota, Narayana (2005). "Zero Expected Wealth Taxes: A Mirrlees Approach to Optimal Dynamic Taxation." Econometrica, Vol 73(5), 1587-1621.

[22] Koehne, Sebastian (2018). "On the Taxation of Durable Goods." International Economic Review, Vol 59(2), 825-857.

[23] Koehne, Sebastian and Moritz Kuhn (2015). "Optimal Taxation in a Habit Formation Economy." Journal of Public Economics, Vol 122, 31-39.

[24] Mirrlees, James A. (1971). "An Exploration in the Theory of Optimal Income Taxation." The Review of Economic Studies, Vol 38(2), 175-208.

[25] Pavan, Alessandro, Ilya Segal, and Juuso Toikka (2014). "Dynamic Mechanism Design; A Myersonian Approach." Econometrica, Vol 82(2), 601-653.

[26] Phelan, Christopher (2006). "Opportunity and Social Mobility." The Review of Economic Studies, Vol 73(2), 487-504.

[27] Rogerson, William (1985). "Repeated Moral Hazard." Econometrica, Vol 53(1), 69-76.

[28] Saez, Emmanuel (2001). "Using Elasticities to Derive Optimal Income Tax Rates." The Review of Economic Studies, Vol 68(1), 205-229.

[29] Stantcheva, Stefanie (2017). "Optimal Taxation and Human Capital Policies over the Life Cycle." Journal of Political Economy, Vol 125(6), 1931-1990.

[30] Stantcheva, Stefanie (2020). "Dynamic Taxation." Annual Review of Economics, Vol 12, 801-831.

[31] Thomas, Jonathan, and Tim Worrall (1990). "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem." Journal of Economic Theory, Vol 51(2), 367-390.

9 Appendix: Proofs

Proof of Proposition 1:

The ODE that characterizes $\mu_t\left(\theta^t\right)$ satisfies

$$\dot{\mu}_t\left(\theta^t\right)+\mu_t\left(\theta^t\right)\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}=\left(\lambda_t-\eta_t\mathcal{J}\left(\theta_t,\theta^{t-1}\right)-\frac{R^{-1}}{U_C\left(\theta^t\right)}\right)f\left(\theta_t|\theta^{t-1}\right)$$

along with boundary conditions $\mu_t\left(\underline{\theta}\right)=\mu_t\left(\overline{\theta}\right)=0$. Define $\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}=\frac{\hat{m}_{\theta_t}\left(\theta^t\right)}{\hat{m}\left(\theta^t\right)}$, or $\hat{m}\left(\theta^t\right)=e^{-\int_{\theta_t}^{\overline{\theta}}\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}d\theta_t}$. Substituting into the above ODE and integrating out yields

$$\mu_t\left(\overline{\theta}\right)\hat{m}\left(\overline{\theta}\right)-\mu_t\left(\theta^t\right)\hat{m}\left(\theta^t\right)=\int_{\theta_t}^{\overline{\theta}}\left(\lambda_t-\eta_t\mathcal{J}\left(\theta_t,\theta^{t-1}\right)-\frac{R^{-1}}{U_C\left(\theta^t\right)}\right)f\left(\theta_t|\theta^{t-1}\right)\hat{m}\left(\theta^t\right)d\theta_t,$$

or

$$\frac{\mu_t\left(\theta^t\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)}=\frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)}\left\{\hat{\mathbb{E}}\left(\frac{R^{-1}}{U_C\left(\theta',\theta^{t-1}\right)}|\theta'\geq\theta_t,\theta^{t-1}\right)+\eta_t\hat{\mathbb{E}}\left(\mathcal{J}\left(\theta_t,\theta^{t-1}\right)|\theta'\geq\theta_t,\theta^{t-1}\right)-\lambda_t\right\}.$$

From the boundary conditions it follows that

$$\int_{\underline{\theta}}^{\overline{\theta}} \left(\lambda_t - \eta_t \mathcal{J}\left(\theta_t, \theta^{t-1}\right) - \frac{R^{-1}}{U_C\left(\theta^t\right)}\right) f\left(\theta_t|\theta^{t-1}\right) d\theta_t = 0$$

or $\lambda_t = R^{-1}\hat{\mathbb{E}}\left(1/U_C\left(\theta^t\right)|\theta^{t-1}\right) + \eta_t \hat{\mathbb{E}}\left(\mathcal{J}\left(\theta_t, \theta^{t-1}\right)|\theta^{t-1}\right)$. Rewrite $\hat{\mathbb{E}}\left(\mathcal{J}\left(\theta_{t+1}, \theta^t\right)|\theta^t\right)$ as

$$\begin{aligned}
\hat{\mathbb{E}}\left(\mathcal{J}\left(\theta_{t+1}, \theta^t\right)|\theta^t\right) &= \int_{\underline{\theta}}^{\overline{\theta}} \mathcal{J}\left(\theta_{t+1}, \theta^t\right) \frac{\hat{m}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{m}\left(\theta^{t+1}\right)|\theta^t\right)} f\left(\theta_{t+1}|\theta^t\right) d\theta_{t+1} \\
&= \int_{\underline{\theta}}^{\overline{\theta}} \int_{\theta_{t+1}}^{\overline{\theta}} \mathcal{J}\left(\theta', \theta^t\right) f\left(\theta'|\theta^t\right) d\theta' \frac{\hat{m}'\left(\theta_{t+1}, \theta^t\right)}{\hat{m}\left(\theta_{t+1}, \theta^t\right)} \frac{\hat{m}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{m}\left(\theta^{t+1}\right)|\theta^t\right)} f\left(\theta_{t+1}|\theta^t\right) d\theta_{t+1} \\
&= \int_{\underline{\theta}}^{\overline{\theta}} \varrho_{t+1}\left(\theta^{t+1}\right) \frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)} \frac{\hat{m}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{m}\left(\theta^{t+1}\right)|\theta^t\right)} f\left(\theta_{t+1}|\theta^t\right) d\theta_{t+1} \\
&= \hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right) \frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)
\end{aligned}$$

Substituting into the solution for $\mu_t\left(\theta^t\right)$ and rearranging terms yields

$$\frac{\mu_t\left(\theta^t\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)} = R^{-1}\frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)}\left\{\hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta', \theta^{t-1}\right)}|\theta' \geq \theta_t, \theta^{t-1}\right) - \hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta^t\right)}|\theta^{t-1}\right)\right\} + \eta_t \hat{\varrho}_t\left(\theta^t\right)$$

where

$$\begin{aligned}
\hat{\varrho}_t\left(\theta^t\right) &= \frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)}\left\{\hat{\mathbb{E}}\left(\mathcal{J}\left(\theta_t, \theta^{t-1}\right)|\theta' \geq \theta_t, \theta^{t-1}\right) - \hat{\mathbb{E}}\left(\mathcal{J}\left(\theta_t, \theta^{t-1}\right)|\theta^{t-1}\right)\right\} \\
&= \frac{1}{\hat{f}\left(\theta_t|\theta^{t-1}\right)} \int_{\theta_t}^{\overline{\theta}} \mathcal{J}\left(\theta', \theta^{t-1}\right) \frac{\hat{m}'\left(\theta'\right)}{\mathbb{E}\left(\hat{m}\left(\theta^t\right)|\theta^{t-1}\right)} f\left(\theta'|\theta^{t-1}\right) d\theta' - \hat{\mathbb{E}}\left(\mathcal{J}\left(\theta_t, \theta^{t-1}\right)|\theta^{t-1}\right) \frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)} \\
&= \frac{1}{\hat{f}\left(\theta_t|\theta^{t-1}\right)} \int_{\theta_t}^{\overline{\theta}} \mathcal{J}\left(\theta', \theta^{t-1}\right) f\left(\theta'|\theta^{t-1}\right) d\theta \frac{\hat{m}\left(\theta_t\right)}{\mathbb{E}\left(\hat{m}\left(\theta^t\right)|\theta^{t-1}\right)} - \hat{\mathbb{E}}\left(\mathcal{J}\left(\theta_t, \theta^{t-1}\right)|\theta^{t-1}\right) \frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)} \\
&\quad + \frac{1}{\hat{f}\left(\theta_t|\theta^{t-1}\right)} \int_{\theta_t}^{\overline{\theta}} \int_{\theta}^{\overline{\theta}} \mathcal{J}\left(\theta', \theta^{t-1}\right) f\left(\theta'|\theta^{t-1}\right) d\theta' \frac{\hat{m}'\left(\theta\right)}{\mathbb{E}\left(\hat{m}\left(\theta^t\right)|\theta^{t-1}\right)} d\theta \\
&= \varrho_t\left(\theta^t\right) + \frac{1-\hat{F}\left(\theta_t|\theta^{t-1}\right)}{\hat{f}\left(\theta_t|\theta^{t-1}\right)}\left\{\hat{\mathbb{E}}\left(\varrho_t\left(\theta, \theta^{t-1}\right) \frac{U_{\theta C}\left(\theta, \theta^{t-1}\right)}{U_C\left(\theta, \theta^{t-1}\right)}|\theta \geq \theta_t, \theta^{t-1}\right) - \hat{\mathbb{E}}\left(\varrho_t\left(\theta^t\right) \frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}|\theta^{t-1}\right)\right\}
\end{aligned}$$

The same steps apply to the ODE

$$\dot{\mu}_t\left(\theta^t\right) + \mu_t\left(\theta^t\right) \frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)} = \left(\lambda_t - \eta_t \mathcal{J}\left(\theta_t, \theta^{t-1}\right) - \frac{R^{-1}}{U_Y\left(\theta^t\right)}\right) f\left(\theta_t|\theta^{t-1}\right)$$

to arrive at the characterization linking $\mu_t\left(\theta^t\right)$ to the inverse marginal utilities of earnings.

Proof of Proposition 2:

For Part (i), notice that

$$\begin{aligned}
\frac{\frac{d}{d\theta_t}U_C\left(\theta^t\right)}{U_C\left(\theta^t\right)} &= \frac{U_{C\theta}\left(\theta^t\right)}{U_C\left(\theta^t\right)} + \frac{U_{CC}\left(\theta^t\right)}{U_C\left(\theta^t\right)}C'\left(\theta^t\right) + \frac{U_{CY}\left(\theta^t\right)}{U_C\left(\theta^t\right)}Y'\left(\theta^t\right) \\
&= \frac{U_{C\theta}\left(\theta^t\right)}{U_C\left(\theta^t\right)} + \left(\frac{U_{CC}\left(\theta^t\right)}{U_C\left(\theta^t\right)} - \frac{U_{CY}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}\right)C'\left(\theta^t\right) \\
&= \frac{U_{C\theta}\left(\theta^t\right)}{U_C\left(\theta^t\right)} - \mathcal{E}_C\left(\theta^t\right)\frac{C'\left(\theta^t\right)}{C\left(\theta^t\right)}
\end{aligned}$$

where the second equality substituted the local IC constraint $U_C\left(\theta^t\right)C'\left(\theta^t\right) = -U_Y\left(\theta^t\right)Y'\left(\theta^t\right)$, the third equality the definition of $\mathcal{E}_C\left(\theta^t\right)$. Integrating out yields

$$\hat{m}\left(\theta^t\right) = KU_C\left(\theta^t\right)e^{\int_{\underline{\theta}}^{\theta_t}\mathcal{E}_C\left(\theta',\theta^{t-1}\right)\frac{C'\left(\theta',\theta^{t-1}\right)}{C\left(\theta',\theta^{t-1}\right)}d\theta'}$$

where K is a constant of proportionality that can be set equal to 1 as a normalization (given that the change in probability measure is defined by $\hat{m}\left(\theta\right)/\hat{\mathbb{E}}\left(\hat{m}\left(\theta\right)\right)$.

The steps for Part (ii) are identical: combining

$$\begin{aligned}
\frac{\frac{d}{d\theta_t}\left(-U_Y\left(\theta^t\right)\right)}{-U_Y\left(\theta^t\right)} &= \frac{U_{Y\theta}\left(\theta^t\right)}{U_Y\left(\theta^t\right)} + \frac{U_{YY}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}Y'\left(\theta\right) + \frac{U_{CY}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}C'\left(\theta^t\right) \\
&= \frac{U_{Y\theta}\left(\theta^t\right)}{U_Y\left(\theta^t\right)} + \left(\frac{U_{YY}\left(\theta^t\right)}{U_Y\left(\theta^t\right)} - \frac{U_{CY}\left(\theta^t\right)}{U_C\left(\theta^t\right)}\right)Y'\left(\theta^t\right) \\
&= \frac{U_{Y\theta}\left(\theta^t\right)}{U_Y\left(\theta^t\right)} + \mathcal{E}_Y\left(\theta^t\right)\frac{Y'\left(\theta^t\right)}{Y\left(\theta^t\right)}
\end{aligned}$$

from which it follows that $\widetilde{m}\left(\theta^t\right) = -U_Y\left(\theta^t\right)e^{-\int_{\underline{\theta}}^{\theta_t}\mathcal{E}_Y\left(\theta',\theta^{t-1}\right)\frac{Y'\left(\theta',\theta^{t-1}\right)}{Y\left(\theta',\theta^{t-1}\right)}d\theta'}$.

Proof of Theorem 1:

Rearranging the inter-temporal optimality condition and substituting

$$\frac{R\mu_t\left(\theta^t\right)}{f\left(\theta_t|\theta^{t-1}\right)} = \frac{1}{U_C\left(\theta\right)}\frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\left(\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} - \frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}\right)^{-1}$$

yields

$$\frac{1}{U_C\left(\theta\right)}\left(1 + \frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\frac{\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} - \hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)\frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)}{\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} - \frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}}\right) = \left(\beta R\right)^{-1}\hat{\mathbb{E}}\left(\frac{1}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)$$

It follows that $\hat{s}_C\left(\theta^t\right) \gtreqless 0$ if and only if $\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)} \gtreqless \hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)\frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)$. Re-arranging terms,

$$\begin{aligned}\hat{s}_C\left(\theta^t\right) &= \frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\frac{\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}-\hat{\mathbb{E}}\left(\varrho_{t+1}\left(\theta^{t+1}\right)\frac{U_{\theta C}\left(\theta^{t+1}\right)}{U_C\left(\theta^{t+1}\right)}|\theta^t\right)}{\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}-\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}} \\ &= \frac{\tau_t\left(\theta^t\right)}{1-\tau_t\left(\theta^t\right)}\frac{\frac{1}{U_C\left(\theta^t\right)}\left(U_{\theta C}\left(\theta^t\right)-\beta R\left(1+\hat{s}_C\left(\theta^t\right)\right)\mathbb{E}\left(\varrho_{t+1}\left(\theta^{t+1}\right)U_{\theta C}\left(\theta^{t+1}\right)\frac{\hat{M}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{M}\left(\theta^{t+1}\right)|\theta^t\right)}|\theta^t\right)\right)}{\frac{U_{\theta C}\left(\theta^t\right)}{U_C\left(\theta^t\right)}-\frac{U_{\theta Y}\left(\theta^t\right)}{U_Y\left(\theta^t\right)}}\end{aligned}$$

which implies that $\hat{s}_C\left(\theta^t\right) \gtreqless 0$ if and only if

$$U_{\theta C}\left(\theta^t\right) \gtreqless \beta R\mathbb{E}\left(\varrho_{t+1}\left(\theta^{t+1}\right)U_{\theta C}\left(\theta^{t+1}\right)\frac{\hat{M}\left(\theta^{t+1}\right)}{\mathbb{E}\left(\hat{M}\left(\theta^{t+1}\right)|\theta^t\right)}|\theta^t\right).$$

Proof of Proposition 3:

The result follows directly from the characterization of $\hat{\varrho}_t\left(\theta^t\right)$ and $\widetilde{\varrho}_t\left(\theta^t\right)$ in proposition 1.

Proof of Theorem 2:

From the representation of $\widehat{MB}_t\left(\theta^t\right)$ and $\widetilde{MB}_t\left(\theta^t\right)$ it follows that

$$\frac{\widetilde{MB}_t\left(\theta^t\right)}{\widehat{MB}_t\left(\theta^t\right)} = \frac{\widetilde{B}\left(\theta^t\right)}{\left(1-\tau_t\left(\theta^t\right)\right)\widehat{B}\left(\theta^t\right)}.$$

It follows from $\widehat{MB}_t\left(\theta^t\right)+\hat{\varrho}_t\left(\theta^t\right)\beta R\cdot MC_{t-1}\left(\theta^{t-1}\right)=\widetilde{MB}_t\left(\theta^t\right)+\widetilde{\varrho}_t\left(\theta^t\right)\beta R\cdot MC_{t-1}\left(\theta^{t-1}\right)$ that $\widehat{MB}_t\left(\theta^t\right) \gtreqless \widetilde{MB}_t\left(\theta^t\right)$ or equivalently $1-\tau_t\left(\theta^t\right) \gtreqless \frac{\widetilde{B}\left(\theta^t\right)}{\widehat{B}\left(\theta^t\right)}$ if and only if $\hat{\varrho}_t\left(\theta^t\right) \lesseqgtr \widetilde{\varrho}_t\left(\theta^t\right)$.

Proof of Proposition 4:

(i) Notice that $\widehat{MB}\left(\cdot\right)U_C\left(\cdot\right)$ can be rewritten as

$$R\widehat{MB}\left(\theta\right)U_C\left(\theta\right)=\frac{1}{f\left(\theta\right)}\left\{\hat{F}\left(\theta\right)\int_{\theta}^{\overline{\theta}}\frac{\hat{M}\left(\theta'\right)}{\hat{M}\left(\theta\right)}f\left(\theta'\right)d\theta'-\left(1-\hat{F}\left(\theta\right)\right)\int_{\underline{\theta}}^{\theta}\frac{\hat{M}\left(\theta'\right)}{\hat{M}\left(\theta\right)}f\left(\theta'\right)d\theta'\right\}$$

Since $\frac{\hat{M}(\theta')}{\hat{M}(\theta)}=e^{\int_{\theta}^{\theta'}\mathcal{E}_C(\theta'')dlnC(\theta'')}$, it follows that for a given allocation, $\hat{M}\left(\theta'\right)/\hat{M}\left(\theta\right)$ is increasing in θ and becomes steeper when $\mathcal{E}_C\left(\cdot\right)$ goes up, but does not change with $\mathcal{E}_Y\left(\cdot\right)$. Therefore, holding fixed $\hat{F}\left(\theta\right)$, $\widehat{MB}\left(\theta\right)U_C\left(\theta\right)$ is increasing with $\mathcal{E}_C\left(\cdot\right)$ and bounded below by $R\widehat{MB}\left(\theta\right)U_C\left(\theta\right) \geq \frac{\hat{F}(\theta)-F(\theta)}{f(\theta)}$.

In addition, $\hat{F}(\theta)$ does not depend on $\mathcal{E}_C(\cdot)$ or $\mathcal{E}_Y(\cdot)$, but an increase in χ'/χ or a reduction in $-\mathcal{U}''/\mathcal{U}'$ both increase $U_{\theta C}/U_C$, which is independent of a'/a. The results then follow by noting that an increase in $U_{\theta C}/U_C$ results in a FOSD shift in $\hat{F}(\cdot)$ which lowers $\widehat{MB}(\cdot)$.

(ii) Rewrite $\widetilde{MB}(\cdot)(-U_Y(\cdot))$ as

$$R\widetilde{MB}(\theta)(-U_Y(\theta)) = \frac{1}{f(\theta)}\left\{\widetilde{F}(\theta)\int_\theta^{\overline{\theta}} \frac{\widetilde{M}(\theta')}{\widetilde{M}(\theta)} f(\theta')\,d\theta' - \left(1-\widetilde{F}(\theta)\right)\int_{\underline{\theta}}^{\theta} \frac{\widetilde{M}(\theta')}{\widetilde{M}(\theta)} f(\theta')\,d\theta'\right\}$$

Since $\frac{\widetilde{M}(\theta')}{\widetilde{M}(\theta)} = e^{-\int_\theta^{\theta'} \mathcal{E}_Y(\theta'') dlnC(\theta'')}$, it follows that for a given allocation $\widetilde{M}(\theta')/\widetilde{M}(\theta)$ is decreasing in θ and becomes steeper when $\mathcal{E}_Y(\cdot)$ goes up, but does not change with $\mathcal{E}_C(\cdot)$. Therefore, holding fixed $\widetilde{F}(\theta)$, $\widetilde{MB}(\cdot)(-U_Y(\cdot))$ is decreasing with $\mathcal{E}_Y(\cdot)$ and bounded above by $R\widetilde{MB}(\cdot)(-U_Y(\cdot)) \le \frac{\widetilde{F}(\theta)-F(\theta)}{f(\theta)}$.

In addition, $\widetilde{F}(\theta)$ does not depend on $\mathcal{E}_C(\cdot)$ or $\mathcal{E}_Y(\cdot)$, but an increase in a'/a or a reduction in $-\mathcal{U}''/\mathcal{U}'$ both increase $U_{\theta Y}/U_Y$, which is independent of χ'/χ. The results then follow by noting that an increase in $U_{\theta Y}/U_Y$ results in a FOSD shift in $\widetilde{F}(\cdot)$ which lowers $\widetilde{MB}(\cdot)$.

(iii) Rewrite $\widehat{MB}(\cdot)U_C(\cdot)$ and $\widetilde{MB}(\cdot)(-U_Y(\cdot))$ as

$$R\widehat{MB}(\cdot)U_C(\cdot) = \frac{\frac{\tau(\theta)}{1-\tau(\theta)}}{U_{\theta C}/U_C - U_{\theta Y}/U_Y} = \frac{R\widetilde{MB}(\theta)(-U_Y(\theta))}{1-\tau(\theta)}$$

and which only depend on $U_{\theta C}/U_C - U_{\theta Y}/U_Y = \chi'/\chi - a'/a$.

Proof of Theorem 3:

Multiply both sides of $MC_t(\theta^t) = \widehat{MB}_t(\theta^t) + \hat{\varrho}_t(\theta^t)\beta R\cdot MC_{t-1}(\theta^{t-1})$ by $\frac{U_{\theta C}(\theta^t)}{U_C(\theta^t)} - \frac{U_{\theta Y}(\theta^t)}{U_Y(\theta^t)}$, divide by $\hat{\mathbb{E}}\left(1/U_C(\theta^t)|\theta^{t-1}\right)$, and take expectations with respect to the incentive-adjusted probability measure to find

$$\hat{\mathbb{E}}\left(\frac{\tau_t(\theta^t)}{1-\tau_t(\theta^t)}\frac{1/U_C(\theta^t)}{\hat{\mathbb{E}}(1/U_C(\theta^t)|\theta^{t-1})}|\theta^{t-1}\right) = \hat{\mathbb{E}}\left(\left(\frac{U_{\theta C}(\theta^t)}{U_C(\theta^t)} - \frac{U_{\theta Y}(\theta^t)}{U_Y(\theta^t)}\right)\frac{\widehat{MB}_t(\theta^t)}{\hat{\mathbb{E}}(1/U_C(\theta^t)|\theta^{t-1})}|\theta^{t-1}\right)$$
$$\frac{\beta R}{U_C(\theta^{t-1})}\frac{1}{\hat{\mathbb{E}}(1/U_C(\theta^t)|\theta^{t-1})}\cdot\hat{\mathcal{R}}(\theta^{t-1})\frac{\tau_{t-1}(\theta^{t-1})}{1-\tau_{t-1}(\theta^{t-1})}$$

By Theorem 1, $\frac{\beta R}{U_C(\theta^{t-1})}\frac{1}{\hat{\mathbb{E}}(1/U_C(\theta^t)|\theta^{t-1})} = \frac{1}{1+\hat{s}_C(\theta^{t-1})}$. At the same time,

$$\hat{\mathbb{E}}\left(\left(\frac{U_{\theta C}(\theta^t)}{U_C(\theta^t)} - \frac{U_{\theta Y}(\theta^t)}{U_Y(\theta^t)}\right)\frac{\widehat{MB}_t(\theta^t)}{\hat{\mathbb{E}}(1/U_C(\theta^t)|\theta^{t-1})}|\theta^{t-1}\right) = \hat{\mathbb{E}}\left(\left(\frac{\frac{\partial\hat{m}(\theta^t)}{\partial\theta_t}}{\hat{m}(\theta^t)} - \frac{\frac{\partial\widetilde{m}(\theta^t)}{\partial\theta_t}}{\widetilde{m}(\theta^t)}\right)\frac{\widehat{MB}_t(\theta^t)}{\hat{\mathbb{E}}(1/U_C(\theta^t)|\theta^{t-1})}|\theta^{t-1}\right)$$
$$= \widehat{Cov}\left(log\left(\frac{\hat{m}(\theta^t)}{\widetilde{m}(\theta^t)}\right), \frac{1/U_C(\theta^t)}{\hat{\mathbb{E}}(1/U_C(\theta^t)|\theta^{t-1})}|\theta^{t-1}\right)$$

Applying the same steps to $MC_t\left(\theta^t\right) = \widetilde{MB}_t\left(\theta^t\right) + \widetilde{\varrho}_t\left(\theta^t\right)\beta R \cdot MC_{t-1}\left(\theta^{t-1}\right)$ leads to the analogous representation based on redistribution through earnings.